HHS Public Access

Author manuscript

Cancer Causes Control. Author manuscript; available in PMC 2024 March 01.

Published in final edited form as:

Cancer Causes Control. 2022 May ; 33(5): 701–710. doi:10.1007/s10552-022-01555-0.

Geographic hotspot detection for late-stage hepatocellular carcinoma: novel approach to cancer control

Kali Zhou[1,2], **Laura K. Thompson**[3], **Lihua Liu**[3,4], **Norah A. Terrault**[1,2], **Myles G. Cockburn**[3,4]

[1]Division of Gastrointestinal and Liver Diseases, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

[2]USC Research Center for Liver Disease, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

[3]Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA, USA

[4]Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA

Abstract

Importance—As hepatocellular carcinoma (HCC)-associated mortality continues to rise in the United States, there is a crucial need for strategies to shift diagnoses from late to early stage in order to improve survival.

Objective—To describe a population-based geospatial approach to identifying areas with high late-stage HCC burden for intervention.

Design—Cross-sectional study between 2008 and 2017.

Setting—Los Angeles County.

Participants—All incident cases of HCC with residential address at diagnosis in Los Angeles County were identified from a population-based cancer registry. Late stage included AJCC 7th Edition stages III-IV and unstaged cases.

Exposure—Sociodemographic factors.

Main outcome(s)—Geographic “hotspots” or areas with a high density of late-stage HCC, identified using kernel density estimation in ArcMap 10.3.1.

Results—51.8% of 7,519 incident cases of HCC were late stage. We identified a total of 23 late-stage hotspots, including 30.0% of all late-stage cases. Cases within hotspots were more

Kali Zhou, kalizhou@usc.edu.

Author contributions KZ is the guarantor of the article and all authors approved the final version of the manuscript. The authors made the following contributions: Concept and design: KZ, LT, and MC. Acquisition, analysis, or interpretation of data: KZ, LT, and MC. Statistical analysis: LT and MC. Drafting of the manuscript: KZ. Critical revision of the manuscript for important intellectual content: All authors.

Conflict of interest Dr. Zhou and Dr. Terrault report institutional grant support from Gilead Sciences. The other authors report no disclosures.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10552-022-01555-0.

often racial/ethnic minorities, foreign-born, under or uninsured, and of lower socioeconomic status. The age-adjusted incidence rate of late-stage HCC was twofold higher within hotspots (6.85 per 100,000 in hotspots vs 3.38 per 100,000 outside of hotspots). The calculated population-attributable risk was 43%, suggesting that a substantial proportion of late-stage HCC burden could be averted by introducing interventions in hotspot areas. We mapped the relationship between hotspots and federally qualified health centers primary care clinics and subspecialty clinics in Los Angeles County to demonstrate how clinic partnerships can be selected to maximize impact of interventions and resource use. Hotspots can also be utilized to identify “high-risk” neighborhoods that are easily recognizable by patients and the public and to facilitate community partnerships.

Conclusion and relevance—Reducing late-stage HCC through geographic late-stage hotspots may be an efficient approach to improving cancer control and equity.

Keywords

Disparities; Geospatial methods; Geographic information system; Liver cancer; Implementation science

Introduction

Hepatocellular carcinoma (HCC) is one of few remaining cancers with rising mortality in the United States (US) [1]. Survival is closely linked to tumor stage at diagnosis, with 5-year survival of 33% with localized versus 2% with metastatic cancer [2]. The wide gap in survival may be attributed to curative therapies, such as radiofrequency ablation, resection, and transplantation, for select patients with early localized disease, which increase median survival from 4 to 6 months to 29 to 59 months [3-5]. Shifting diagnoses from late to early stage for a substantial number diagnosed is critical for improvements in survival to be realized at a population-level.

Existing strategies to increase the early detection of HCC have not taken a population-based or spatially oriented approach [6, 7]. As individuals with similar health-related risk factors and access-to-care parameters often live in geographically close communities [8], directing interventions toward geographic areas, in addition to individuals, can induce cancer equity through multiple pathways. This is of particular relevance to HCC, for which surveillance is only recommended in specific clinical conditions (e.g., patients with cirrhosis). Geographic hotspot detection, in the context of our study, is a term ascribed to geostatistical methods for identifying areas of elevated disease burden without the limitations inherent in use of arbitrary administrative boundaries [9]. Applying this approach to cancer control allows the strategic re-allocation of resources to a smaller subset of highest risk areas and populations to maximize the impact of delivered interventions and provide a rationale for the practical selection of locations (e.g., clinics, neighborhoods) to perform such interventions.

In this study, we use a high-quality population-based cancer registry to detect areas with high density of late-stage HCC, termed “late-stage hotspots”, in a defined metropolitan setting (single county). We demonstrate the utility of hotspots in recognizing priority targets (populations, clinics, and neighborhoods) for intervention and quantify the potential impact of modifying risk of late-stage cancer within hotspot areas.

Methods

Study setting

Los Angeles County (LAC) is the most populous county in the US, with a multi-ethnic urban population of over 10 million constituents (49% Hispanic, 26% White, 15% Asian, and 9% Black) [10]. Compared to other counties, LAC has one of the highest age-adjusted incidence rates of HCC across the nation and has the highest absolute number of cases diagnosed per year [11]. The LAC Cancer Surveillance Program (LACSP) is the population-based cancer registry for all of LAC and contributes to the California Cancer Registry (CCR) and the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) database.

