

# Integrating geo-linked electronic health records with neighborhood data to identify place-based barriers for diabetes management in a public healthcare system

**Riya Desai**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Kim Hanh Nguyen**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Pushkar Inamdar**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Anna D. Rubinsky**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Mindy C. Hebert-DeRouen**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Jessica Fields**

Division of General Internal Medicine at Zuckerberg San Francisco General Hospital, University of California San Francisco, San Francisco, CA

**Kirsten Bibbins-Domingo**

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

**Courtney R. Lyles** (✉ [Courtney.Lyles@ucsf.edu](mailto:Courtney.Lyles@ucsf.edu))

Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA

---

## Research Article

**Keywords:** Diabetes, Neighborhood, Geospatial Analysis, Disease Hot Spot, Social Determinants of Health, Health Disparities, Electronic Medical Records, Population Health

**Posted Date:** October 8th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-913499/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Despite the body of research on place-based disparities in diabetes, there is room for advancing methods in this space by leveraging multiple datasets, including population census and clinical datasets. Specifically, healthcare systems can leverage geo-linked electronic health records (EHRs) to investigate geospatial clustering and identify social determinants of health.

## Methods

We examined geographic clustering of poor diabetes control and corresponding structural and social determinants of health within San Francisco's safety-net healthcare system. Specifically, we used EHR data to identify individual patient-level addresses and clinical outcomes for diabetes, using hot spot analysis to determine significant hot and cold clusters of diabetes control throughout San Francisco. In addition, by linking patient addresses to public datasets, we described the neighborhood-level conditions associated with hot spots.

## Results

Our hot spot analysis showed 18.8% of patients with poor glycemic control ( $n = 2,126$ ) clustered in hot spots, and 12.8% of patients with controlled diabetes clustered in cold spots. Neighborhoods of patients in hot spots had more Black, Latinx, Native American, and Native Hawaiian or Pacific Islander residents and more residents receiving SNAP benefits, experiencing unemployment, food insecurity, greater rent burden, and low grocery store access.

## Conclusions

We present methodological advantages of hotspot analyses utilizing EHR data, as well as the value of combining these approaches with additional public population census data about neighborhood conditions. Moving forward, healthcare systems must partner with public health agencies to utilize EHR data as a means to evaluate structural interventions and eventually target investment in place-based health-enabling resources.

## Background

It is well known that fundamental causes of place-based health disparities are rooted in structural racism and other structural determinants of health.[1–3] Numerous factors related to where people live—access to healthy food,[4–7] access to spaces for recreation and walkability,[5, 8, 9] built environment,[10, 11] neighborhood safety,[12] and neighborhood socioeconomic status[13, 14]—impact diabetes-related health

outcomes. Furthermore, these fundamental place-based factors also directly influence multiple other domains of health, from social and interpersonal interactions to individual health behaviors.[15]

Prior social epidemiological studies have utilized various methods to identify place-based disparities in diabetes prevalence and management, including county-, census tract-, or neighborhood-level prevalence measures, and corresponding hot spot analysis of prevalence.[16–19] While these studies contribute to an evidence-base for prioritizing place-based disparities in diabetes, there is an opportunity to leverage these approaches using clinical data captured in electronic health records (EHR), including objective health outcomes based on laboratory measures, linked to public population-level datasets that include information on neighborhood conditions.[20]

Beyond facilitating clinical data linkages and affording nuanced objective measures of health outcomes, healthcare systems themselves are increasingly prioritizing social determinants of health (SDOH) and considering new ways to address upstream and structural factors.[21–23] Financial incentive structures and value-based payment models have emphasized population and public health.[24, 25] Often, this focus on SDOH has increased the screening and referral for social resources for patient populations, and documenting these actions within electronic health records (EHRs).[23, 26, 27] However, most health systems have not effectively leveraged EHRs to understand the impact of place, specifically patient residence, on patient health outcomes.[23] Health systems have a specific opportunity to utilize geocoded EHR data at the level of individual patient address to understand the implications of patient neighborhood factors as they relate to healthcare utilization and health outcomes.

Therefore, we linked EHR and publicly available population data on neighborhood conditions to examine geographic clustering of poor glycemic control and corresponding place-based structural and social drivers of poor glycemic control within the public safety-net healthcare delivery system in San Francisco. Our ultimate goal was to enable data-driven approaches for future place-based interventions for diabetes management.

## Methods

**Study Population, Study Setting and Clinical Outcome.** The San Francisco Health Network (SFHN), including Zuckerberg San Francisco General Hospital, is an integrated safety-net healthcare system serving publicly insured and underinsured patients. We used SFHN EHR data from Jan 1, 2013 to Dec 31, 2017 to identify our study population and obtain individual-level patient characteristics including sociodemographic (e.g., race/ethnicity, insurance type) and clinical information (e.g., diabetes control). Our study population included SFHN patients who had an outpatient visit 2016-2017 and at least one additional outpatient visit within the prior two years, an ICD-9-CM or ICD-10-CM diagnosis of diabetes, at least one HbA1c lab result subsequent to diagnosis, and a residential address in San Francisco.[28]

We defined patients as having poor diabetes control if they had a glycosylated hemoglobin (HbA1c) level greater than 9% at their most recent lab test during the study period. We only considered HbA1c values

after a recorded diabetes diagnosis to ensure the lab result was capturing control of an active diabetes diagnosis instead of a lab result potentially leading to a diagnosis, prior to clinical treatment.

