Cancer Epidemiology 38 (2014) 248–252

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Interactive effects of individual and neighborhood race and ethnicity on rates of high-grade cervical lesions

Christina Waggaman [a], Pamela Julian [b], Linda M. Niccolai [a,b,*]

[a] Yale School of Public Health, Department of Epidemiology of Microbial Diseases, 60 College Street, New Haven, CT 06520, USA
[b] Yale School of Public Health, Connecticut Emerging Infections Program, One Church Street, New Haven, CT 06520, USA

ARTICLE INFO

Article history:
Received 18 September 2013
Received in revised form 27 February 2014
Accepted 10 March 2014
Available online 4 April 2014

Keywords:
Health disparities
Race
Ethnicity
High grade cervical lesions
Interaction
Synergy

ABSTRACT

We estimated the main and interactive effects of individual race/ethnicity (black, Hispanic, white) and area race, ethnicity, and poverty (proportions of the female population black, Hispanic, and living below the federal poverty level at the census tract level, respectively) on rates of high-grade cervical lesions among young women. Using data from a statewide surveillance system during 2008–2011, we found a marginally significant interaction ($P < 0.05$) between individual race/ethnicity and area race, with black and Hispanic women living in areas with ≥20% of the female population black having elevated rates compared to those living in areas with <20% of the female population black. These findings indicate a possible synergistic effect between individual race/ethnicity and racial composition in neighborhoods on precancerous cervical lesions.

1. Introduction

Disparities in cervical cancer incidence and mortality have persisted in the United States despite overall declines in recent decades [1,2]. High-grade cervical lesions (HGCL) are precursors to invasive cervical carcinoma and an important public health problem for several additional reasons including their high burden among young women, associated health care costs, and psychological distress [3,4]. Over half of HGCL will persist or progress in the absence of treatment which typically involves an outpatient excisional surgical procedure to remove abnormal tissue [5]. Previous work has shown disparities in the burden of HGCL by area measures of race and poverty [6], but we are not aware of any studies that have assessed possible synergy between area and individual measures of race/ethnicity on cervical cancer or its precursors.

Combining census data with geocoded surveillance data can be used to explore geographic differences in health outcomes to better elucidate socioeconomic disparities [7]. Specifically, analyses done at the neighborhood level (e.g., census tracts) can inform targeted public health interventions for specific communities [8,9]. Area characteristics such as the physical environment, quality of housing, availability of services, socio-cultural factors, and perceived reputation by outsiders can lead to disparities in health outcomes for resident individuals [10]. However, studies that solely assess risk at the area-level are unable to disentangle health risks caused by the area itself from those caused by the characteristics of the individuals who inhabit it. As such, multilevel analysis has increasingly become a valuable method for analyzing geographic health differences, as it allows for the amount of risk "explained" by area characteristics to be measured while controlling for individual level risk factors [11].

Previous research using surveillance data in multilevel models has shown neighborhood socioeconomic variables to be risk factors for health outcomes such as all-cause mortality, obesity, risky sexual behavior, and cancer survival after controlling for individual risk factors [12–17]. Although less researched, interactive effects between area and individual measures may also be present and are potentially important when considering public health interventions; significant interaction between individual demographic variables and neighborhood socioeconomic variables has been shown to be associated with risk of low birth-weight and ischemic stroke [18,19].

While no research thus far has explored synergy between individual and area-level race in development of HGCL, studies have used multilevel models to demonstrate that both area and individual levels of race/ethnicity and socioeconomic status are

* Corresponding author at: Yale School of Public Health, Department of Epidemiology of Microbial Diseases, 60 College Street, New Haven, CT 06520, USA. Tel.: +1 203 785 7834.
E-mail address: linda.niccolai@yale.edu (L.M. Niccolai).

http://dx.doi.org/10.1016/j.canep.2014.03.004
1877-7821/

associated with cervical cancer mortality. One study discovered that the measure of proportion of black residents was associated with mortality after controlling for individual race/ethnicity, age at diagnosis, cancer stage, and surgery [20]; another study showed that neighborhood deprivation was a significant risk factor for cervical cancer mortality after controlling for individual socioeconomic variables [21]. The goal of the present analysis was to examine the interaction between individual race/ethnicity and area measures of racial, ethnic, and income composition at the census tract level specifically on the rates of HGCL.

2. Methods

Methods for this project have been previously described [6]. Briefly, the Connecticut Department of Public Health added HGCL including cervical intraepithelial neoplasia grades 2 and 3 and adenocarcinoma in situ to the list of mandatory reportable diseases in 2008. All pathology laboratories in the state report diagnostic information for all cases as well as patient demographics including residential address using a one-page case report form. Cases were geocoded to the census tract level and linked to 2010 US Census data. Census tracts are widely used as proxies for neighborhood, and were thus used as the area unit of analysis [22]. Enhanced surveillance activities for women residing in New Haven County between the ages of 18 and 39 years include brief telephone interviews and medical record reviews to collect missing demographic information and vaccination histories. All data undergo extensive quality control procedures.