Cohort selection

We included incident cases of HCC using International Classification of Diseases for Oncology, 3rd edition topography code C22.0 (liver) and restricted to histology codes 8170–8175 (HCC subtypes, excluding intrahepatic cholangiocarcinoma, hepatic adenoma, and metastatic lesions). Cases diagnosed between 2008 and 2017 were included. Exclusion criteria included residence outside of LAC and missing residential address (< 1% missing in LACSP). We defined late stage as American Joint Committee on Cancer (AJCC) 7th Edition stages III-IV, which includes tumors with large vessel involvement, regional or distant spread, and any unstaged cases [12].

Hotspot detection

To identify areas in LAC with the highest density of late-stage HCC, we created a Kernel Density surface in ArcMap 10.3.1 (Redlands, CA), based on the latitude/longitude coordinates of residential address at diagnosis. Kernel density estimation is a geostatistical method that produces a smooth surface to represent the density of spatial point events across space [13]. Each cell is given a density value that is based on a user-determined bandwidth centered over each point location (= coordinates of late-stage HCC case); points closer to the focus point are weighted more heavily than those farther away [14]. To prevent edge effect errors at county boundaries, we included cases from neighboring counties identified in the CCR database during the creation of the density surface. Because we were interested in targeting our interventions in areas with the highest occurrence (i.e., raw number of cases) of late-stage cancer, we did not consider the underlying population at this stage of analysis.

We iteratively divided our continuous density surface into 'hotspots' (areas meeting or exceeding the density cutpoint) and surrounding lower-density areas using each density value integer present in the data (range 1–241). Any small, highly localized hotspots meeting the density value threshold but containing fewer than 20 cases (any stage) were excluded. We qualitatively reviewed maps and statistics for each of the 241 cutpoints (potential cutpoints and corresponding maps are available in online Appendix Figs. 1 and 2), selecting a cutpoint that (1) maximized areas with the highest proportion of late-stage cases, (2) optimized difference in proportion late stage within and outside of hotspots, and (3) captured at least 20% of cases (any stage) within hotspots. The rationale for these specific considerations was to optimize the geographic dimensions of hotspots such that

any intervention delivered to hotspots would impact a sufficient number of individual cases to have an appreciable effect on overall late-stage cancer. Yet, the number of discrete hotspots and size needed to be minimized such that a given intervention could be reasonably delivered with finite resources.

Characteristics of hotspot cases

The following covariates were available in the LACSP registry: age, sex, race/ethnicity, birthplace, marital status, insurance, comorbidity index, and socioeconomic status (SES). Mutually exclusive racial/ethnic groups included Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian/Asian Pacific Islander (API), and non-Hispanic Other. Birthplace was categorized as US-born and foreign-born. Insurance was categorized as insured (including private and Medicare), Medicaid, and not insured/unknown. The comorbidity index grouped individuals into low, moderate, and high comorbidity as previously described [15]. We used the well-validated update of the Yost index called the Yang index as a measure for census tract-level socioeconomic status (SES), based on principal components analysis of the following variables from the American Community Survey: 1) Education Index; 2) Percent persons with a ratio of household income to poverty line 2 or higher (percent persons above 200% poverty line); 3) Percent persons with a blue-collar job; 4) Percent persons employed; 5) Median rental; 6) Median value of owner-occupied housing unit; 7) Median household income; 8) Percent occupied housing unit; and 9) Percent owner-occupied housing unit [16, 17]. Yang SES index was categorized into low to high SES quintiles (Q1: lowest; Q2: lower-middle, Q3: middle, Q4: upper-middle, and Q5: upper) and assigned to the census tract in which each case resided.

Clinics and neighborhoods

To visualize the geographic availability of primary services and specialty services (for liver disease) and identify specific clinics for intervention, we overlaid the locations of federally qualified health centers (FQHCs) and gastroenterology/hepatology (GI/HEP) clinics on our hotspot map. Primary care clinics with FQHC designation in LAC were obtained from the Health Center Service Delivery Sites database provided by the Health Resources and Services Administration (https://data.hrsa.gov/data/download). Active GI/HEP clinics in LAC were obtained from the publicly available Medicare Physician Compare Database. A clinic was considered inside a hotspot if its latitude/longitude coordinates fell directly inside of hotspot boundaries. Due to some clinics being proximate but not inside a hotspot, we examined the impact of applying 1- and 2-km buffers around clinics on reducing the number of hotspots without a clinic. We also identified neighborhoods within hotspot areas to provide an example of how results can be translated to community leaders and stakeholders. LAC is divided into 272 unofficial neighborhoods that are well established and recognizable to the general public [18]. A neighborhood was considered inside a hotspot if the centroid of the neighborhood fell within hotspot boundaries.

Statistical analysis

Descriptive characteristics of cases were compared within and outside of hotspots with chi-square testing for categorical variables. We performed univariate and multivariable logistic regression (with inclusion of variables with univariate $p < 0.05$) to examine association

between covariates and residence in a late-stage hotspot as outcome. We calculated the age-adjusted incidence rate (AAIR) per 100,000 persons of late-stage HCC by hotspot status and overall. We then calculated attributable risk using the equation ($AAIR_{hotspot} - AAIR_{nothotspot}$) / $AAIR_{hotspot}$, followed by population-attributable risk (PAR) using the equation ($AAIR_{hotspot} - AAIR_{overall}$)/$AAIR_{hotspot}$. Hotspot status for each census tract was assigned based on whether or not the centroid of the census tract fell within a hotspot boundary [19]. Population denominators for AAIR and PAR were based on the 2010 US Census [19], linearly interpolated on a yearly basis using the 2000 and 2010 Decennial Census tract population counts (interpolated through 2017) [19-21]. All statistical analyses were performed using SAS 9.3 (SAS Institute, Inc., Cary, North Carolina).