**Ethics and consent to participate.** This study was approved by the University of California, San Francisco Institutional Review Board. IRB approval allows for use of clinical patient health data for analysis. Human subjects were not involved in this study and therefore written or verbal informed consent was not required

**Neighborhood characteristics.** Structural and social determinants of health across several domains (racial/ethnic and language composition, socioeconomic context including poverty and unemployment, food environment and access, housing) were compared across patient cluster groupings. Data for neighborhood-level characteristics were downloaded from the UCSF Health Atlas, an interactive map with a catalog of characteristics that illustrate social, economic, and built environments in California.[29] Specific data sources for each characteristic are described below.

*1. Racial/ethnic and language composition of neighborhoods.*[3] At the census tract level, we extracted percent White, percent Black, percent Asian, percent Native Hawaiian or Pacific Islander, percent Native American (alone or in combination with other races), and percent Latinx residents, sourced from 2013-2017 American Community Survey (ACS) data.[30] We also measured limited English proficiency, defined as the percent of the population at the census tract level that speaks English less than "very well," and percent of the population that speaks a language other than English at home, also sourced from ACS.[30]

*2. Socioeconomic Context and Neighborhood Built Environment.* To measure several additional indicators of structural and social determinants of health at the neighborhood level, we examined:

*a. Poverty and unemployment.* Percent poverty from ACS data was defined as the percent of the population with income below 100% of the federal poverty level in the past 12 months,[30] which we included along with the Housing and Urban Development's Extremely Low-Income (ELI) measure, defined as below 30% of the area median income (relevant for high-income locations, such as San Francisco).[31] We also included percent unemployment from ACS.[30]

*b. Socioeconomic Indices.* We used a composite measure of the Healthy Places Index (combining economic, education, housing, healthcare access, neighborhood, clean environment, transportation, and social factors), where a higher percentile indicates less healthy neighborhood conditions.[32]

*c. Food environment and access.* Percent low-income, low-food access tracts was obtained from the US Department of Agriculture Food Access Research Atlas,[33] defined as low-income tracts where at least 500 people or 1/3 of the population lives more than half a mile away from the nearest supermarket. Percent of the population with Supplemental Nutrition Assistance Program (SNAP) benefits in the past 12 months was sourced from ACS.[30] Finally, food insecurity measures were census-tract level modeled estimates of percent food insecurity sourced from Feeding America.[29]

*d. Housing.* Housing data was sourced from HUD Comprehensive Housing Affordability Strategy Data.[31] Renter-occupied households were defined as the percent of housing units within a census tract that

are lived in by a renter. Severe rent burden is defined as the percentage of renter-occupied households in a census tract for whom housing costs are over 50% of household income.

**Statistical Analyses.** We conducted descriptive analyses of patient characteristics by uncontrolled diabetes, overall and by sociodemographic characteristics.

*Geospatial Analyses.* We geocoded residential addresses of patients in our study population, using patients' most recently recorded address in the EHR as of June 13, 2019. We used ArcGIS Pro Version 2.6 (Environmental Systems Research Institute, Inc., Redlands, CA, USA) for all geospatial analysis.

*Census Tract Prevalence.* We calculated the prevalence of poor glycemic control among diabetic patients in our study population by census tract and categorized census tracts into tertiles of high (between 18% - 47.1%), medium (11.9% - 17.9%), or low prevalence (less than 11.8%). Rates of small case counts are less reliable and therefore census tracts with fewer than 10 patients were excluded in rate calculation.

*Hot Spot Analysis.* We then conducted a hot spot analysis to identify hot and cold spots of poor diabetes control in San Francisco. We used the Getis-Ord  $G_i^*$  statistic to assess randomness of the spatial distribution of high (poor glycemic control) and low (good glycemic control) values using the "Hot Spot Analysis (Getis-Ord  $G_i^*$ )" tool in ArcGIS Pro 2.6.

The tool defines a "neighborhood" for each patient as the set of patients within a fixed distance band. The fixed distance band is determined using an incremental spatial autocorrelation test to assess the likelihood that spatial distribution patterns of high and low values are random. Using the Global Moran's  $I$  statistic and z-scores generated from the test for a range of fixed distances, we identified distance bands with the greatest likelihood of having a non-random spatial distribution in order to identify the distance band at which spatial patterns of diabetes control are most likely to be clustered. For the incremental autocorrelation test, we used the Euclidean distance method to examine 15 distance bands each 20 meters (0.03 miles) apart over a range of 500-800 meters (0.31 - 0.49 mi). We identified the distance band at 620 meters or 0.385 miles as having the maximum spatial autocorrelation with a z-score of 5.83 and p-value <0.001.

