



Published in final edited form as:

*Int J Public Pol.* 2010 ; 5(2-3): 237–258. doi:10.1504/IJPP.2010.030606.

## Healthcare access, socioeconomic factors and late-stage cancer diagnosis: an exploratory spatial analysis and public policy implication

**Fahui Wang,**

Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA 70803 USA

**Lan Luo, and**

Department of Geography, University of Illinois, Urbana-Champaign, Urbana, IL 61801-3671 USA

**Sara McLafferty**

Department of Geography, University of Illinois, Urbana-Champaign, Urbana, IL 61801-3671 USA

Fahui Wang: fwang@lsu.edu; Lan Luo: lanluoni@yahoo.com; Sara McLafferty: smclaff@uiuc.edu

### Abstract

Patients diagnosed with late-stage cancer have lower survival rates than those with early-stage cancer. This paper examines possible associations between several risk factors and late-stage diagnosis for four types of cancer in Illinois: breast cancer, prostate cancer, colorectal cancer, and lung cancer. Potential risk factors are composed of spatial factors and nonspatial factors. The spatial factors include accessibility to primary healthcare and distance or travel time to the nearest cancer screening facility. A set of demographic and socioeconomic variables are consolidated into three nonspatial factors by factor analysis. The Bayesian model with convolution priors is utilised to analyse the relationship between the above risk factors and each type of late-stage cancer while controlling for spatial autocorrelation. The results for breast cancer suggest that people living in neighbourhoods with socioeconomic disadvantages and cultural barriers are more likely to be diagnosed at a late stage. In regard to prostate cancer, people in regions with low socioeconomic status are also more likely to be diagnosed at a late stage. Diagnosis of late-stage colorectal or lung cancer is not significantly associated with any of the abovementioned risk factors. The results have important implications in public policy.

### Keywords

late-stage cancer; healthcare access; Poisson regression; spatial autocorrelation; Bayesian model with convolution priors; USA

## 1 Introduction

Cancer has a tremendous impact on the health of the nation. About 1,444,920 new cancer cases are expected to be diagnosed, and 559,650 are expected to die of cancer in the USA in 2007 [American Cancer Society, (2007), pp.1–2]. Cancer stage of development at the time of diagnosis plays a critical role in determining the prognosis of patients. Detection at an early stage helps cancer patients live longer and maintain a better quality of life. Persons

diagnosed with late-stage cancer suffer from a lower survival rate. Four types of cancer, namely breast, prostate, colorectal and lung cancer, have ranked highest in incidence and mortality in Illinois for years. But the definitive risk determinants for late-stage diagnosis of these four leading types of cancer have not yet been established. This paper examines possible associations between several risk factors and late-stage diagnosis by zip code for these four types of cancer in Illinois.

Breast cancer is the most common cancer among women. Early diagnosis of breast cancer by mammography screening has become a critical way to reduce mortality (Sheehan and DeChello, 2005). Women without adequate accessibility to timely mammography screening are more likely to develop late-stage breast cancer (Mandelblatt et al., 1991). Racial and ethnic disparities in mortality are significant: the chance of African-American women dying from breast cancer within five years of diagnosis is twice as great as for Caucasian women, and the chance of Hispanic women is 1.5 times greater than for Caucasian women (Eley et al., 1994). Women in low-income neighbourhoods are also more likely to be diagnosed with late-stage breast cancer than women in more affluent neighbourhoods (Barry and Breen, 2005).

Prostate cancer is the most frequently diagnosed cancer in men. Prostate cancer incidence rates have a strong geographical variation, not only at an international level but also within regions of a country. Early diagnosis depends partly on accessibility to medical facilities (Klassen et al., 2005). A larger proportion of African-American men and foreign-born Asian-American males are diagnosed with late-stage prostate cancer (Oakley-Girvan et al., 2003). However, our understanding of causes for late-stage prostate cancer diagnosis remains very limited.

Colorectal cancer is the third leading cause of cancer mortality for both men and women in the US. The survival rate depends on the tumor stage as well as age at time of diagnosis (Rushton et al., 2004), and is significantly higher for poor African-Americans (Freeman and Alshafie, 2002). Geographic variation of colorectal cancer mortality rates is evident in the US (Devesa et al., 1999). This variation in the rates of late-stage colorectal cancer from place to place is important because the greatest proportion of mortality (90.2%) occurs among those patients diagnosed at a late stage.

Lung cancer is the leading cause of cancer death among both men and women in the USA. Since lung cancer has a very long latency period, it is often diagnosed at late stage (Blagojevich and Whitaker, 2006). Due to the lack of effective early detection methods for lung cancer, the majority of research about lung cancer focuses on factors influencing its mortality rather than on late-stage diagnosis (e.g., Haynes, 1988; Brown et al., 1994).

In summary, there are two groups of potential risk factors for late-stage cancer diagnosis: access to healthcare, and socioeconomic and demographic characteristics. The former emphasises spatial elements and thus is referred to as spatial factors; and the latter nonspatial factors. On the spatial factors, this research uses two measures: a comprehensive index of spatial accessibility to primary healthcare, and distance or travel time to the nearest cancer screening facility. The spatial accessibility index accounts for complex interaction between primary healthcare supply and demand locations and the distance and ease of travel between them (Luo and Wang, 2003). To measure spatial access to cancer screening, the distance and travel time from the nearest mammography facility is calculated for breast cancer screening; distance and travel time from the nearest hospital with any colonoscopy capacity is calculated for colorectal cancer. No such measure is feasible in this research for prostate or lung cancer because of the lack of data on specific kinds of screening facilities. On the nonspatial factors, selected socioeconomic and demographic variables are

consolidated into three major factors by factor analysis, which are used in models for all four types of cancer. An earlier study (Wang et al., 2008) used the same set of variables to explain late-stage breast cancer diagnosis and reported some promising results.

On the methodological realm, this research employs a Bayesian model with convolution priors to examine possible association between the aforementioned risk factors and late-stage cancer diagnosis. This is an improved method over the Poisson regression used in the earlier study (Wang et al., 2008). The regular Poisson regression is widely used in modelling count data (Cameron and Trivedi, 1998), but does not control for spatial autocorrelation or spatial dependence, which occurs when attributes of nearby areas are more similar (or dissimilar) than those of distant ones. The problem has long plagued statistical analyses of spatial data. Different from the regular Poisson regression, the Bayesian model with convolution priors incorporates spatial random effects and any other unstructured random effects. Previous studies suggest that the Bayesian model has more advantages than the regular Poisson regression in analysis of rare events with small study areas (Law et al., 2005; Mollie, 1996). This research uses WinBUGS (Best and MRC Biostatistics, 2004) to implement the Bayesian model, and the results are compared with those from the regular Poisson regression.

The remainder of this paper is organised as follows. Section 2 provides an expanded literature review. Section 3 describes the data and definitions for variables. Section 4 explains the Bayesian model for examining the risk factors of late-stage cancer diagnosis. Section 5 presents and discusses the results. The paper is concluded with a brief summary and discussion of implications in public policy.

## 2 Literature background

Much research has focused on possible causes for late-stage cancer diagnosis. This section reviews the literature on primary risk factors for advanced-stage diagnosis of the aforementioned four types of cancer except for lung cancer, which is often diagnosed at advanced stages because of lack of effective early detection methods (Blagojevich and Whitaker, 2006).

### 2.1 Spatial accessibility

In regard to breast cancer, accessibility to mammography facilities is poorer for women in rural settings than women in urban regions (Coughlin et al., 2002; Cummings et al., 2002). Low accessibility to mammography facilities in rural regions delays early detection of breast cancer and allows the cancer to develop to an advanced stage. This accounts for the higher proportion of late-stage breast cancer diagnosis in rural areas (Montella et al., 1995; Amey et al., 1997; Hoffman et al., 2000; Menck and Mills, 2001), particularly among African-American women in rural districts (Coughlin et al., 2002; Cummings et al., 2002). The densities of people and facilities are much lower in rural areas than in urban regions; therefore, distance or travel time is longer between patients' homes and appropriate facilities in rural areas. Even in an urban setting, women who live farther from facilities are less likely to receive mammography screening than those closer to facilities (Zhang et al., 1999; Rohan et al., 2000). By the same token, those dependent on public transportation spend more travel time to healthcare facilities than those with a private automobile and thus are less likely to utilise the screening facilities (Zenk et al., 2006). With respect to prostate cancer, Liff et al. (1991) found that rural patients were more likely to have advanced stages of prostate cancer than urban patients and concluded that accessibility might contribute to this difference. Rushton et al. (2004) concluded that high rates of late-stage colorectal cancer diagnosis appeared in places where the average distances to diagnostic facilities were lengthy.