Results

Overview of late-stage hotspots

A total of 7,519 incident cases of HCC were included, 51.8% of which were late stage. 23 discrete hotspots were identified (Fig. 1). Hotspots captured 26.7% of all cases in LAC, including 30.0% of all late-stage cases. The proportion of late-stage cases within hotspots was 59.1%, compared to 49.1% across the rest of LAC, a 10.0% difference or an absolute difference of 1,517 late-stage cases. In coldspots (the lowest density cutpoint that includes 30% of late-stage HCC cases), 42.9% were late stage, a difference in proportion of late-stage cases of 16.2% (data not shown). Hotspots spanned 430 (18.3%) of 2,344 census tracts in LAC.

Comparison of cases diagnosed within and outside of hotspots

There was no difference in age, sex, and year of diagnosis by hotspot residence ($p > 0.05$) (Table 1). Over 50% were diagnosed under the age of 65 and 74% were male in both groups. There were striking differences, however, in sociodemographic composition. 86.3% of individuals within hotspots were racial/ethnic minorities (43.8% Hispanic; 30.0% Asian/API; 11.6% Black; 0.9% Other) compared to 69.5% (36.7% Hispanic; 21.8% Asian/API; 10.0% Black; 1.0% Other) in all other areas ($p < 0.01$). 48.6% within hotspots compared to 36.9% outside were born outside of the US ($p < 0.01$). There was a gradient in proportion within each SES quintile in hotspot areas, from 1.1% in the highest SES quintile to 44.8% in the lowest SES quintile (see Fig. 2). In comparison, there was an even distribution across SES quintiles in non-hotspot areas with 19.1% in the lowest SES quintile and 15.1% in the highest quintile.

Factors associated with residence in a late-stage hotspot

Age, sex, and comorbidity index were not associated with residence in a late-stage hotspot in univariate models (Table 2). While all non-White racial/ethnic groups were associated with hotspot residence on univariate testing, only Asian/Pacific Islanders had independently increased odds of residence in a hotspot compared to Whites (OR 2.1, 95% CI 1.7–2.6) after adjustment for all other covariates, including SES. In multivariable models, hotspot residents were also more likely to be foreign-born (OR 1.3, 95% CI 1.2–1.5), not married (OR 1.2, 95% CI 1.1–1.3), and Medicaid insured (OR 1.3, 95% CI 1.1—1.4). Likelihood of residence in a hotspot sequentially increased with each quintile decrease in SES (OR 3.1,

9.7, 20.3, and 32.5 for upper-middle, middle, lower-middle, and lowest quintile, respectively, compared to highest quintile, $p < 0.01$).

Clinic and neighborhoods within late-stage hotspots

Mapped relationship between location of primary (FQHCs) and specialty (GI/HEP clinics) services and hotspots is presented in Fig. 3. 178 (37.0%) of 481 FQHCs were within hotspots and 5 hotspots contained no FQHC clinics. A total of 229 unique GI/HEP clinics in LAC were identified. Of these, 36 (16%) GI/HEP clinics fell directly within hotspot areas and 14 of the 23 hotspots contained no GI/HEP clinic. We were able to reduce the number of hot spots without a GI/HEP clinic to 8 using a 1-km buffer around clinics and to 3 using a 2-km buffer around clinics. However, the addition of a buffer led to an increase in total GI/HEP clinics within hotspot areas to 67 (with 1-km buffer) and 103 clinics (with 2-km buffer), demonstrating the trade-off between number of clinics within hotspots and number of hotspots without a clinic. There are 272 total neighborhoods in LAC [18]. Hotspots spanned 60 (22.1%) of LAC neighborhoods (online Appendix Fig. 3). The median number of neighborhoods within hotspots was 2 (IQR 1–2); the largest hotspot spanned 17 neighborhoods.

Population-attributable risk of late-stage disease in hotspots

The AAIR of incident late-stage HCC in LAC overall was 3.93 per 100,000 population (95% CI 3.25–3.51). The AAIR within hotspots was 6.85 per 100,000 (6.43–7.27) compared to an AAIR outside hotspots of 3.38 per 100,000 (95% CI 3.80–4.05). The attributable risk of late-stage HCC to residence in a hotspot was 0.51, with a population-attributable risk of 0.43.

Discussion

In this study, we present a population-based, geospatial approach to identifying priority geographic areas and population subgroups to achieve cancer control for HCC, a cancer with rising mortality in the US [1]. As resources in any healthcare system are finite, data-driven methods to allocate resources, particularly those that take advantage of existing population-based databases, are attractive. Critically, our analysis identifies precisely where the greatest *absolute* risk of late-stage HCC occurs, which is where interventions should be targeted for maximal impact. Moreover, we demonstrate the efficiency inherent to our approach, such that modification of risk of late-stage disease in hotspot areas has the potential to reduce late-stage HCC overall by up to 43%.

