Risk factors for persistent fatal opioid-involved overdose hotspots in Massachusetts 2011-2021: A spatial statistical analysis with socio-economic, access and prescription factors

Sumeeta Srinivasan (sumeeta.srinivasan@tufts.edu)
Tufts University

Jennifer Pustz
Tufts University School of Medicine

Elizabeth Marsh
Institute of Health Metrics and Evaluation

Leonard D. Young
Massachusetts Department of Public Health

Thomas J. Stopka
Tufts University School of Medicine

Research Article

Keywords: Opioids, hotspots, opioid death, overdose, social determinants of health, public health, built environment, spatial statistics

Posted Date: September 11th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3249650/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** Fatal opioid-involved overdose rates increased precipitously from 5.0 per 100,000 population to 33.5 in Massachusetts between 1999 and 2022.

**Methods:** We use spatial rate smoothing techniques to identify persistent opioid overdose fatality clusters at the ZIP Code Tabulation Area (ZCTA) level. Rate smoothing techniques were effective in reducing variance common with zero-inflated rates locations where population counts are low such as rural and suburban areas which were affected by the epidemic in Massachusetts such as Worcester, Fall River, New Bedford, and Wareham. We use Getis-Ord hotspot analyses with the smoothed incidence rates to identify locations of persistent risk from 2011-2021. We constructed measures of the socio-built environment and potentially inappropriate prescribing (PIP) using principal components analysis (PCA). The resulting measures were used as covariates in autologistic, zero-inflated Poisson, negative binomial and Conditional Autoregression (CAR) Bayesian regression models to predict if a ZCTA was part of an opioid-involved smoothed hotspot cluster for fatal overdose rates as well as the number of times that it was part of a hotspot.

**Results:** Persistent hotspot clusters in Massachusetts had higher mean percentages of Black and Hispanic residents, and residents experiencing poverty. PCA helped in identifying unique socio-environmental factors, such as poverty and minority presence by combining socioeconomic, built environment and prescription variables that were highly correlated with each other. Fatal opioid-involved overdose hotspots were found to be significantly more likely to be ZCTA with high poverty levels and high percentages of people from minoritized populations. Regressions models that corrected for spatial autocorrelation were necessary to avoid model misspecification.

**Conclusion:** Conducting spatially robust analyses may help inform policies to identify community-level risks for opioid-involved overdose deaths. The results can help inform policy makers and planners about locations of persistent risk.

Introduction

The opioid overdose crisis continues as one of the most significant public health challenges of the past two decades in the US. Between 1999 and 2019, the national age-adjusted opioid-involved overdose death rate increased from 2.9 per 100,000 population to 15.5 per 100,000.\textsuperscript{1} Rates have increased steeply during the COVID-19 pandemic. In 2020, the rate climbed to 21.4 per 100,000 and to 24.7 per 100,000 in 2021.\textsuperscript{2} Commonly understood drivers of the overdose crisis include drug supply and demand side pressures. Supply-side drivers have evolved over multiple waves, from prescription opioids, to heroin, to illicitly manufactured fentanyl.\textsuperscript{3,4} Demand side drivers include deindustrialization and concentrated poverty, pain arising from work-related injuries,\textsuperscript{4} income inequality,\textsuperscript{5} and added stress, isolation, and economic disadvantage connected to the COVID-19 pandemic.\textsuperscript{6,7}
More studies have begun to consider the association between drug use, opioid-related mortality, and the built-environment, defined by Ezell and colleagues as “the purposeful creation and spatial arrangement of housing, sidewalks, roadways, retail and institutional buildings, public transit, and green spaces.” 8–10 Research has already suggested that the built environment influences health and health behaviors,11 including substance use,8,9,12 and opioid-related mortality.10 For example, analgesic opioid-involved overdose fatalities were found more likely to occur in “fragmented” neighborhoods than higher-income neighborhoods in New York City.13 Chichester et al. found that bus stops and public schools were associated with increased risk of opioid overdose in rural areas of an Alabama county and that inpatient treatment centers, transitional living facilities, express loan establishments, and liquor vendors were associated with increased opioid overdose risk in urban areas of the same county.14 Inequality and racial composition of neighborhoods have also been associated with increased opioid mortality.15,16

The connection between substance use and built environment variables (access to public restrooms, access to pharmacies, and driving distance to services, defined in our study as fast-food restaurants, gas stations, and highway exits) are also important. Public restrooms are associated with people who inject drugs (PWID) because many people (one study estimates 48%) use drugs in these spaces.8,17,18 Pharmacies represent an important access variable for several reasons. During the initial wave of the overdose crisis, pharmaceutical prescriptions, either legitimate, diverted, or potentially inappropriate, fed the opioid supply.3,16,19 In addition, naloxone (a medication to reverse an opioid overdose) is available at pharmacies without a prescription,20,21 although this provision may vary by neighborhood socio-demographic levels.22 For example, prescription opioid poisoning increased more in postal codes with greater pharmacy density in California.23 Road access to services may mediate several factors related to substance use, such as access to meeting places to buy illicit substances, as well as access to harm reduction services such as syringe services programs.8

