Applied Geography 75 (2016) 28–35

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Modelling of geographic cancer risk factor disparities in US counties☆

Kyle D. Buck

US EPA/ORD/NHEERL/Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561, USA

ARTICLE INFO

Article history:
Received 18 August 2015
Received in revised form
29 July 2016
Accepted 2 August 2016

Keywords:
Cancer
Mortality to incidence ratio
Geographically weighted regression
Social
Economic
Index

ABSTRACT

The goal of this research is to create a theoretical framework for the identification of cancer risk factor disparities and address the recognition of geographic patterns in these factors. 34 secondary variables covering the entire US at the county level in 2010 were analyzed, both individually and grouped (theoretically and statistically), in relation to the mortality to incidence ratio (MIR) for all cancer sites. An a priori assessment and a principal components analysis (PCA) were used to group variables to test societal constructs. OLS and geographically weighted regressions (GWRs) were used to assess influence of both individual and grouped variables against the MIR. The theoretical grouping of variables showed little change in predictive capability of OLS models. In GWR model, there was marked improvement over the OLS. Maps produced using local R2 showed clear regional patterns of influence between the indicators and the MIR. Both the theoretical model and the justification for a spatial approach to cancer risk factor disparities were shown to be effective in this paper. The link between this suite of indicators and the health outcomes is clear, and supports the idea that a full representation of the SES landscape should be used to both predict health outcomes and to assess policy options for improving these outcomes. With the presence of definitive regional patterns and clear connections between the MIR and societal groupings, the findings from this research suggest a need to shift to a more comprehensive and spatial approach to cancer disparities research.

1. Introduction

The impact of cancer is enormous and takes a toll on both the individual and societal level. The total US economic impact of cancer in 2014 is estimated at $216.6 billion dollars, with nearly 13.7 million people living with cancer, over 1.6 million diagnoses, and more than half a million deaths (Howlander, Noone, & Krapcho, 2012; ACS, 2012). There is good news amidst the bad, however. Cancer incidence and mortality rates have been dropping in recent years according to the American Cancer Society (ACS) along with 5-year survival rates, due in part to lifestyle improvements, more advanced treatment options, and earlier detection of many cancer types (ACS, 2010).

Although the overall impact of cancer in the US looks to be headed in the right direction, the effect is not felt equally among all groups in the US. Cancer disparities, defined by National Cancer Institute (NCI), as "adverse differences in cancer incidence, cancer prevalence, cancer death, cancer survivorship, and burden of cancer or related health conditions that exist among specific population groups in the United States", are becoming an increasing focus (National Health Disparities Act, 2000). As a result, NCI funded programs and research initiatives have aimed at the lack of cohesive analysis and clear frameworks by which disparities are assessed (Harper & Lynch, 2010). This paper proposes both a theoretical framework as well as a method of analysis intended to fill this identified gap.

In order to effectively address the cancer health disparities issue, a theoretical model is proposed that takes a more holistic approach to the assessment of social and economic constructs as they relate to cancer outcomes. This approach builds on previous research, which has concentrated predominantly on socioeconomic status (SES), race, ethnicity and gender differences as they relate to cancer outcomes (Calo, Suarez, Soto-Sal, gado, Quintana, & Ortiz, 2015; Cook et al., 2015; Hess, Lee, Fish, Daly, Cress, & Mayadev, 2015; Rizzo, Sherman & Arciero, 2015; Kim, Paik, Yoon, Lee, Kim & Sung, 2015). Additional studies investigate the interaction of societal variables that exists across communities and how other health behaviors influence specific cancer outcomes (Goovaerts et al., 2015; Kuo, Mobley & Anselin, 2011; Oliver, Smith, Siadaty, Hauck,

☆ A theoretical framework and geographic approach to the identification of cancer risk factor disparities in the United States.

E-mail address: buck.kyle@epa.gov.

http://dx.doi.org/10.1016/j.apgeog.2016.08.001
0143-6228/© 2016 Published by Elsevier Ltd.

Pickle, 2006; Xiao, Gwede & Milla, 2007). Using a geographic approach in the analysis of disparities, the aim of this research is ultimately on the identification of regional trends and changes in societal influence that lead to these differential impacts across all cancer types.

2. Materials and methods

2.1. Conceptual background

In order for any type of analysis to be successful, a solid theoretical framework is required. In the case of cancer health disparities, the framework proposed here will be based on the merging of two separate fields. The conceptual model of place-based health vulnerability, shown in Fig. 1, forms the backbone of this research and is significant in its combination of spatial methodologies adopted from hazards geography and health disparities models (Cutter, 1996; Roux, 2012). By breaking apart each of the components of health risk, operationalization is possible along with measurement of each component's influence.

A big piece of this research lies in the correspondence of health disparities and hazards geography fields and what they are attempting to measure. Establishing the connection based on the concept of vulnerability provides justification for the combination of fields as well as the formation of a conceptual model merging the two. The link between cancer outcomes and geography has provided further impetus into the development of new models for risk assessment (Lin, Schootman, & Zhan, 2015). In addition to this link, the ability to operationalize the model is of key concern, as it allows for the identification and measurement of cancer disparities based on place and the measurement and comparison of the constructed factors to the places with identified disparities.

Within the field of hazards geography, a great deal of research has been conducted on drivers of social vulnerability, with great attention paid to the interaction of variables in space and time (Adger, 2006; Cutter, Mitchell, & Scott, 2000; Cutter, Boruff, & Shirley, 2003). What the hazards research has revealed is an intricate social structure with a high geographic dependence, where one social factor does not always exert the same level of influence on vulnerability. Utilizing the knowledge gained in the hazards field provides a much better metric for assessment of vulnerability to negative cancer outcomes. The outcomes as well as the drivers of vulnerability between cancer and hazards are very similar and treating the analysis of them similarly is a logical progression in the advancement of cancer outcomes prediction.