Geographic patterns of disease are well described in many types of cancer, including HCC [22, 23]. However, many approaches for mapping point-based disease occurrence require the aggregation of point records to an existing areal unit (e.g., census tracts, ZIP Codes.), enabling (1) the de-identification of secure data and (2) the detection of spatial patterns across these areal units, but may obscure the true underlying pattern of disease, known as the Modifiable Areal Unit Problem (MAUP) [9]. The advantage of kernel density estimation of hotspots is that it overcomes MAUP by mapping disease observations without aggregating points into predetermined areal units, a method not previously applied to cancer

registry data. Furthermore, we purposefully chose to examine areas of highest density without adjusting for underlying population. Our approach differs from conventional hotspot analyses that are looking for areas with *higher than expected* disease parameters, which often result in the identification of very small areas with few cases and therefore is not practical or meaningful for changing disease burden across a population (an example is provided in online Appendix Fig. 4).

Our hotspots also highlight the disproportionate burden of late-stage HCC in low-income, minority, and immigrant communities. Racial/ethnic and socioeconomic disparities in HCC are frequently described, including greater likelihood of late-stage cancer at diagnosis and lower survival among Hispanics and Blacks as compared to Asians and Whites [24, 25]. There are little data to suggest host or tumor genetics underlie these differences [26], while factors related to healthcare utilization and delivery including insurance, access to and utilization of curative therapies, and surveillance uptake very likely play a role. The overlap of late-stage hotspots with areas of lowest SES (33-fold higher likelihood of residence in a hotspot) strongly supports the hypothesis that social determinants of health and access-to-care are primary drivers of differential stage at diagnosis for HCC and interventions targeting these pathways or policies to expand access will ultimately have the greatest impact on cancer equity. While our hotspots only cover 30% of HCC cases in LAC, they capture a much larger proportion of cancer within LAC's lowest SES populations, who experience the lowest survival (15% 5-year HCC survival in lowest SES quintile versus 25% in highest quintile in LAC). Thus, any intervention to improve early detection targeting these areas will have outsized influence on downstream outcomes. The hotspot approach can therefore be a useful tool to address the clear social and racial injustice in the disproportionate burden of late-stage HCC in low SES populations.

The core of early detection of HCC is guideline-recommended imaging and tumor marker surveillance every six months in at-risk patients [27]. Uptake of timely surveillance in the US is suboptimal, with an overall pooled compliance of only 20%, with lowest rates seen in low-income individuals [28]. Multilevel early detection interventions are needed to combat this disparity. A geographic hotspot approach aggregates multiple levels of influence (e.g., patient, provider, and community) to support the development of multilevel interventions, which are often more effective [29]. To that end, we visualized the macro-geographic relationship between availability of primary/specialty clinics and hotspots—showing here that FQHCs are well-aligned clinic targets to reach hotspot populations. We also demonstrate the potential number of clinics available for feasibility assessments in the pre-implementation planning phase. In hotspots where a substantial fraction of residents are immigrants, culturally tailored and language-concordant interventions may be especially impactful if implemented within the FQHCs that primarily serve these communities. Existing interventions that have demonstrated efficacy such as mailed outreach and patient navigators can be packaged in such a way and delivered to hotspot FQHCs [6, 7]. Further, ascertainment of hotspot neighborhoods can be used to educate and raise awareness in the general public, media, and among stakeholders, particularly in metropolitan areas like Los Angeles that has distinct and well-recognized neighborhoods. This knowledge can also be leveraged toward establishing high-yield community partnerships and deploying community-based interventions tailored to the unique needs of the most highly impacted

neighborhoods. It is important to note that early detection is a start but alone is insufficient to address poor outcomes in vulnerable populations; post-diagnosis treatment pathways, resources, and support across the entire cancer continuum are also needed. We envision this hotspot method as a tool to support effective multi-faceted implementation work—a framework to identifying the characteristics and locations of those at greatest need and the partnerships and stakeholders needed for success.

There are a few limitations of our approach. The disparities found in LAC may not be generalizable to other geographic locales, although we anticipate similar findings in large diverse metropolitan counties across the US. An advantage of our approach is that high-quality cancer registries are available in all US states and our method of hotspot detection can easily be replicated for individualized, region-specific, and cancer-specific results. Comparisons across all SEER registries is not yet feasible since the smallest level of geography available currently is census tracts. We do not account for residential mobility and temporal changes in environment and neighborhoods over time. Migration both in and out of LAC is generally low, estimated at 1–2% of the total population per year [30]. Further, comparing density maps of the cohort separated into 5-year intervals (2008–2012 vs 2013–2017), the highest density areas were relatively stable across the two time periods (online Appendix Fig. 5).

Registry databases lack the granularity in clinical data to determine staging more relevant to clinical practice, thus individuals categorized as “early stage” in this study may in actuality have few or no therapeutic options. However, we believe the likelihood that this would introduce a geography-specific pattern that substantially influences the location and boundaries of hotspots is low. Our definition of late stage captures those in which HCC surveillance, if performed, likely would have resulted in earlier diagnosis. Lastly, registry data do not include etiology of liver disease, which if known, would further support prevention efforts and upstream identification of individuals at-risk for HCC to enter surveillance programs. At present, we can hypothesize likely etiology based on racial/ethnicity composition in a given hotspot (e.g., chronic hepatitis B if predominantly Asian), but this is a crude approximation and principal etiologic causes of liver disease by race/ethnicity has changed over time [31]. Future directions include registry linkage to claims databases or changes in data abstraction to include these crucial variables within registries to bridge this gap.