The Hot Spot Analysis (Getis-Ord  $G_i^*$ ) tool compares the prevalence expected value of diabetes control for all patients in the study population with the prevalence value of diabetes control within a patient's "neighborhood" and calculates a z-score and p-value for each patient. A patient is classified as a hot spot if there are statistically significantly more high values in the patient's "neighborhood" than in the full study area (San Francisco). A patient is classified as a cold spot if there are statistically significantly more low values in the patient's "neighborhood" than in the full study area. A high z-score indicated clustering of higher levels of poor diabetes control and a low negative z-score indicated clustering of lower levels of poor diabetes control, where the higher or lower the z-score indicates the intensity of clustering. We categorized the z-score into 3 categories—hot spots, cold spots, and not statistically significant—based on a 90% confidence level.

To mask point locations of patient residences and protect patient privacy, we used inverse distance weighting interpolation to visualize geographic areas of hot and cold spots.

*Associations between patient diabetes control clustering and census tract characteristics.* Finally, we summarized structural and social determinants of health indicators by patient cluster groupings (hot spots, cold spots, and not significant) to compare values and describe observed differences across these cluster groups. We assigned the census tract values for all characteristics to each patient and then averaged the census tract values of each characteristic for patients within each cluster classification (hot spot, cold spot, and not significant).

## Results

Our study sample included 11,333 SFHN patients with diabetes living in San Francisco. Average patient age was 59.6 and 51.1% were women (Table 1). Most patients were insured through public insurance plans: 30.9% with Medicare, 31.8% with Medi-Cal, and 6.9% with Healthy SF, a county insurance program for low-income San Francisco residents ineligible for Medicare or Medi-Cal. Nearly a third (32.6%) of patients were Asian, 31.9% were Latinx, 16.4% were Black, 11.7% were White, and 2.1% were Native Hawaiian or Pacific Islander. Overall, 16.3% of patients had uncontrolled diabetes; uncontrolled diabetes was most common among Black, Latinx, and Native Hawaiian or Pacific Islander patients, and patients with Healthy SF insurance, Medi-Cal, or private insurance and patients who were uninsured (Table 1).

Table 1  
Sociodemographic Characteristics of San Francisco Health Network Patients with Diabetes between 2013-2017

	<b>Total N=11,333 (%)</b>	<b>% with Uncontrolled Diabetes (16.3% overall)</b>
Mean Age (SD)	59.6 (12.4)	-
Sex		
Female	5,794 (51.1%)	14.4%
Male	5,539 (48.9%)	18.3%
Race/ethnicity		
White (Non-Hispanic)	1,327 (11.7%)	16.0%
Black (Non-Hispanic)	1,864 (16.4%)	17.3%
Latinx/Hispanic	3,610 (31.9%)	23.2%
Asian (Non-Hispanic)	3,693 (32.6%)	8.7%
Native Hawaiian or Pacific Islander (Non-Hispanic)	235 (2.1%)	28.1%
Other	564 (5%)	14.5%
Insurance Type		
Medicare	3,506 (30.9%)	11.4%
Medi-Cal	3,599 (31.8%)	20.0%
Uninsured	219 (1.9%)	25.6%
Private	134 (1.2%)	24.6%
Other Coverage	1,045 (9.2%)	9.0%
Healthy SF	787 (6.9%)	26.1%
Missing	2,043 (18.0%)	16.5%

**Census tract-level prevalence of uncontrolled diabetes.** Patients in our study population were dispersed across 194 of the 196 San Francisco census tracts. In calculating prevalence of poor glycemic control at the census tract level, we excluded 33 census tracts because they each had fewer than 10 patients. Census tracts identified in the tertile with the highest prevalence of uncontrolled diabetes are shown in Figure 1.

**Hotspot analysis findings.** Within the study population, 18.8% of patients (n = 2,126) were clustered in hot spots (high levels of uncontrolled diabetes), 12.8% (n = 1,448) were clustered in cold spots (low levels of

uncontrolled diabetes), and 68.5% (n = 7,759), were not significantly spatially clustered (Figure 2). Hot spots were primarily within the southeast quadrant of San Francisco, whereas cold spots were primarily within the western and north eastern neighborhoods of San Francisco.[34] As expected, very few census tracts had both hot and cold spot patients.

**Census tract SDOH characteristics and clustering of uncontrolled diabetes.** Census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, had much higher prevalence on average of Black residents (20.1% vs. 3.7%), Latinx residents (31.1% vs. 8.0%), Native Hawaiian or Pacific Islander residents (3.2% vs. 0.6%) and Native American residents (1.4% vs. 0.6%), and lower prevalence on average of Asian residents (26.6% vs. 54.7%) and White residents (21.3% vs. 32.7%) (Table 2). Census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, also had lower prevalence on average of residents with limited English proficiency (25.0% vs. 32.6%) and residents who spoke a language other than English at home (49.3% vs. 53.7%).