In addition to distance or travel time from cancer screening facilities, it is also important to consider access to primary care. Primary care physicians typically provide the first point of contact with the healthcare system, and are critical for the success of preventive care and cancer screening (Lee, 1995). An earlier study by Wang et al. (2008) indicated that spatial access to primary care is an important predictor of late-stage breast cancer diagnosis. Very little has been reported on its role in the diagnosis stage of other cancer types. It is also important to point out the need to use improved measures of spatial accessibility to cancer screening or primary healthcare (Guagliardo, 2004), which is considered in this research.

## 2.2 Socio-economic status

Many studies consistently show that socioeconomic status (SES) plays an important role in explaining late-stage cancer diagnosis. In those studies, SES is typically comprised of income, health insurance status, and transportation means.

Lower SES has been associated with a higher rate of late-stage breast cancer diagnosis (Joslyn and West, 1999; Baquet and Commiskey, 1999; Bradley et al., 2002) because of less frequent mammography screening among women of lower SES. SES is a characteristic not only of individuals but is also tied to a neighbourhood context. Several studies found that women living in areas with low SES were more likely to be diagnosed with late-stage breast cancer than women in areas with high SES (Nosek and Howland, 1997; Merkin et al., 2002; Barry and Breen, 2005). A study by Merkin et al. (2002) confirmed the finding and further suggested that living in areas with low SES increased the odds of advanced-stage breast cancer more for Caucasian women than for black women. Also, women having Medicaid insurance, similar to uninsured women, had more advanced stages of breast cancer compared with women possessing other insurance plans (Roetzheim et al., 1999; McWilliams et al., 2003; Thomas and Carlin, 2003).

SES is also important in the diagnosis of prostate and colon cancer. Research suggests that men with low SES are more likely to be diagnosed with late-stage prostate cancer (Klassen et al., 2005; Liu, 1998; Morris et al., 1999). People of lower SES, often without health insurance or having only Medicaid coverage, are less likely to receive colonoscopy screening and thus are more likely to be diagnosed with late-stage colorectal cancer (Roetzheim et al., 1999, 2000; Palmer and Schneider, 2005). Particularly for elderly women, health insurance status is one of the most important factors influencing their utilisation of colonoscopy screening (Hsia et al., 2000).

However, SES has its own limitation as a risk factor. First of all, SES is a poorly understood concept, and there are no standard principles for accurately defining SES or determining which variables need to be contained in SES (Baquet and Commiskey, 1999). Furthermore, the resolution of measuring SES has varied by the level of data that is available. Given the difficulty in collecting data at an individual level due to privacy rights, SES data are often measured at an aggregated level, such as a census tract level (Bigby and Holmes, 2005). Therefore, researchers suggest taking caution when using SES as a predictor for late-stage cancer diagnosis (Oakley-Girvan et al., 2003).

## 2.3 Cultural disparity

Previous studies of cancer diagnosis stage consider two factors of cultural disparity: cultural bias existing in some cultures that tends to cause negative opinions toward cancer screening, and cultural/linguistic isolation as a barrier in access to cancer screening or healthcare in general.

According to research about breast cancer, after controlling for other risk factors, cultural bias does affect the decision whether or not to utilise mammography screening, especially in

African-American women (Trock et al., 1993; Burns et al., 1996; Lannin et al., 2002). Coker et al. (1997) pointed out that cultural bias, such as fear of the side effects of radiation, commonly existed in African-American women, preventing them from seeking cancer screening. Another study suggested that after adjustment for other factors, cultural/linguistic isolation was also a significant barrier in receiving cancer screening (Goel et al., 2003).

In regard to prostate cancer, cultural/language isolation is especially important for immigrants (Williams and Horm, 1977; Blair and Fraumeni, 1978; Hakky et al., 1979). In particular, immigrants often have low usage of healthcare facilities. Cultural/language isolation might be the main reason, for people not born in the USA and not speaking English as their mother tongue, to have relatively lower screening rates than native-born Americans (Johnson, 2004).

A recent study suggests that cultural bias may be responsible for a lower colonoscopy screening rate and thus a higher late-stage diagnosis rate of colorectal cancer in Asian-Americans than Caucasians (Wen, 2007).

## 2.4 Race

Studies suggest that African-American women are more likely to present with late-stage breast cancer at time of diagnosis than Caucasian women in the USA (Ownby et al., 1985; Hunter et al., 1993; Eley et al., 1994). Breast cancer diagnosis in African-American women is also related to younger age and more advanced stages (Yost et al., 2001). In comparison with other ethnic groups, such as Asians and Hispanics, some studies found that African-American women are more consistently diagnosed with advanced stages of breast cancer (McCarthy et al., 1998; Lannin et al., 2002; Swanson et al., 2003). Other research indicates that African-American women and Hispanic women are both more inclined to be diagnosed with late-stage breast cancer (Bradley et al., 2002). However, after adjustment for other possible risk factors such as socioeconomic status (hereafter referred to as SES) and age of patient at time of diagnosis, there is conflicting evidence as to whether or not race actually plays a role in stage of breast cancer diagnosis: Richardson et al. (1992) and Lyman et al. (1997) found that race remained significant; and Bradley et al. (2001) concluded that it was not significant.

Prostate cancer is similar to breast cancer in that late-stage diagnosis is typically charted differently for race, depending on whether race is introduced as the only risk predictor or other risk predictors are involved. Several studies find that the odds of diagnosis with late-stage prostate cancer are much higher in African-American men, particularly in young men (Pienta et al., 1995; Powell et al., 1995; Brawley et al., 1998). Dayal and Chiu (1982) and Bennett et al. (1998) found that race did not have an important influence on late-stage prostate cancer diagnosis after adjustment for other risk factors.

For colorectal cancer, there are also contradictory conclusions about racial disparities in late-stage diagnosis. The findings of Roetzheim et al. (2000) are consistent with those of Wu et al. (2001) in that African-Americans were more likely to be diagnosed with late-stage colorectal cancer than Caucasians, given that colorectal cancer mortality rates have been declining among Caucasian patients but have been increasing in African-American patients at the national level in recent years. However, Dominitz et al. (1998) suggested that no racial disparity existed when equal accessibility to related healthcare was provided, suggesting that racial disparities are related to access to healthcare. Yet another study concluded that Caucasians tended to present with more advanced stages of colorectal cancer than African-Americans at time of diagnosis (Haas and Sark, 1997). Colonoscopy screening has been regarded as the major way to detect colorectal cancer, but several studies generated contrasting conclusions about whether different ethnic groups received equal opportunity for

colonoscopy screening. According to Paskett et al. (1997), African-Americans probably received colonoscopy screening more frequently than Caucasians. However, others found that African-Americans had less chance of receiving colonoscopy screening than Caucasian counterparts after other factors were controlled (Liff et al., 1991; Mandelblatt et al., 1996; Wu et al., 2001).

The studies reviewed above suggested that race as an independent variable in late-stage cancer diagnosis has not been sufficiently substantiated. It is also important to acknowledge that race is not a strict biological category, but rather a social category that encompasses social, economic and cultural variations (Williams, 1996; Liu, 1998). While some find that race plays a role in late-stage cancer diagnosis, they cannot decisively concur on why the results differ by race.

Researchers also have considered other factors in analysing late-stage cancer diagnosis. Some find that obesity is positively associated with late-stage breast cancer diagnosis (Ownby et al., 1985), but obesity can be partially explainable by SES (Cui et al., 2002). Several studies examine how age plays a role in women receiving mammograms and thus stage of breast cancer diagnosis (Polednak, 1986; Yancik et al., 1989; Mandelblatt et al., 1991; Randolph et al., 2002; Peek and Han, 2004). A couple of studies focused on the impact of gender on late-stage colorectal cancer diagnosis, but produced contradictory outcomes (Mandelblatt et al., 1996; Wu et al., 2001).