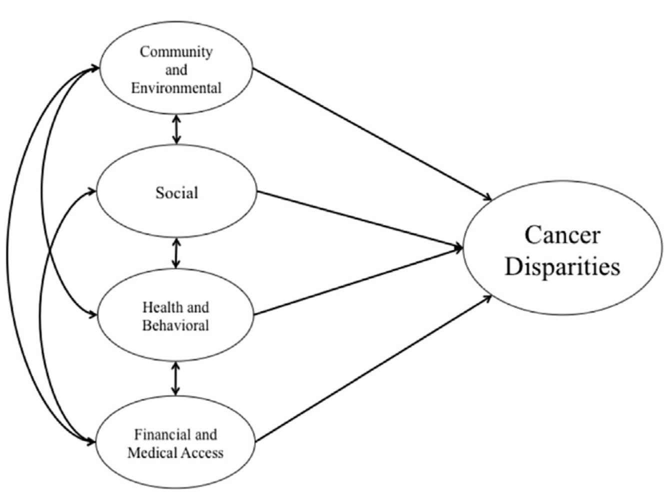

Fig. 1. Place-based health vulnerability model.

In this conceptualization, vulnerability begins with the access, and health/behavior, and community/environmental characteristics, which interact to yield a baseline health risk. Variables used to measure these constructs are shown in the breakout boxes. The resulting health risk is then filtered through the local social fabric to yield community health vulnerability, which will result in certain cancer outcomes and lead to potential disparities. Each factor in this model has the potential to influence the other, and contribute to changes in the health vulnerability of a place. In this model, the shift in terminology from risk to vulnerability marks the change to a place-based measurement, rather than an individual-based measure.

Health disparities can stem from ethnic, gender, income, and age divisions. In order to accurately reflect the influence of these, the analysis must account for multiple combinations of variables that can exist amongst groups. Combinations of factors have been utilized in a few studies, but the scale has remained limited and only a small number of variables are used in each case (Wagner et al., 2012; Li, Sunquist & Sunquist, 2012; Harper & Lynch, 2010). It is not necessarily accurate to say a group is of a certain social class, and therefore more vulnerable. Other social indicators may exist, making them more or less vulnerable. For example, an individual may be vulnerable due to their age, but this vulnerability could be decreased if the individual is a wealthy, married female. Access to healthy food options and green space can also influence the overall vulnerability (Bader, Purciel, Yousefzadeh, Neckerman, 2010; Dai, 2011). Determining the relative impact of all cancer drivers in addition to how these drivers interact with each other will allow for a much more thorough and accurate assessment of the social landscape and lead to better measurement of the drivers.

Cancer mortality-to-incidence ratios (MIR) are chosen as health disparity outcomes for a multitude of reasons. Cancer as an outcome is relevant due to the large burden along with a well-researched history and established patterns of disparities among certain populations. The MIR measure represents potentially avoidable cancer deaths and has proven to be effective in controlling for latency periods and relocation. It also helps to capture the early detection of cancer and any effective treatment outcomes. Also, due to the interest in cancer disparities, the MIR is used to help isolate counties that are not receiving appropriate care, most likely due to differences in SES (Hebert et al., 2009; Wan, Zhan, Zou & Wilson, 2013).

The geographic analysis of cancer disparities is carried out in this research using a geographically weighted regression (GWR) due to the demonstrated improvement in predictive ability of these models in landscapes where characteristics are clustered (Kupfer & Farris, 2007; Zhao, Gao, Wang, Liu & Li, 2015; Fotheringham, Brundson, & Charlton, 2002). A GWR model allows for regression coefficients to vary by location, and thus helps to control for spatial non-stationarity (Fotheringham et al., 2002; Legendre, 1993). The causes of cancer disparities will likely not be the same for all locations, resulting in poor predictive models over the large spatial extent of the U.S. By using a GWR in addition to the proposed theoretical framework for assessing cancer vulnerability, a picture can be created that demonstrates large scale trends across the US. The regions where disparities are known to exist can be examined in this larger context to better inform decisions related to the causes of the disparities.

2.1.1. Data sources

All data collected for this research is freely available and accessible. The temporal availability of each variable lies in the range of 2005–2010, with every attempt made to match the date for accuracy of statistical analysis. Details for data sources along with dates can be found in Table 1. The data for outcome measures

Table 1
Detailed list of 34 variables identified as potential cancer disparity drivers including collection years and sources.

