Impact of Urban Neighborhood Disadvantage on Late Stage Breast Cancer Diagnosis in Virginia

Pam Baker DeGuzman · Wendy F. Cohn · Fabian Camacho ·
Brandy L. Edwards · Vanessa N. Sturz · Anneke T. Schroen

Published online: 13 March 2017
© The New York Academy of Medicine 2017

Abstract Research suggests that residents of inner-city urban neighborhoods have higher rates of late stage cancer diagnosis. Identifying urban neighborhoods with high rates of both concentrated disadvantage and late stage cancer diagnosis may assist health care providers to target screening interventions to reduce disparities. The purposes of this study were to (1) create an index to evaluate concentrated disadvantage (CD) using non-racial measures of poverty, (2) determine the impact of neighborhood CD on late stage breast cancer diagnosis in US cities, and (3) to understand the role of obesity on this relationship. We used census block group- (CBG) level poverty indicators from five Virginia cities to develop the index. Breast cancer cases of women aged 18–65 who lived in the five cities were identified from the 2000–2012 Virginia Cancer Registry. A logistic regression model with random intercept was used to evaluate the impact of disadvantage on late stage breast cancer diagnosis. CBG-level maps were developed to geographically identify neighborhoods with both high rates of CD and late breast cancer staging. Over 900

CBGs and 6000 breast cases were included. Global fit of the concentrated disadvantage model was acceptable. The effect of disadvantage on late stage was significant ($OR = 1.0083$, $p = 0.032$). Inner-city poverty impacts risk of late stage breast cancer diagnosis. Area-level obesity is highly correlated with neighborhood poverty ($\rho = 0.74$, $p < 0.0001$) but the mediating direct and indirect effects are non-significant. Intervening in these high poverty neighborhoods may help combat disparities in late stage diagnosis for urban poor and for minorities living in these underserved neighborhoods, but more study is needed to understanding the complex relationship between concentrated neighborhood poverty, obesity, and late stage diagnosis.

Keywords Neighborhood · Concentrated disadvantage · Poverty · Urban health · Breast cancer · Stage of diagnosis

US disparities in early breast cancer screening may be related to urban concentrated neighborhood poverty. Research has suggested that distance to screening facilities may be an important component of timely access to breast cancer screening [1], with rural women more likely to be diagnosed with late stage cancer than urban women [2]. However, research further suggests a nonlinear J-shaped curve, in which late cancer screening is associated with both those living at very long distances from screening locations, and those living very close to screening locations, in inner-city urban neighborhoods [3]. Although the actual distance to screening facilities may not be a factor, there is mounting evidence that living in a poor neighborhood may be an

P. B. DeGuzman (✉)
University of Virginia School of Nursing, PO Box 800782,
Charlottesville, VA 22908-0782, USA
e-mail: deguzman@virginia.edu

W. F. Cohn · F. Camacho · V. N. Sturz
Department of Public Health, University of Virginia School of
Medicine, Charlottesville, VA, USA

B. L. Edwards · A. T. Schroen
Department of Surgery, University of Virginia School of
Medicine, Charlottesville, VA, USA

important component of access to timely breast cancer screening, regardless of geographic distance from screening location [4, 5].

In urban US neighborhoods, clusters of poverty characteristics have been termed concentrated disadvantage [6]. Although concentrated disadvantage is often defined as areas with racial segregation and multiple poverty indicators, corollaries of high levels of segregation within poor urban neighborhoods [7] suggest that it may be the neighborhood-level socioeconomic characteristics, rather than the racial characteristics that are limiting access to health.

To date, studies that have attempted to evaluate the relationship between area poverty and late stage diagnosis have utilized zip code or census tract-level data [5, 8–10]. Zip code or zip code tabulation areas are area-level US Postal Divisions that are divided up based on postal delivery. Census tracts are small (average of 4000 people), relatively stable county subdivisions for which the US Census Bureau collects socioeconomic and demographic data, whereas Census block groups are even smaller (typically 600–3000 people), and the smallest statistical region for which area socioeconomic and demographic data are available [11]. Both census tract and census block group are preferable to zip code for identifying gradations in area health outcomes [8, 12]; however, census block group (CBG) population and health data are preferable for research that attempts to examine small area relationships, such as understanding how proximal neighborhood social and built environment conditions affect health [13]. In low-income urban areas of the USA, built and social environmental conditions can be closely related because neighborhood housing, access to care, race, and poverty characteristics are clustered in many highly segregated urban neighborhoods, and these neighborhood characteristics are related to health [14, 15]. Using cancer registry data identified at the CBG level, and CBG-level census data, the primary purpose of this study is to create an index to evaluate concentrated disadvantage using non-racial measures of poverty and use that index to determine the impact of neighborhood concentrated disadvantage on late stage breast cancer diagnosis in US cities.