In this research, we group the risk factors for stage of cancer diagnosis into spatial and nonspatial factors. Spatial factors are factors related to patients' residential location in relation to healthcare services that affect the risk of late-stage cancer diagnosis. One spatial factor is distance or travel time to the nearest cancer screening facilities (data not available for prostate or lung cancer in this study). Another spatial factor is a comprehensive measure of spatial accessibility to primary healthcare. Nonspatial factors represent risk determinants related to population characteristics at the individual or aggregated level. Due to the limitations of our data, the nonspatial factors used in this research are demographic and socioeconomic attributes extracted from census data at an aggregated level (initially the census tracts and later interpolated to the zip code areas). Therefore, the study attempts to explain late-stage cancer diagnosis by neighbourhood (zip code) characteristics rather than individual attributes, and thus is ecological in nature (Robinson, 1950). Both spatial and nonspatial factors are included in the models to examine whether any factor plays a role in late-stage diagnosis of the four types of cancer.

### 3 Data issues and variable definitions

#### 3.1 Cancer data

The cancer incidence dataset is obtained from the Illinois State Cancer Registry (ISCR), Illinois Department of Public Health. This study uses the cancer data in 1998–2000 to match the census data in 2000. The ISCR dataset has the records of all Illinois patients, including those diagnosed in neighbouring states, such as Missouri and Wisconsin. Lehnerr and Havener (2002) estimated that the case ascertainment reached 98% completeness. The dataset includes individual records of cancer incidence, geocoded to zip code areas, with variables such as cancer type, age group, sex, race, diagnosis stage and year. Privacy and confidentiality prevent access to relevant cancer data from areas smaller than zip codes. Cases of late-stage cancer are defined as cases with cancer stage in the range from two to seven at time of diagnosis. This research focuses on the four leading types of cancer: breast, prostate, colorectal and lung. There were 1,245 zip code areas in Illinois in 2000.



Table 1 shows the basic statistics of these four types of late-stage cancer across zip code areas in Illinois during 1998–2000. From Table 1, prostate cancer has the lowest late-stage incidence, followed by breast cancer, colorectal cancer and lung cancer. As explained earlier, the high late-stage rate of lung cancer was due to lack of an effective early-stage diagnostic method.

### 3.2 Defining spatial access to primary care and cancer screening

In analysing spatial access to healthcare, we consider two types of healthcare services: primary care physicians and cancer screening facilities.

Primary care physicians include family physicians, general practitioners, general internists, obstetricians-gynecologists, and physician specialists such as oncologists (Cooper, 1994). Since physicians in primary healthcare play an important role in preventive care and cancer screening, accessibility to their facilities is a necessary consideration (Lee, 1995). Measuring accessibility to primary healthcare involves evaluating supply and demand of these physicians. On the supply side, the data for Illinois primary care physicians in 2000 were extracted from the physician master file of the American Medical Association (AMA). Because geocoding to specific streets cannot be done due to lack of geographic accuracy in the AMA data, we use the population-weighted centroid of each zip code to represent sites of primary healthcare whose general addresses locate within that zip code area. On the demand side, the Census 2000 dataset is used to compute the population-weighted centroids of census tracts representing locations of residents. This research utilises the result from a previous study focusing on measuring accessibility to primary healthcare in Illinois (Luo and Wang, 2003; Wang and Luo, 2005). That prior study developed and implemented a two-step floating catchment area (abbreviated as 2SFCA) method to account for the supply and demand match ratio within a region and the complex interaction between them through a transportation network. In essence, accessibility in a region is equivalent to the traditional measure of physicians-to-residents ratio. However, the contribution of physicians at a place to accessibility is discounted by the number of residents within a reasonable range from the physicians (i.e., crowdedness effect); and in the meantime, the total accessibility score of residents consists of cumulated effects of physicians within a reasonable range from the residents.

We use the shortest distance or travel time to the closest facility to measure spatial access to cancer screening services. We start with the Euclidean distance as a baseline, and estimate travel time based on the road networks as an improved measure. This approach does not differentiate the capacity or quality of the service providers and assumes that the service is available for all residents. With respect to breast cancer, the distance or travel time is from the nearest mammography screening facilities. There were 386 mammography screening facilities in Illinois in 2000. After eliminating duplicated addresses, 380 of those facilities are geocoded according to their address information. For colorectal cancer, the distance or travel time is from the closest hospital with colonoscopy screening. There were 196 hospitals with colonoscopy screening in Illinois in 2000, and 194 of those can be geocoded according to their addresses. There is no standard or routine screening test for prostate cancer ([www.cancer.gov/cancertopics/pdq/screening/prostate/Patient](http://www.cancer.gov/cancertopics/pdq/screening/prostate/Patient)) or lung cancer, so this measure of spatial access to cancer screening services is not available for these two types of cancer.

### 3.3 Consolidating nonspatial factors

The literature review indicates that nonspatial factors such as demographic and socioeconomic characteristics also affect healthcare access and likely late-stage cancer diagnosis. The great variation in a single risk factor's influence on late-stage cancer

diagnosis depends on whether or how many other risk factors are integrated into the research. It is critical to include as many variables as possible to capture the many dimensions of socio-economic and cultural variation. Again, we use the result from the previous study on healthcare access (Wang and Luo, 2005), which considered 11 variables, all from the census data. These include demographic variables (i.e., population with high healthcare needs including seniors of ages above 65, children of ages 0–4 and women of child-bearing ages 15–44), socioeconomic status (e.g., population in poverty, female-headed households, home ownership and median income), environment (e.g., households with an average of more than one person per room and housing units lack of basic amenities), linguistic barriers and education (e.g., non-white population, population without a high-school diploma and households linguistically isolated), and transportation mobility (e.g., households without vehicles).

Factor analysis is used to consolidate these variables. Table 2 shows how the 11 original variables are loaded on three major factors. Based on the loadings, the three factors are labelled as: socio-economic disadvantages (Factor 1), socio-cultural barriers (Factor 2), and healthcare needs (Factor 3). Each echoes the risk factors as highlighted in the literature review. Socioeconomic disadvantages may account for a large proportion of late-stage cancer cases at time of diagnosis. Cultural disparities and low education attainment may serve to decrease awareness of necessary cancer screening. Certain groups of people, depending on age, sex or others, may have more healthcare needs than other persons.

### 3.4 Transforming data to zip code areas

As discussed above, the cancer data are for zip code areas, and the spatial and nonspatial risk factors are defined at the census tract level. A basic geographic information systems (GIS) function, areal interpolation, is used to change datasets with various zonal systems into datasets having a common zonal system (Sadahiro, 1999). In this research, the two spatial access measures and the three nonspatial factors at the census tract level are transformed to those at the zip code area level by a simple interpolation method, area weighting interpolator [Goodchild and Lam, 1980; Wang, (2006), pp.47].

The choice of using zip code area as the geographic unit of analysis is dictated by the data available, as explained earlier. However, several limitations with zip code as the analysis unit need to be noticed. Boundaries of a zip code area are imprecise and probably vary over time (Kirby, 1996). The small size can easily result in small area problems when dealing with rare cases (Best et al., 2000). Moreover, a zip code area may not always be a social or neighbourhood area (Grubestic, 2006). Despite those limitations, zip code can provide a delicate spatial resolution and has been utilised in many health studies (Knapp and Hardwick, 2000; Ng et al., 1993; Parker and Campbell, 1998).

## 4 The models

This section describes the models for examining the association of late-stage diagnosis in each of the four types of cancer and risk factors, including spatial factors and nonspatial factors.

Poisson models, discovered nearly two centuries ago, have been proven effective for analysis of count data of rare events (Cameron and Trivedi, 1998). Poisson models have been used in many cancer studies (e.g., Frome, 1983; Whittemore and Gong, 1989; Tango, 1994; Wang et al., 1996; Kulldorff et al., 1997; Sheehan and DeChello, 2005), and Poisson regression is ‘the basic workhorse of statistics for analysing relationships’ [Griffith and Haining, (2005), pp.133]. In the Poisson model, because the total number of cancer cases in a zip code constrains the number of late stage cases, the total number of cases were included



as an offset variable (with a coefficient always equal to 1). The regular Poisson regression is described below:

$$O_i \sim \text{Poisson}(\mu_i) \quad (1)$$

where  $O_i$  represents the observed late-stage cancer counts in zip code  $i$ ;  $\mu_i$  indicates expected values of dependent variables in zip code  $i$ ; and observed late-stage cancer cases are assumed to fit the Poisson distribution in each zip code  $i$ . After that, expected late-stage cancer cases  $\mu_i$  are calculated as follows:

$$\log \mu_i = \log E_i + \alpha + \sum_k (\beta_k \times X_{ki}) \quad (2)$$

where  $E_i$  is the offset calculated as all stage cases at zip code  $i$ ,  $X_k$  represents the  $k$ -th risk factor, and  $\alpha$  and  $\beta_k$ 's are parameters to estimate. The model is calibrated by the *PROC GENMOD* module in SAS.