Predictor variables			
Variable name (influence on MIR)	Data source	Year available	Calculations
Outcome measure: MIR	CDC-NPCR	05–09	Calculated as the mortality rate divided by the incidence rate
Income (−)	Census - ACS	05–09	Mean household income in last 12 months
Income inequality (+)	County Health Rankings	05–09	GINI Index – Income inequality Range (0–1)
Unemployed (+)	Census - ACS	05–09	Percentage unemployed
Population growth (abs)	Census 2010	2010	Percentage population change (2000–2010)
Renters (+)	Census - ACS	05–09	Tenure- calc. percentage of renters
Race-Non-white (+)	Census - ACS	05–09	Percentage of population not classified as white
Religious affiliation (−)	US Religious Census	2010	County level congregation membership
Married population (−)	Census - ACS	05–09	Percentage of pop (>18) now married
Single-parent household (+)	Census - ACS	05–09	Male householder + female householder
Number of dependents (+)	Census - ACS	05–09	Percentage of families with >1 dependent (<18 or >65 years old)
Educational attainment (−)	County Health Rankings	2010	Percentage of population have a college diploma
Language isolation (+)	Census - ACS	05–09	Percentage of population not speaking English
Parks per thousand (−)	USDA-ERS	2010	Count of all parks standardized by county population
Recreation Facilities (−)	County Health Rankings	2010	Per capita count of recreational facilities in a county.
Natural Amenities Scale (−)	USDA-ERS	1999	Index for livability of area based on climate factors
Environmental hazards (+)	EPA-TRI Locator	2009	Total amount of emissions from TRIs in county per capita
Rural population (+)	Census 2010	2010	% Living in rural areas - calculated
Particulate matter days (+)	EPA- County Health Rankings	2010	Number of days the particulate latter exceeded safe limits
Ozone days (+)	EPA- County Health Rankings	2010	Number of days the level of ozone exceeded safe levels
Liquor store density (+)	County Health Rankings	2010	Density of liquor stores per square mile in the county
Fast food access (+)	USDA-ERS	2010	Number of fast food restaurants per 1000 population
High risk occupation (+)	Economic Census	2009	Percentage working in high risk professions
Health food access (−)	County Health Rankings	2010	Percentage of zip codes in county with healthy food options
Population density (+)	Census - ACS	05–09	Number of people per square mile – calculated
Smoking (+)	BRFSS – Cnty Health Rank	2010	Percentage of population (>18) who smoke
Alcohol (+)	BRFSS – Cnty Health Rank	2010	Percentage of population (>18) who consume > 5 (male) or 4 (female) alcoholic beverages at a time
Exercise (−)	BRFSS – Cnty Health Rank	2010	Percentage with less than daily recommended exercise
Obesity (+)	BRFSS – Cnty Health Rank	2010	Percentage of population (>20) with BMI > 25
Mammography units (−)	FDA	2010	Number of certified mammography units per 1000 women in county
"poor" general health (+)	BRFSS – Cnty Health Rank	2010	Percentage of population ranking health as "poor"
Low birth weight (+)	BRFSS – Cnty Health Rank	2010	Percentage of live births with babies weighing less than 5 pounds.
No social support (+)	BRFSS – Cnty Health Rank	2010	Percentage of population reporting no social support
Number of doctors (−)	Area Resource File	2010	Number of practicing doctors per 100,000 population
Number of internal MDs (−)	Area Resource File	2010	Number of internal Medicine DRs per 1000 population
Hospitals with oncology service (−)	Area Resource File	2010	Hospitals with oncology services per 1000 population
Mammogram/pap smear <2yrs (−)	BRFSS	2009	Percentage of population getting recommended screening in last 2 years
Uninsured population (+)	SAHIE	2010	Percentage of population without health insurance

is obtained from the State Cancer Profiles Database, obtained from Center for Disease Control's National Program of Cancer Registries (NPCR) and the National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) program, with rate reported at a 95% confidence level and age-standardized based on 5-year age groups to the 2000 U.S. standard million population (Hebert et al., 2009).

2.1.2. *Statistical analysis*

Of the potential 3143 counties in the US, 2868 counties in 46 states are used for analysis. Certain counties are thrown out due to lack of cancer reporting data as well as specific population based SES measures. The CDC removes county level incidence and mortality data where there are three or less cases, while the states of Kansas and Minnesota are absent due to state laws repressing cancer data collection. In addition, the state of Alaska and Hawaii, Petroleum County, Montana; Arthur and Bradford County, Nebraska; and Kenedy, King, Loving, McMullen, Roberts, Sterling, and Terrell Counties, Texas are removed due to population numbers limiting statistical power.

Statistical tests and the generation of MIR values is accomplished using IBM SPSS software version 22.0 (IBM SPSS Statistics for Windows, Armonk, NY). GWR analysis and maps were generated using a geographic information system (ArcMAP Software, version 10.2; ESRI, Redlands, Calif).

Theoretically grouped variables are combined into an index value through an additive method involving z-score standardization followed by averaging each group to accommodate different

Table 2
Comparison of OLS Regression models used in aspatial analysis ($\alpha = 0.05$).

Independent Variables	N	R	R2	Adj. R2	Std. Error of the estimate	Durbin-Watson
Individual variables	34	0.596	0.355	0.347	0.791720	2.012
Theoretically grouped	4	0.570	0.325	0.324	0.805752	2.017
PCA grouped	10	0.578	0.355	0.332	0.807530	2.008