Over the past two decades, the Massachusetts’s opioid-involved mortality rate has often been higher than the national rate and, at times, twice as high.24 The most recent data released by the Massachusetts Department of Public Health indicated that the provisional opioid-involved mortality rate reached a new high of 33.5 per 100,000 population in 2022.25 Studies focusing on the overdose crisis in Massachusetts have addressed several intersections of the built environment and opioid overdose. A spatial analysis of potentially inappropriate prescribing (PIP) identified several overdose and PIP clusters, but did not find a significant overlap between the two.19 Other research findings include the identification of one rural county in Massachusetts with both good access to harm reduction measures and high overdose rates,26 and that a majority of overdose deaths in the state occurred at home between 2015-2017.27 A recent study of opioid-related deaths in Massachusetts incorporated psychosocial, economic, built environment, and health-related variables using multilevel mixed-effects regression models, and found that none of the built environment variables had a statistically significant association with opioid-related mortality.28

While these studies across the US and within Massachusetts have contributed importantly to our understanding of mediating factors between substance use and the built environment, many often fail to
use spatial statistical methods that properly account for excess zero counts for overdose outcomes and spatial autocorrelation, two of the most common confounding factors in spatial epidemiology. With spatial analysis of rare events, such as overdose deaths, the distribution of death counts is often “zero-inflated” (i.e., a large number of locations have no deaths while a few locations have many). Several methods have been used to address this issue, including use of: small area estimation techniques; a rate smoothing technique to examine characteristics of prescription opioid poisoning; smoothing rates in ZCTAs with small populations; and Empirical Bayes smoothing in rural Illinois. These methods include data from adjacent spatial units to refine the estimates for locations that have low populations. Two models that account for both spatial autocorrelation and for excess zeros include zero-inflated Poisson regression models, such as the Besag-York-Mollie (BYM) model, and a Bayesian hierarchical space–time misalignment Poisson model.

The problem of excess zeros skewing true counts and rates exists in Massachusetts analyses because the geographical distribution of opioid overdose deaths in Massachusetts spans urban, suburban, and sparsely populated rural areas that are likely to have very low or zero counts. To date, we are only aware of a few studies of opioid-involved overdose deaths using Massachusetts data that have employed techniques that account for the zero-inflation and spatial autocorrelation at smaller spatial units (such as the ZCTA level) but these techniques could greatly help in precisely understanding geographic variation in overdose mortality across Massachusetts.

The goal of our study was to identify fatal overdose clusters as well as socioeconomic and built environment factors associated with opioid-related overdose rates in Massachusetts from 2011 to 2021 accounting for spatial autocorrelation and for zero-inflated rates. We aimed to: 1) address the issue of zero-inflated opioid-involved overdose death rates by comparing three spatial methods (raw rates, Empirical Bayes, and Empirical Bayes Spatial); 2) utilize the three rate types to identify statistically significant fatal opioid overdose clusters in Massachusetts between 2011 and 2021; 3) derive socio-environmental and pharmacological variables using PCA in those clusters; and 4) model community-level factors associated with overdose mortality rates using autologistic regression models to predict if a ZCTA was part of a hotspot and then zero-inflated regression models to predict how many times it was part of a hotspot during the time period while accounting for spatial autocorrelation of the outcome (Fig. 1).

Data and Methods

Data Sources. We obtained fatal opioid-related overdose data by address of residence for decedents between 2011 and 2021, and by address of recorded death for 2015 to 2021, from the Massachusetts Registry of Vital Records and Statistics (RVRS). These data included sociodemographic characteristics of decedents, including sex, race, ethnicity, and age at the time of death. We obtained opioid prescription data from the Massachusetts Prescription Monitoring Program (MA PMP), aggregated at the ZCTA level. We obtained address level data for services (gas stations, fast food restaurants, pharmacies), and “access to infrastructure” measures (highway exits, major roads) from Data Axle and from MassGIS. We obtained pharmacy addresses from the Massachusetts Board of Registration in Pharmacy. Finally, we compiled
population level sociodemographic data, at the ZCTA level, from the American Community Survey (ACS) 2011–2015 and the 2017–2021 5-year estimates for people aged ≥ 10 years.44

Outcome. Whether and the number of times a ZCTA was within fatal overdose hotspot cluster based on annual rates between 2011 and 2021. We calculated the Getis-Ord hotspots for each year between 2011 and 2021 for smoothed rates using the Empirical Bayes and Spatial Empirical Bayes rates at the ZCTA level.45 We used 2011–2015 population estimates for the years 2011–2017 and 2017–2021 population estimates for the later years included in our study: 2017–2021.