Table 2  
Census Tract Characteristics by Patient Clustering Types

		Average Values by Patient Clustering Types		
Structural and Social Determinants		Hot Spot (n= 2,126 patients)	Cold Spot (n=1,448 patients)	Not Significant (n=7,759 patients)
Demographics	White (%)	21.3	32.7	29.4
	Black (%)	20.1	3.7	9.6
	Asian (%)	26.6	54.7	39.4
	Native Hawaiian and Pacific Islander (%)	3.2	0.6	0.8
	Native American (%)	1.4	0.6	1.4
	Latinx (%)	31.1	8.0	21.5
	Limited English Proficiency (%)	25.0	32.6	27.3
	Speaks a language other than English at home (%)	49.3	53.7	52.5
Food	SNAP (%)	15.2	5.1	8.8
	Low-income low-access tracts (%)	32.1	8.6	10.2
	Food insecure (%)	19.9	17.3	17.8
Poverty	Poverty (%)	19.7	13.9	16.2
	Extremely Low-income households (%)	36.5	28.8	31.3
	Percent unemployed (%)	5.8	3.8	4.2
Housing	Renter-occupied (%)	67.7	55.9	62.6
	Rent burden (%)	16.1	13.3	14.3
Neighborhood Indices	Healthy Places Index (Percentile)	55.9	32.1	43.0
	Social Vulnerability Index (Percentile)	62.5	42.7	53.5

Additionally, census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, had a higher prevalence of residents living below poverty (19.7% vs. 13.9%), extremely low-income households (36.5% vs. 28.8%), and unemployment (5.8% vs. 3.8%).

Census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, had a higher prevalence of residents receiving SNAP benefits (15.2% vs 5.1%), residents experiencing food insecurity (19.9% vs. 17.3%), and low-income, low-food access tracts (32.1% vs. 8.6%).

Census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, had a higher prevalence of renter-occupied households (67.7% vs. 55.9%) and households experiencing high rent burden (16.1% vs. 13.3%).

As indicators of overall neighborhood socioeconomic status, census tracts of residence for patients in hot spot clusters, compared to patients in cold spot clusters, had a higher Healthy Places Index percentile score (55.9% vs. 32.1%), indicating less healthy neighborhood conditions.

## Discussion

Our study supports the use of geospatial analysis to identify place-based diabetes disparities in a safety-net healthcare system. Hot spot analysis showed that nearly 1/5 of diabetes patients in a safety-net healthcare system were clustered in hot spots, and hot spots were primarily located within the neighborhoods of San Francisco known to be more underserved in terms of social service investment prioritization from the city and county government.

Based on the public availability of neighborhood-level data at the census tract level, we also compared these hot spots to census tract characteristics of patients. This revealed associations between our geospatial analysis and neighborhood variables reflecting structural racism and its health-harming impacts on glycemic control, such as the racial/ethnic composition of neighborhoods and multiple other neighborhood socioeconomic indicators. While many research studies are often exclusively using either geospatially-driven associations or place-based analyses using census tracts or other fixed boundaries, we feel that it is important to combine these methods – especially since different scientific disciplines and different leaders (e.g., healthcare system executives, policymakers) have different sets of data available to them as they make critical decisions about interventions and programming.

A strength of our study was our ability to link EHR data with information about place-based social determinants of health. Most publicly available data sources do not typically include robust clinical data on uncontrolled diabetes, given that survey data often relies on self-reported diagnosis at a cross-sectional time point, and are limited to associations using pre-defined geographic boundaries. Thus, 1) the ability to use patient-level lab test results of clinical control verified within the EHR is a significant improvement on outcome ascertainment, and 2) our primary hotspot analysis was not limited from the outset by administrative geographic boundaries such as census tract definitions. Moving forward, healthcare systems must be at the table when examining place-based disparities in their regions and catchment areas – particularly in the U.S. where the clinical data stored in EHRs can be fragmented across many different healthcare institutions and not available for public health planning or programming. More specifically, with the increased use of clinical outcomes, public datasets, and hot spot analysis in future studies, we may be better able to bridge the health and healthcare priorities across

healthcare systems, public health departments, and city agencies to align and guide investment in local health-enabling resources and programs.

Beyond working with other public health partners in their region, healthcare systems can also leverage EHRs to integrate population health informatics to conduct their own programs and interventions, including those at the structural or neighborhood level.[21, 35, 36] Population-level data infrastructure within healthcare systems can be a tool for surveillance of health disparities, programmatic decision-making, integration of health outcomes and social determinants of health, and evaluation of interventions.[21, 23, 26, 35–38] Furthermore, this work is underscored by new reimbursement approaches prioritizing value and incentivizing health plans and systems to more holistically address social and medical needs, reduce healthcare disparities, and build infrastructural and organizational capacity for structural change.[24, 25, 27]

**Limitations.** Our study population was comprised of patients from only one healthcare system in San Francisco, yet we chose to focus on the public delivery system which serves the majority of underserved patients in our city. In addition, the results of a spatial analysis do not capture the lived experiences of residents with uncontrolled diabetes, nor does it reveal what resources, support, and changes patients with uncontrolled diabetes want and need in order to improve their health.[39] Integration of data-driven approaches with patient and community-centered partnerships presents an important opportunity for healthcare systems to meaningfully invest in place-based health interventions and evaluate programs by tracking geospatial health outcomes. Finally, while we compared our geospatial hotspot analysis to census tract level prevalence of key neighborhood characteristics, we also understand that this approach was descriptive to show the intersection of these methods.

