Published in final edited form as:

Ann Epidemiol. 2014 February ; 24(2): 104–110. doi:10.1016/j.annepidem.2013.11.006.

Local Geographic Variation in Chronic Liver Disease and Hepatocellular Carcinoma: Contributions of Socioeconomic Deprivation, Alcohol Retail Outlets, and Lifestyle

Jacqueline M. Major, PhD^{1,2}, James D. Sargent, MD³, Barry I. Graubard, PhD¹, Heather A. Carlos³, Albert R. Hollenbeck, PhD⁴, Sean F. Altekruse, DVM, PhD⁵, Neal D. Freedman, PhD¹, and Katherine A. McGlynn, PhD¹

¹Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

²Office of Surveillance and Epidemiology, US Food and Drug Administration, Silver Spring, MD

³Cancer Control Research Program, Dartmouth Medical School, Lebanon, NH

⁴AARP, Washington DC

⁵Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

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 healthcare.

Methods—Using the NIH-AARP 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, healthcare 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. Additionally accounting for educational attainment and health-risk factors, associations for HCC incidence were no longer significant; associations for CLD mortality remained (HR=1.78, 95% CI: 1.34–2.36). SES differences in alcohol outlet density and health behaviors explained the largest proportion of SES-CLD mortality association, 10% and 29%, respectively. No associations with healthcare 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.

Corresponding author: Jacqueline M. Major, PhD, Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20993. jacqueline.major@fda.hhs.gov.

Conflict of interest: None declared.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

neighborhood; socioeconomic disparities; healthcare; liver cancer; liver disease; cohort; kernel density estimation; multilevel; census

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 U.S. Hepatocellular carcinoma (HCC) is the major histological subtype, accounting for up to 80% of all cases [3, 4]. In the U.S., major risk factors include chronic infection with hepatitis B and hepatitis C viruses, excessive alcohol consumption, obesity, and type 2 diabetes mellitus [5].

While 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. Individual-level socioeconomic status (SES) related factors (e.g., educational attainment) as well as area-level SES conditions (i.e., neighborhood socioeconomic deprivation and availability of healthcare) 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 utilization of surgical therapy [10] have been reported. Geographic disparity in liver disease might reflect inequalities in healthcare infrastructure [11] and socioeconomic status [12, 13], as well as differences in individual-level characteristics such as age, race, sex, educational attainment and 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 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 U.S. [18]. Yet, only one study to our knowledge has specifically examined local area-level (census tract) socioeconomic deprivation and HCC [19]. Though this study found an association using the Surveillance, Epidemiology and End Results (SEER) 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 be assessed, and additional area-level SES measures were not examined. Further, 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 healthcare, account for small-area variation after adjustment for individual-level risk factors and to assess the extent to which health behaviors and obesity may contribute to socioeconomic disparities in HCC incidence and CLD mortality using the large, prospective 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 1995–96, baseline questionnaires including a food frequency questionnaire were mailed to AARP members living in six US states (California, Florida, Louisiana, New Jersey, North Carolina, and Pennsylvania) and two metropolitan areas (Atlanta, Georgia; Detroit, Michigan). Out 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 major 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 (ICD-O) 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 utilized 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, non-white, 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 (ARF) (www.arf.hrsa.gov), to examine measures of healthcare 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 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 life-style factors, including alcohol consumption, cigarette use, physical activity, and dietary intake, were ascertained at baseline through a mailed questionnaire sent to members of the cohort in 1995. A validated 124-item food frequency questionnaire (<http://riskfactor.cancer.gov/DHQ/>) was used to assess a participant's usual dietary intake in the prior 12 months, as well as to estimate energy and nutrient intakes. Daily intake of food was energy-adjusted using the nutrient density method [27]. Total meat included beef, pork, chicken, turkey, and fish. Self-reported body weight (lbs) and height (ft-inches) were used to derive body mass index [weight (kg)/height (m²)]. Vigorous physical activity was defined as 5 or more times of vigorous activity lasting at least 30 minutes per week.