Covariates. PIP measures were aggregated by ZCTA. We defined PIP measures, obtained from the Massachusetts PMP, based on the following criteria, established through our previous research, for the years 2011–2017 at the ZCTA scale: high MME ( > = 90), co-prescribing of benzodiazepines and opioids, poly-pharmacy opioid prescription fills, multiple provider episodes (i.e., doctor shopping), >=3 cash purchases of opioid prescriptions and opioid prescriptions without a pain diagnosis.

Socioeconomic Measures. We selected socioeconomic measures from the ACS at the ZCTA-level based on the literature and our previous research.19,46,47 Poverty as percentage of households living below the poverty threshold as defined by the 2011–2015 and 2017–2021 five-year ACS estimates. We also included median age, and population percentages by race and ethnicity for White, Black, Asian, American Indian and Alaskan Native, and Hispanic communities.

Built Environment Measures. We compiled built environment measures based on the literature and our previous research.19,46,47 Specifically, we selected gas stations and fast-food restaurant locations as a proxy for access to public restrooms, locations where overdoses often occur.17,18,42,48,49 We used pharmacy addresses to analyze the spatial distribution of access to sources of over-the-counter naloxone.21 We compiled highway exits and major roads from state of Massachusetts GIS agency (MassGIS) as a proxy for access to services. For each ZCTA we calculated the mean distance to the nearest exit, major road, pharmacy, fast food restaurant, and gas station from the centroid of the ZCTA. We also calculated the average gas station density for each ZCTA.

Methods

Figure 1 provides an overview of the data and methods used in this paper.

Death Rate Mapping. Using ArcGIS Pro 3.1 (ESRI, Redlands, CA) we created descriptive and maps of raw and smoothed overdose death rates at the ZCTA level across the state.

Spatial smoothing techniques. Smoothing techniques work by detrending the overdose rate within target polygons (ZCTAs), by using data from neighboring polygons thus allowing for the calculation of a local average overdose measure that is less susceptible to variation due to outlier values. We employed two spatial smoothing techniques to derive more stable estimates for spatial measures (fatal overdose rates). These included: 1) an Empirical Bayes Method; and, 2) an Empirical Bayes Spatial method.50–52
The Empirical Bayes (EB) smoothing for rate calculations relies on calculating a weighted average of the raw rate for each ZCTA and the state average, with weights proportional to the underlying population at risk. Therefore, ZCTAs with a small population at risk will tend to have their rates adjusted considerably, while rates for ZCTAs with larger populations at risk will not change much. The second method was Spatial EB smoothing, which is like EB smoothing, except that the reference rate is computed for a group of neighboring ZCTAs that share a boundary (i.e., queen's contiguity) with each individual ZCTA, rather than taking the same overall reference rate for all ZCTAs. We used GeoDa to calculate EB and Spatial EB smoothed rates.\textsuperscript{51,53} We then used the Getis-Ord Gi* hotspot statistic to identify and map clusters (i.e., ZCTAs with high overdose rates surrounded by ZCTAs with high rates).\textsuperscript{51} Each hotspot identified is statistically significant at the p < 0.05 level. This was calculated in the software GeoDa. For estimating the Getis-Ord clusters, we used the queen's criterion to define our spatial weights matrix.

**Principal Components Analysis (PCA).** To avoid multicollinearity amongst the PIP, access, and socioeconomic variables due to significant cross correlation, we used PCA with a Varimax rotation to extract new variables that summarized the covariance in the PIP, socioeconomic, and access variables. PCA is a dimension reduction technique that is commonly used to reduce the number of variables used in analyses while preserving the information from them. We used a cut point of 0.25 for the included factor loadings as recommended in other studies.\textsuperscript{54} The psych and dplyr packages in R were used for PCA and data management [https://personality-project.org/r/psych/, 2023].

**Statistical Modeling.** We estimated logistic regressions with an autologistic term to predict if a ZCTA was in a hotspot (0/1) as identified by the Getis-Ord Gi* statistic as the outcome and the components derived from the PCA as the explanatory variables. We assessed potential multicollinearity among explanatory variables through variance inflation factors (VIF) in the regressions and did not find multicollinearity to be an issue in the final models since it was below 2.5.\textsuperscript{55} To address potential misestimation of parameters and error terms in logistic regression results due to spatial dependency, an autologistic term was used to correct for spatial dependence.\textsuperscript{51,56,57} The logistic and autologistic models were estimated in R using the spatialeco package. We used negative binomial, zero-inflation, and Conditional Auto Regression (CAR) Bayes Poisson models to predict the number of times the ZCTA was located in a hotspot as identified by the smoothing methods in the 11 years between 2011–2021. The negative binomial and zero inflation models are suitable when the outcome is a count. CAR models corrected for spatial dependency of the outcome in a Bayesian framework. The PCA components were the explanatory variables for these models. The packages CARBayes, MASS, and pscl were used to estimate these models in R.