The most important concern with Poisson regression models is overdispersion, emerging when the prediction of models does not match realistic observations. The reasons are as follows: heterogeneity of cases between study areas; important independent variables missing from the model; and spatial autocorrelation between areas [Griffith and Haining, (2005), pp.133]. For the purpose of effectively reducing overdispersion, the Bayesian model with convolution priors has been designed to examine the association between rare events and independent variables (Besag et al., 1991; Mollie, 1996). The Bayesian model with convolution priors is similar to model (1) with a different formulation for the expected late-stage cancer cases  $\mu_i$ :

$$\log \mu_i = \log E_i + \alpha + \sum_k (\beta_k + X_{ki}) + b_i + h_i \quad (3)$$

where  $b_i$  can be viewed as a surrogate for unknown or unobserved variables with spatial structures, such as spatial autocorrelation between neighbourhoods, assigned a conditional autoregressive (hereafter referred to as CAR) prior;  $h_i$  captures the influence of all unknown or unobserved variables, which are assumed to be an exchangeable normal prior; and other notations are the same as in equation (2).

The major difference of this Bayesian model in (3) from the regular Poisson model in (2) is the inclusion of term  $b_i$  for controlling for spatial autocorrelation. Therefore, the model may be also referred to as a 'Poisson regression model controlling for spatial autocorrelation,' or simply 'spatial Poisson regression model'. Besag et al. (1991) argued that this improved model was more flexible than a model containing only CAR prior, given that the extra-Poisson variation can be divided into two parts: one that is spatially structured ( $b_i$ ) and the other representing unstructured variables ( $h_i$ ).

From equation (3), relative risks in zip code area  $i$ ,  $RR_i$ , can also be computed for mapping:

$$RR_i = \exp\left(\alpha + \sum_k (\beta_k + X_{ki}) + b_i + h_i\right) \quad (4)$$

$RR_i$  is the estimated place-specific late-stage cancer rate (per 100 of all stage cases) when risk factors and spatial autocorrelation are considered.

WinBUGS 1.4.2 is used to implement the Bayesian model with convolution priors (Spiegelhalter et al., 2002; Best and MRC Biostatistics, 2004). In particular, GeoBUGS 1.2.1 offers an important connection between the model and WinBUGS. GeoBUGS is a

carry-on module in WinBUGS, and it acts as an interface that has two functions: creating adjacency matrixes for inputting models in WinBUGS and mapping the outcome of those models.

Based on previous studies (Mollie, 1996; Law et al., 2005), the Bayesian model with convolution priors (or the spatial Poisson regression) shows advantages over the regular Poisson regression for rare events with small study unit. The Bayesian model considers spatial random effects which often contribute to overdispersion in Poisson regression models. Meanwhile, the utilisation of convolution priors can tell researchers whether or not spatial autocorrelation remains in the model. Specifically, if spatial random effects dominate in the Bayesian model, the analysis outcomes can alert researchers that errors exist in spatial variables or that some important spatial variables have been omitted.

For comparing performances of the two models, this research uses the deviance information criterion (hereafter referred to as DIC) as a generalisation of Akaike's information criterion (Spiegelhalter et al., 2002). The DIC is a natural way to compare complex models with prior distributions in that it is based on the posterior distribution of the log-likelihood, following the Bayesian model framework built by Dempster (1974). The DIC builds a trade-off between the data fit of the model and the complexity of the model. A smaller DIC value indicates a better data fit and less complicated model (Best and MRC Biostatistics, 2004). This model comparison criterion has been successfully applied in the field of medical statistics (Zhu and Carlin, 2000). For each type of late-stage cancer in this paper, the DIC value for the two models is obtained from WinBUGS 1.4.2.

## 5 Results and discussions

Table 3 presents the results of analysing late-stage diagnosis for the four leading types of cancer in Illinois using the Bayesian model with convolution priors and the regular Poisson regression model. In the table, the variables that are statistically significant are denoted by asterisks, and the smaller DIC values between the two models are also highlighted (in italics). We have tested the models for breast and colorectal cancer using both distance and travel time from the nearest cancer screening facility. The results for distance are consistent with the ones using travel time. Only the models using travel time are presented as travel time is considered a more accurate measure of spatial barriers than distance. The findings are summarised as follows.

### 5.1 Is there any difference between the two models?

Yes, there is a difference between the two models, but that difference is minimal. First, the regression coefficients and corresponding statistical significance levels are generally consistent between the two models. This shows that our observations on the relationship between late-stage cancer diagnosis and each risk factor as well as its significance (to be discussed next) are robust. Secondly, the deviation information criterion (DIC) is calibrated to test the fit and complexity of each model. A smaller value of DIC indicates a model better suited for the data with less complexity. In all cases, the differences in DIC between the two models are small. Except for the case of prostate cancer, the DIC values for the regular Poisson regression are slightly lower than the ones for the Bayesian model with convolution priors. That is to say, the regular Poisson regression is slightly favoured over the Bayesian model with convolution priors for breast, colorectal and lung cancer; but the opposite is true for prostate cancer. In other words, the gain from the attempt to control for spatial autocorrelation by the Bayesian model with convolution priors is minimal in this study. This does not necessarily mean that spatial autocorrelation does not exist in the data or that the Bayesian model is inferior to the regular Poisson model. The Bayesian model is much more complex and requires more coding and computation power to implement. For our study, the

results do not support its advantages over the regular Poisson regression after considering the balance between data fit and model complexity.

## 5.2 Does spatial access to primary care physicians or to cancer screening matter in stage of cancer diagnosis?

Neither the spatial access to primary care physicians or travel time from the nearest cancer screening facility is statistically significant in the model for any type of cancer. An earlier study (Wang et al., 2008) showed that spatial access for either primary care or cancer screening was mainly an issue for rural areas. Since the data set is dominated by samples from the urban areas, the effect of spatial access is not evident in this study.

## 5.3 Do nonspatial factors affect late-stage cancer diagnosis?

Yes, but the results vary among cancer types. Both socioeconomic disadvantages and socio-cultural barriers have a significant influence on late-stage breast cancer diagnosis. The coefficients of the two risk factors are positive, suggesting that people residing in disadvantaged areas, such as neighbourhoods with low SES or high concentration of minorities or immigrants, be more likely to be diagnosed with late-stage breast cancer. This observation is consistent with previous studies that reveal a strong positive association between late-stage breast cancer diagnosis and low SES or cultural barriers (Burns et al., 1996; Lannin et al., 2002; McWilliams et al., 2003; Roetzheim et al., 1999; Trock et al., 1993). Only the factor of socioeconomic disadvantages has a statistically significant influence on late-stage prostate cancer diagnosis. People residing in neighbourhoods with lower socioeconomic status are more likely to be diagnosed with late-stage prostate cancer, supporting the findings from Morris et al. (1999) and Klassen et al. (2005). None of the three nonspatial factors is significant for the late-stage diagnosis of colorectal or lung cancer. In other words, the risk of late diagnosis is not associated with socioeconomic and cultural characteristics of neighbourhoods for colorectal and lung cancer. From the earlier discussion, majorities of colorectal and lung cancer cases are diagnosed at late stages in Illinois, reflecting national trends. For lung cancer, it is explainable by lack of effective early detection methods. However, most colorectal cancer cases have missed the opportunity for early detection for two possible reasons. One may be lack of awareness of screening services for colorectal cancer, and the other may be the nature of colonoscopy screening procedure. Colonoscopy is considered the most effective diagnosis method for colon cancer but is also perceived by many as costly and inconvenient. Implications for public policy are discussed in the conclusion section.

## 5.4 Can we predict high risk areas for late-stage cancer diagnosis?

Not really, and the confidence of predicting the high risk areas is the lowest for colorectal and lung cancer. As explained earlier, WinBUGS uses the function in equation (4) to compute and map area-specific relative risks. Like in any regression model, the reliability of prediction depends on the explanatory power of independent variables. From Table 3, none of the risk factors is significant for colorectal or lung cancer, and thus estimated risks for these two types of cancer would not be very meaningful. Accordingly, the confidence of risk estimates based on the model is the highest for breast cancer, and then for prostate cancer.