Statistical analysis

We applied a multilevel Cox proportional hazards regression model with a frailty that includes a census-tract-level random effect to estimate the geographic variations in, and fixed effects of, census tract socioeconomic deprivation on HCC incidence and CLD mortality with person-years as the underlying time metric [28]. Participants were followed from date baseline questionnaire was returned to the date of HCC diagnosis (i.e., through December 31, 2006 for HCC incidence), moving out of study area, death, or end of study follow-up (December 31, 2008 for vital status), whichever came first. The proportional hazards assumption was assessed using cross-product terms for interaction between deprivation and follow-up time; no significant deviations from proportionality were found. The hazard ratio (HR) and 95% confidence interval (CI) were calculated for each variable in the multivariable Cox models. We first examined risk estimates adjusted for age at study entry (continuous), sex, and race. Fully adjusted models additionally included baseline measures of education, marital status, history of diabetes, smoking (never, former 20 cigarettes/d, former >20 cigarettes/d, current 20 cigarettes/d, current >20 cigarettes/d, missing), self-reported health status (excellent/great, good, fair/poor), body mass index (<18.5, 18.5 to <25, 25 to <30, 30 kg/m²), alcohol consumption (none, 0 to <1, 1 to 3, 3 drinks/d), energy from non-alcohol sources, energy-adjusted total meat, fruit, vegetable intake, and saturated fat, all of which altered risk estimates by 10% or more. Tests for interaction between socioeconomic indicators and race and sex, were performed by additionally including cross-product terms in the model. Tests for trend were based on quantile-specific median values entered as a continuous term in regression model.

We further examined the magnitude by which alcohol outlet density, diabetes, obesity, and behavioral risk factors explained SES-related differences in risk of chronic liver disease

mortality. For these analyses, neighborhood socioeconomic deprivation (quintiles of the standardized index scores) was analyzed as a continuous variable. The percent contribution was calculated according to methodology for intermediate endpoints proposed by Freedman, Graubard, and Schatzkin [29]. Leaving-one-out jackknife variance estimation with 75 random groups was used to estimate the standard errors of the percent change and further used to obtain 95% CIs [30]. All tests were twosided and P -values <0.05 were considered statistically significant. All analyses were performed using SAS® (version 9.3, SAS Institute Inc., Cary, NC).

RESULTS

In the present study, 434 individuals in 4,814,247 person-years of follow-up were diagnosed with HCC through December 2006 and 805 individuals (in 5,832,355 person-years) died of chronic liver disease during the maximum follow-up period of 13 years, which was consistent with expectations. Compared to non-cases, individuals with HCC and those dying of CLD were more likely to be men, have diabetes, consume 3 drinks per day, be obese, currently smoke, and be less likely to have completed college or be physically active (Table 1). In addition, those with CLD were less likely to be married than non-cases. Those dying of CLD had rates of smoking and alcohol use that were close to twice those of the HCC cases (26.4% vs. 14.5% current smokers and 27.2% vs. 13.6% consuming 3 drinks/d of alcohol for CLD and HCC cases, respectively).

Among non-cases, participant characteristics by SES indicators and healthcare infrastructure are shown in Table 2. Individuals living in most deprived neighborhoods (i.e., upper quintile of the socioeconomic deprivation index), were more likely to be female, African American, currently smoke, and have diabetes, and less likely to be married and to have completed college than individuals residing in the least deprived neighborhoods. An unexpected finding was that reported alcohol consumption of 3 drinks/d of alcohol was lower among residents of the most deprived neighborhoods compared to those residing in less deprived neighborhoods (5.8 vs. 8.1%, $P<0.001$) and consumption was similar across areas with high vs. low alcohol outlet density.

Associations with risk of HCC and CLD mortality

A suggested increased risk of HCC was observed for several socioeconomic measures, including neighborhood socioeconomic deprivation, % families below the poverty level, % unemployed, % no high school diploma, and areas with more uninsured residents, after adjustment for age, sex, and race (Table 3). Compared to least deprived neighborhoods (lowest quintile), there is a 48% increase risk in HCC (HR=1.48, 95% CI: 1.03–2.14) among those residing in the most deprived areas (highest quintile). Individuals residing in areas with a higher proportion of non-insured individuals had a 23% increased risk of HCC compared to individuals residing in other neighborhoods (HR=1.23, 95% CI: 1.01–1.51). These associations, however, were attenuated and no longer statistically significant after further adjustment of education, marital status, diabetes, body mass index, cigarette smoking, perceived health status, alcohol consumption, energy (non-alcohol), total meat intake, total fruit intake, total vegetable intake, and saturated fat intake. The strongest confounders were female sex (HR=0.32; 95% CI: 0.24–0.43) and history of diabetes (HR=3.01; 95% CI: 2.35, 3.84) (data not shown).