## Conclusions

Combining methods, such as geospatial analyses with neighborhood-level associations, as well as datasets, such as EHR-based demographics and clinical outcomes with publicly available neighborhood indicators, can demonstrate multiple lenses for place-based and structural health and healthcare disparities research. Leveraging these datasets in future strategic and collaborative local partnerships, healthcare systems can be better positioned to provide structural interventions that align with ongoing public health programs.

## Abbreviations

SDOH

Social Determinants of Health

PRIME

Public Hospital Redesign and Incentives in Medi-Cal

EHR

Electronic Health Record

SFHN  
San Francisco Health Network  
HbA1c  
Glycosylated Hemoglobin  
ELI  
Extremely Low-Income  
SNAP  
Supplemental Nutrition Assistance Program

## Declarations

**Ethics approval:** This study was approved by the University of California, San Francisco Institutional Review Board (RB#: 17-22277). IRB approval allows for use of clinical patient health data for analysis. Human subjects were not involved in this study and therefore written or verbal informed consent was not required. All methods were carried out in accordance with relevant guidelines and regulations.

**Consent for publication:** Not applicable.

**Availability of data and materials:** The electronic health record data that support the findings of this study are not publicly available due to containing protected health information that could compromise research participant privacy and HIPAA regulations. Publicly available population-level data sources accessed for this study are collated and available from: [www.healthatlas.ucsf.edu](http://www.healthatlas.ucsf.edu)

**Conflicting and Competing Interests:** None to report.

**Funding statement:** This work was funded by the Population Health Data Initiative within the University of California San Francisco School of Medicine Dean's Office and the University of California San Francisco Executive Vice Chancellor's Office. CRL also received support from a UCSF Mid-Career Development Award.

**Authors' contributions:** All authors contributed to the paper as follows: study conception and design: RD, KHN, PI, ADR, MCHD, JF, KBD, CRL; data collection: RD, ADR; analysis and interpretation of results: RD, KHN, PI, ADR, MCHD, JF, KBD, CRL; draft manuscript preparation: RD, KHN, ADR, JF, CRL. All authors reviewed the results and approved the final version of the manuscript.

**Acknowledgments:** We would like to acknowledge Alice Fishman for her administrative support with this project.

## References

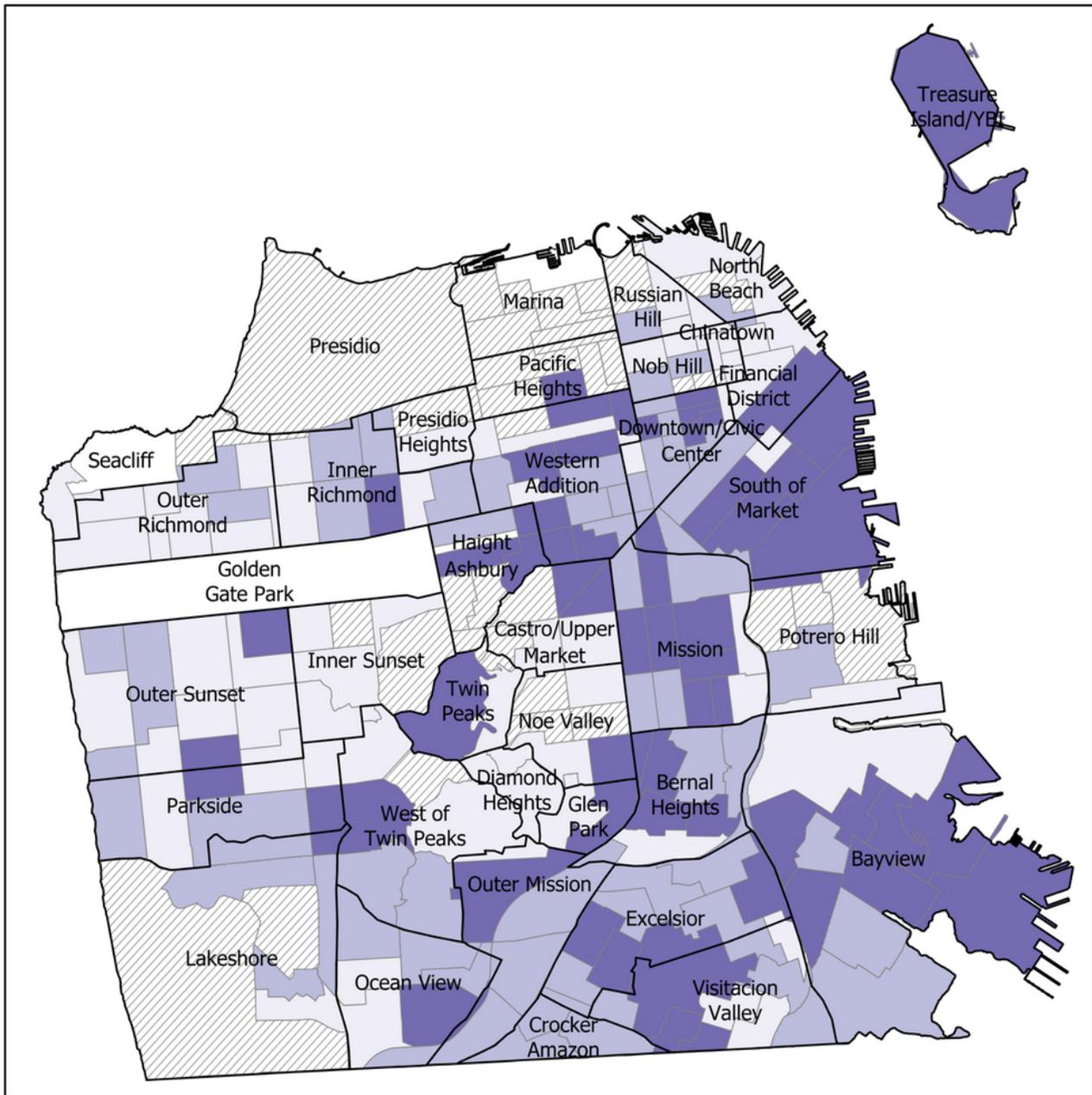
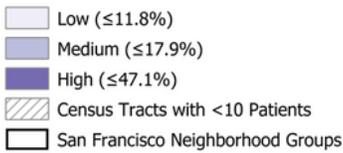
1. Phelan JC, Link BG. Is Racism a Fundamental Cause of Inequalities in Health? *Annual Review of Sociology*. 2015;41(1):311–330. doi:10.1146/annurev-soc-073014-112305