**Results**

Using rate smoothing techniques resulted in lower variances in the rate of overdose deaths because the denominator was expanded to compensate for lower population in rural and suburban ZCTAs (Fig. 2). This was especially useful in identifying high incidence ZCTAs surrounding Springfield, Fall River, and New Bedford, which show high death rates in 2021 based on the smoothed data that are not evident in maps of the raw rates for the same year.
Comparing the results from Getis-Ord analyses with the calculated rates, the smoothed rates allowed us to identify statistically significant hotspot clusters of overdose rates. Figure 3 shows that the impact varied over time. For example, in 2011 the smoothed rates identified clusters in the Boston area, in addition to the area south of Worcester, Fall River, New Bedford, and Cape Cod (Fig. 3). Likewise, in 2014, using the raw rate alone resulted in fewer and more scattered clusters (38) than the spatial EB methods, which identified 74 ZCTAs as hotspots. The EB and spatial EB methods also depicted more geographic variability, facilitating identification of clusters near Plymouth in the southeast and Haverhill in the northeast. In 2017, the numbers of clusters identified based on the raw rates was the lowest (32) when compared to the ZCTA clusters identified based on rates that relied on smoothing techniques (44). The hotspot results suggested that overdose deaths persisted in Worcester, Fall River, and Boston. By 2020, the number of identified hotspots using raw rates was still 32 but, using smoothing techniques, the number of identified ZCTA in hotspot clusters was 75 in 2020 and 80 in 2021. Figure 4 highlights the locations that were persistent hotspots, with the most notable ongoing clusters in Worcester, New Bedford, Fall River, and Wareham. A corridor along Interstates 95 and 495, from Everett, north of Boston, to Fall River, in southeastern Massachusetts, is notable as an area of persistent hotspots, as is the area of southwestern Massachusetts, close to Springfield, and the Greater Worcester Area.

We compared sociodemographic and built environment access variables in the statistically significant ZCTA hotspot clusters, which illuminated important differences in sociodemographic factors. In ZCTAs that were always identified as hotspots, the mean percentages of the population living in poverty, and residents who were non-Hispanic Black or Hispanic, were higher than in ZCTAs that never appeared on fatal overdose hotspots (Table 1). In terms of built environment access variables, decedents in hotspots were closer to major roads and highway exits. Both cold spots and hotspots had similar gas station density and proximity to fast food restaurants and pharmacies.
Table 1
Mean (SD) of demographic and built environment variables by number of years in which a ZIP Code tabulation area (ZCTA) was within a smoothed rate Getis-Ord hotspot*