In regard to CLD mortality, a much greater association of area measures was observed than for HCC. Although the associations were attenuated, those with the largest effects (HRs >1.38) in the minimally adjusted models remained statistically significant after additionally accounting for education, marital status, diabetes, BMI, smoking, alcohol, energy, and dietary intake, with the exception of % unemployed (HR=1.45, 95% CI: 1.13–1.86 adjusting

for age, sex, and race; HR=1.17, 95% CI: 0.89–1.53 in fully adjusted model; P -trend=0.18). The risk of CLD mortality was 78% higher among individuals residing in the most deprived census tracts compared to those residing in the least deprived (HR=1.78, 95% CI: 1.34–2.36; P -trend<0.001), which was supported by similar associations for individuals residing in areas with the highest % families below the poverty level (HR=1.75, 95% CI: 1.34–2.28; P -trend<0.001) and highest % without a high school diploma (HR=1.65, 95% CI: 1.25–2.18; P -trend<0.001). The risk estimates for CLD mortality across quintile of socioeconomic deprivation index scores are provided in Figure 1. For every 1-unit increase in deprivation index scores, the risk of dying from CLD increased 23% (HR=1.23, 95% CI: 1.12–1.35). Individuals residing in census tracts with a higher density of alcohol outlets (i.e., above the median) had a 26% higher risk of dying from CLD than individuals residing in areas with fewer alcohol outlets. We found no evidence of an effect modification for risk of HCC and CLD mortality with sex, obesity and diabetes (P -interaction>0.22 for all).

Contribution to SES-disparities by individual-level risk factors and alcohol outlet density

We analyzed the contribution of individual-level risk factors (e.g., diabetes, obesity, lifestyle) and the density of alcohol outlets within census tracts to differences in the risk of CLD mortality by SES, after accounting for age, sex, and race (Table 4). Differences in the prevalence of individual health-risk factors by SES explained 5.1% (obesity) and 8.6% (diabetes) of the association between neighborhood SES and CLD mortality, whereas the individual behavioral factors explained 29.1% (95% CI: 17.6–40.6%) of the higher risk of CLD associated with higher neighborhood SES. The percentage mediated was 9.5% (95% CI: 3.9–15.1%) for alcohol outlet density. Results of a sensitivity analysis showed that the contribution of individual-level education was 5.3% (data not shown).

DISCUSSION

In this large prospective study of U.S. adults, we examined area socioeconomic deprivation and healthcare infrastructure to determine if these socioeconomic-related conditions explain small-area differentials in HCC incidence and CLD mortality after accounting for individual-level behaviors and other health risk factors. No associations were observed between measures of healthcare infrastructure and HCC incidence or CLD mortality. No other study, to our knowledge, has examined the multilevel CLD mortality-associations within a population-based prospective study in the United States. Our results show that CLD mortality, but not HCC incidence, varied across census-tracts and that higher census tract-level SES deprivation increased the risk of CLD mortality beyond that explained by individual education and other health risk factors (i.e., diabetes, obesity, high alcohol consumption). In addition to observing an association between neighborhood SES and CLD, our study showed that the association of census-tract socioeconomic deprivation could be partly explained by individual-level characteristics including history of type II diabetes, obesity, and behavioral factors such as smoking, diet, and alcohol consumption. The independent association observed between alcohol outlets and CLD mortality after accounting for alcohol consumption and other health behaviors is not clear. The proportion of variation explained by area-level alcohol outlet density may be due in part to unmeasured factors (i.e., residual confounding) or insufficient control for individual-level alcohol consumption because of a possible underestimation of self-reported alcohol intake on the study questionnaire (i.e., measurement error).

The preliminary findings of the current study are consistent with results from SEER registry database, which indicated that HCC patients were more likely to live in deprived-SES conditions [19]. However, after additionally adjusting for individual education attainment, diabetes, alcohol consumption and other health behaviors, our observed association with

HCC risk disappeared suggesting that area-level socioeconomic associations may result from the geographic distribution in the prevalence of individual health-risk factors. No previous studies that we are aware of have quantified the extent to which obesity and behavioral factors explain CLD mortality disparities by SES or explored the influence of residential area on CLD mortality. In one study, however, findings suggest that geographic variation in rates of surgical therapy for HCC could partially explain the disparity in HCC survival [10]. We did not have information available on surgical therapy to examine this association.

