Annals of Epidemiology 24 (2014) 104–110

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Local geographic variation in chronic liver disease and hepatocellular carcinoma: contributions of socioeconomic deprivation, alcohol retail outlets, and lifestyle

Jacqueline M. Major PhD [a,b,*], James D. Sargent MD [c], Barry I. Graubard PhD [a], Heather A. Carlos MS [c], Albert R. Hollenbeck PhD [d], Sean F. Altekruse DVM, PhD [e], Neal D. Freedman PhD [a], Katherine A. McGlynn PhD [a]

[a] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[b] Office of Surveillance and Epidemiology, US Food and Drug Administration, Silver Spring, MD
[c] Cancer Control Research Program, Norris Cotton Cancer Center, Dartmouth College, Lebanon, NH
[d] AARP, Washington, DC
[e] Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

ARTICLE INFO

Article history:
Received 3 May 2013
Accepted 11 November 2013
Available online 21 November 2013

Keywords:
Neighborhood
Socioeconomic disparities
Health care
Liver cancer
Liver disease
Cohort
Kernel density estimation
Multilevel
Census

ABSTRACT

Purpose: Hepatocellular carcinoma (HCC) incidence rates continue to increase in the United States. Geographic variation in rates suggests a potential contribution of area-based factors, such as neighborhood socioeconomic deprivation, retail alcohol availability, and access to health care.
Methods: Using the National Institutes of Health-American Association of Retired Persons Diet and Health Study, we prospectively examined area socioeconomic variations in HCC incidence ($n = 434$ cases) and chronic liver disease (CLD) mortality ($n = 805$ deaths) and assessed contribution of alcohol outlet density, health care infrastructure, diabetes, obesity, and health behaviors. Hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated from hierarchical Cox regression models.
Results: Area socioeconomic deprivation was associated with increased risk of HCC incidence and CLD mortality (HR, 1.48; 95% CI, 1.03–2.14 and HR, 2.36; 95% CI, 1.79–3.11, respectively) after accounting for age, sex, and race. After additionally accounting for educational attainment and health risk factors, associations for HCC incidence were no longer significant; associations for CLD mortality remained significant (HR, 1.78; 95% CI, 1.34–2.36). Socioeconomic status differences in alcohol outlet density and health behaviors explained the largest proportion of socioeconomic status-CLD mortality association, 10% and 29%, respectively. No associations with health care infrastructure were observed.
Conclusions: Our results suggest a greater effect of area-based factors for CLD than HCC. Personal risk factors accounted for the largest proportion of variance for HCC but not for CLD mortality.

Published by Elsevier Inc.

Introduction

The prevalence of chronic liver disease (CLD) has stabilized in the United States [1], yet the incidence rates of liver cancer have tripled over the past three decades [2]. Due to a very poor prognosis, liver cancer is the seventh most common cause of cancer death in the United States. Hepatocellular carcinoma (HCC) is the major histologic subtype, accounting for up to 80% of all cases [3,4]. In the United States, major risk factors include chronic infection with hepatitis B and hepatitis C viruses, excessive alcohol consumption, obesity, and type 2 diabetes mellitus [5].

Although the risk of HCC is known to vary by sex, age, race, and ethnic group, relatively little is known about whether there is variability by socioeconomic status (SES). Individual-level SES-related factors (e.g., educational attainment) as well as area-level SES conditions (i.e., neighborhood socioeconomic deprivation and availability of health care) may contribute to disparities in the incidence and mortality rates of the disease [6,7]. Geographic disparities in HCC incidence [8] and mortality [9] as well as in HCC survival and use of surgical therapy [10] have been reported. Geographic disparity in liver disease might reflect inequalities in health care infrastructure [11] and SES [12,13], as well as differences in individual-level characteristics such as age, race, sex, educational attainment, obesity [14], and health risk behaviors (heavy alcohol consumption, physical inactivity, and smoking) [15]. Few studies, however, have quantified small-area variation (geographic disparity) in HCC incidence and examined the extent to which area-level socioeconomic conditions

Conflict of interest statement: The authors declare no conflict of interest.

* Corresponding author. Division of Epidemiology, Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20993. Tel.: +1 2404020342.
E-mail address: jacqueline.major@fda.hhs.gov (J.M. Major).

1047-2797/$ – see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.annepidem.2013.11.006

account for geographic variation in liver disease. Minority and underserved communities often share an unequal burden of liver disease [16,17]. A recent study found that median household income was associated with better survival among treated HCC patients from high- versus low-income counties in the United States. [18]. Yet, only one study to our knowledge has specifically examined local area-level (census tract) socioeconomic deprivation and HCC [19]. Although this study found an association using the Surveillance, Epidemiology and End Results registry database, the socioeconomic characteristics of the area were measured close to the time of cancer diagnosis, which may not reflect earlier socioeconomic deprivation, data on individual educational attainment, and important HCC risk factors, such as history of diabetes and alcohol consumption were not to be assessed, and additional area-level SES measures were not examined. Furthermore, no multilevel studies have quantified the extent to which obesity and behavioral factors and residential area factors explain CLD mortality and HCC incidence disparities by SES. The purpose of this study was to investigate whether area socioeconomic factors, including neighborhood deprivation, alcohol outlet density, and availability of health care, account for small-area variation after adjustment for individual-level risk factors and assess the extent to which health behaviors and obesity may contribute to socioeconomic disparities in HCC incidence and CLD mortality using the large, prospective National Institutes of Health-American Association of Retired Persons (NIH-AARP) Diet and Health Study cohort.