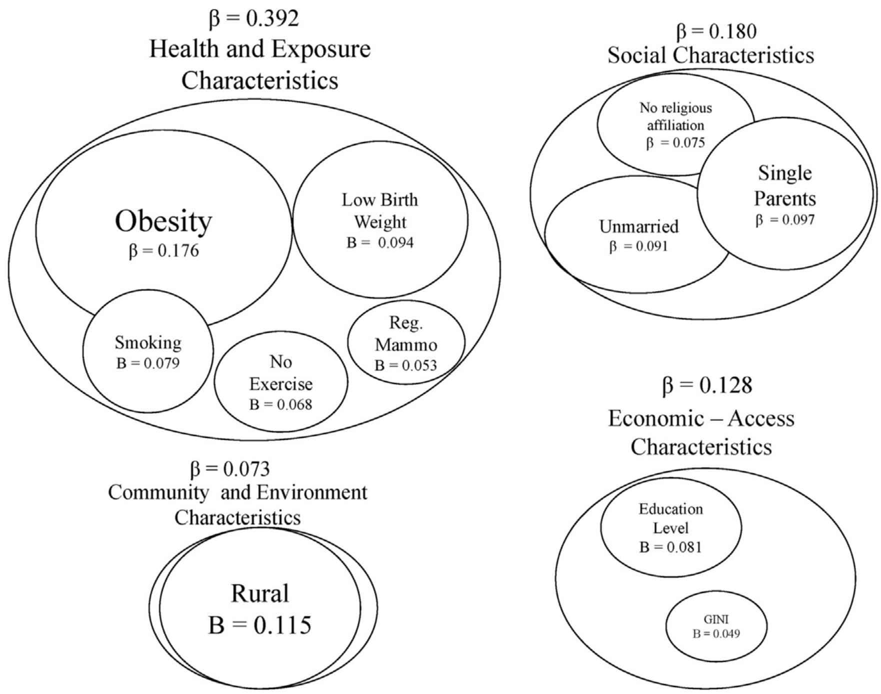

Fig. 2. The 4 theoretical groupings along with their contribution to the MIR (shown as beta value) and the relative influence of the significant values they contain (ranked by correlation to MIR). Ovals are all proportionate to their beta values in the regression model. Note: diagram not drawn to exact scale.

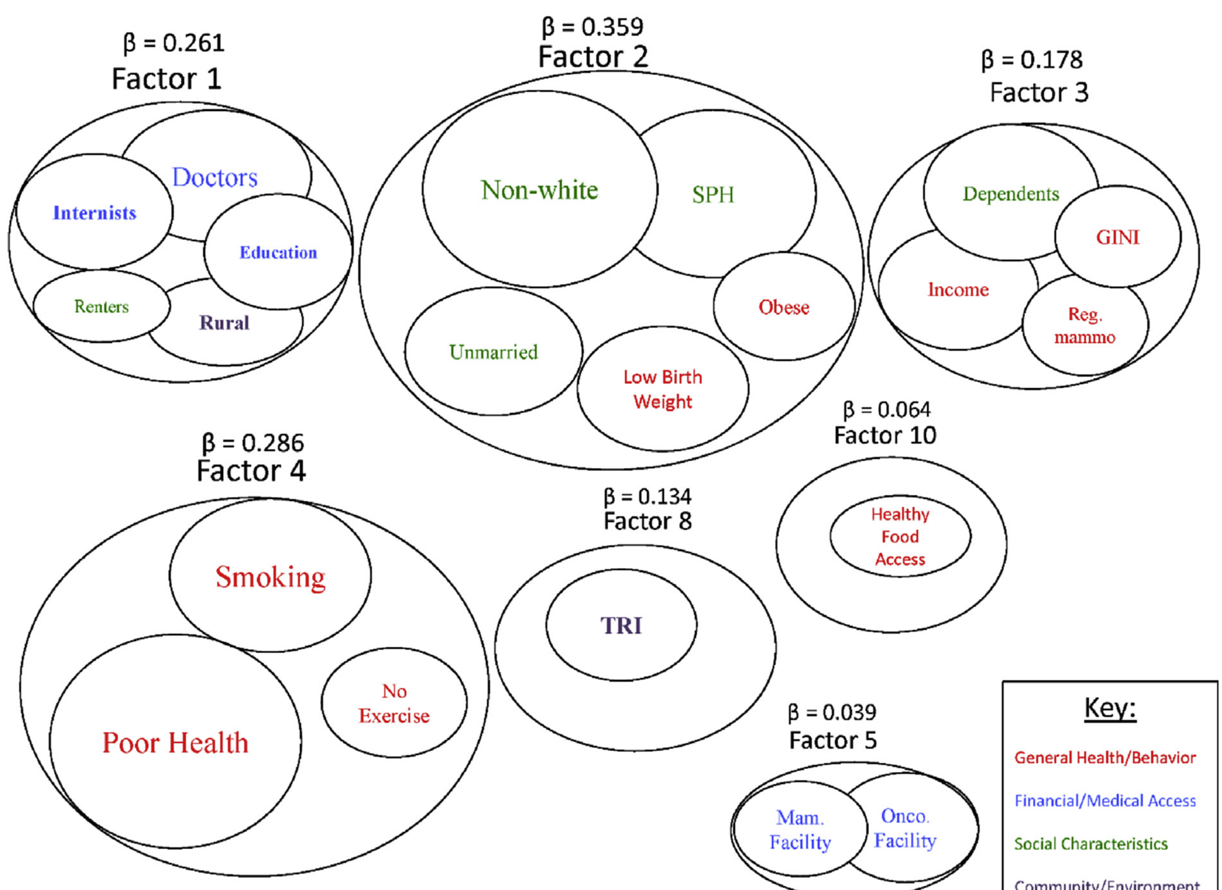

Fig. 3. Diagram depicting the 10 PCA constructed factors. Also shown are the significant variables contributing to each factor, if there are any. Ovals are all proportionate to their beta values in the regression model. Variables are color coded to denote the theoretical groups to which each variable belongs. Note: diagram not drawn to exact scale.

Table 3
Results from OLS and GWR models using theoretically grouped variables versus the MIR ($\alpha = 0.05$).

Model	N	R2	Adjusted R2	AIC
OLS	4	0.325	0.324	−1178.714
GWR	4	0.525	0.414	−1663.346

numbers of variables. The result is a single indicator score representing each construct, with higher values equaling better conditions. For the PCA grouped variables, the indicators are constructed in SPSS by using the scaled factor scores from the PCA output.