This study had major strengths in that most prior studies have examined a single socioeconomic indicator, such as percentage of population below the federal poverty rate, when investigating area-level deprivation. The use of a composite index likely captures more broadly the concept of deprivation than poverty rate alone. In addition, our study utilizes a multilevel Cox proportional hazard model which models the random effect (e.g., geographic variation) and accounts for the clustering of participants within neighborhoods, which allowed us to simultaneously examine relevant neighborhood- and individual-level characteristics that may be associated with HCC incidence and CLD-related mortality. Among the study limitations, other unmeasured individual-level and area-level characteristics may also play a role, including chronic hepatitis B and C infection [31], neighborhood social support and neighborhood disorganization [32, 33]. HBV and HCV infection status was not known in the study and therefore, further studies are needed to determine whether the associations observed for CLD mortality in this study are independent of HBV and HCV. The analysis used county-level variables such as population density of hospitals and physicians (general practitioners), which may be considered too coarse to measure “local-area” variation; however, this information was not available at the census tract level. Another potential limitation is that the study cohort was predominantly older, middle class Caucasians; therefore, results may not apply to other populations.

In conclusion, local-area geographic variations in CLD mortality, but not HCC incidence, were present in the NIH-AARP Diet and Health Study cohort. Census-tract-level deprivation is associated with higher risk of CLD death after adjustment for individual-level characteristics and known disease risk factors and health behaviors. This suggests that targeting populations living in socioeconomically deprived neighborhoods may be an effective means of decreasing CLD mortality. Our findings, however, are not definitive and need to be confirmed; future studies should investigate these associations in more socioeconomically-diverse study populations.

Acknowledgments

This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute. Cancer incidence data from the Atlanta metropolitan area were collected by the Georgia Center for Cancer Statistics, Department of Epidemiology, Rollins School of Public Health, Emory University. Cancer incidence data from California were collected by the California Department of Health Services, Cancer Surveillance Section. Cancer incidence data from the Detroit metropolitan area were collected by the Michigan Cancer Surveillance Program, Community Health Administration, State of Michigan. The Florida cancer incidence data used in this report were collected by the Florida Cancer Data System (FCDC) under contract with the Florida Department of Health (FDOH). The views expressed herein are solely those of the authors and do not necessarily reflect those of the FCDC or FDOH. Cancer incidence data from Louisiana were collected by the Louisiana Tumor Registry, Louisiana State University Medical Center in New Orleans. Cancer incidence data from New Jersey were collected by the New Jersey State Cancer Registry, Cancer Epidemiology Services, New Jersey State Department of Health and Senior Services. Cancer incidence data from North Carolina were collected by the North Carolina Central Cancer Registry. Cancer incidence data from Pennsylvania were supplied by the Division of Health Statistics and Research, Pennsylvania Department of Health, Harrisburg, Pennsylvania. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions. Cancer incidence data from Arizona were collected by the Arizona Cancer Registry, Division of Public Health Services, Arizona Department of Health Services. Cancer incidence data from Texas were collected by the Texas Cancer Registry, Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services. Cancer incidence data from

Nevada were collected by the Nevada Central Cancer Registry, Center for Health Data and Research, Bureau of Health Planning and Statistics, State Health Division, State of Nevada Department of Health and Human Services. We are indebted to the participants in the NIH-AARP Diet and Health Study for their outstanding cooperation. The views expressed in this manuscript do not necessarily represent those of the DHHS or the US Food and Drug Administration.