Obesity has also been linked to late stage diagnosis [16–18] and has a strong association with area-level poverty [19–21]. In the context of neighborhood poverty, it is important to understand how obesity may affect this relationship. The mechanism through which obesity leads to later stage cancer detection may be through lower accuracy of diagnosis tests in obese patients [22], as well as

hypothesized lower functional status in these patients. Thus, a secondary purpose is to determine the mediating effect of obesity on the relationship between concentrated neighborhood disadvantage and breast cancer stage of diagnosis.

Methods

Study Area, Population, and Data

Breast cancer cases from the Virginia Cancer Registry (VCR) from 2000 to 2012 were identified. The VCR is a statewide registry of demographic, cancer diagnostic, and treatment data of individuals diagnosed or treated in Virginia or Virginia residents who received cancer care out of state [23]. The analysis included women aged 18–65 who lived in one of five cities in Virginia classified as a large central metropolitan counties by the National Center for Health Statistics' *Urban-Rural Classification Scheme for Counties*: Arlington, Alexandria, Norfolk, Richmond, or Virginia Beach City [24]. Virginia county geographic classification is unique in that major cities are considered independent of the counties that surround them, with independent local governing bodies. This separation allows the examination of urban neighborhoods without developing independent criteria for separation of the urban core from its surrounding county, as they may differ greatly from each other with respect to health [24].

Individual-Level Variables

Individual address was captured to develop the concentrated disadvantage index (CDI) for the first aim. To accomplish the second aim, we captured individual breast cancer stage from both in situ and invasive cases. Cancers diagnosed as regional or distant using registry-standard SEER summary staging variables were considered late stage; all others (in situ and localized) were considered to be diagnosed at an early stage. Other variables included in the analysis were race, insurance status, age at diagnosis, and year of diagnosis.

Data Analysis

Development of Neighborhood-Level Concentrated Disadvantage Index

A CBG-level disadvantage score (CDI) was created for each individual, based on home address, using a

longitudinal structural equation model (SEM) fit to the CBG data set with American Community Survey (ACS) socioeconomic status (SES) variables. For each CBG, the percent of female-headed households and percent of persons living below the poverty line, on public assistance, unemployed, and with less than a high school education were extracted from the ACS to develop the CDI. SAS software was used to develop the disadvantage factor (v. 9.4 SAS Institute Inc., Cary, NC, USA). We estimated the disadvantage factor by examining the latent factors at two time points (2000 and 2013), using CALIS regression scores.

CDI scores were calculated for 2000 and 2013 and compared using a nested model. Both time points were included in order to examine the SES variables for longitudinal measurement invariance, to assess how stable the disadvantage construct was over time. Disadvantage scores from both time points were expected to be very highly correlated, resulting in a near perfect rank ordering of census block groups using either of the two time points. We followed Widaman's recommendation that factorial invariance must hold across times of measurement [25]. Longitudinal invariance was defined as the case in which the conditional distributions of observed SES values were unchanging given the same latent variables values over time [26, 27]. Fit and details for the unrestricted model which relaxed longitudinal invariance and the nested model assuming longitudinal invariance are covered in the "Results" section. The nested model resulted in a significant worse fit and it was rejected in favor of the unrestricted model. Since the conditions for longitudinal measurement invariance were not met, a summary CDI score was created for each CBG by taking the average of factor scores at both time points. Each individual was assigned a disadvantage score based on home address.

Evaluation of Concentrated Disadvantage Impact on Late Stage Breast Cancer Rates

Choropleth maps at the CBG level were created for late diagnosis breast cancer rates and for SES CDI score using ArcMap mapping software (v. 10.2). Due to their adjacency, maps of Alexandria and Arlington cities are combined in Northern Virginia maps, and Norfolk and Virginia Beach City maps were combined in Norfolk maps. Because late stage rates at the census block group may be imprecise due to small sample size, the rates were adjusted using a spatial Empirical Bayes (EB)

smoother available in the GeoDa spatial analysis software [28, 29]. The EB smoother borrows information from local regions (for this application contiguous regions were chosen) and weighs the smoothed estimated more towards the regional mean as its variance increases due to small samples.

A spatial association between smoothed late diagnosis breast cancer rates and CDI score at a geographical level was then investigated using bivariate local indicators of spatial association (LiSA) [30]. This technique calculates a measure of association between one variable with another variable's average values in neighboring geographical units (chosen as contiguous regions) determined by a weighting matrix. Output from the Geoda software includes a scatter plot showing the correlation between standardized values of disadvantage at a location and the standardized late stage mean of the surrounding region. Additional plots include cluster maps showing combinations of high disadvantage values with high neighboring late stage rates, and combinations of low disadvantage with low rates, as well as high-low and low-high combinations. The cluster maps are color coded to identify the various combinations. The regression slope is shown in the scatter plot and is equivalent to Moran's I spatial autocorrelation statistic.