For example, Figure 1 shows the relative risks for late-stage breast cancer across zip code areas in Illinois. A relative risk value larger than 1 indicates that on average, considering the risk factors (i.e., the spatial access and non spatial factors), the estimated late-stage breast cancer rate among people in that zip code area is greater than the statewide average rate. In Figure 1, there are a total of 425 zip code areas having the value of relative risk larger than one. This map shows that zip code areas having the above-average relative risk are spread out over both urban and rural areas. The highest values of relative risk have the largest

cluster of zip code areas in the southern Chicago area, as shown in the inset. The southern Chicago area has the state's highest concentrations of disadvantaged population groups (minorities and population in poverty).

## 6 Conclusions

For various types of cancer, tumor stage at time of diagnosis plays a vital role in determining health outcomes. This paper examines the risk factors for four leading types of cancer in Illinois, namely breast, prostate, colorectal and lung cancer. The related risk factors are divided into two groups: one group is spatial factors of a geographic nature, including accessibility to primary healthcare, and distance or travel time from the nearest screening facilities; the other group is nonspatial factors consolidated into three factors such as socioeconomic disadvantages, socio-cultural barriers, and healthcare needs. This study uses the Illinois State Cancer Registry data at the zip code area level in 1998–2000. The research uses two models to analyse the relationship between late-stage cancer diagnosis and risk factors: the regular Poisson regression that is commonly used for analysis of count data of rare events, and the Bayesian model with convolution priors to control for possible spatial autocorrelation. The results from the two models are generally consistent with each other. The Bayesian model is much more complex and requires more coding work and computation power to implement. Our results do not support its advantages over the regular Poisson regression after considering the balance between data fit and model complexity. However this conclusion is based only on the analysis of late-stage cancer in Illinois. For other kinds of health issues in other settings, the more complex spatial Poisson model may be justified.

The results for breast cancer suggest that people living in neighbourhoods with socioeconomic disadvantages and cultural barriers are more likely to be diagnosed at a late stage. In regard to prostate cancer, people in regions with low socioeconomic status are more likely to be diagnosed at a late stage. Diagnosis of late-stage colorectal or lung cancer is not significantly associated with any of the aforementioned risk factors. Due to the limitation of the data, our analysis examines the risk factors at the zip code area level, not at the individual level. Therefore, we cannot confidently assert that any areal relationships we find between neighbourhood characteristics and cancer stage are reflective of individual relationships, commonly known as 'ecological fallacy'. In any case, the results suggest that residents in disadvantaged neighbourhoods (foremost socio-economically deprived, and secondarily socio-culturally isolated) in general suffer from a higher-risk for diagnosis of more advanced stages clearly in breast cancer, to a lesser degree in prostate cancer, but not necessarily in colorectal or lung cancer.

To achieve ultimate equality in healthcare access, utilisation and outcome is an ambitious goal that will take time to reach. Once an early cancer detection method becomes available, its awareness and adoption are a gradual process probably resembling the innovation adoption curve of Rogers (2003). The ones with better healthcare access and awareness are the most likely to be early beneficiaries. Among the four types of cancer, early detection for breast cancer is perhaps most mature and has the longest history in terms of availability of medical technology and public awareness, and the research shows that people of more privileged have taken better advantage of the screening opportunities. A similar trend seems to emerge for prostate cancer. However, this is not the case for colorectal cancer even though the technology for colorectal cancer screening (e.g., colonoscopy) has been available for some years. It is possibly attributable to lack of public awareness and education of early detection of colorectal cancer. Finally, the high rate of late-stage lung cancer and the high mortality rate associated with it certainly call for the need of developing effective early detection methods for lung cancer.

The drawbacks of this paper also need to be recognised. Using zip code area as the analysis unit, artifact results affected by the heterogeneous zip code size are inevitable. Also limited by the areal unit, there is no accurate way to calibrate travel time or distance used in measures of spatial access to cancer screening facilities and to primary healthcare because each zip code area is approximated by its centroid. Travel time is also estimated for a planning purpose, and thus in no way represents actual trip time that may be affected by real travel elements such as rush hours, days of the week, and emergent traffic conditions. As for residents depending on public transportation, estimate of travel time would require additional data such as public transit routes, which are not feasible for this study. Additionally, if the study area is divided into urban versus rural areas, Chicago versus non-Chicago areas, more localised relationships between late-stage cancer diagnosis and risk factors might be revealed. Most importantly, this research uses aggregate data at the zip code area level and thus is ecological in nature. When cancer data of individual cases with more demographic and socioeconomic attributes become available, a multilevel model may be used to analyse the risk factors of individuals and possible interaction between individual and neighbourhood characteristics.

## Acknowledgments

Financial support from the National Cancer Institute (NCI), National Institutes of Health (NIH), under Grant 1-R21-CA114501-01, is gratefully acknowledged. Points of view or opinions in this paper are those of the authors, and do not necessarily represent the official position or policies of NCI. When this research was conducted, Wang and Luo were at Northern Illinois University.

## References

- American Cancer Society. Cancer Facts & Figures 2007. National Home Office: American Cancer Society; Atlanta, Georgia: 2007. available at <http://www.cancer.org/downloads/STT/CAFF2007PWSecured.pdf>
- Amey CH, Miller MK, Albrecht SL. The role of race and residence in determining stage at diagnosis of breast cancer. *Journal of Rural Health*. 1997; 13:99–108. [PubMed: 10169323]
- Baquet CR, Commiskey P. Socioeconomic factors and breast carcinoma in multicultural women. *American Cancer Society*. 1999; 88(5):1256–1264.
- Barry J, Breen N. The importance of place of residence in predicting late-stage diagnosis of breast or cervical cancer. *Health and Place*. 2005; 11:15–19. [PubMed: 15550353]
- Bennett CL, Ferreira MR, Davis TC, Kaplan J, Weinberger M, Kuzel T, Seday MA, Sartor O. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. *Journal of Clinical Oncology*. 1998; 16(9):3101–3104. [PubMed: 9738581]
- Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. 1991; 1:1–59.
- Best, NG. MRC Biostatistics Unit. WinBUGS. Cambridge, UK: 2004. available at <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>
- Best NG, Ickstadt K, Wolpert RL. Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association*. 2000; 95(452):1076–1088.
- Bigby JA, Holmes MD. Disparities across the breast cancer continuum. *Cancer Causes and Control*. 2005; 16:35–44. [PubMed: 15750856]
- Blagojevich, RR.; Whitaker, EE. Every Breath Counts. Illinois Department of Public Health; Springfield, Illinois: 2006. Cancer in Illinois — Lung Cancer Facts. at [http://www.idph.state.il.us/cancer/types/publications\\_lung.htm](http://www.idph.state.il.us/cancer/types/publications_lung.htm)
- Blair A, Fraumeni JF. Geographic patterns of prostate cancer in the United States. *Journal of the National Cancer Institute*. 1978; 61:1379–1384. [PubMed: 281545]
- Bradley CJ, Given CW, Caralee R. Disparities in cancer diagnosis and survival. *American Cancer Society*. 2001; 91(1):178–188.