Efficient strategies to combat high rates of late-stage HCC at time of cancer diagnosis are urgently needed. In summary, population-based geographic late-stage hotspot detection is a novel, rational, and practical approach to addressing this need. Identified hotspots can be incorporated into the development and implementation of multilevel early detection interventions with the ultimate goal of improving HCC cancer control and outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by a University of Southern California Research Center for Liver Diseases grant 5P30DK048522 to Dr. Zhou. This work was also supported in part by the Population Research Core of the USC/Norris Comprehensive Cancer Center funded by the National Cancer Institute grant P30CA014089. The collection of cancer incidence data used in this study was supported by the California Department of Public Health pursuant to California Health and Safety Code Section 103885; Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries, under cooperative agreement 5NU58DP003862-04/DP003862; the National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute. The ideas and opinions expressed herein are those of the author(s) and do not necessarily reflect the opinions of the State of California, Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors. The funders/sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data availability

The data that support the findings of this study are available from the Los Angeles County Cancer Surveillance Registry. The data are not publicly available due to restrictions from the registry as they contain information that could compromise confidentiality of cancer patients.

Abbreviations

AAIR	Age-adjusted incidence rate
AJCC	American Joint Committee on Cancer
API	Asian Pacific Islander
FQHC	Federally Qualified Health Center
GI/HEP	Gastroenterology and hepatology
HCC	Hepatocellular carcinoma
LAC	Los Angeles County
LACSP	Los Angeles Cancer Surveillance Program
PAR	Population-attributable risk
SES	Socioeconomic status
US	United States

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424 [PubMed: 30207593]
2. American Cancer Society (2019) Liver Cancer Survival Rates. https://www.cancer.org/cancer/liver-cancer/detection-diagnosis-staging/survival-rates.html. Accessed 28 Oct 2020

3. Mathur AK, Osborne NH, Lynch RJ, Ghaferi AA, Dimick JB, Sonnenday CJ (2010) Racial/ethnic disparities in access to care and survival for patients with early-stage hepatocellular carcinoma. Arch Surg 145(12):1158–1163 [PubMed: 21173289]
4. Livraghi T, Meloni F, Di Stasi M et al. (2008) Sustained complete response and complications rates after radiofrequency ablation of very early hepatocellular carcinoma in cirrhosis: is resection still the treatment of choice? Hepatology 47(1):82–89 [PubMed: 18008357]
5. Kim WR, Lake JR, Smith JM et al. (2019) OPTN/SRTR 2017 Annual Data Report: liver. Am J Transplant 19(Suppl 2):184–283 [PubMed: 30811890]
6. Singal AG, Tiro JA, Marrero JA et al. (2017) Mailed outreach program increases ultrasound screening of patients with cirrhosis for hepatocellular carcinoma. Gastroenterology. 152(3):608–615 [PubMed: 27825963]
7. Aby ES, Winters AC, Lin J et al. (2020) A telephone and mail outreach program successfully increases uptake of hepatocellular carcinoma surveillance. Hepatol Commun 4(6):825–833 [PubMed: 32490319]
8. Baicker K, Chandra A, Skinner JS (2005) Geographic variation in health care and the problem of measuring racial disparities. Perspect Biol Med 48(1 Suppl):S42–53 [PubMed: 15842086]
9. Carlos HA, Shi X, Sargent J, Tanski S, Berke EM (2010) Density estimation and adaptive bandwidths: a primer for public health practitioners. Int J Health Geogr 9:39 [PubMed: 20653969]
10. DATA USA (2020) Los Angeles County, CA. https://datausa.io/profile/geo/los-angeles-county-ca. Accessed 22 Oct 2020
11. White DL, Thrift AP, Kanwal F, Davila J, El-Serag HB (2017) Incidence of hepatocellular carcinoma in all 50 United States, from 2000 through 2012. Gastroenterology 152(4):812–820 [PubMed: 27889576]
12. Edge SB, Compton CC (2010) The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 17(6):1471–1474 [PubMed: 20180029]
13. De SMJ, Goodchild MF, Longley P (2007) Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Matador, Leicester
14. Gibin M, Longley P, Atkinson P (2007) Kernel density estimation and percent volume contours in general practice catchment area analysis in urban areas. Paper presented at Proceedings of Geographical Information Science Research Conference
15. Lichtensztajn DY, Giddings BM, Morris CR, Parikh-Patel A, Kizer KW (2017) Comorbidity index in central cancer registries: the value of hospital discharge data. Clin Epidemiol 9:601–609 [PubMed: 29200890]
16. Yang JSC, Harrati A, Clarke C, Keegan THM, Gomez SL (2014) Developing an area-based socioeconomic measure from American Community Survey data. Cancer Prevention Institute of California, Fremont
17. Yost K, Perkins C, Cohen R, Morris C, Wright W (2001) Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. Cancer Causes Control 12(8):703–711 [PubMed: 11562110]
18. Mapping LA Neighborhoods (2020) Los Angeles Times. http://maps.latimes.com/neighborhoods/. Accessed 12 Aug 2020
19. United States Census Bureau (2010) Census. U.S. Census Bureau. http://www.census.gov/data/tables.html. Published 2010. Accessed 1 Feb 2020
20. United States Census Bureau (2000) Census. U.S. Census Bureau. http://www.census.gov/data/tables.html. Published 2000. Accessed 1 Feb 2020
21. United States Census Bureau. TIGER/Line Shapefiles. https://www.census.gov/geo/maps-data/data/tiger-line.html. Published 2010. Accessed 1 Feb 2020
22. Andrilla CHA, Moore TE, Man Wong K, Evans DV (2019) Investigating the Impact of Geographic Location on Colorectal Cancer Stage at Diagnosis: A National Study of the SEER Cancer Registry. J Rural Health. 36(3):316 [PubMed: 31454856]
23. Altekruse SF, Petrick JL, Rolin AI et al. (2015) Geographic variation of intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, and hepatocellular carcinoma in the United States. PLoS One. 10(3):e01250574