Using the Stata procedure *xtmelogit*, a logistic regression model with random intercept was fit to the data to assess the relationship between disadvantage and breast cancer late stage rates, adjusting for other proximal predictors of late stage. Covariates in the model consisted of individual-level patient data and the census block group disadvantage treated as a continuous variable. The random intercept was varied at the census block group level. Several extensions to this model were considered. The presence of a random disadvantage coefficient was tested with likelihood ratio tests and removed if no statistical evidence was found. Subgroup analysis was conducted by testing for the interactions between the CDI and each other confounder. Potential nonlinearity was assessed by entering disadvantage as a categorical variable as well as using likelihood ratio test comparing fit of the model with cubic smoothing spline to the model with only the linear trend. Finally, the average predictions for levels of the predictors were calculated using the estimated model by calculating late stage probabilities for each patient holding the level of interest at a particular counterfactual value and averaging over the data set of patients.

Determining Mediating Effects of Obesity

Individual body mass index (BMI) at the patient level was not available in our registry data, so obesity was included in this study by calculating ecological level of obesity, defined as BMI ≥ 30 kg/m², at the census tract level using the 2014 CDC 500-City project small area estimates 30 [31]. Some census tracts had missing data (18% of 337 census tracts); thus, missing BMI data were imputed using county-level Behavioral Risk Factor Surveillance System estimates.

The CDI score and area obesity estimates were highly correlated in the sample of 6225 patients ($\rho=0.74$, $p<0.0001$), so while we did not incorporate obesity into our final regression models to determine the effect of CDI on breast cancer staging rates, we evaluated the models including both CDI and obesity as predictors solely to guide the mediation analysis. In these models, both CDI and obesity were not significantly predictive of late stage diagnosis for either invasive cancers (CDI

$p=0.32$, obesity $p=0.38$) or for all cancers (CDI $p=0.31$, obesity $p=0.32$). We interpreted these findings using two plausible scenarios. In the first scenario, obesity is along a confounding pathway to late stage detection, in which case failure to adjust for obesity (as in our primary analysis) may lead to a biased effect estimate for CDI. In the second scenario, CDI is instead a driver of obesity rates, so that obesity can be seen as mediating the effect of CDI on late stage. In this case, adjusting for obesity may lead to over control of the effect.

We evaluated the second mediation scenario, using the formulas and SAS macro provided in Valeri [32], assuming the identification conditions described therein. This mediation analysis consists of estimating two simultaneous equations, a regression on the late stage outcome as a function of obesity, CDI, and fixed covariates (presented in Table 1) and then a regression on the mediator (obesity) as a function of CDI and Table 1 fixed covariates. The total effect of CDI as well as direct and indirect effects on

Table 1 Individual sample characteristics

	Invasive female breast cancers N= 4938	All female breast cancers N= 6225
Mean age	51.6 (8.8) [21,65]	51.6 (8.6) [21,65]
Age quartile groups	N (%)	N (%)
21–46	1328 (26.9)	1674 (26.9)
47–52	1139 (23.1)	1480 (23.8)
53–59	1379 (27.9)	1733 (27.8)
60–65	1092 (22.1)	1338 (21.5)
Race		
White	3171 (64.2)	3972 (63.8)
Black	1337 (27.1)	1702 (27.3)
Other	234 (4.7)	301 (4.8)
Hispanic	196 (4.0)	250 (4.0)
Insurance (SEER recode)		
Uninsured	294 (6.0)	348 (5.6)
Medicaid	239 (4.8)	282 (4.5)
Insured	4405 (89.2)	5595 (89.9)
Year of diagnosis		
2000–2003	1469 (29.8)	1782 (28.6)
2004–2007	1455 (29.5)	1835 (29.5)
2008–2012	2014 (40.8)	2608 (41.9)
Stage of diagnosis (SEER staging)		
Early (in situ)	—	1287 (20.7)
Early (local)	2848(57.7)	2848(45.8)
Late (regional)	1817(36.8)	1817(29.2)
Late (distant)	273 (5.5)	273 (4.4)
Concentrated disadvantage score	−1.1 (8.5)[−10.1,52.4]	−1.1(8.5)[−10.1,52.4]

late stage diagnosis were then derived. The model for late stage diagnosis was changed to a log-linear model instead of a logistic regression model, and the effects were interpreted on the risk ratio scale rather than the odds ratio scale in order to use available formulas for the total and direct effects with non-rare outcomes [32]. Obesity and CDI were included as continuous variables, and statistical interactions were also included if significant.