Surveillance data were analyzed from January 1, 2008, through December 31, 2011, years for which reporting is complete, for women aged 20–39 years in New Haven County. Women ages 18 and 19 were excluded because different cervical cancer screening guidelines may result in different rates of detection, and women aged 20 years were included to create age categories consistent with surveillance reports as described below. Data were restricted to cases from New Haven County because of more rigorous data collection (patient interviews and medical record reviews) for demographic information. New Haven County has a total population of 862,477 including 14% black and 15% Hispanic residents.

Individual race/ethnicity measures obtained from surveillance reports, medical record reviews, and interviews were combined to form a single race/ethnicity variable, which consisted of three categories: non-Hispanic black (subsequently referred to as 'black'), Hispanic, and non-Hispanic white (subsequently referred to as 'white'). We used a hierarchical classification scheme when race/ethnicity was available from more than one data source that prioritized interview responses over medical records over surveillance reports. Of women with non-missing race/ethnicity, interview data were most complete (64%) followed by medical records (53%) and surveillance reports (23%). Concordance between each pair of the three data sets was >90%. Age, an important determinant of rates of HGCL [3], was obtained from surveillance reports and grouped into the following five-year categories: 20–24, 25–29, 30–34, and 35–39 years.

Area measures of race, ethnicity, and poverty at the census tract level were obtained from 2010 US Census data and included percentage of the female population black (≥20% black or <20% black), percentage of the female population Hispanic (≥20% Hispanic or <20% Hispanic), and percentage of the female population living below the federal poverty level (≥20% in poverty or <20% in poverty), respectively. These cut-points were chosen based on the Public Health Disparities Geocoding Project with the three lower levels combined to maximize differences for comparison [19]. Of the 189 census tracts in the county, 23% had ≥20% black and 29% had ≥20% Hispanic, and 20% had ≥20% in poverty.

The data were fit to a multilevel Poisson random effects model with the following variables: individual race/ethnicity, area race, area ethnicity, area poverty, individual age, and census tract. Age was included to adjust for potential confounding and census tract was included as a random effect to control for correlation between strata in the same census tract. The four age and three individual racial/ethnic groups were stratified so that each of the 189 census tracts was divided into 12 strata. Stratum numerators were the counts of cases of HGCL for each age and race/ethnicity combination. Stratum denominators were the total female population in each census tract for each age and race/ethnicity combination (the 12 strata) and multiplied by four, the number of years during which surveillance data were collected.

To assess the main effect of individual and area measures, the following model was fit (Model 1):

$$\log(\lambda_{ij}) = \log(\eta_{ij}) + \beta_0 + \beta_i(\text{age}) + \beta_i(\text{race/ethnicity}) + \beta_j(\text{CT race}) + \beta_j(\text{CT ethnicity}) + \beta_j(\text{CT poverty}) + \mu_j$$

where i represents stratum membership, j represents variables at the census tract-level, η_i represents the person-time in each stratum, and υ_j represents the random effect due to each census tract. To assess interaction between individual race/ethnicity and area race, ethnicity, and poverty, three interaction terms were added and the model was fit as (Model 2):

$$\begin{aligned}\log(\lambda_{ij}) = {} & \log(\eta_{ij}) + \beta_0 + \beta_i(\text{age}) + \beta_i(\text{race/ethnicity}) \\ & + \beta_j(\text{CT race}) + \beta_j(\text{CT ethnicity}) + \beta_j(\text{CT poverty}) \\ & + \beta_{ij}(\text{race/ethnicity})(\text{CT race}) \\ & + \beta_{ij}(\text{race/ethnicity})(\text{CT ethnicity}) \\ & + \beta_{ij}(\text{race/ethnicity})(\text{CT poverty}) + \mu_j\end{aligned}$$

All of the statistical analysis was conducted using SAS v 9.2 (SAS Institute Inc., Cary, NC).

3. Results

During 2008–2011, 2338 cases of HGCL were reported among women aged 20–39 years in New Haven County and considered for inclusion in this analysis. Of those, 1359 (58.1%) had non-missing individual race/ethnicity information; cases with missing race/ethnicity were more likely to be from census tracts with low percentages of the female population black and Hispanic but were not different regarding area poverty and individual age. The analyzed sample included 21.7% (n = 295) black, 29.1% (n = 395) Hispanic, and 49.2% (n = 669) white cases. All cases had complete information on census tract residence; 37%, 39%, and 26% of cases were residents of census tracts with ≥20% black, Hispanic, and in poverty, respectively. Black women had the highest rates (475 cases per 100,000) followed by Hispanic women (456 cases per 100,000); white women had the lowest rates (253 cases per 100,000). Rates for race/ethnicity groups by area measures of race, ethnicity, and poverty are presented in Table 1.

In the analysis of main effects (Table 2), black and Hispanic women had significantly higher rates of HGCL as estimated by the rate ratio (RR) than white women (RR = 1.70, 95% confidence interval (CI): 1.42, 1.99 and RR = 1.64, 95% CI: 1.40, 1.87, respectively). Women residing in census tracts with ≥20% black had significantly higher rates compared to women in census tracts with <20% black (RR = 1.23, 95% CI: 1.02, 1.47). Rates did not differ significantly by area ethnicity (RR = 1.07, 95% CI: 0.91, 1.27) or area poverty (RR = 0.90, 95% CI: 0.74, 1.09).