List of Abbreviations and Acronyms

AARP	American Association of Retired Persons
ARF	Area Resource File
BMI	body mass index
CI	confidence intervals
CLD	chronic liver disease
HCC	Hepatocellular carcinoma
HR	hazard ratios
ICD	International Classification of Diseases
ICD-O	International Classification of Diseases for Oncology
Kcal	kilocalories
NIH-AARP	National Institutes of Health-AARP Diet and Health Study
PI	Pacific Islander
SEER	Surveillance, Epidemiology and End Results (SEER) registry database
SES	socioeconomic status

References

1. Younossi ZM, Stepanova M, Afendy M, et al. Changes in the prevalence of the most common causes of chronic liver disease in the United States from 1998 to 2008. *Clin Gastroenterol Hepatol*. 2011; 9(6):524–30. [PubMed: 21440669]
2. Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. *J Clin Oncol*. 2009; 27(9):1485–91. [PubMed: 19224838]
3. Perz JF, Armstrong GL, Farrington LA, et al. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*. 2006; 45(4):529–38. [PubMed: 16879891]
4. London, T.; McGlynn, K. Liver Cancer. In: Schottenfeld, D.; Fraumeni, JF., editors. *Cancer epidemiology and prevention*. New York: Oxford University Press; 2006.
5. El-Serag HB. Epidemiology of hepatocellular carcinoma in USA. *Hepatol Res*. 2007; 37 (Suppl 2):S88–94. [PubMed: 17877502]
6. Singal AG, Yopp A, Skinner CS, et al. Utilization of hepatocellular carcinoma surveillance among american patients: a systematic review. *J Gen Intern Med*. 2012; s1:1606–11.
7. Nguyen GC, Thuluvath PJ. Racial disparity in liver disease: biological, cultural, or socioeconomic factors. *Hepatology*. 2008; 47(3):1058–1066. [PubMed: 18302296]
8. Davila JA, Petersena NJ, Nelson HA, et al. Geographic variation within the United States in the incidence of hepatocellular carcinoma. *J Clin Epidemiol*. 2003; 56(5):487–93. [PubMed: 12812824]
9. Devesa, SS.; Grauman, DJ.; Blot, WJ., et al. *Atlas of Cancer Mortality in the United States: 1950–94*. Washington, DC: U.S. Gov. Printing Office; 1999. DHHS Publ. No. (NIH) 99–4564
10. Sonnenday CJ, Dimick JB, Schulick RD, et al. Racial and geographic disparities in the utilization of surgical therapy for hepatocellular carcinoma. *J Gastrointest Surg*. 2007; 11(12):1636–46. [PubMed: 17912593]

11. Radley DC, Schoen C. Geographic variation in access to care - the relationship with quality. *N Engl J Med*. 2012; 367(1):3–6. [PubMed: 22693955]
12. Faggiano, F.; Partanen, T.; Kogevinas, M., et al. Socioeconomic differences in cancer incidence and mortality. In: Kogevinas, M.; Pearce, N.; Susser, M.; Boffetta, P., editors. *Social inequalities and cancer*. Lyons: IARC Scientific; 1997. Publications No. 138
13. 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. *J Natl Cancer Inst*. 1977; 58(3):525–47. [PubMed: 557114]
14. Davila JA, Morgan RO, Shaib Y, et al. Diabetes increases the risk of hepatocellular carcinoma in the United States: A population based case control study. *Gut*. 2005; 54:533–539. [PubMed: 15753540]
15. Diez Roux AV. Residential environments and cardiovascular risk. *J Urban Health*. 2003; 80(4): 569–89. [PubMed: 14709706]
16. Chang ET, Yang J, Alfaro-Velcamp T, et al. Disparities in liver cancer incidence by nativity, acculturation, and socioeconomic status in California Hispanics and Asians. *Cancer Epidemiol Biomarkers Prev*. 2010; 19(12):3106–18. [PubMed: 20940276]
17. Sloane D, Chen H, Howell C. Racial disparity in primary hepatocellular carcinoma: tumor stage at presentation, surgical treatment and survival. *J Natl Med Assoc*. 2006; 98(12):1934–9. [PubMed: 17225837]
18. Artinyan A, Mailey B, Sanchez-Luege N, et al. Race, ethnicity, and socioeconomic status influence the survival of patients with hepatocellular carcinoma in the United States. *Cancer*. 2010; 116(5): 1367–77. [PubMed: 20101732]
19. Shebl FM, Capo-Ramos DE, Graubard BI, et al. Socioeconomic status and hepatocellular carcinoma in the United States. *Cancer Epidemiology Biomarkers & Prevention*. 2012
20. Schatzkin A, Subar AF, Thompson FE, et al. Design and serendipity in establishing a large cohort with wide dietary intake distributions: the National Institutes of Health-American Association of Retired Persons Diet and Health Study. *Am J Epidemiol*. 2001; 154(12):1119–25. [PubMed: 11744517]
21. Michaud DS, Midthune D, Hermansen S, et al. Comparison of cancer registry case ascertainment with SEER estimates and self-reporting in a subset of the NIH-AARP Diet and Health Study. *J Registry Manag*. 2005; 32:70–5.
22. Fritz, A.; Percy, C.; Shanmugaratnam, K., et al. *International Classification of Diseases for Oncology: ICD-O. 3*. Geneva, Switzerland: World Health Organization; 2000.
23. Kung HC, Hoyert DL, Xu J, et al. Deaths: final data for 2005. *Natl Vital Stat Rep*. 2008; 56(10):1–120. [PubMed: 18512336]
24. Major JM, Doubeni CA, Freedman ND, et al. Neighborhood socioeconomic deprivation and mortality: NIH-AARP diet and health study. *PLoS One*. 2010; 5(11):e15538. [PubMed: 21124858]
25. Gastwirth JL. The estimation of the Lorenz Curve and Gini index. *The Review of Economics and Statistics*. 1972; 54:11.
26. Berke EM, Tanski SE, Demidenko E, et al. Alcohol retail density and demographic predictors of health disparities: a geographic analysis. *Am J Public Health*. 2010; 100:1967–1971. [PubMed: 20724696]
27. Willett, WC. *Nutritional Epidemiology*. 2. New York, NY: Oxford University Press; 1998.
28. Banerjee S, Wall MM, Carlin BP. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics*. 2003; 4(1):123–42. [PubMed: 12925334]
29. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. *Stat Med*. 1992; 11(2)
30. Korn, EL.; Graubard, BI. *Analysis of health surveys*. New York, NY: Wiley; 1999.
31. Nordenstedt H, White DL, El-Serag HB. The changing pattern of epidemiology in hepatocellular carcinoma. *Dig Liver Dis*. 2010; 42 (Supl 3):S206–14. [PubMed: 20547305]
32. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997; 277(5328):918–24. [PubMed: 9252316]