<table>
<thead>
<tr>
<th>Number of years</th>
<th>ZCTA Count</th>
<th>Poverty percentage</th>
<th>White Percentage</th>
<th>Black Percentage</th>
<th>Hispanic Percentage</th>
<th>Male Percentage</th>
<th>Median Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>259</td>
<td>8.1 (9.3)</td>
<td>87.8 (12.8)</td>
<td>3.5 (8.2)</td>
<td>4.1 (4.3)</td>
<td>48.9 (6)</td>
<td>43.1 (8.6)</td>
</tr>
<tr>
<td>1</td>
<td>107</td>
<td>8.5 (6.5)</td>
<td>91.8 (9.8)</td>
<td>2.3 (3.6)</td>
<td>4.6 (8.2)</td>
<td>49 (9.5)</td>
<td>44 (10.5)</td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>10.2 (7.7)</td>
<td>86.2 (17.9)</td>
<td>4.5 (10)</td>
<td>7.1 (12.6)</td>
<td>47.7 (6.9)</td>
<td>43.1 (8.4)</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>14.1 (10.5)</td>
<td>80.8 (22.3)</td>
<td>7.8 (13.8)</td>
<td>9 (11.1)</td>
<td>50.2 (4.4)</td>
<td>40.5 (7.4)</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>15.4 (13.8)</td>
<td>80.3 (18.6)</td>
<td>6.6 (8.6)</td>
<td>16.7 (22.8)</td>
<td>48.5 (3.1)</td>
<td>43.1 (10.2)</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>16.1 (15.2)</td>
<td>73.7 (22.7)</td>
<td>10 (11.7)</td>
<td>25.5 (30.8)</td>
<td>49.9 (4)</td>
<td>38.8 (9)</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>13.8 (12.1)</td>
<td>78.1 (20.4)</td>
<td>7.4 (8.4)</td>
<td>11.4 (11.9)</td>
<td>48.1 (2.3)</td>
<td>38.9 (5.9)</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>16.8 (6.9)</td>
<td>77.4 (16.7)</td>
<td>6.8 (5.2)</td>
<td>13.5 (13)</td>
<td>48.2 (1.6)</td>
<td>38.5 (4)</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>20 (10.6)</td>
<td>73 (24.6)</td>
<td>8.2 (9.1)</td>
<td>15.7 (17.1)</td>
<td>47.6 (2.4)</td>
<td>41 (5.2)</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>23.8 (16.5)</td>
<td>75.3 (14.5)</td>
<td>10.9 (7.8)</td>
<td>22.3 (19.9)</td>
<td>48.6 (3.6)</td>
<td>33.7 (8.5)</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>12.6 (2.5)</td>
<td>85.3 (15.3)</td>
<td>0.6 (0.8)</td>
<td>1.1 (1.6)</td>
<td>45.6 (5.7)</td>
<td>50 (3.3)</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>26.8 (0)</td>
<td>83.6 (0)</td>
<td>3.8 (0)</td>
<td>11.3 (0)</td>
<td>47.1 (0)</td>
<td>37.2 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distance**</th>
<th>Gas station</th>
<th>Distance fast food restaurant</th>
<th>Distance Pharmacy</th>
<th>Distance Highway exit</th>
<th>Gas station density***</th>
<th>Distance major road</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>259</td>
<td>425.5 (1264.6)</td>
<td>1709.2 (5133.3)</td>
<td>1317.6 (2562.3)</td>
<td>7425.2 (13933.3)</td>
<td>3846.1 (6296.1)</td>
</tr>
<tr>
<td>1</td>
<td>107</td>
<td>526.1 (1544.2)</td>
<td>1691.1 (2965.3)</td>
<td>1817.4 (2931.7)</td>
<td>12137.1 (15700.4)</td>
<td>2804.2 (5978.1)</td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>495.5 (1530.9)</td>
<td>1434 (3395.6)</td>
<td>1299.1 (2686.7)</td>
<td>8266.3 (12663.2)</td>
<td>5784.4 (9050.3)</td>
</tr>
</tbody>
</table>

*Empirical Bayes (EB) rate smoothed and spatial EB smoothed rates

**All distances in meters.

***Density per square mile at the ZCTA level.

Abbreviations: SD = standard deviation, ZCTA = ZIP Code Tabulation Area
Because many of the prescription, socioeconomic and access variables were highly correlated with each other, we derived composite variables using PCA. The results of the PCA suggested that the four components we used in this analysis explained about 62% of the variance (Table 2). The loadings on the first PCA component suggested that it was a measure of potentially inappropriate prescribing because it had a positive loading (indicating positive correlation) with all the PIP measures. This component had high scores in the central parts of the state along Interstate 90 and along interstate 95 in the south (Fig. 5). The second component was correlated with poverty, non-Hispanic Black, American Indian and Alaska Native (AIAN) and Hispanic population percentages and we refer to it as the “poverty and minority” component. Locations that had high values for the poverty and minority component were found across Massachusetts in locations like Worcester, Springfield, New Bedford, Fall River, and in the southern neighborhoods of Boston within the Dorchester neighborhood (Fig. 5).
Table 2
Principal Components Analysis using a Varimax rotation: Loadings of the opioid prescription, socioeconomic, and infrastructure variables for the first four components

<table>
<thead>
<tr>
<th></th>
<th>RC1: Potentially inappropriate prescribing (PIP)</th>
<th>RC2: Minority and Poverty</th>
<th>RC3: Poor Infrastructure</th>
<th>RC4: Rurality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent living in Poverty</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Percentage</td>
<td>-0.80</td>
<td></td>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>Black Percentage</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIAN Percentage</td>
<td>0.82</td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Hispanic Percentage</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian Percentage</td>
<td></td>
<td></td>
<td></td>
<td>-0.85</td>
</tr>
<tr>
<td>Male Percentage</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Age</td>
<td>-0.38</td>
<td>-0.48</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Mean Distance to Gas Station</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Distance to Fast food restaurant</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Distance to Pharmacy</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Distance to Highway Exit</td>
<td></td>
<td></td>
<td></td>
<td>0.40</td>
</tr>
<tr>
<td>Mean Distance to Major Road</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Gas Station Density</td>
<td>0.49</td>
<td></td>
<td></td>
<td>-0.60</td>
</tr>
<tr>
<td>Opioid prescriptions without a pain diagnosis (Rate)</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;=3 cash purchases of opioid prescriptions (Rate)</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poly-pharmacy opioid prescription fills (Rate)</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple prescriber (Rate)</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-prescribing of benzodiazepines and opioids (Rate)</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion Variance explained</td>
<td>24</td>
<td>16</td>
<td>11</td>
<td>10</td>
</tr>
</tbody>
</table>