24. Rich NE, Hester C, Odewole M et al. (2019) Racial and ethnic differences in presentation and outcomes of hepatocellular carcinoma. Clin Gastroenterol Hepatol. 17(3):551–559 [PubMed: 29859983]

25. Artinyan A, Mailey B, Sanchez-Luege N et al. (2010) Race, ethnicity, and socioeconomic status influence the survival of patients with hepatocellular carcinoma in the United States. Cancer 116(5):1367–1377 [PubMed: 20101732]

26. Nathani P, Gopal P, Rich N et al. (2021) Hepatocellular carcinoma tumour volume doubling time: a systematic review and meta-analysis. Gut 70(2):401–407 [PubMed: 32398224]

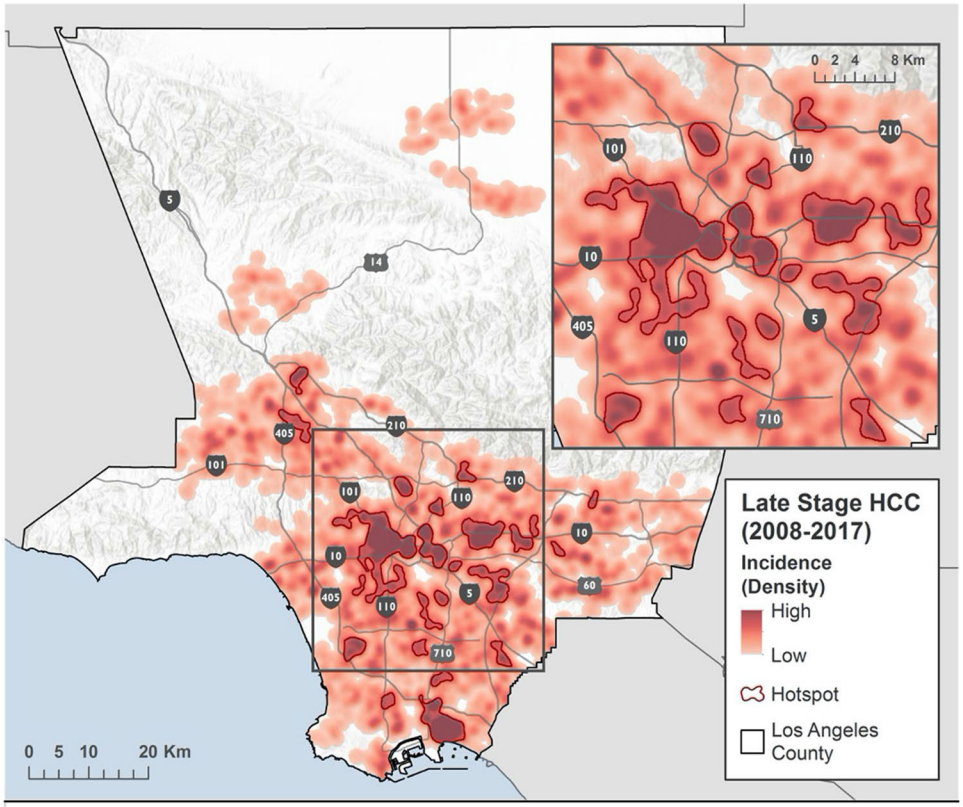
27. Marrero JA, Kulik LM, Sirlin CB et al. (2018) Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. Hepatology 68(2):723–750 [PubMed: 29624699]

28. Singal AG, Yopp A, Skinner CS, Packer M, Lee WM, Tiro JA (2012) Utilization of hepatocellular carcinoma surveillance among American patients: a systematic review. J Gen Intern Med 27(7):861–867 [PubMed: 22215266]

29. Gorin SS, Badr H, Krebs P, Prabhu DI (2012) Multilevel interventions and racial/ethnic health disparities. J Natl Cancer Inst Monogr 2012(44):100–111 [PubMed: 22623602]

30. Kerns K, Locklear L. Three New Census Bureau Products Show Domestic Migration at Regional, State, and County Levels. United States Census Bureau. https://www.census.gov/library/stories/2019/04/moves-from-south-west-dominate-recent-migration-flows.html. Published 2019. Accessed 8 Nov 2019

31. Moon AM, Singal AG, Tapper EB (2020) Contemporary epidemiology of chronic liver disease and cirrhosis. Clin Gastroenterol Hepatol 18(12):2650–2666 [PubMed: 31401364]



Cutpoint 147 (n=23 Hotspots)	Los Angeles County	Within Hotspots	Outside Hotspots	Difference
Number of Cases	7519	2011	5508	-3497
Number of Late Stage	3893	1188	2705	-1517
Proportion Late Stage	51.8%	59.1%	49.1%	10.0%

Fig. 1.