- Bradley CJ, Given CW, Caralee R. Race, socioeconomic status and breast cancer treatment and survival. *Journal of the National Cancer Institute*. 2002; 94:490–496. [PubMed: 11929949]
- Brawley OM, Knopf K, Merrill R. The epidemiology of prostate cancer part I: descriptive epidemiology. *Seminars in Urology Oncology*. 1998; 16(4):187–192.
- Brown HS, Gobe R, Kirschner H. Social and environmental factors in lung cancer mortality in post-war Poland. *Environmental Health Perspectives*. 1994; 103(1):64–70.
- Burns RB, McCarthy EP, Freund KM, Sandra L, Marwill MD, Shwartz M, Ash A, Moskowitz MA. Black women receive less mammography even with similar use of primary care. *Annals of Internal Medicine*. 1996; 125:173–82. [PubMed: 8686974]
- Cameron, A.; Trivedi, P. *Regression Analysis of Count Data*. Cambridge University Press; Cambridge: 1998.
- Coker AL, Crane MM, Sticca RP, Sepkovic DW. Re: ethnic differences in estrogen metabolism in healthy women [Letter]. *Journal of the National Cancer Institute*. 1997; 89(89)
- Cooper RA. Seeking a balanced physician work-force for the 21st century. *Journal of the American Medical Association*. 1994; 272:680–687. [PubMed: 8064983]
- Coughlin SS, Thompson TD, Hall HI, Logan P, Uhler RJ. Breast and cervical carcinoma screening practices among women in rural and nonrural areas of the United States, 1998–1999. *Cancer*. 2002; 94:2801–2812. [PubMed: 12115366]
- Cui Y, Whiteman MK, Langenberg P, Sexton M, Tkczuk KH, Flaws JA, Bush TL. Can obesity explain the racial differences in stage of breast cancer diagnosis between black and white women? *Journal of Women's Health & Gender-Based Medicine*. 2002; 11(6):527–536.
- Cummings DM, Whitestone LM, Earp JA, Mayne L. Disparities in mammography screening in rural areas: analysis of county differences in North Carolina. *Journal of Rural Health*. 2002; 18:77–83. [PubMed: 12043758]
- Dayal HH, Chiu C. Factors associated with racial differences in survival for prostatic carcinoma. *Journal of Chronic Disease*. 1982; 35:553–560.
- Dempster, AP. *Proceedings of the Conference on Foundational Questions in Statistical Inference*. Department of Theoretical Statistics, University of Aarhus; 1974. The direct use of likelihood for significance testing; p. 335-352.
- Devesa, SS.; Grauman, DJ.; Blot, WJ.; Pennello, GA.; Hoover, RN.; Fraumeni, JF. *Atlas of Cancer Mortality in the United States: 1950–94*. National Institute of Health, National Cancer Institute; 1999. (NIH Publication No.99-4564)
- Dominitz JA, Gregory PS, Lansman P, Provenzale D. Race, treatment, and survival among colorectal carcinoma patients in an equal-access medical system. *American Cancer Society*. 1998; 82(12): 2312–2320.
- Eley JW, Hill HA, Chen VW, Austin DF, Wesley MN, Muss HB, Greenberg RS, Coates RJ, Correa P, Redmond CK. Racial differences in survival from breast cancer, results of the National Cancer Institute black/white cancer survival study. *The Journal of the American Medical Association*. 1994; 272(12):947–954.
- Freeman HP, Alshafie TA. Colorectal carcinoma in poor blacks. *Cancer*. 2002; 94:2327–2332. [PubMed: 12015756]
- Frome EL. The analysis of rates using Poisson regression models. *Biometrics*. 1983; 39:665–674. [PubMed: 6652201]
- Goel MS, Wee CC, McCarthy EP, Davis RB, Ngo-Metzger Q, Phillips RS. Racial and ethnic disparities in cancer screening: the importance of foreign birth as a barrier to care. *Journal of General Internal Medicine*. 2003; 18:1028–1035. [PubMed: 14687262]
- Goodchild MF, Lam NS. Areal interpolation: a variant of the traditional spatial problem. *Geoprocessing*. 1980; 1:297–331.
- Griffith DA, Haining R. Beyond mule kicks: the Poisson distribution in geographical analysis. *Geographical Analysis*. 2005; 38:123–139.
- Grubestic, TH. Zip codes and spatial analysis: problems and prospects. Paper presented at the Association of American Geographers Annual Meeting; Chicago, Illinois. 2006.
- Guagliardo MF. Spatial accessibility of primary care: concepts, methods and challenges. *International Journal of Health Geography*. 2004; 3(1):3.



- Haas GP, Sark WA. Epidemiology of prostate cancer. *Cancer Journal for Clinicians*. 1997; 47(5):273–287.
- Hakky SI, Chiholm GD, Skeet RG. Social class and carcinoma of the prostate. *The International British Journal of Urology*. 1979; 51:393–396.
- Haynes R. The urban distribution of lung cancer mortality in England and Wales 1980–1983. *Urban Studies*. 1988; 25:497–506.
- Hoffman M, Pinho DH, Cooper D, Sayed R, Dent DM, Gudgeon A, Zyl VJ, Rosenberg L, Shapiro S. Breast cancer incidence and determinants of cancer stage in the Western Cape. *South African Medical Journal*. 2000; 90:1212–1216. [PubMed: 11234652]
- Hsia J, Kemper E, Kiefe C, Zapka J, Sofaer S, Pettinger M, Bowen D, Limacher M, Lillington L, Mason E. The importance of health insurance as a determinant of cancer screening: evidence from the women's health initiative. *American College of Preventive Medicine*. 2000; 31(3):261–270.
- Hunter CP, Redmond CK, Chen VW. members of the Black/White Cancer Survival Study Group . Breast cancer: factors associated with stage at diagnosis in black and white women. *Journal of the National Cancer Institute*. 1993; 85:1129–1137. [PubMed: 8320742]
- Johnson GD. Small area mapping of prostate cancer incidence in New York State (USA) using fully Bayesian hierarchical modelling. *The International Journal of Health Geographics*. 2004; 3(29):1–12.
- Joslyn SA, West MM. Racial differences in breast carcinoma survival. *Cancer*. 1999; 88(1):114–123. [PubMed: 10618613]
- Kirby R. Toward congruence between theory and practice in small area analysis and local public health data. *Statistics in Medicine*. 1996; 15:1859–1866. [PubMed: 8888478]
- Klassen AC, Kulldorff M, Curriero F. Geographic clustering of prostate cancer grade and stage at diagnosis, before and after adjustment for risk factors. *International Journal of Health Geographics*. 2005; 4(1):1–16. [PubMed: 15649329]
- Knapp KK, Hardwick K. The availability and distribution of densities in rural zip codes and primary care health professional shortage areas (PC-HPSA) zip codes: comparisons with primary care providers. *Journal of Public Health Dentistry*. 2000; 60:43–49. [PubMed: 10734616]
- Kulldorff M, Feuer JE, Miller BA, Freedman LS. Breast cancer clusters in the Northeast United States: a geographic analysis. *American Journal of Epidemiology*. 1997; 146(2):161–170. [PubMed: 9230778]
- Lannin DR, Mathews HF, Mitchell J, Swanson MS. Impacting cultural attitudes in African-American women to decrease breast cancer mortality. *American Journal of Surgery*. 2002; 18:418–423. [PubMed: 12433605]
- Law J, Haining R, Maheswaran R, Pearson T. Analyzing the relationship between smoking and coronary heart disease at the small area level: a Bayesian approach to spatial modeling. *Geographical Analysis*. 2005; 38:140–159.
- Lee P. Health system reform and generalist physician. *Academic Medicine*. 1995; 70:S10–S13. [PubMed: 7826450]
- Lehnerr, M.; Havener, L. Assessment of Interstate Exchange of Cancer Data: Illinois, 1986–1998. Illinois State Department of Public Health; Springfield, Illinois: 2002.
- Liff JM, Chow WH, Greenberg RS. Rural-urban differences in stage at diagnosis. Possible relationship to cancer screening. *Cancer*. 1991; 67(5):1454–1459. [PubMed: 1991313]
- Liu ET. The uncoupling of race and cancer genetics. *Cancer*. 1998; 83:1765–1769.
- Luo W, Wang FW. Measure of spatial accessibility to healthcare in a GIS environment: synpaper and a case study in the Chicago region. *Environmental and Planning B: Planning and Design*. 2003; 30:865–884.
- Lyman GH, Kuderer NM, Lyman SL, Cox CE, Reintgen D, Baekey P. Importance of race on breast cancer survival. *Annals of Surgical Oncology*. 1997; 4(1):80–87. [PubMed: 8985521]
- Mandelblatt J, Andrews H, Kao R, Wallace R, Kerner J. The late-stage diagnosis of colorectal cancer: demographic and socioeconomic factors. *American Journal of Public Health*. 1996; 86:1794–1797. [PubMed: 9003140]

