Spatial and Spatio-temporal Epidemiology 26 (2018) 143–151

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

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

Original Research

Kernel density analysis reveals a halo pattern of breast cancer incidence in Connecticut

Natalya Rybnikova[a], Richard G. Stevens[b,*], David I. Gregorio[b], Holly Samociuk[b], Boris A. Portnov[a]

[a] Department of Natural Resources and Environment Management, Faculty of Management, University of Haifa, Haifa, Israel
[b] Department of Community Medicine, School of Medicine, University of Connecticut, Farmington, CT 06030, United States

ARTICLE INFO

Article history:
Received 7 August 2017
Revised 30 April 2018
Accepted 8 June 2018
Available online 19 June 2018

Keywords:
Breast cancer (BC)
Artificial light-at-night (ALAN)
Double kernel density (DKD) analysis
Connecticut

ABSTRACT

Breast cancer (BC) incidence rates in Connecticut are among the highest in the United States, and are unevenly distributed within the state. Our goal was to determine whether artificial light at night (ALAN) played a role. Using BC records obtained from the Connecticut Tumor Registry, we applied the double kernel density (DKD) estimator to produce a continuous relative risk surface of a disease throughout the State. A multi-variate analysis compared DKD and census track estimates with population density, fertility rate, percent of non-white population, population below poverty level, and ALAN levels. The analysis identified a "halo" geographic pattern of BC incidence, with the highest rates of the disease observed at distances 5-15 km from the state's major cities. The "halo" was of high-income communities, with high ALAN, located in suburban fringes of the state's main cities.

1. Introduction

Breast cancer (BC) is the second leading cause of cancer-related death among women in the US after lung cancer (ACS, 2017a). Several US states have age-adjusted BC rates above 130 per 100,000, with the state of Connecticut being among them (see Fig. 1).

Socio-demographic and life-style factors, such as alcohol consumption, excess body weight, lack of physical activity, age, ethnicity, family history and shift work are associated with BC incidence (ACS, 2017; Gorham et al., 1988; Brody et al., 2007). An emerging theory is that BC is also associated with exposure to artificial light at night (ALAN) (see *inter alia* Stevens, 1987; Stevens & Rea, 2001; Navara & Nelson, 2007; Stevens, 2009; Blask et al., 2011; Haim and Portnov, 2013). Possible mechanisms behind this association include the suppression of nocturnal melatonin production (Blask et al., 2011; Stevens & Zhu, 2015); daily rhythms disruption due to night-time activities enabled by ALAN (Haim & Portnov, 2013; Navara & Nelson, 2007; Stevens & Rea, 2001), and general stress associated with ALAN exposure (Haim & Portnov, 2013; Navara & Nelson, 2007; Zubidat et al., 2007).

Portnov et al. (2016) investigated factors behind BC incidence in Connecticut. The study revealed cumulated BC incidence rates in 2005–2009 ranged across census tracts of the state from 0 to 890 cases per 100,000, although the analysis detected no clear geographic pattern of BC incidence in the study area. Variation was found to be linked to ALAN exposure, percentage of urban population, percent population below poverty and town fertility rates.

The failure of epidemiological studies that utilize disease rates calculated for statistical areas established for purposes other than health investigations is not surprising (Torre et al., 2012). Data averaging associated with the use of these statistical divisions often results in considerable information loss such that no clear patterns of disease incidence can be detected (Portnov et al., 2009).

* Corresponding author.
E-mail address: bugs@uchc.edu (R.G. Stevens).