2. Williams DR, Lawrence JA, Davis BA. Racism and Health: Evidence and Needed Research. *Annual Review of Public Health*. 2019;40(1):105–125. doi:10.1146/annurev-publhealth-040218-043750
3. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *The Lancet*. 2017;389(10077):1453–1463. doi:10.1016/s0140-6736(17)30569-x
4. Phillips AZ, Rodriguez HP. Adults with diabetes residing in “food swamps” have higher hospitalization rates. *Health Serv Res*. 2019;54:217–225. doi:10.1111/1475-6773.13102
5. Auchincloss AH. Neighborhood Resources for Physical Activity and Healthy Foods and Incidence of Type 2 Diabetes Mellitus: The Multi-Ethnic Study of Atherosclerosis. *Arch Intern Med*. 2009;169(18):1698. doi:10.1001/archinternmed.2009.302
6. Zhang YT, Mujahid MS, Laraia BA, et al. Association Between Neighborhood Supermarket Presence and Glycated Hemoglobin Levels Among Patients With Type 2 Diabetes Mellitus. *American Journal of Epidemiology*. 2017;185(12):1297–1303. doi:10.1093/aje/kwx017
7. Kraft AN, Thatcher EJ, Zenk SN. Neighborhood Food Environment and Health Outcomes in U.S. Low-Socioeconomic Status, Racial/Ethnic Minority, and Rural Populations: A Systematic Review. *Journal of Health Care for the Poor and Underserved*. 2020;31(3):1078–1114. doi:10.1353/hpu.2020.0083
8. Booth GL, Creatore MI, Moineddin R, et al. Unwalkable Neighborhoods, Poverty, and the Risk of Diabetes Among Recent Immigrants to Canada Compared With Long-Term Residents. *Diabetes Care*. 2013;36(2):302–308. doi:10.2337/dc12-0777
9. Creatore MI, Glazier RH, Moineddin R, et al. Association of Neighborhood Walkability With Change in Overweight, Obesity, and Diabetes. *JAMA*. 2016;315(20):2211. doi:10.1001/jama.2016.5898
10. Christine PJ, Auchincloss AH, Bertoni AG, et al. Longitudinal Associations Between Neighborhood Physical and Social Environments and Incident Type 2 Diabetes Mellitus: The Multi-Ethnic Study of Atherosclerosis (MESA). *JAMA Intern Med*. 2015;175(8):1311. doi:10.1001/jamainternmed.2015.2691
11. Gordon-Larsen P. Inequality in the Built Environment Underlies Key Health Disparities in Physical Activity and Obesity. *PEDIATRICS*. 2006;117(2):417–424. doi:10.1542/peds.2005-0058
12. Tamayo A, Mujahid MS, Laraia B, et al. Police-Recorded Crime and Perceived Stress among Patients with Type 2 Diabetes: the Diabetes Study of Northern California (DISTANCE). *J Urban Health*. 2016;93(5):745–757. doi:10.1007/s11524-016-0069-2
13. Nazmi A, Diez Roux A, Ranjit N, Seeman TE, Jenny NS. Cross-sectional and longitudinal associations of neighborhood characteristics with inflammatory markers: Findings from the multi-ethnic study of atherosclerosis. *Health & Place*. 2010;16(6):1104–1112. doi:10.1016/j.healthplace.2010.07.001
14. Laraia BA, Karter AJ, Warton EM, Schillinger D, Moffet HH, Adler N. Place matters: Neighborhood deprivation and cardiometabolic risk factors in the Diabetes Study of Northern California (DISTANCE). *Social Science & Medicine*. 2012;74(7):1082–1090. doi:10.1016/j.socscimed.2011.11.036