Results

Data from the five cities used to calculate the CDI scores were comprised of 928 CBGs, all of which were included in the analysis, as well as 10 variables, and 56 parameters, which included intercepts, loadings, and correlation coefficients. The path diagram resulting from the unrestricted SEM is shown in Fig. 1. The SEM expressed each manifest SES variable as a linear function of disadvantage latent factor, intercept, and residual term. Figure 1 additionally shows the standardized factor models and estimated correlation between factors for each time point. The figure does not show the residual terms, which were allowed to correlate across time points. Global fit of the model was considered acceptable, with standardized root mean square residual of 0.0162 (<0.10), root mean square

error of approximation of 0.0849 (<0.10), comparative fit index of 0.99 (>0.90), and non-normed fit index of 0.95 (>0.90). The variances for disadvantage factors were set to 1 to identify the model and expected factor score was set to zero.

A chi-square test comparing the nested model assuming longitudinal invariance to the unrestricted model resulted in a change-in-chi-square of 1007.73 with change-in-degrees-of freedom of 10, which was statistically significant ($p < 0.0001$), indicating worse fit. Measures which assess the practical significance of misfit also did not satisfy thresholds of good fit to accept invariance. A Cohen chi-square effect size measure was 0.33, interpreted as a medium effect size. A comparison between McDonald Centrality indices yielded a difference of 0.4026 (0.9681–0.5655), where a difference of >0.02 implies practical significance [33].

Figures 2, 3, and 4 map CDI scores and breast cancer late stage rates for each of the 928 CBGs in the study area. Figure 2 shows the empirically smoothed late stage cancer rate and the CDI scores by CBG for all female breast cancer cases in the dataset. Figure 3 shows clusters of high disadvantage and late stage rates, particularly located in southern and eastern Richmond. There is a net positive association between disadvantage and neighboring late stage rates when the three study regions are combined. The Moran Scatter plot (Fig. 4) shows a positive and highly significant correlation ($p = 0.001$)

Fig. 1 Structural equation model for census block group disadvantage ($N = 928$). Standardized factor loadings are shown. Residual terms affecting SES variables (not shown in figure) were made

to correlate across time points. To enable identification, the factor variances were set to 1. Factor means were set to zero

Fig. 2 Empirically smoothed late stage rate for all female breast cancers vs. Disadvantage Score. Choropleth maps to the *left* are for late stage rates (*darker areas* indicate higher late stage rates).

Choropleth maps on *right* are for disadvantage score (*darker areas* indicate greater disadvantage)

between CDI score and average neighboring late stage rates, with a regression slope of 0.30.

Table 1 shows the patient level characteristics for both invasive tumor cases and all female breast cancer cases. The mean age of women included was 52 years. Most (64%) individuals were White; 27% were Black. Most of

the cases had insurance (90%). Seventy-eight percent of all breast cancer cases were diagnosed at an early stage breast cancer compared to 57.7% of early diagnoses among invasive-only cancer cases.

Table 2 displays characteristics by quartiles of the CDI scores. A noted monotonic increase across quartiles is

Fig. 3 Bivariate LiSa cluster map: disadvantage values and late stage rates. Low-high is low disadvantage values with high late stage rates; high-low is high disadvantage values with low late stage rates

observed for all disparity indicators, with relative risks compared to Q1 ranging from 1.5 to 3.3 for Q2, 2.5 to 7.6 for Q3, and 5.1 to 26 for Q4. Furthermore, there was a late stage trend increase from 34 to 49% for invasive cancers and from 28 to 39% for all cancers. Tests of trends using logistic regression were statistically significant ($OR=1.02$, $p<0.001$) for CDI score association with invasive cancers and identical for all cancers ($OR=1.02$, $p<0.001$; Hosmer Ledeshow test of fit >0.05). In addition to those neighborhood characteristics shown, the percent

of African American in 2000 and 2013 was fairly stable (26.9 and 28.4%) in all CBGs, and increased monotonically by quartile as disadvantage increased (2000 ACS: 4.9, 11.7, 26.0, 63.1; 2013 ACS: 4.8, 12.4, 25.6, 69.1).

Table 3 shows results from the random intercept model predicting late stage rates among invasive cancers. Younger women had higher rates compared to older women (groups 21–46, compared to 47–52, 53–59, 65+ had $OR=0.72$, 0.68 , 0.61 , $p<0.001$); Blacks had higher rates compared to Whites ($OR=1.32$, $p=0.001$), and cases with insurance had lower rates compared to cases with no insurance ($OR=0.73$, $p=0.011$) as well as cases with Medicaid ($OR=0.58$, $p<0.001$). Effect of disadvantage on late stage was attenuated by control for the other variables and approached significance ($OR=1.0082$, $p=0.052$). The increase of adjusted predictions by unit of CDI score, approximating with a linear trend, was 0.19% percentile points compared to 0.58%, unadjusted. Inclusion of individual-level race into the model was the main contributor for the decrease. No interactions of predictors in the model with CDI score were found to be significant and thus were not included into the model.