- Mandelblatt J, Andrews H, Kerner J, Zauber A, Burnett W. Determinants of late-stage diagnosis of breast and cervical cancer: the impact of age, race, social class, and hospital type. *American Journal of Public Health*. 1991; 81:646–649. [PubMed: 2014871]
- 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(9):729–736. [PubMed: 9556466]
- McWilliams JM, Zaslavsky AM, Meara E, Ayanian JZ. Impact of Medicare coverage on basic clinical services for previously uninsured adults. *The Journal of the American Medical Association*. 2003; 290:757–764.
- Menck HR, Mills PK. The influence of urbanization, age, ethnicity, and income on the early diagnosis of breast carcinoma: opportunity for screening improvement. *Cancer*. 2001; 92:1299–1304. [PubMed: 11571746]
- Merkin SS, Stevenson L, Powe N. Geographic socioeconomic status, race, and advantaged-stage breast cancer in New York City. *American Journal of Public Health*. 2002; 92(1):64–70. [PubMed: 11772763]
- Mollie, A. Bayesian mapping of disease. In: Gilks, WR.; Richardson, S.; Spiegelhalter, DJ., editors. *Markov Chain Monte Carlo in Practice*. Chapman & Hall; New York: 1996. p. 359–379.
- Montella M, Biondi E, Marco DM, Botti G, Tatangelo F, Capasso I, Marone A. Sociodemographic factors associated with the diagnostic staging of breast cancer in Southern Italy. *Cancer*. 1995; 76:1585–1590. [PubMed: 8635062]
- Morris CR, Snipes KP, Schlag R, Wright WE. Sociodemographic factors associated with prostatectomy utilization and concordance with the physician data query for prostate cancer (United States). *Cancer Causes and Control*. 1999; 10:503–511. [PubMed: 10616820]
- Ng E, Wilkins R, Perras A. How far is it to the nearest hospital? Calculating distances using the statistics Canada postal code conversion file. *Health Report*. 1993; 5:179–188.
- Nosek MA, Howland CA. Breast and cervical cancer screening among women with physical disabilities. *American Journal of Physical Medicine & Rehabilitation*. 1997; 78:39–44.
- Oakley-Girvan I, Kolonel N, Gallagher RP, Wu AH, Felberg A, Whittemore AS. Stage at diagnosis and survival in a multiethnic cohort of prostate cancer patients. *American Journal of Public Health*. 2003; 93(10):1753–1759.
- Ownby HE, Frederick J, Russo J, Brook SC, Swanson GM, Heppner GH, Brennan MJ. Racial differences in breast cancer patients. *Journal of National Cancer Institute*. 1985; 75(1):55–60.
- Palmer RC, Schneider EC. Social disparities across the continuum of colorectal cancer: a systematic review. *Cancer Causes and Control*. 2005; 16:55–61. [PubMed: 15750858]
- Parker EB, Campbell JL. Measuring access to primary medical care: some examples of the use of geographical information systems. *Health and Place*. 1998; 4:183–193. [PubMed: 10671022]
- Paskett RJ, Agostino DR, Tatum C, Velez A. Cancer screening behaviours of low-income women: the impact of race. *Women's Health*. 1997; 3:203–226.
- Peek ME, Han JH. Disparities in screening mammography. Current status, interventions, and implications. *The Journal of General Internal Medicine*. 2004; 19:184–94.
- Pienta KJ, Demeers R, Hoff M, Kau TY, Montie JE, Severson RK. Effect of age and race on the survival of men with prostate cancer in the metropolitan Detroit tricounty area, 1973 to 1987. *Urology*. 1995; 45:93–102. [PubMed: 7817485]
- Polednak AP. Breast cancer in black and white women in New York State. Case distribution and incidence rates by clinical stage at diagnosis. *Cancer*. 1986; 58:807–815. [PubMed: 3731032]
- Powell IJ, Schwartz K, Hussain M. Removal of the financial barrier to healthcare: does it impact on prostate cancer at presentation and survival? A comparative study between black and white men in a veterans affairs system. *Urology*. 1995; 46:825–830. [PubMed: 7502424]
- Randolph WM, Goodwin JS, Mahnken JD, Freeman JL. Regular mammogram use is associated with elimination of age-related disparities in size and stage of breast cancer. *Annual of International Medicine*. 2002; 137:783–790.

- Richardson JL, Langholz B, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. *British Journal of Cancer*. 1992; 65(6):922–926. [PubMed: 1616865]
- Robinson WS. Ecological correlations and the behaviour of individuals. *American Sociological Review*. 1950; 15:351–357.
- Roetzheim RG, Naazneen P, Eduardo GC, Ferrante JM, Van DDJ, Krischer JP. Effects of health insurance and race on colorectal cancer treatments and outcomes. *American Journal of Public Health*. 2000; 90(11):1746–1754. [PubMed: 11076244]
- Roetzheim RG, Naazneen P, Colleen T, Voti L, Ayanian JZ, Schwabe A, Krischer JP. Effects of health insurance and race on early detection of cancer. *Journal of the National Cancer Institute*. 1999; 91(16):1409–1415. [PubMed: 10451447]
- Rogers, EM. *Diffusion of Innovations*. 5. Free Press; New York: 2003.
- Rohan TE, Jain MG, Howe GR, Miller AB. Dietary folate consumption and breast cancer risk. *Journal of the National Cancer Institute*. 2000; 92:266–269. [PubMed: 10655445]
- Rushton G, Peleg I, Banerjee A, Smith G, West M. Analyzing geographic Patterns of disease incidence: rates of late-stage colorectal cancer in Iowa. *Journal of Medical Systems*. 2004; 28(3): 223–235. [PubMed: 15446614]
- Sadahiro Y. Accuracy of areal interpolation: a comparison of alternative methods. *Journal of Geographical Systems*. 1999; 1(4):323–346.
- Sheehan TJ, DeChello LM. A space-time analysis of the proportion of late-stage breast cancer in Massachusetts, 1988 to 1997. *International Journal of Health Geographics*. 2005; 4(15):1–9. [PubMed: 15649329]
- Spiegelhalter D, Best N, Carlin B, Linde AVD. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*. 2002; 64:583–640.
- Swanson GM, Haslam SZ, Azzouz F. Breast cancer among young African-American women. A summary of data and literature and of issues discussed during the summit meeting on breast cancer among African American women, Washington DC, September 8–10, 2000. *Cancer*. 2003; 97:273–279. [PubMed: 12491491]
- Tango T. Effect of air pollution on lung cancer: a Poisson regression model based on vital statistics. *Environmental Health Perspective*. 1994; 102(8):41–45.
- Thomas A, Carlin BP. Late detection of breast and colorectal cancer in Minnesota counties: an application of spatial smoothing and clustering. *Statistics in Medicine*. 2003; 22:113–127. [PubMed: 12486754]
- Trock B, Rimer BK, King E, Balshem A, Cristinazio CS, Engstrom PF. Impact of an HMO-based intervention to increase mammography utilization. *Cancer Epidemiology Biomarkers & Prevention*. 1993; 2:151–156.
- Wang, F. *Quantitative Methods and Applications in GIS*. CRC Press; London: 2006.
- Wang F, Luo W. Assessing spatial and nonspatial factors in healthcare access in Illinois: towards an integrated approach to defining health professional shortage areas. *Health and Place*. 2005; 11(2): 131–146. [PubMed: 15629681]
- Wang F, McLafferty S, Escamilla V, Luo L. Late-stage breast cancer diagnosis and healthcare access in Illinois. *The Professional Geographer*. 2008; 60(1):54–69. [PubMed: 18458760]
- Wang P, Puterman ML, Cockburn I, Le H. Mixed Poisson regression models with covariates dependent rates. *Biometrics*. 1996; 52:381–400. [PubMed: 10766499]
- Wen, M. Low rates of cancer screening in Asian Americans compared with whites: evidence from California. Paper presented at the research seminar for the Huntsman Cancer Institute Cancer Control and Population Sciences group; March 21, 2007; 2007.
- Whittemore AS, Gong G. Poisson regression with misclassified counts: application to cervical cancer mortality rates. *Applied Statistics*. 1989; 40(1):81–93. [PubMed: 12157994]
- Williams DR. Race/ethnicity and socioeconomic status: measurement and methodological issues. *International Journal of Health Service*. 1996; 26:483–505.
- Williams RR, Horm JW. Association of cancer sites with tobacco and alcohol consumption and socioeconomic status of patients: interview study from the third national cancer survey. *Journal of the National Cancer Institute*. 1977; 58:525–547. [PubMed: 557114]