Abbreviations: RC = Rotated Component; AI/AN = American Indian/Alaskan Native
### Table

<table>
<thead>
<tr>
<th>Component</th>
<th>RC1: Potentially inappropriate prescribing (PIP)</th>
<th>RC2: Minority and Poverty</th>
<th>RC3: Poor Infrastructure</th>
<th>RC4: Rurality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative variance explained</td>
<td>24</td>
<td>40</td>
<td>52</td>
<td>62</td>
</tr>
</tbody>
</table>

Abbreviations: RC = Rotated Component; AI/AN = American Indian/Alaskan Native

The third component, which we termed “poor infrastructure,” was related to being relatively far from fast food restaurants, pharmacies, and gas stations. The ZCTAs with high values for the poor infrastructure component were located not only in high poverty locations such as urban neighborhoods in Boston, Worcester, New Bedford, and Springfield, but also in the more rural areas of Western Massachusetts, and on Cape Cod and the Islands (Fig. 5).

The fourth component, which we referred to as “rurality” loaded positively on White resident percentage, median age, and distance to a highway while it loaded negatively on percentage Asian, and mean gas station density. As can be seen in Fig. 5, high scores for the rurality component are more likely in western Massachusetts and Cape Cod and low scores are in urban areas within and near urban neighborhoods in Boston, Springfield, New Bedford, and Worcester.

We used these four derived components as explanatory variables to predict whether a ZCTA was a hotspot or not (derived from the Getis-Ord Gi* statistic on the smoothed death rate) for each year from 2011 to 2021 using logistic, autologistic and CAR Bayesian regressions (Table 3). We include only significant coefficients from the autologistic and CAR Bayesian regressions in Table 3 for brevity. The results of the autologistic as well as CAR Bayes logistic model, both of which corrected for significant spatial dependencies, suggested that for most years between 2011 and 2021, fatal opioid-involved overdose hotspot ZCTA were more likely to have a high minority and poverty component. The odds ratio, which indicates the factor by which the probability of being part of a hotspot is likely to increase based on a unit change in each predictor, controlling for additional model correlates of the posterior quantities reported by the CAR Bayes model in the credible interval were lowest in 2012 between [2.13–4.47], rising to between [4.1–51.9] in 2021. In the autologistic model, the minority and poverty component was significant from 2012-14 and then more recently 2018–2021 with the 95% confidence intervals of [1.2–1.9] in 2012 and [1.3–2.10] in 2021. The PIP component was significant only for two years (2014 and 2018) in the CAR Bayesian model suggesting higher odds in 2014 but an opposite effect in 2018. The poor infrastructure and rural component, when significant, had a protective effect in that ZCTAs with high scores on these variables were less likely to be in hotspots just like the rurality component. The odds ratio on the poor infrastructure component varied from a high of [0.29–0.89] in 2021 to lower effects in 2011 of [0-0.06] in 2018. The 95% confidence intervals for the rurality component odds ratio varied from a lower range more recently in 2019 [0.05–0.67] to a wider range of odds [0.07–0.99] in 2014.
Table 3
Significant coefficients with $p = 0.05$ (autologistic model) and within the 95% credible interval (for Bayesian model) associated with a ZCTA appearing within a Getis-Ord hotspot cluster in Massachusetts, 2011–2021.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of ZCTA hotspots</th>
<th>Autologistic Model</th>
<th>CAR Leroux Bayes Logistic Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Significant Coefficient Odds Ratio ($p = 0.05$)</td>
<td>Auto- covariate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RC1</td>
<td>RC2</td>
</tr>
<tr>
<td>2011</td>
<td>72</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2012</td>
<td>82</td>
<td>-</td>
<td>1.51</td>
</tr>
<tr>
<td>2013</td>
<td>70</td>
<td>-</td>
<td>1.35</td>
</tr>
<tr>
<td>2014</td>
<td>74</td>
<td>-</td>
<td>1.45</td>
</tr>
<tr>
<td>2015</td>
<td>67</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2016</td>
<td>39</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2018</td>
<td>82</td>
<td>-</td>
<td>1.60</td>
</tr>
<tr>
<td>2019</td>
<td>74</td>
<td>-</td>
<td>1.65</td>
</tr>
<tr>
<td>2020</td>
<td>75</td>
<td>-</td>
<td>1.52</td>
</tr>
<tr>
<td>2021</td>
<td>80</td>
<td>-</td>
<td>1.66</td>
</tr>
<tr>
<td>N</td>
<td>535</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