Table 4 shows results from the random intercept model predicting late stage rates among all cancers. Younger aged women had higher rates compared to older aged women (21–46 vs. 47–52, 53–59, 65+ had $OR=0.72$, 0.71 , 0.68 , $p<0.001$); diagnosis from 2004 to 2007 had higher rates compared to 2008–2012 ($OR=1.18$, $p=0.012$); Blacks had higher rates compared to Whites ($OR=1.20$, $p=0.001$); and

Fig. 4 Bivariate LiSa scatter plot

Table 2 Disadvantage score quartiles by block group characteristics ($N=918$ census block groups)

	Q1 -10.6, -7.1	Q2 -7.1, -3.3	Q3 -3.3, 3.0	Q4 3.0, 52.4	Overall
Mean	-8.0	-5.3	-0.6	13.1	-0.02
ACS 2000					
% Single female households	3.5	8.2	14.1	28.6	14.0
% Poverty	1.2	3.1	9.1	26.4	10.0
% Public assistance	0.4	0.7	2.0	5.8	2.2
% Unemployed	2.0	4.5	8.2	15.3	7.7
% No high school	1.9	5.0	10.8	24.2	10.8
ACS 2013					
% Single female households	7.5	13.7	21.9	45.4	22.6
% Poverty	2.3	5.2	10.5	28.7	11.8
% Public assistance	0.3	1.0	2.2	7.8	2.8
% Unemployed	1.3	1.9	3.2	6.6	3.3
% No high school	3.0	6.3	10.1	20.2	10.2
Late stage rates [invasive BC only]	34.0%	41.6%	44.2%	48.8%	42.8%
Late stage rates [all female BC]	28.0%	33.3%	34.8%	38.9%	34.1%

Disadvantage score determined by factor analysis of % Female HH, % Poverty, % Public Assistance, % Unemployed, % No high school at time periods 2000 and 2013. Disadvantage score was calculated by scoring the latent variables, and taking the average score between the two time points

BC breast cancer, ACS US Census American Community Survey

cases with insurance had lower rates compared to cases with no insurance ($OR=0.67$, $p<0.001$) as well as cases with Medicaid ($OR=0.55$, $p<0.001$). Effect of disadvantage on late stage was attenuated by control for the other variables but was significant ($OR=1.0083$, $p=0.032$). An approximate linear increase of adjusted predictions by unit of CDI score was 0.18% percentile points compared to 0.44% unadjusted. No interactions of predictors in the model with disadvantage were found to be significant.

From the mediation analysis, the total relative risk effect of CDI comparing highest vs. lower quartiles late stage rates was 1.13 (95% CI 1.02, 1.24) which is close to the risk ratio of adjusted predictions in Tables 3 and 4. The controlled direct effect, which is equal to the natural direct effect in this case (after excluding the non-significant interaction term between CDI and obesity), and can be interpreted as the effect of CDI after controlling for obesity, was 1.08 (95% CI 0.95, 1.22). The indirect effect, interpreted as the effect of the CDI that operates through changing obesity, was 1.04 (95% CI 0.95, 1.15). For the subset of invasive cancers, the total effect of CDI is 1.10 (95% CI 1.00, 1.20) with controlled direct effect of 1.05 (95% CI 0.93, 1.18) and indirect effect of 1.05 (95% CI 0.96, 1.14).

Discussion

Using CBG data, we found that inner-city poverty was related to a higher risk of late stage diagnosis. The findings are consistent with and support other studies that have evaluated the impact of poverty on breast cancer staging. A national study of breast cancer staging in the USA found significant odds of being diagnosed with late stage breast cancer if living in an area of high poverty [5]. However, that study measured poverty using a simple measure of census tract poverty rates. Our analysis applies an index using multiple poverty metrics available from the US census on a CBG level, allowing for a finer granularity of both poverty and geography, which will allow investigators to develop more targeted interventions.