Spatial analysis is conducted using a GWR with the theoretically grouped variables. This regression uses adaptive kernel determination method utilizing the Akaike Information Criterion (AIC). The projection used to run this analysis is North American Albers Equal Area Conic in order to minimize distortion of data.

3. Results

3.1. Aspatial analysis

Three OLS regression models are run initially to serve as both a test of the theoretical groupings and as a baseline for comparisons to the geographic analysis. All models are tested to determine overall fit as well as influence of each variable or grouping on the MIR. Tests for independence of residuals, collinearity and variance inflation are all run to confirm the adequacy of each model for predicting the MIR. Model results, shown in Table 2, reveal little change in predictive power of the three based on the adjusted R^2. This finding provides evidence to support use of the proposed theoretical constructs.

Two figures, using proportional ovals to represent correlation coefficients, are created to help summarize the groupings as well as the relative influence of each variable and grouping on the MIR. Fig. 2 displays the relative influence of each theoretically grouped factor along with beta scores and the individual variables with significant correlations to the MIR, while Fig. 3 displays the relative contribution of each indicator to the variance determined from the PCA model. The divergence between the theoretical and PCA grouped variables is also significant to note. While the theoretical groupings are meant to represent identifiable human societal constructs, the PCA groupings are created based on variables responsible for the most variance in the set. In the measurement of disparities, the use of a PCA methodology may be beneficial in identifying the characteristics that vary most between different geographic areas. If these characteristics are also proven to contribute heavily to cancer outcomes, their importance in disparities would be even more evident.

Going back to the proposed model, health outcomes in a place result from a combination of all factors present in that place. Certain factors may correlate highly with other factors due to a path of influence. In other words, the presence of one characteristic may increase the likelihood of another set of characteristics being present. For example, a higher education level (access factor) may lead to higher income levels (also access), which will lead to a higher

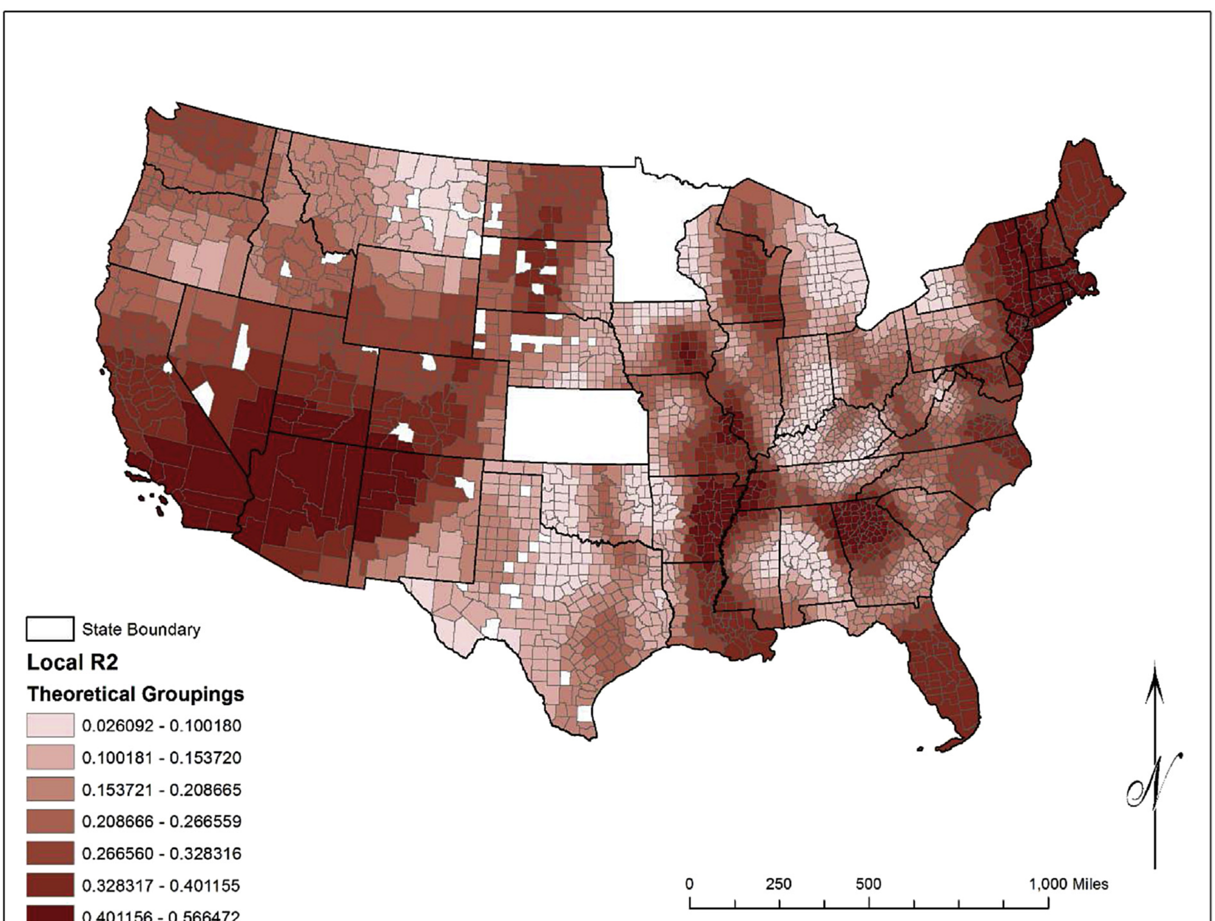

Fig. 4. Local R^2 results from GWR with theoretically grouped variables.