Materials and methods

Study population

Details of the NIH-AARP study have been previously described [20]. Briefly, the NIH-AARP Diet and Health study is a prospective cohort study of AARP members. In 1995–1996, baseline questionnaires including a food frequency questionnaire were mailed to AARP members living in six U.S. states (CA, FL, LA, NJ, NC, and PA) and two metropolitan areas (Atlanta, GA; Detroit, MI). Of the 566,402 participants who completed the baseline questionnaires, we excluded proxy respondents and those who were previously diagnosed with cancer. To allow for appropriate energy-adjustment of reported dietary intake, we further excluded those with total energy intake more than twice the interquartile range of log-transformed energy intake (<1% of the sample). The resulting cohort included 494,988 participants: 294,965 men and 199,432 women. The NIH-AARP Diet and Health Study was approved by the Special Studies Institutional Review Board of the U.S. National Cancer Institute; participants gave informed consent by virtue of completing and returning the questionnaire.

Case ascertainment

Cancer cases in the study population were identified through linkage with the cancer registries in the states of residence and in Arizona and Texas, which were the major sites of relocation. The cancer registries were certified by the North American Association of Central Cancer Registries as being at least 90% complete within 2 years of cancer occurrence and the National Death Index Plus [21]. Primary liver cancer cases (C22.0) were identified from study entry through December 2006. HCC morphology was determined using *International Classification of Diseases for Oncology, Third Edition* [22]. HCC was identified using morphology codes 8170–8175.

Ascertainment of vital status

Vital status and cause of death were obtained by linking cohort participants to the U.S. Social Security Administration Death Master File and National Death Index [20]. Study participants were followed for overall vital status from recruitment (1995–1996) to December 31, 2008. Underlying cause of death codes were provided as *International Classification of Diseases* (*ICD*)-*9* and *ICD-10* codes. For the identification of deaths due to CLD, we used the classification scheme of the National Center for Health Statistics [23]. Participants ($n = 805$) were classified as having died from CLD if their *ICD* codes included those for CLD, liver fibrosis and cirrhosis, alcoholic liver disease, and chronic hepatitis (*ICD-9*: 571.0, 571.2–571.9; *ICD-10*: K70, K73, and K74). Persons who died from HCC were not included in the CLD mortality analyses.

Contextual measures

We linked the addresses of study participants to census data from 2000, obtained from the U.S. Census Bureau at the census tract level. As previously reported [24], we used the census tract data to generate a composite index of neighborhood socioeconomic deprivation for tracts covered in the NIH-AARP Study based on 10 census variables (percentage of persons in the census tract who had less than high-school education, were unemployed, on public assistance, in managerial jobs by gender, as well as the percent of households headed by a female, nonwhite, without a car, had annual income of <$30,000, or with incomes below 1999 federal poverty levels) shown to explain the most variance using a principal components analysis approach. The index score was derived for census tracts using the weighted sum of the coefficients for each variable and standardized to have a mean of 0 and standard deviation of 1. In our analyses, the index was categorized into quintiles based on the distribution of the study's census tracts ($n = 18{,}603$); the upper quintile being the highest level of neighborhood deprivation. We also examined measures of socioeconomic deprivation individually as a sensitivity check, including percent families below the poverty level, percent unemployed, percent with no high-school diploma, and the Gini coefficient as a measure of income inequality [25]. Because census tracts do not necessarily correspond to boundaries that affect alcohol-purchasing patterns, we calculated a nationwide density surface for alcohol outlets using adaptive bandwidth kernel density estimation and the LandScan Global Population Database (National Laboratory, Oak Ridge, TN), which estimates the underlying population based on satellite imagery. Adaptive bandwidth kernel density estimation allows the influence (the bandwidth) of each alcohol outlet to be limited to a surrounding population or in our case, 1000 people (about the number of people that would be needed to support a small retailer). We then used this density surface to calculate the mean alcohol outlet density within each polygonal census tract. The resultant density unit was then scaled to alcohol outlets per 1000 people [26].

The NIH-AARP data were also linked to a national county-level health resource information database, the Area Resource File (www.arf.hrsa.gov), to examine measures of health care access and availability and other socioeconomic-related indicators, including % foreign born, % language isolation, number of hospitals and physicians, % without health insurance, and % Medicare beneficiaries. Population densities of hospitals and physicians were calculated by dividing the number of hospitals and physicians by the population of the county according to the United States Census Population Estimates (per 10,000 individuals). Quantiles were based on the distribution of the study's counties ($n = 401$).

Individual characteristics

Information on demographic, medical history, and lifestyle factors, including alcohol consumption, cigarette use, physical activity, and dietary intake, were ascertained at baseline through a mailed