Late-stage HCC hotspots in Los Angeles County. Abbr: HCC = hepatocellular carcinoma 23 discrete hotspots were identified (outlined in red) representing areas of highest density of late-stage hepatocellular carcinoma. Cancer data came from the California Cancer Registry, Sites: C220 (8,170–8,175, males and females, all ages, 2008–2017). Data were managed, analyzed, and mapped in ArcGIS 10.3.1. Density distribution is based on latitude/longitude of patient address at diagnosis. Areas with sparse data have been suppressed

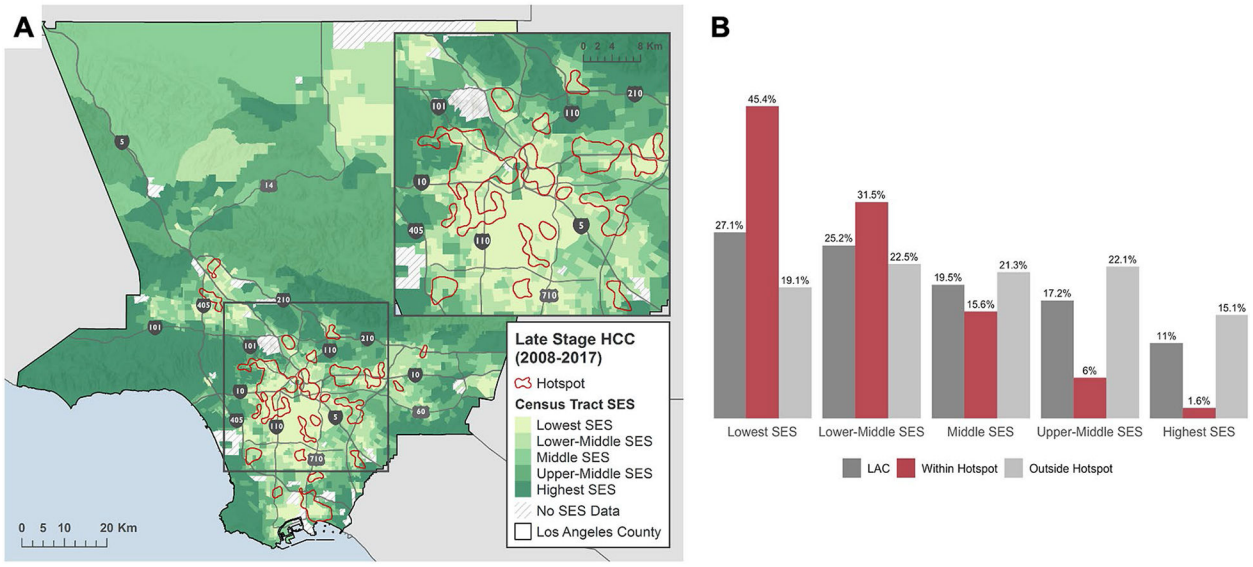
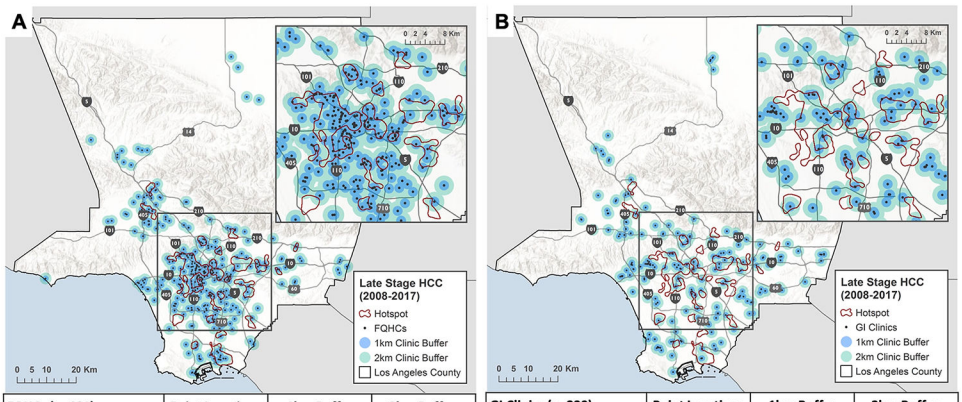


Fig. 2.

Relationship between socioeconomic status and hotspots by (**A**) mapping hotspots overlaid on census tract SES quintiles and (**B**) distribution of SES quintiles by hotspot status. *SES* socioeconomic, *HCC* hepatocellular carcinoma, *LAC* Los Angeles County. Each census tract in LAC was assigned a single SES quintile as represented in the left panel by a color scale: lightest green = lowest SES to darkest green = highest SES. Hotspot status for each census tract was assigned based on whether or not its centroid fell within a hotspot boundary



FQHCs (n=481)	Point Locations	1km Buffer	2km Buffer
Clinics in Hotspots	178	290	336
Hotspots without a Clinic	5	0	0

GI Clinics (n=229)	Point Locations	1km Buffer	2km Buffer
Clinics in Hotspots	36	67	103
Hotspots without a Clinic	14	8	3

Fig. 3.