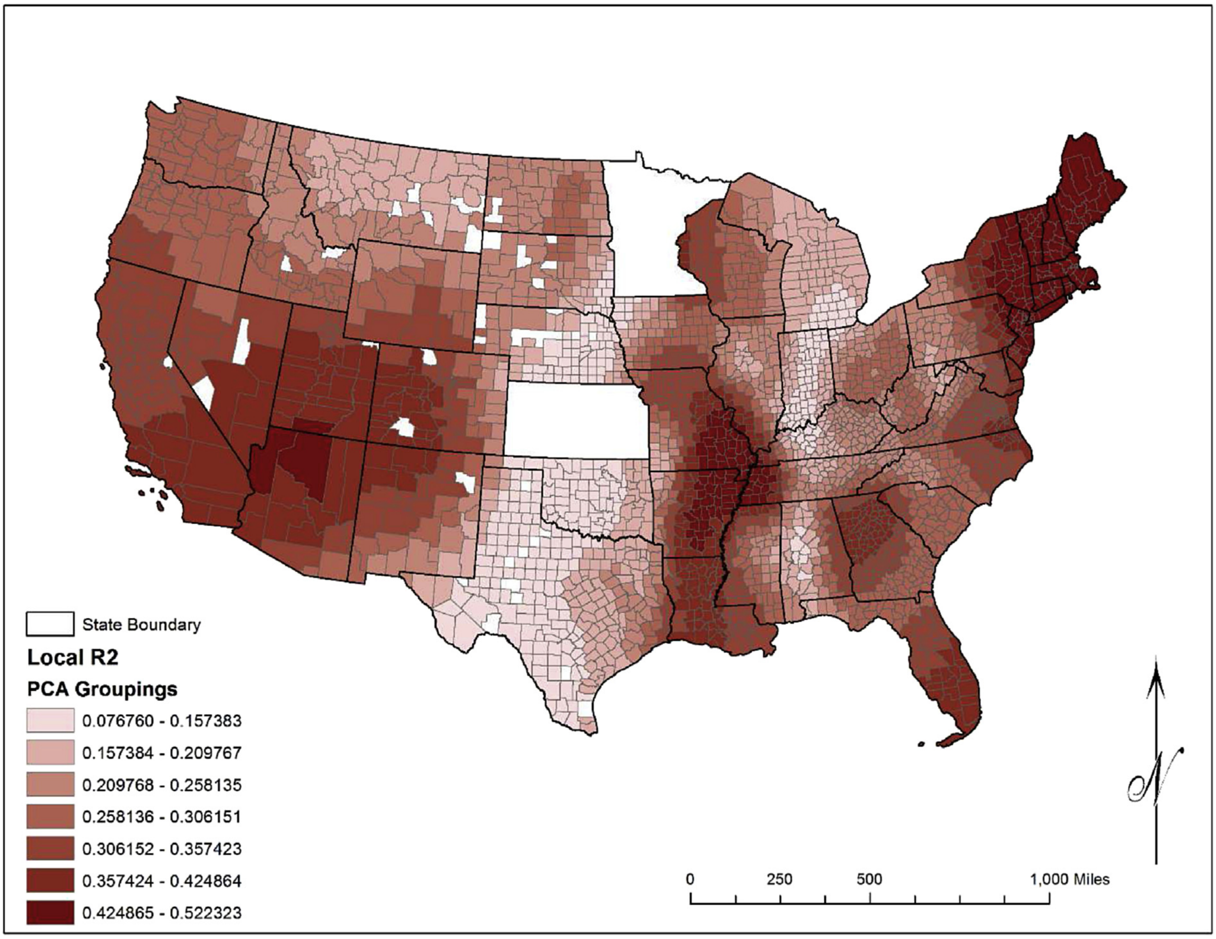

Fig. 5. Local R^2 results from GWR with theoretically grouped variables.

likelihood of access to recreational facilities and healthy food. All of this may result in a higher chance of exercising and eating well and a lower vulnerability to negative cancer outcomes. To summarize, even though the model constructs (factors) may be assembled accurately, the relationship between the factors may be directionally dependent, meaning that one or more factors are ultimately responsible for starting a chain of events that leads to higher cancer fatality rates. Ultimately, the results support the use of this theoretical framework in assessing the influence of societal structures via theoretically grouped variables on cancer MIRs.

3.1.1. *Spatial analysis*

The outcomes of the GWR models for both the theoretically and PCA grouped variables reveal improvements over the OLS models (Table 3). The adjusted R^2 of the theoretically grouped OLS model was 0.325, whereas the GWR model produced an adjusted R^2 of 0.414. For the PCA grouped OLS model, the adjusted R^2 was 0.332, while for the GWR model the adjusted R^2 improved to a 0.417. This improvement provides further evidence that spatial patterns are a better predictor of cancer outcomes than using only aspatial methods.

A major benefit of the GWR is the ability to visually represent the varying strength of relationship between the dependent and independent variables by mapping the local R^2 values. In this way, the explanatory strength can be tied to places. In Fig. 4, the variation in local R^2 values for the theoretically grouped variables is evident. The locations where the predictive values are strongest correspond with areas of both higher and lower MIR values as well, indicating that this trend is not related solely to a better or worse outcome. What it does signify is a strong regional trend. Fig. 5 reveals a similar pattern in the PCA grouped variables, tying in the areas where variances are highest as well.