https://doi.org/10.1016/j.sste.2018.06.003
1877-5845/

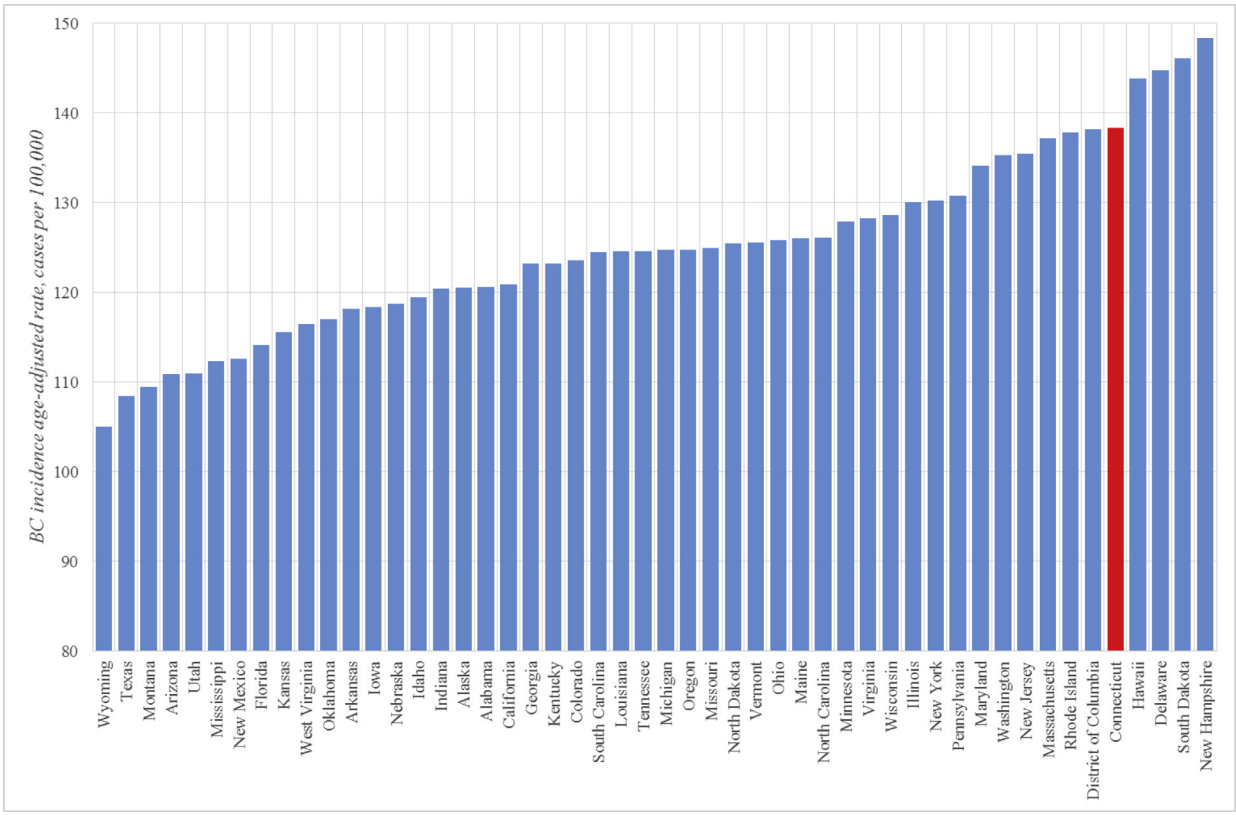

Fig. 1. BC incidence age-adjusted rates in the US states in 2013 (cases per 100,000) (Source: NCI, 2017).

Recently developed geo-information tools provide alternative solutions for investigating associations of environmental factors with disease (Zusman et al., 2016). Double kernel Density (DKD) analysis is one such tool widely used in environmental studies. The DKD is a nonparametric smoothing technique that uses event point data and produces a continuous surface of geographically varying relative risk estimates by normalizing the aerial density of the event according to the density of population at risk from which the disease events are drawn (Kloog et al., 2009; Portnov et al., 2009; Zusman et al., 2012; Zusman et al., 2016). In recent studies, DKD estimators have been used for cancer analysis (Kloog et al., 2009; Portnov et al., 2009; Shi, 2009; Zusman et al., 2012), analysis of asthma (Svechkina and Portnov, 2017), avian influenza (Minh et al., 2009) and foot-and-mouth disease (Wilesmith et al., 2003).

This paper is a follow-up to a recent ecological study by Portnov et al. (2016). Using a similar set of research variables to investigate factors affecting BC incidence, we apply a variety of alternative geo-informatics tools, attempting to identify the environmental and locational factors behind the visually "patched" and seemingly randomized patterns of BC incidence in the study area.

2. Research methods

2.1. Data sources

We obtained BC incidence data from the Connecticut Tumor Registry (CTR) for 829 census tracts for the period 2005 to 2009. Presuming a 10–15 year latency period for BC onset (Hoover et al., 1976; Aschengrau et al., 1998; Rybnikova et al., 2015), we matched these data with ALAN intensity levels obtained from the US-DMSP database for 1996–97 (NOAA, 2017). In addition, we retrieved information on potential BC confounders, including population density, fertility rate, percentage of non-white population, urban population and population below poverty from the U.S. Census database for the year 2000. Table 1 reports descriptive statistics of the research variables.

2.2. BC rate calculations and data processing

Two alternative BC incidence measures were calculated for each census tract: (1) observed BC rates and (2) simulated DKD estimates of BC incidence rates. We calculated the former as the ratio between the total number of BC cases recorded in 2005-2009 and the total number of women in the census tract (Zusman et al., 2012). To obtain the DKD estimates, we first calculated kernel densities of BC cases and of female population per km^2 using the *kernel density* tool in ArcGIS™10.x (ArcGIS, 2017*a*).

To those values, we applied the *extract multi values to points* tool (ArcGIS, 2017b), to assign BC kernel densities and density of female population to census tracts' centroids. For each of those points, we calculated ratios between BC kernel densities and non-zero female population densities, following the methodology described in Zusman et al. (2012). Both observed and simulated rates,