15. National Institute on Minority Health and Health Disparities Research Framework. Published online 2017. Accessed December 21, 2020. <https://www.nimhd.nih.gov/about/overview/research-framework.html>
16. Lee DC, Jiang Q, Tabaei BP, et al. Using Indirect Measures to Identify Geographic Hot Spots of Poor Glycemic Control: Cross-sectional Comparisons With an A1C Registry. *Dia Care*. 2018;41(7):1438–1447. doi:10.2337/dc18-0181
17. Curtis AJ, Lee W-AA. Spatial patterns of diabetes related health problems for vulnerable populations in Los Angeles. *Int J Health Geogr*. 2010;9(1):43. doi:10.1186/1476-072X-9-43
18. S. Shrestha S. Spatial Clusters of County-Level Diagnosed Diabetes and Associated Risk Factors in the United States. *TODIAJ*. 2012;5(1):29–37. doi:10.2174/1876524601205010029
19. Kolak M, Abraham G, Talen MR. Mapping Census Tract Clusters of Type 2 Diabetes in a Primary Care Population. *Prev Chronic Dis*. 2019;16:180502. doi:10.5888/pcd16.180502
20. Schinasi LH, Auchincloss AH, Forrest CB, Diez Roux AV. Using electronic health record data for environmental and place based population health research: a systematic review. *Annals of Epidemiology*. 2018;28(7):493–502. doi:10.1016/j.annepidem.2018.03.008
21. Galea S, Abdalla SM, Sturchio JL. Social determinants of health, data science, and decision-making: Forging a transdisciplinary synthesis. *PLOS Medicine*. 2020;17(6):e1003174. doi:10.1371/journal.pmed.1003174
22. Braveman P, Egerter S, Williams DR. The Social Determinants of Health: Coming of Age. *Annual Review of Public Health*. 2011;32(1):381–398. doi:10.1146/annurev-publhealth-031210-101218
23. Gamache R, Kharrazi H, Weiner J. Public and Population Health Informatics: The Bridging of Big Data to Benefit Communities. *Yearbook of Medical Informatics*. 2018;27(01):199–206. doi:10.1055/s-0038-1667081
24. Chuang E, Pourat N, Haley LA, O'Masta B, Albertson E, Lu C. Integrating Health And Human Services In California's Whole Person Care Medicaid 1115 Waiver Demonstration. *Health Affairs*. 2020;39(4):639–648. doi:10.1377/hlthaff.2019.01617
25. Pourat N. *California Public Hospitals Improved Quality of Care Under Medicaid Waiver Program*. UCLA Center for Health Policy Research; 2017:10.
26. Horwitz LI, Chang C, Arcilla HN, Knickman JR. Quantifying Health Systems' Investment In Social Determinants Of Health, By Sector, 2017–19: Study analyzes the extent to which US health systems are directly investing in community programs to address social determinants of health. *Health Affairs*. 2020;39(2):192-198. doi:10.1377/hlthaff.2019.01246
27. Saulsberry L, Peek M. Financing Diabetes Care in the U.S. Health System: Payment Innovations for Addressing the Medical and Social Determinants of Health. *Curr Diab Rep*. 2019;19(11):136. doi:10.1007/s11892-019-1275-6
28. Khokhar B, Jette N, Metcalfe A, et al. Systematic review of validated case definitions for diabetes in ICD-9-coded and ICD-10-coded data in adult populations. *BMJ Open*. 2016;6(8):e009952. doi:10.1136/bmjopen-2015-009952