RC1: Potentially inappropriate prescribing; RC2: Minority Poverty; RC3: Poor Infrastructure; RC4: Rurality

CAR = Conditional Autoregression

The negative binomial, zero-inflated and CAR Bayesian models predicting the number of times a ZCTA was a hotspot had a positive and significant coefficient for the poverty and minority component suggesting that ZCTAs with high scores on this component such as ZCTAs in Worcester, Springfield, New Bedford, Brockton, Boston, and Fall River were more likely to have been a persistent hotspot over the 11-year period. The odds ratio confidence intervals are shown in Table 4 for the negative binomial and zero-inflated model. They vary from [1.1–1.5] in 95% credible intervals for the CAR Bayesian model. The negative coefficient for the poor infrastructure component was not significant once spatial dependency was accounted for in the CAR Bayes model (Table 4).
Table 4
Negative Binomial, Zero-inflated Poisson regression and CAR Bayesian Poisson regression coefficients predicting the number of times the ZCTA was within a Getis-Ord hotspot cluster, Massachusetts, 2011–2021.

<table>
<thead>
<tr>
<th></th>
<th>Negative Binomial</th>
<th>Zero Inflation</th>
<th>CAR Leroux model (Poisson)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff</td>
<td>Std Error</td>
<td>Conf Interval (Odds)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.25</td>
<td>0.06</td>
<td>1.1, 1.46</td>
</tr>
<tr>
<td>RC1</td>
<td>0.10</td>
<td>0.05</td>
<td>0.00, 1.22</td>
</tr>
<tr>
<td>RC2</td>
<td>0.37*</td>
<td>0.06</td>
<td>1.29, 1.61</td>
</tr>
<tr>
<td>RC3</td>
<td>-0.15*</td>
<td>0.07</td>
<td>0.75, 0.99</td>
</tr>
<tr>
<td>RC4</td>
<td>0.13</td>
<td>0.06</td>
<td>0.99, 1.29</td>
</tr>
<tr>
<td>AIC</td>
<td>1724.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>1699.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Coefficients that are significant at p value of 0.05 are boldfaced as are the Bayesian coefficients that are within the 95% credible interval

Discussion

We identified overdose clusters using rate smoothing techniques and found persistent and significant hotspots in Massachusetts. We estimated PCA components to account for multicollinearity while incorporating opioid prescription, socio-environmental and built environment variables; and we corrected for spatial autocorrelation in our models in the presence of spatial dependency in both the location of hotspots and their persistence over the study period. Using PCA, we identified four unique components: PIP regression coefficient 1 (RC1), Minority and Poverty (RC2), Poor infrastructure (RC3), and Rurality (RC4) in predicting both presence and persistence of hotspots. In the autologistic models predicting if a ZCTA was a hotspot, we found our poverty and minority presence factor (RC1), was a significant predictor of a hotspot in most years. In predicting the count of the number of times a ZCTA was a hotspot, we found that the poverty and minority presence factor (RC1) was again a significant predictor. Many studies use socioeconomic, opioid prescription and access variables, but by using our approach of combining them and accounting for spatial autocorrelation, which is inevitable in spatial analytic studies, we better highlight the role that location plays in persistent hotspots of opioid decedents.

Zero inflation of disease rates is common in spatial epidemiological studies, which has resulted in several techniques to address this issue.\textsuperscript{30–38,58} We chose EB, and EB spatial smoothing methods, which reduce variance resulting in better estimates of hotspots in ZCTAs that have relatively low population counts.
These rate smoothing techniques offer considerable promise in identifying regional clusters of persistent risk for fatal overdoses in Massachusetts. Our application of the EB methods supports the use of this technique as a method to detect patterns at smaller spatial scales in which zero-inflation is an issue.

The results of the autologistic, and the CAR Bayes models demonstrate the importance of correcting for spatial autocorrelation in regressions. For example, in the model predicting the number of times a ZCTA was a hotspot the zero inflation and negative binomial models suggested that poor infrastructure (RC3) had a protective effect and rurality had a significant risk but the CAR model suggests that poverty and minority was the most significant predictor. Our results may also indicate that other research, which has identified rurality as an important risk factor for fatal overdose, are not as important as poverty and minority presence in a ZCTA in predicting hotspots in Massachusetts especially in more recent years. Recent studies have also noted the role of poverty and race in predicting opioid incidence in other states.