Census tract-level obesity rates were found to have a high correlation with neighborhood-level CDI, similar to previous research findings [19–21]. In particular, a Canadian study evaluating the effects of area-level socioeconomic differences in obesity measures found a significant relationship between area-level average dwelling value for females but not for males, with the difference in waist circumference for females living in

Table 3 Invasive female breast cancers ($n = 4935$): main effects logistic regression with random intercept

Fixed effects	B_0	p value	Odds ratio	Adjusted predictions
Outcome: late stage				
Intercept	0.14	0.318		
Age category (reference: 21–46)		All: <0.0001		
21–46	Ref	Ref	1.00	50.1
47–52	−0.32	<0.001	0.72	42.2
53–59	−0.39	<0.001	0.68	40.5
65+	−0.50	<0.001	0.61	38.1
Year of rating (reference: 2008–2012)		All: 0.0745		
2008–2012	Ref	Ref	1.00	42.2
2004–2007	0.13	0.062	1.14	45.4
2000–2003	−0.03	0.696	0.97	41.6
Race/ethnicity (reference: White)		All: 0.0039		
White	Ref	Ref	1.00	40.8
Black	0.28	0.001	1.32	47.4
Other	0.16	0.262	1.17	44.5
Hispanic	0.24	0.117	1.27	46.5
Insured (reference: no insurance)		All: <0.0001		
No insurance	Ref	Ref	1.00	49.8
Medicaid	0.22	0.227	1.24	55.1
Other insured	−0.32	0.011	0.73	42.0
Disadvantage score (per unit increase)	0.0082	0.052	1.0082	
At −8.0 [Q1]				41.0
At −5.3 [Q2]				41.5
At −0.6 [Q3]				42.4
At 13.1 [Q4]				45.1
Standard deviation of random intercept	~0			

Likelihood ratio test comparing this model vs. nested logistic regression without random intercept was not significant, $p = 1.000$

the lowest value areas having a waist circumference 2.95 cm larger than those living in the highest value areas [19]. Our mediation analysis detected a significant total effect of CDI on late stage, but is inconclusive with regards to whether the effect is fully or partially mediated by obesity. Thus, while high obesity rates are likely to be found in poor neighborhoods, and high neighborhood poverty is associated with high rates of late stage breast cancer, the directionality and paths of these relationships remain unclear. Additional contextual factors need to be investigated, such as exposure to neighborhood violence, which via a stress pathway can affect both obesity and vulnerability to cancer [15], and the impact of social diffusion of health behaviors, which can be affected by interconnected social networks [34].

Reducing rates of late stage breast cancer diagnoses in poor, minority populations has been a relatively intractable problem in some locations. In California for

example, between 1990 and 2000, although improvements in early detection were made in some groups over the decade, no changes in detection rates were made in both Black and White women in the lowest income quartile, with low-income Black women consistently having the highest rates of late stage diagnosis and the lowest rate of early stage diagnosis among all racial and ethnic groups [9]. Rates of late stage breast cancer diagnoses in low-income Black women may be historically unchanged because in the USA Blacks are more likely than any other racial or ethnic group to live in areas of highly concentrated poverty [7]. Targeting interventions to identified at-risk neighborhoods could allow for precision reduction in these disparate rates. However, we need more understanding of how the complex relationships between area-level poverty, obesity, and cancer diagnosis function before developing these interventions.

Table 4 All female breast cancers ($n = 6221$): main effects logistic regression with random intercept

Fixed effects	B_0	p value	Odds ratio	Adjusted predictions
Outcome: late stage				
Intercept	-0.021	0.111		
Age category (reference: 21–46)		All: <0.0001		
21–46	Ref	Ref	1.00	39.9
47–52	-0.32	<0.001	0.72	32.5
53–59	-0.34	<0.001	0.71	32.2
65+	-0.39	<0.001	0.68	31.1
Year of rating (reference: 2008–2012)		All: 0.0428		
2008–2012	Ref	Ref	1.00	32.5
2004–2007	0.16	0.012	1.18	36.2
2000–2003	0.08	0.214	1.09	34.3
Race/ethnicity (reference: White)		All: 0.0764		
White	Ref	Ref	1.00	32.7
Black	0.18	0.013	1.20	36.9
Other	0.10	0.438	1.10	34.9
Hispanic	0.16	0.258	1.17	36.3
Insured (reference: no insurance)		All: <0.0001		
No insurance	Ref	Ref	1.00	42.4
Medicaid	0.19	0.238	1.21	47.1
Other insured	-0.41	<0.0001	0.67	32.9
Disadvantage score (per unit increase)	0.0083	0.032	1.0	
At -8.0 [Q1]				32.2
At -5.3 [Q2]				32.7
At -0.6 [Q3]				33.6
At 13.1 [Q4]				36.1
Standard deviation of random intercept	0.15 (95% CI 0.04, 0.60)			

Likelihood ratio test comparing this model vs. nested logistic regression without random effect was not significant, $p = 0.2329$

Limitations

There are several limitations to our approach. The poverty index, while it identifies areas of risk, may be limited in some localities. Northern Virginia overall is a wealthy area relative to the rest of Virginia. Using a standardized evaluation for percent living in poverty may have caused our CDI score in these neighborhoods to overlook some areas that are, in fact, disadvantaged, resulting in missed opportunities for interventions to be targeted. We selected five variables from the ACS to include in our CDI score. We intentionally did not include race, since some areas of disadvantage may have varying concentrations of different races. However, other researchers have included race in disadvantage indices [6]. This may be more appropriate in cities that have clearly segregated neighborhoods. Finally, our obesity

measures relied on census tract-level data rather than neighborhood-level data. Applying more granular data to these analyses may help identify the complex relationships between neighborhood poverty and late stage breast cancer diagnosis.