Looking at these two figures provides a clear depiction of the varying relationship that exists between the SES of an area and the cancer outcomes for better or for worse. Regardless of the MIR values, there seems to be a presence of higher predictive statistics in areas where populations are more concentrated. The northeast, for example, shows up as a place with higher predictive values. In addition to this, the Mississippi river corridor, Florida, the Atlanta metro area, and the Southwest US all have higher local R^2 values. This is significant, especially in the Southeast, due to the higher numbers of cancer fatalities that exist in this area. This region possesses many of the characteristics associated with poor cancer rates, and this shows up very clearly.

The most promising output from the GWR analysis comes from the individual regression outputs. Every county in the analysis has its own regression equation, complete with local R^2 values and coefficients for each independent variable. These equations can provide details on both the MIR as well as the level to which the societal structure influences this value. Even more detail may be gleaned from the coefficients, which indicate the relative influence of each construct on the MIR in that county. This output yields numerous possibilities for future research. Table 4 provides a small example of data extracted from the analysis. In this table is the predicted MIR as well as the measured MIR, along with all relevant information revealing strength of relationships between the

Table 4
Example from GWR model showing 4 counties. The observed MIR and model predicted MIR are shown along with the local R2 and the coefficients for each of the 4 theoretically grouped variables.

Observed MIR	LocalR2	Predicted MIR	Social Coeff.	Health Coeff.	Economic Coeff.	Community Coeff.
3.903199	0.45625	2.70978	0.83920	0.58787	0.61050	−0.13738
3.565063	0.09086	0.24843	0.15098	0.66697	−0.00978	0.54660
−1.351778	0.45395	−0.93240	0.16859	0.4052	0.64910	0.21508
−2.449035	0.09192	−0.21783	0.55487	0.17388	−0.65367	0.81945

indicators and the MIR. Two counties with high and two with low MIR are shown. Each group has one county with high and one with low local R^2 as well. It is evident in just this small sample that the influence of each theoretical group on the MIR varies greatly depending on location. This must be considered a major change in the way these influences are analyzed.

4. Discussion

The primary outcome of this research is the establishment of regional trends between MIRs and SES factors. Just as there is not a single SES factor explaining cancer incidence or mortality rates, there also exists no single correlation among the regions. In other words, the linkage between SES factors and MIR in one place cannot be assumed to exist in other places. There does appear to be regional clustering of the relationships, however, which implies that MIR outcomes and SES factors tend to vary in a manner that could be predicted. Given this information it would be possible to better identify the specific characteristics of a community that potentially drive poor cancer outcomes. In addition, there is reason to look at more localized patterns, given appropriate data.

In addition to the spatial distribution among the MIR outcomes and SES factors, this research also demonstrates the benefit of a theoretical model for place-based assessment of cancer disparities. The model shows promise as a way to account for multiple components of social structure existing in a specific geography. In addition, it proves capable of operationalization for US counties, making possible the testing of multiple societal characteristics in a cohesive manner. Dependent on data availability, this could also translate to scaling at different levels.

Geographic regression models were shown to improve the predictive capabilities by accounting for spatial non-stationarity that existed in the data. This proves a definitive link between the characteristics of a place and the change in how certain predictor variables influence cancer outcomes and implies that smaller case studies are not applicable to other places where influences may not be the same. Having a country wide analysis of this regional variance should prove very helpful in making comparisons across case studies in the future.

Lastly, this research shows promise in the identification of spatial cancer disparities and the ability to identify sets of community characteristics linked to them. The ability to break down and analyze individual places and quantify the link between a multitude of characteristics and cancer outcomes could be very helpful in justifying the location of specific services known to help decrease cancer rates.

5. Conclusion

The policy implications of this research are broad reaching and have the potential to aid in the identification of places where not only disparities exist, but also the reasons why they exist. A major goal that has carried through each iteration of the Healthy People initiative is the reduction/elimination of disparities. In order to accomplish this goal, both the location of the disparities as well as and understanding of the drivers is necessary. Removing obstacles to proper health care and equitable health outcomes is critical, and understanding how these obstacles present themselves in a place is essential to achieving the goal of Healthy People, 2020 (Healthy People, 2015).

Hopefully the findings in this paper will begin to address the need for identification of both the fundamental drivers of cancer differences and the regional patterns that impact how societal drivers actually affect cancer outcomes in a specific place. As opposed to working from the bottom up, starting with smaller spatial studies, the work presented here attempts a top down approach by creating a country-wide look at disparities and their drivers. With this overview, smaller scale studies can be better situated in relation to others and better predictions made.

References

ACS, A. C. S. (2010). *Cancer facts & figures 2010*. Atlanta, GA: American Cancer Society ACS.

ACS, A. C. S. (2012). *Cancer facts & figures 2012*. Atlanta, GA.

Adger, W. N. (2006). Vulnerability. *Global Environmental Change-Human and Policy Dimensions, 16*(3), 268–281.

Bader, M. D. M., Purciel, M., Yousefzadeh, P., & Neckerman, K. M. (2010). Disparities in neighborhood food environments: Implications of measurement strategies. *Economic Geography, 86*(4), 409–430.

Calo, W. A., Suarez, E., Soto-Salgado, M., Quintana, R. A., & Ortiz, A. P. (2015). Assessing lung cancer incidence disparities between Puerto Ricans and other Racial/Ethnic Groups in the United States, 1992-2010. *Journal of Immigrant and Minority Health, 17*(3), 971–975. http://dx.doi.org/10.1007/s10903-014-0153-1.

Cook, M. B., Rosenberg, P. S., McCarty, F. A., Wu, M. X., King, J., Eheman, C., et al. (2015). Racial disparities in prostate cancer incidence rates by census division in the United States, 1999-2008. *Prostate, 75*(7), 758–763. http://dx.doi.org/10.1002/pros.22958.