Our hotspot cluster analysis found that hotspots were comprised of higher percentages of Black, AIAN and Hispanic populations, as well as higher percentages of people living in poverty. In 2011-2015, the overdose epidemic in Massachusetts and elsewhere was characterized in the media as affecting people who were middle income and non-Hispanic White as part of the “deaths of despair” narrative. The role of poverty in opioid-involved overdoses in minority people who use drugs (PWUD) should be surveilled carefully as opioid overdoses continue to rise in these groups.

Our findings about the persistence of hotspots in New Bedford, Fall River and along the I-95 corridor as well as in Worcester and the protective effect suggested by poor infrastructure in the models, suggests that proximity to highways may have played a role in supplying opioids. This finding suggests that other measures of built environment and access to opioid supply may be important to incorporate into future analyses as reported in other studies.

Our findings have several limitations. For privacy reasons, the prescribing data and the census data in this study were only available at the ZCTA scale and, therefore, our results are subject to ecological fallacy. Models that are at the individual level may show different relationships than our findings, which were aggregated. Also, the associations that we identified may only be applicable to Massachusetts. Several laws and regulations were enacted in the state during the opioid overdose epidemic that may not be generalizable to other states. In 2019, for instance, the Governor of Massachusetts enacted a law that allowed no more than a seven-day supply of prescribed opioids (with certain exceptions), which, along with other policies, resulted in a notable decline in the number of opioid prescriptions dispensed. The PIP data were obtained from the MA PMP, and they are also subject to ecological fallacy due to aggregation at the ZCTA level.

**Conclusion**

We employed a range of spatial analytical and statistical modeling approaches to better understand the opioid overdose crisis in Massachusetts and improve targeting of public health interventions. Smoothing
methods allowed us to derive more stable estimates for spatial measures and enhanced our understanding of spatial clustering patterns related to fatal opioid-involved overdoses. We recommend using EB and EB smoothing methods for zero-inflated rates especially for spatial analysis and spatial model estimation. PCA facilitated a better understanding of unique community-level spatial and socioeconomic variables that contribute to overdose mortality patterns. By using the components that combined several socioeconomic, built environment and prescription variables in the regression models which also accounted for spatial autocorrelation, we were able to characterize the biggest contributing factors to overdose-related deaths in Massachusetts. Future research should investigate access, poverty and minority variables and their potential proxies. Many factors impact the cascade of opioid use, opioid use disorder, and opioid-involved death and by identifying factors at the beginning of the cascade, as we have done here, we can inform policies to intervene early in the cascade and prevent opioid-involved deaths.

Declarations

*Ethics Approval and Consent to Participate*

Not applicable.

*Consent for Publication*

Not applicable.

*Availability of Data and Materials*

The datasets used and/or analyzed during the current study are available from the corresponding author on request.

*Competing Interests*

The authors declare that they have no competing interests.

*Funding*

None.

*Author’s Contributions*

SS and TJS designed the study. SS and TJS reviewed the data. SS conducted the statistical analysis. JP and EM collated the literature. All authors interpreted the analyses and wrote the manuscript.

*Acknowledgements*

Not applicable.
References


**Figures**
Figure 1

Flow diagram visualizing data and methods described in this paper. Note: PIP is potentially inappropriate opioid prescribing. ACS represents data from the US Census Bureau's American Community Survey. Data Axle was used to for built environmental variables such as gas stations and pharmacies; MassGIS is the Massachusetts State GIS Data provider; Getis Ord statistics generate hotspots and cold spot statistics.
Figure 2

Fatal opioid overdose rate quintiles in 2021 in Massachusetts by ZIP Code Tabulation Area (ZCTA): 2a) Raw rate; 2b) Empirical Bayes (EB) smoothed rates; 2c) Spatial EB smoothed rates 2d) Reference map. This comparison of maps displaying raw and smoothed overdoseroates highlights that identification of local hotspots is easier in the EB smoothed rate maps than in the raw rate map’s patchwork of high and low values.
Figure 3

Fatal opioid overdose Getis-Ord hotspot and coldspot clusters (p<0.05) for 2011, 2014, 2017 and 2021: 3a) Raw, 3b) Empirical Bayes (EB) smoothed, 3c) spatial EB smoothed death rates by ZIP Code Tabulation Area (ZCTA), Massachusetts.
Figure 4

Number of times (i.e., years) the ZCTA was a Getis-Ord hotspot cluster between 2011-2021 using Empirical Bayes (EB) spatial smoothing.
Figure 5

Principal Components Analysis scores mapped in quintiles for Massachusetts by Zip Code Tabulation Area (ZCTA): 4a) RC1: potentially inappropriate prescribing (PIP) with highway lines; 4b) RC2: Poverty and minority; 4c) RC3: Poor infrastructure 4d) RC4: Rurality.