Conclusion

Since the millennium, areas of concentrated poverty in the USA have increased. In the USA, the fastest growing populations of high poverty neighborhoods have occurred in those cities with between 250,000 and 500,000 people [35]. Because of the relationship between socioeconomic characteristics and late stage breast cancer diagnosis, combating disparities in breast cancer screening may be aided by identifying those neighborhoods with high rates of concentrated poverty

in US cities. Micro-targeting interventions towards individuals in urban neighborhoods with the highest risk cannot only combat disparities in diagnosis, but also morbidity and mortality due to earlier treatment. On a population level, reaching people earlier in the treatment cycle may decrease healthcare dollars spent overall, and has the potential to lessen both the treatment and financial burdens on an individual level.

Acknowledgments This research was supported by resources within the Cancer Control and Population Health program at the University of Virginia Cancer Center.

References

- Dickens C, Joffe M, Jacobson J, et al. Stage at breast cancer diagnosis and distance from diagnostic hospital in a periurban setting: a South African public hospital case series of over 1,000 women. *Int J Cancer*. 2014; 135(9): 2173–82. doi:10.1002/ijc.28861.
- Nguyen-Pham S, Leung J, McLaughlin D. Disparities in breast cancer stage at diagnosis in urban and rural adult women: a systematic review and meta-analysis. *Ann Epidemiol*. 2014; 24(3): 228–35. doi:10.1016/j.annepidem.2013.12.002.
- McLafferty S, Wang F. Rural reversal? Rural-urban disparities in late-stage cancer risk in Illinois. *Cancer*. 2009; 115(12): 2755–64. doi:10.1002/cncr.24306.
- Schootman M, Jeffe DB, Baker EA, Walker MS. Effect of area poverty rate on cancer screening across US communities. *J Epidemiol Community Health*. 2006; 60(3): 202–7. doi:10.1136/jech.2005.041020.
- Henry KA, Boscoe FP, Johnson CJ, Goldberg DW, Sherman R, Cockburn M. Breast cancer stage at diagnosis: is travel time important? *J Community Health*. 2011; 36(6): 933–42. doi:10.1007/s10900-011-9392-4.
- Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997; 277(5328): 918–24. doi:10.1126/science.277.5328.918.
- Jargowsky PA. Architecture of segregation: civil unrest, the concentration of poverty, and public policy. New York: The Century Foundation; Camden: Center for Urban Research and Education, Rutgers; 2015.
- Dai D. Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. *Heal Place*. 2010; 16(5): 1038–52. doi:10.1016/j.healthplace.2010.06.012.
- Flores YN, Davidson PL, Nakazono TT, Carreon DC, Mojica CM. Neighborhood socio-economic disadvantage and race/ethnicity as predictors of breast cancer stage at diagnosis. *BMC Public Health*. 2013; 13(1): 1. doi:10.1186/1471-2458-13-1061.
- Risser DR, Miller EA. Cancer in relation to socioeconomic status. *South Med J*. 2012; 105(10): 508–12. doi:10.1097/SMJ.0b013e318268c752.
- U.S. Census Bureau. 2010 Census Redistricting Data Summary File. Geographic Terms and Concepts. <https://www.census.gov/geo/reference/terms.html>. 2015. Accessed 20 Feb 2017.
- Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter? The public health disparities geocoding project. *Am J Epidemiol*. 2002; 156(5): 471–82. doi:10.1093/aje/kwf068.
- Rushton G, Armstrong MP, Gittler J, Greene BR, Pavlik CE, West MM, et al. Geocoding in cancer research: a review. *Am J Prev Med*. 2006; 30(2): S16–S24. doi:10.1016/j.amepre.2005.09.011.
- WILLIAMS DR. Race, socioeconomic status, and health the added effects of racism and discrimination. *Ann N Y Acad Sci*. 1999; 896(1): 173–88. doi:10.1111/j.1749-6632.1999.tb08114.x.
- DeGuzman PB, Schminkey DL. Influencing genomic change and cancer disparities through neighborhood chronic toxic stress exposure: a research framework. *Public Health Nurs*. 2016; 33(6): 547–57. doi:10.1111/phn.12290.
- Lipscomb J, Fleming ST, Trentham-Dietz A, et al. What predicts an advanced-stage diagnosis of breast cancer? Sorting out the influence of method of detection, access to care, and biologic factors. *Cancer Epidemiol Biomarkers Prev*. 2016; 25(4): 613–23. doi:10.1158/1055-9965.EPI-15-0225.
- Leung J, Martin J, McLaughlin D. Rural-urban disparities in stage of breast cancer at diagnosis in Australian women. *Aust J Rural Health*. 2016; 24(5): 326–32. doi:10.1111/ajr.12271.
- Cui Y, Whiteman MK, Flaws JA, Langenberg P, Tkaczuk KH, Bush TL. Body mass and stage of breast cancer at diagnosis. *Int J Cancer*. 2002; 98(2): 279–83. <http://www.ncbi.nlm.nih.gov/pubmed/11857420>. Accessed January 9, 2017.
- Harrington DW, Elliott SJ. Weighing the importance of neighbourhood: a multilevel exploration of the determinants of overweight and obesity. *Soc Sci Med*. 2009; 68(4): 593–600. doi:10.1016/j.socscimed.2008.11.021.
- Drewnowski A, Rehm CD, Solet D. Disparities in obesity rates: analysis by ZIP code area. *Soc Sci Med*. 2007; 65(12): 2458–63. doi:10.1016/j.socscimed.2007.07.001.
- Huang R, Moudon AV, Cook AJ, Drewnowski A. The spatial clustering of obesity: does the built environment matter? *J Hum Nutr Diet*. 2015; 28(6): 604–12. doi:10.1111/jhn.12279.
- Kim YA, Lee CW. Effects of obesity on breast cancer stage at diagnosis in Korean women. *Eur J Cancer Prev*. 2004; 13(1): 13–17.
- Virginia Department of Health. Virginia Cancer Registry. <http://www.vdh.virginia.gov/livewell/data/records/cancerregistry/about.html>. 2016. Accessed 19 May 2016.
- Ingram DD, Franco SJ. NCHS urban-rural classification scheme for counties. *Vital Health Stat*. 2012; 154(1): 1–65.
- Widaman KFK, Ferrer E, Conger RD. Factorial invariance within longitudinal structural equation models: measuring the same construct across time. *Child Dev Perspect*. 2010; 4(1): 10–8. doi:10.1111/j.1750-8606.2009.00110.x.