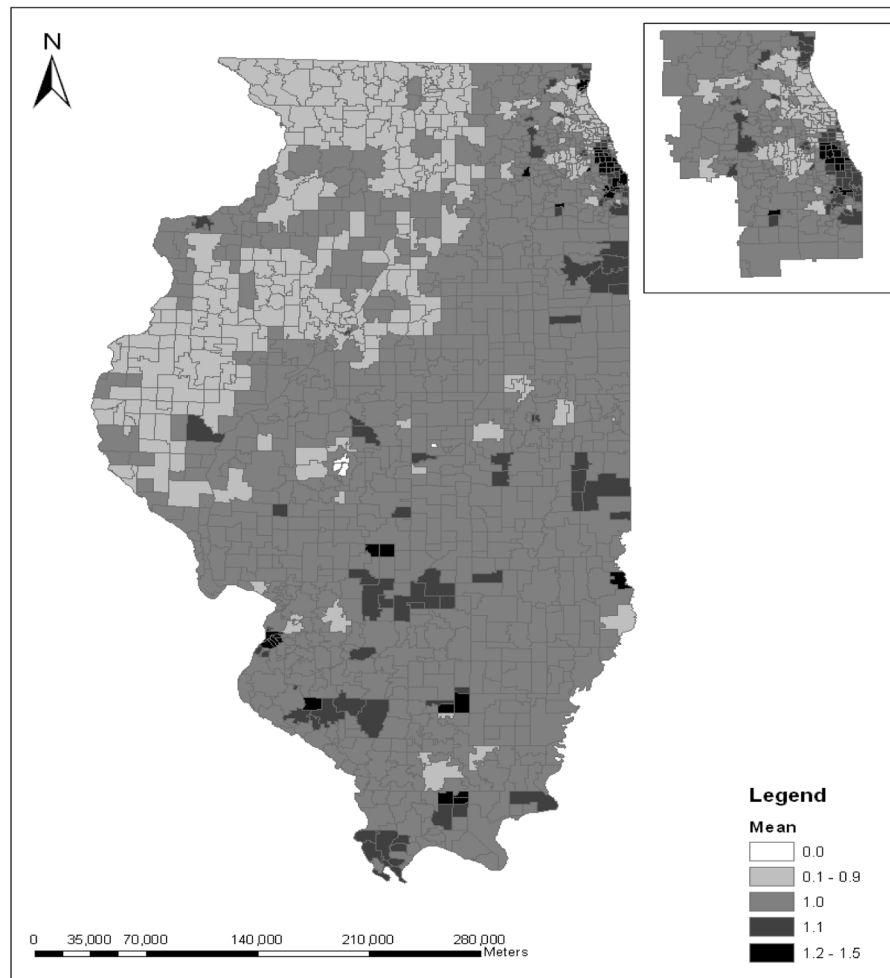
- Wu XC, Chen VW, Steele B, Ruiz B, Fulton J, Liu L, Carozza SE, Greenlee R. Subsite-specific incidence rate and stage of disease in colorectal cancer by race, gender, and age group in the United States, 1992–1997. *American Cancer Society*. 2001; 92(10):2547–2554.
- Yancik R, Ries LG, Yates JW. Breast cancer in aging women. A population-based study of contrasts in stage, surgery, and survival. *Cancer*. 1989; 63:976–981. [PubMed: 2914302]
- Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes and Control*. 2001; 12:703–711. [PubMed: 11562110]
- Zenk SN, Tarlov E, Sun J. Spatial equity in facilities providing low – or no-fee screening mammography in Chicago neighbourhoods. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*. 2006; 83(2):195–210. [PubMed: 16736369]
- Zhang S, Hunter DJ, Hankinson SE. A prospective study of folate intake and the risk of breast cancer. *The Journal of the American Medical Association*. 1999; 281:1632–1637.
- Zhu L, Carlin BP. Comparing hierarchical models for spatio-temporally misaligned data using the deviance information criterion. *Statistics in Medicine*. 2000; 19:2265–2278. [PubMed: 10960852]

## Biographies

**Biographical notes:** Fahui Wang is a Professor in the Department of Geography and Anthropology at Louisiana State University, Baton Rouge, LA. His research interests include GIS and spatial analysis applications in health and crime research, urban and regional analysis, and China studies.

Lan Luo is a PhD student in the Department of Geography at the University of Illinois at Urbana-Champaign, Urbana, IL. Her research interests include GIS, spatial analysis, and medical geography.

Sara McLafferty is a Professor in the Department of Geography at the University of Illinois at Urbana-Champaign, Urbana, IL. Her research interests include the geographies of health and healthcare, urban geography, and spatial analysis methods/GIS.



**Figure 1.** Relative risk of late-stage breast cancer in Illinois (inset for Chicago region)

**Table 1**

Four types of late-stage cancer in Illinois 1998–2000 at zip code level

Cancer type	Total late-stage cases (%)	No. late-stage cases across zip code areas		
		Mean	Std. error	Minimum Maximum
Breast	8,240 (28.57)	6.62	11.7	0 81
Colorectal	9,864 (58.60)	7.92	15.47	0 109
Prostate	2,884 (15.36)	2.32	4.79	0 47
Lung	16,925 (68.31)	13.59	22.23	0 145



**Table 2**

Loadings of variables on three factors

<b>Contents</b>	<b>Socio-economic disadvantages</b>	<b>Socio-cultural barriers</b>	<b>Healthcare needs</b>
Female-headed households (%)	<i>0.9089</i>	-0.0058	0.0504
Population in poverty (%)	<i>0.8662</i>	0.1642	0.2405
Non-white minorities (%)	<i>0.8481</i>	0.2153	-0.0153
Households w/o vehicles (%)	<i>0.8231</i>	0.1905	0.2699
Home ownership (%)	<i>-0.6686</i>	-0.3362	-0.3922
Housing units lacking basic amenities (%)	<i>0.4278</i>	0.2703	-0.0323
Households with linguistic isolation (%)	-0.0479	<i>0.9561</i>	0.0164
Households with >1 person per room (%)	0.4464	<i>0.7966</i>	-0.0631
Population w/o high-school diploma (%)	0.5800	<i>0.6406</i>	0.1219
Population with high healthcare needs (%)	0.0316	-0.1050	<i>0.9186</i>
Median income (\$)	-0.5491	-0.2053	<i>-0.5605</i>
% of variance explained by each factor	<i>54.08</i>	<i>28.09</i>	<i>17.83</i>

**Table 3**  
Results of regular Poisson regression and Bayesian model with convolution priors on late-stage cancer diagnosis

	Risk factors	Breast cancer	Prostate cancer	Colorectal cancer	Lung cancer
Bayesian model with convolution priors	Intercept	0.080* (0.040)	0.007 (0.071)	0.016 (0.040)	-0.016 (0.021)
	Spatial access to primary care	-27.37 (16.12)	-32.64 (30.62)	2.211 (12.47)	-4.543 (11.82)
	Travel time to nearest screening facility	0.0829 (0.001)		-0.0017 (0.001)	
	Socio-economic disadvantages	0.115*** (0.0205)	0.126** (0.040)	-0.0112 (0.0189)	-0.0208 (0.0135)
	Socio-cultural barriers	0.076*** (0.018)	0.007 (0.032)	0.0130 (0.016)	-0.0092 (0.013)
	Healthcare needs	-0.0178 (0.019)	0.0013 (0.034)	-0.0135 (0.017)	-0.0138 (0.014)
	DIC	3346.26	2173.23	2768.59	4269.91
	Intercept	-1.170*** (0.034)	-1.820*** (0.050)	-0.541*** (0.033)	-0.402*** (0.019)
	Spatial access to primary care	-26.084 (11.45)	-16.570 (18.332)	7.207 (10.626)	7.741 (7.434)
	Travel time to nearest screening facility	0.007 (0.001)		-0.0014 (0.001)	
Regular Poisson regression	Socio-economic disadvantages	0.131*** (0.016)	0.115*** (0.024)	-0.015 (0.014)	-0.023 (0.011)
	Socio-cultural barriers	0.075*** (0.016)	0.026 (0.027)	0.016 (0.014)	-0.007 (0.012)
	Healthcare needs	-0.019 (0.016)	-0.034 (0.028)	-0.018 (0.015)	-0.026 (0.012)
	DIC	3340.59	2184.64	2742.79	4245.80

Note: Standard errors in parentheses;

\* significant at 0.05,

\*\* significant at 0.01,

\*\*\* significant at 0.001; smaller DICs between the two models are highlighted in italics.