33. Subramanian SV, Lochner KA, Kawachi I. Neighborhood differences in social capital: a compositional artifact or a contextual construct? *Health Place*. 2003; 9(1):33–44. [PubMed: 12609471]

Figure 1. Association between area-level socioeconomic deprivation quintiles and CLD mortality. Risk estimates account for age, sex, race, education, marital status, diabetes, body mass index, cigarette smoking, perceived health status, alcohol consumption, energy (non-alcohol), total meat intake, total fruit intake, total vegetable intake, and saturated fat intake.

Table 1

Participant characteristics by hepatocellular carcinoma incidence and chronic liver disease mortality in the NIH-AARP Diet and Health Study (N = 494,988)

Variable	Non cases (n=494,477)	Incident HCC (n=434)	CLD deaths (n=805)
Age, years	62.0 (5.4)	64.1 (4.9)	63.3 (4.9)
Sex			
Male, n (%)	294,897 (59.6)	358 (82.5)	565 (70.2)
Education			
Completed college, n (%)	190,602 (38.6)	147 (33.9)	282 (35.0)
Race and/or ethnicity, n (%)			
Non-Hispanic white	451,105 (91.2)	360 (90.0)	744 (92.4)
Non-Hispanic black	19,209 (3.9)	18 (4.2)	19 (2.4)
Hispanic	9,345 (1.9)	24 (5.5)	17 (2.1)
Asian/PI/Native American	8,157 (1.7)	20 (4.6)	10 (1.2)
Other	6,571 (1.3)	12 (2.8)	15 (1.9)
Married			
Yes, n (%)	338,860 (68.5)	313 (72.1)	473 (58.8)
Alcohol consumption			
>3 drinks per day, n (%)	37,149 (7.5)	59 (13.6)	219 (27.2)
Current smokers, n (%)	59,025 (11.9)	63 (14.5)	213 (26.4)
Self-reported diabetes, n (%)	44,602 (9.0)	120 (27.7)	166 (20.6)
Obese, n (%)	106,278 (21.5)	139 (32.0)	235 (29.2)
Body mass index, kg/m ²	27.1 (5.1)	28.4 (5.3)	28.1 (5.4)
Vigorous physical activity 5 times/wk			
5 times/wk, n (%)	94,871 (19.2)	66 (15.2)	100 (12.4)
Self-reported health			
Fair or poor, n (%)	59,876 (12.1)	90 (20.7)	216 (26.8)
Dietary intake			
Energy from non-alcohol, kcal	1,789 (770.1)	1,959 (864.6)	1,777 (839.5)
Saturated fat, g/1000 kcal	10.5 (3.3)	11.1 (3.3)	10.8 (3.7)
Total meat, g/1000 kcal	68.1 (32.3)	66.2 (29.8)	66.0 (34.3)
Fruit, servings/1000 kcal	1.7 (1.2)	1.6 (1.2)	1.6 (1.3)
Vegetables, servings/1000 kcal	2.2 (1.2)	1.9 (1.0)	1.9 (1.1)