Cutter, S. L. (1996). Vulnerability to environmental hazards. *Progress in Human Geography, 20*(4), 529–539. http://dx.doi.org/10.1177/030913259602000407.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly, 84*(2), 242–261.

Cutter, S. L., Mitchell, J. T., & Scott, M. S. (2000). Revealing the vulnerability of people and places: A case study of Geogretown County, South Carolina. *Annals of the Association of American Geographers, 90*(4), 713–737.

Dai, D. (2011). Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? *Landscape and Urban Planning, 102*(4), 234–244.

Fotheringham, A. S., Brundson, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Wiley.

Goovaerts, P., Hong, X., Adunlin, A., Ali, A., Tan, F., Gwede, C. K., et al. (2015). Geographically-weighted regression analysis of percentage of late-stage prostate cancer diagnosis in Florida. *Applied Geography, 62*, 191–200.

Harper, S., & Lynch, J. (2010). Methods for measuring Cancer Disparities: Using data relevant to healthy people 2010 Cancer-Related objectives. In *NCI Cancer surveillance monograph series* (Vol. 6). The National Cancer Institute.

Healthy People 2020 [Internet]. Washington, DC: U.S. Department of Health and Human Services, Office of Disease prevention and health promotion [cited August 11, 2015], 2015, Available from:: [https://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities].

Hebert, J. R., Daguise, V. G., Hurley, D. M., et al. (2009). Mapping cancer mortality-to-incidence ratios to illustrate racial and sex disparities in a high-risk population. *Cancer, 115*(11), 2539–2552.

Hess, C., Lee, A., Fish, K., Daly, M., Cress, R. D., & Mayadev, J. (2015). Socioeconomic and Racial Disparities in the Selection of Chest Wall Boost Radiation Therapy in Californian Women After Mastectomy. *Clinical Breast Cancer, 15*(3), 212–218. http://dx.doi.org/10.1016/j.clbc.2014.11.007.

Howlander, M., Noone, A. M., & Krapcho, M. (2012). *SEER Cancer statistics review, 1975-2009*. National Cancer Institute.

Kim, S. E., Paik, H. Y., Yoon, H., Lee, J. E., Kim, N., & Sung, M. K. (2015). Sex- and gender-specific disparities in colorectal cancer risk. *World Journal of Gastroenterology, 21*(17), 5167–5175. http://dx.doi.org/10.3748/wjg.v21.i17.5167.

Kuo, T. M., Mobley, L. R., & Anselin, L. (2011). Geographic disparities in late-stage breast cancer diagnosis in California. *Health & Place, 17*(1), 327–334.

Kupfer, J. A., & Farris, C. A. (2007). Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. *Landscape Ecology, 22*(6), 837–852. http://dx.doi.org/10.1007/s10980-006-9058-2.

Legendre, P. (1993). Spatial autocorrelation - Trouble or new paradigm. *Ecology, 74*(6), 1659–1673. http://dx.doi.org/10.2307/1939924.

Lin, Y., Schootman, M., & Zhan, F. B. (2015). Racial/ethnic, area socioeconomic, and geographic disparities of cervical cancer survival in Texas. *Applied Geography, 56*, 21–28. http://dx.doi.org/10.1016/j.apgeog.2014.10.004.

Li, X. L., Sunquist, K., & Sunquist, J. (2012). Neighborhood deprivation and prostate cancer mortality: A multilevel analysis from Sweden. *Prostate Cancer and Prostatic Diseases, 15*(2), 128–134.

Minority Health and Health Disparities Research and Education Act of 2000, 106 U.S.C. § 1880, 2000, Retrieved from https://www.congress.gov/bill/106th-congress/senate-bill/1880/.

Oliver, N. M., Smith, E., Siadaty, M., Hauck, F. R., & Pickle, L. W. (2006). Spatial analysis of prostate cancer incidence and race in Virginia, 1990-1999. *American Journal of Preventive Medicine, 30*(2), S67–S76.

Rizzo, J. A., Sherman, W. E., & Arciero, C. A. (2015). Racial disparity in survival from early breast cancer in the department of defense healthcare system. *Journal of Surgical Oncology, 111*(7), 819–823. http://dx.doi.org/10.1002/jso.23884.

Roux, A. V. D. (2012). Conceptual approaches to the study of health disparities. In J. E. Fielding (Ed.), *Annual review of public health* (Vol. 33, pp. 41–58). Palo Alto: Annual Reviews.

Wagner, S. E., Hurley, D. M., Hebert, J. R., McNamara, C., Bayakly, A. R., & Vena, J. E. (2012). Cancer mortalityto-incidence ratios in Georgia describing racial cancer disparities and potential geographic determinants. *Cancer, 118*(6), 4032–4045.

Wan, N., Zhan, F. B., Zou, B., & Wilson, J. G. (2013). Spatial access to health care services and disparities in colorectal cancer stage at diagnosis in Texas. *Professional Geographer, 65*(3), 527–541. http://dx.doi.org/10.1080/00330124.2012.700502.

Xiao, H., Gwede, C. K., & Milla, K. (2007). Analysis of prostate cancer incidence using geographic information system and multilevel modeling. *Journal of the National Medical Association, 99*(3), 218.

Zhao, Z. Q., Gao, J. B., Wang, Y. L., Liu, J. G., & Li, S. C. (2015). Exploring spatially variable relationships between NDVI and climatic factors in a transition zone using geographically weighted regression. *Theoretical and Applied Climatology, 120*(3–4), 507–519. http://dx.doi.org/10.1007/s00704-014-1188-x.