26. Newsom JT. *Longitudinal Structural Equation Modeling: a Comprehensive Introduction*. New York, NY: Taylor & Francis; 2015.
27. Widaman K, Ferrer E. Factorial invariance within longitudinal structural equation models: measuring the same construct across time. *Child Dev*. 2010; 4(1): 10–8. doi:[10.1111/j.1750-8606.2009.00110.x/full](https://doi.org/10.1111/j.1750-8606.2009.00110.x/full). Accessed May 20, 2016.
28. Anselin L, Kim Y, Syabri I. Web-based analytical tools for the exploration of spatial data. *J Geogr Syst*. 2004; 6(2): 197–218. doi:[10.1007/s10109-004-0132-5](https://doi.org/10.1007/s10109-004-0132-5). Accessed May 20, 2016.
29. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. *Geogr Anal*. 2006; 38(1): 5–22. doi:[10.1111/j.0016-7363.2005.00671.x](https://doi.org/10.1111/j.0016-7363.2005.00671.x).
30. Anselin L. Local indicators of spatial association-LISA. *Geogr Anal*. 2010; 27(2): 93–115. doi:[10.1111/j.1538-4632.1995.tb00338.x](https://doi.org/10.1111/j.1538-4632.1995.tb00338.x).
31. Centers for Disease Control and Prevention. 500 cities: local data for better health. U.S. Department of Health & Human Services Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion. 2016. <https://www.cdc.gov/500cities/>. Accessed 20 Feb 2017.
32. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013; 18(2): 137–50.
33. Fan X, Sivo SA. Using Δ Goodness-of-fit indexes in assessing mean structure invariance. *Struct Equ Model Multidiscip J*. 2009; 16(1): 54–69. doi:[10.1080/10705510802561311](https://doi.org/10.1080/10705510802561311).
34. Centola D. The spread of behavior in an online social network experiment. *Science*. 2010;329(5996):1194–7.
35. Jargowsky PA. Concentration of poverty in the new millennium: changes in prevalence, composition, and location in high poverty neighborhoods. The Century Foundation and Rutgers Center for Urban Research and Education; The Century Foundation, New York; NY 2013. p. 1–29.