PI = Pacific Islander, kcal = kilocalories. Means and standard deviations reported unless otherwise indicated.

Table 2
Participant characteristics by area socioeconomic conditions in the NIH-AARP Diet and Health Study

Variable	SES Deprivation* (quintiles)			Alcohol Outlets (median cutpoints)			Healthcare Infrastructure (median cutpoints)		
	Least Deprived	Most Deprived		Fewer Outlets	More Outlets		Fewer Hospitals	More Hospitals	More Physicians
Study participants, n	146,695	36,017		296,074	198,914		38,241	456,747	30,018
Age, years	61.7	62.0		62.0	62.0		62.2	62.0	62.3
Male, %	64.3	48.8		62.1	56.1		59.5	59.7	59.0
Completed college, %	56.1	25.1		38.4	38.7		30.1	39.3	27.6
Race and/or ethnicity, %									
Non-Hispanic white	94.6	65.3		94.1	86.9		94.5	90.9	93.8
Non-Hispanic black	1.1	26.3		2.6	5.8		2.9	4.0	3.4
Hispanic	1.4	3.6		1.1	3.2		0.5	2.0	0.6
Asian/PI/Native American	2.0	2.0		1.1	2.5		0.7	1.7	0.7
Other	1.0	2.9		1.2	1.6		1.3	1.3	1.4
Married, %	74.2	52.6		73.6	61.1		73.1	68.2	72.9
Alcohol, >3 drinks per day, %	8.1	5.8		7.5	7.5		7.0	7.6	6.8
Current smokers, %	9.0	16.0		11.4	12.7		12.7	11.9	13.1
Self-reported diabetes, %	6.7	13.7		8.8	9.5		10.1	9.0	10.6
Body mass index, kg/m ²	26.5	28.2		27.1	27.1		27.3	27.1	27.4
Vigorous physical activity									
5 times/week, %	20.6	16.0		19.4	18.9		19.5	19.2	19.9
Self-reported fair/poor health, %	7.9	19.8		11.7	12.8		14.3	11.9	15.8
Dietary intake									
Energy from non-alcohol, kcal	1739	1873		1808	1761		1851	1784	1864
Saturated fat, g/1000 kcal	10.1	10.6		10.5	10.4		10.7	10.4	10.7
Total meat, g/1000 kcal	69.1	67.3		68.3	67.7		67.0	68.2	66.5
Fruit, servings/1000 kcal	1.7	1.9		1.7	1.8		1.6	1.7	1.6
Vegetables, servings/1000 kcal	2.3	2.1		2.2	2.3		2.1	2.2	2.1

* A subset of the cohort (non-cases) are presented to illustrate the characteristics of the socioeconomic indicator groups (e.g., quintiles 1 vs. 5 for SES deprivation, where the upper quintile indicates the highest level of neighborhood deprivation and < median density vs. median density for alcohol outlets, hospitals and physicians), kcal = kilocalories. Means reported unless otherwise indicated. All chi-square tests (proportions) and T-tests (continuous variables) were significant ($P<0.05$) with the exception of BMI with alcohol outlet density ($P=0.87$) and sex with hospital density ($P=0.57$).

