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
Available online 11 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.

Published by Elsevier Ltd.

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/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.

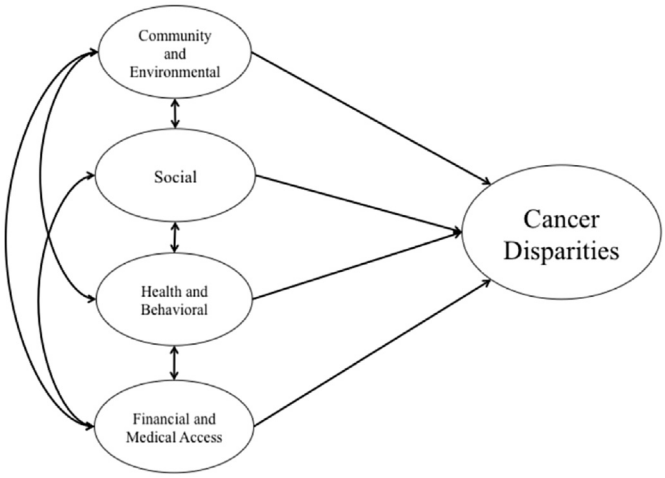

Fig. 1. Place-based health vulnerability model.

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