Relationship between (**A**) primary care clinics and (**B**) specialty care clinics and late-stage hotspots. *FQHC* federally qualified health center, *GI* gastroenterology, *HCC* hepatocellular carcinoma. Primary care clinics with FQHC designation obtained from the Health Center Service Delivery Sites database (Health Resources and Services Administration). GI clinics obtained from Physician Compare Database (Medicare). 1-km (blue) buffer and 2-km (green) buffer added to clinics to examine impact on number of hotspots without a clinic

Table 1

Characteristics of HCC cases within and outside of late-stage hotspots

Characteristic	Overall		Within hotspot		Outside hotspot		*p*-value*
	n	%	*n*	%	*n*	%	
Age at diagnosis							0.78
65 and under	4,014	53.4	1,079	53.7	2,935	53.3	
Over 65	3,505	46.6	932	46.3	2,573	46.7	
Sex							0.84
Female	1,983	26.4	527	26.2	1,456	26.4	
Male	5,536	73.6	1,484	73.8	4,052	73.6	
Year of diagnosis							0.58
2008–2012	3,778	50.2	1,021	50.8	2,757	50.1	
2013–2017	3,741	49.8	990	49.2	2,751	49.9	
Race/ethnicity							< .001
Non-Hispanic White	1,953	26.0	275	13.7	1,678	30.5	
Non-Hispanic Black	786	10.5	234	11.6	552	10.0	
Hispanic (any race)	2,903	38.6	881	43.8	2,022	36.7	
Asian/Pacific Islander	1,805	24.0	603	30.0	1,202	21.8	
Other/unknown	72	1.0	18	0.9	54	1.0	
Nativity							< .001
US-born	2,593	34.5	530	26.4	2,063	37.5	
Foreign-born	3,007	40.0	977	48.6	2,030	36.9	
Unknown	1,919	25.5	504	25.1	1,415	25.7	
Marital status							< .001
Married	3,821	50.8	944	46.9	2,877	52.2	
Single	3,471	46.2	992	49.3	2,479	45.0	
Unknown	227	3.0	75	3.7	152	2.8	
SES [a]							< .001
Highest SES	897	11.9	23	1.1	874	15.9	
Upper-middle SES	1,319	17.5	104	5.2	1,215	22.1	
Middle SES	1,495	19.9	318	15.8	1,177	21.4	

Characteristic	Overall		Within hotspot		Outside hotspot		*p*-value*
	n	%	*n*	%	*n*	%	
Lower-middle SES	1,877	25.0	666	33.1	1,211	22.0	
Lowest SES	1,931	25.7	900	44.8	1,031	18.7	
Insurance							< .001
Insured	5,697	75.8	1,376	68.4	4,321	78.4	
Medicaid	1,346	17.9	480	23.9	866	15.7	
Not Insured/unknown	476	6.3	155	7.7	321	5.8	
Comorbidity index							0.06
Low	814	10.8	212	10.5	602	10.9	
Moderate	1,788	23.8	470	23.4	1,318	23.9	
High	3,476	46.2	903	44.9	2,573	46.7	
Unknown	1,441	19.2	426	21.2	1,015	18.4	

** p*-value from Pearson's chi-square

[a] Socioeconomic status (SES) quintile of Juan Yang's index of socioeconomic status with missing values imputed based on principal components analysis of census tract-level variables from the American Community Survey

Table 2

Individual-level factors associated with residence in a late-stage hotspot

	Univariate			Multivariable		
	OR	**LCL**	**UCL**	**OR**	**LCL**	**UCL**
Age at diagnosis						
65 and under	1.00	–	–	–		
Over 65	0.99	0.89	1.09	–		
Sex						
Female	1.00					
Male	1.01	0.90	1.14	–		
Race/Ethnicity						
Non-Hispanic White	1.00	–	–	1.00	–	–
Non-Hispanic Black	2.59	2.12	3.16	1.23	0.99	1.53
Hispanic (any race)	2.66	2.29	3.09	1.10	0.93	1.31
Asian/Pacific Islander	3.06	2.61	3.60	2.10	1.72	2.57
Other/unknown	2.03	1.18	3.52	1.73	0.94	3.17
Nativity						
US-born	1.00	–	–	1.00	–	–
Foreign-born	1.87	1.66	2.12	1.33	1.15	1.54
Unknown	1.39	1.21	1.59	1.33	1.08	1.64
Marital status						
Married	1.00	–	–	1.00	–	–
Single	1.22	1.10	1.35	1.19	1.06	1.34
Unknown	1.50	1.13	2.00	1.39	1.01	1.91
SES						
Highest SES	1.00	–	–	1.00	–	–
Upper-middle SES	3.25	2.05	5.15	3.09	1.95	4.90
Middle SES	10.27	6.66	15.82	9.73	6.30	15.03
Lower-middle SES	20.90	13.67	31.96	20.31	13.22	31.20
Lowest SES	33.17	21.72	50.67	32.52	21.12	50.06
Insurance						

	Univariate			Multivariable		
	OR	**LCL**	**UCL**	**OR**	**LCL**	**UCL**
Insured	1.00	–	–	1.00	–	–
Medicaid	1.74	1.53	1.98	1.25	1.08	1.44
Not insured/unknown	1.52	1.24	1.85	1.18	0.94	1.47
Comorbidity Index						
Low	1.00	–	–			
Moderate	1.01	0.84	1.22	–		
High	1.00	0.84	1.19	–		
Unknown	1.19	0.98	1.45	–		

LCL lower confidence level, *UCL* upper confidence level