Table 3

Multivariable hazard ratios for area-level socioeconomic factors with hepatocellular carcinoma (HCC) incidence and chronic liver disease (CLD) mortality

	Incident HCC (n=434)		CLD deaths (n=805)	
	Minimally* adjusted HR (95% CI)	Fully** adjusted HR (95% CI)	Minimally* adjusted HR (95% CI)	Fully** adjusted HR (95% CI)
Socioeconomic measures				
Most vs. least deprived (Q5 vs. Q1)	1.48 (1.03, 2.14)	1.01 (0.78, 1.32)	2.36 (1.79, 3.11)	1.78 (1.34, 2.36)
% Families below poverty level (Q5 vs. Q1)	1.09 (0.73, 1.62)	0.94 (0.71, 1.25)	2.31 (1.77, 3.03)	1.75 (1.34, 2.28)
% Unemployed (Q5 vs. Q1)	0.95 (0.67, 1.36)	0.84 (0.65, 1.08)	1.45 (1.13, 1.86)	1.17 (0.89, 1.53)
% No high school diploma (Q5 vs. Q1)	1.38 (0.96, 1.98)	0.96 (0.66, 1.39)	2.11 (1.66, 2.69)	1.65 (1.25, 2.18)
Gini coefficient (0.05 unit)	0.94 (0.81, 1.08)	0.97 (0.82, 1.15)	1.05 (0.96, 1.15)	1.02 (0.93, 1.11)
More uninsured (median value)	1.23 (1.01, 1.51)	1.22 (0.99, 1.50)	0.86 (0.74, 0.98)	0.87 (0.77, 0.99)
Fewer Medicare beneficiaries (< median value)	0.93 (0.58, 1.49)	0.83 (0.52, 1.33)	0.94 (0.70, 1.28)	0.91 (0.68, 1.23)
Acculturation				
More foreign born (median value)	0.75 (0.60, 0.95)	0.89 (0.67, 1.08)	1.10 (0.89, 1.37)	1.10 (0.89, 1.37)
More language isolation (median value)	0.84 (0.67, 1.05)	0.85 (0.67, 1.06)	1.08 (0.92, 1.26)	1.02 (0.88, 1.20)
Healthcare Infrastructure				
Fewer hospitals (< median density value)	1.13 (0.81, 1.58)	1.03 (0.74, 1.45)	0.85 (0.66, 1.11)	0.86 (0.67, 1.10)
Fewer physicians (< median density value)	0.87 (0.57, 1.31)	0.76 (0.50, 1.16)	0.90 (0.67, 1.21)	0.88 (0.65, 1.17)
More alcohol outlets (median density value)	1.14 (0.93, 1.38)	1.10 (0.91, 1.34)	1.41 (1.22, 1.62)	1.26 (1.09, 1.45)

* Minimally adjusted models account for age, sex, and race.

** Risk estimates for area-level socioeconomic factors from fully adjusted models account for age, sex, race, education, marital status, diabetes, body mass index, cigarette smoking, perceived health status, alcohol, energy (non-alcohol), total meat intake, total fruit intake, total vegetable intake, and saturated fat intake.

Census-tract level: socioeconomic deprivation index, % families below poverty level, % unemployed, % with no high school diploma, Gini coefficient, alcohol outlet density; county-level: percent without insurance, Medicare beneficiaries, foreign born, language isolation, population density of hospitals and physicians (general practitioners).

Percentage of the association between neighborhood socioeconomic deprivation and chronic liver disease mortality explained by alcohol outlet density, diabetes, obesity

Table 4

Potential mediators	Chronic Liver Disease	
	Regression coefficient	% Explained (95% CI)
Socioeconomic deprivation (minimally adjusted *)	0.19	reference
Separate effects of alcohol outlet density, diabetes, obesity, and health behaviors		
Effect of alcohol outlet density	0.17	9.51 (3.93, 15.09)
Effect of diabetes	0.17	8.55 (6.98, 10.21)
Effect of obesity	0.18	5.14 (0.14, 10.06)
Effect of other health behaviors and health risk factors †	0.14	29.13 (17.61, 40.59)
Combined effect of diabetes, obesity, and health behaviors and health risk factors	0.13	32.36 (18.63, 46.09)
Combined effect of alcohol outlet density, diabetes, obesity, health behaviors and health risk factors	0.12	39.07 (23.76, 54.38)

* Minimally adjusted for age, sex, and race/ethnicity.

† Includes education, marital status, cigarette smoking, perceived health status, alcohol consumption, energy (non-alcohol), total meat intake, total fruit intake, total vegetable intake, and saturated fat intake.