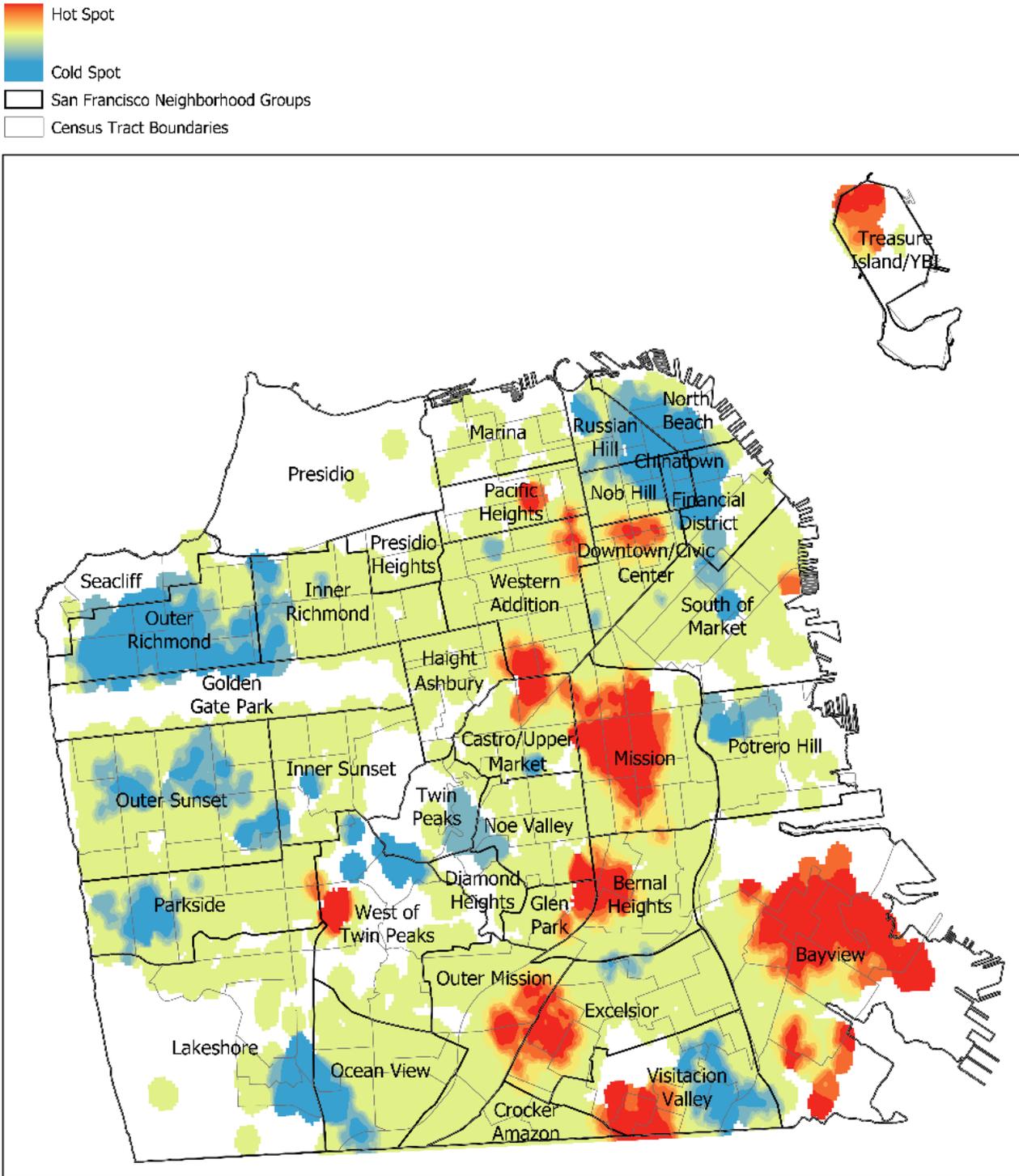
29. UCSF Health Atlas [Internet]. University of California San Francisco, Population Health and Health Equity; Available from: <https://healthatlas.ucsf.edu/>
30. U.S. Census Bureau; American Community Survey, 2013-2017 American Community Survey 5-Year Estimates, Tables DP02, DP03, DP05, S1701, S2201 (10 November 2019).
31. Consolidated Planning/Comprehensive Housing Affordability Strategy Data. U.S. Department of Housing and Urban Development's Office of Policy Development and Research. Accessed December 7, 2020. [https://www.huduser.gov/portal/datasets/cp.html#2006-2017\\_data](https://www.huduser.gov/portal/datasets/cp.html#2006-2017_data)
32. California Healthy Places Index Map. Accessed December 12, 2020. <https://map.healthyplacesindex.org/>
33. USDA ERS - Food Access Research Atlas. Accessed December 7, 2020. <https://www.ers.usda.gov/data-products/food-access-research-atlas/>
34. Neighborhood Groups Map. City and County of San Francisco Planning Department. Accessed December 7, 2020. <https://sfgov.org/sfplanningarchive/neighborhood-groups-map>
35. Alvidrez J, Nápoles AM, Bernal G, et al. Building the Evidence Base to Inform Planned Intervention Adaptations by Practitioners Serving Health Disparity Populations. *American Journal of Public Health*. 2019;109(S1):S94-S101. doi:10.2105/ajph.2018.304915
36. Frier A, Devine S, Barnett F, Dunning T. Utilising clinical settings to identify and respond to the social determinants of health of individuals with type 2 diabetes—A review of the literature. *Health Soc Care Community*. 2020;28(4):1119–1133. doi:10.1111/hsc.12932
37. Minkler M, Estrada J, Dyer S, Hennessey-Lavery S, Wakimoto P, Falbe J. Healthy Retail as a Strategy for Improving Food Security and the Built Environment in San Francisco. *Am J Public Health*. 2019;109(S2):S137-S140. doi:10.2105/AJPH.2019.305000
38. Pastor M, Morello-Frosch R. Integrating Public Health And Community Development To Tackle Neighborhood Distress And Promote Well-Being. *Health Affairs*. 2014;33(11):1890–1896. doi:10.1377/hlthaff.2014.0640
39. Petteway R, Mujahid M, Allen A. Understanding Embodiment in Place-Health Research: Approaches, Limitations, and Opportunities. *Journal of Urban Health*. 2019;96(2):289–299. doi:10.1007/s11524-018-00336-y

## Figures



**Figure 1**

Census-tract level prevalence of poor glycemic control among diabetic patients in the San Francisco Health Network between 2013-2017 categorized into tertiles (excluding census tracts with fewer than 10 patients).



**Figure 2**

Inverse distance weighted interpolation of hot spot analysis at a distance band of 620 meters (.385 miles) with patients categorized as hot spot, cold spot, or not statistically significant, based on a 90% confidence interval.