



Published in final edited form as:

Int J Drug Policy. 2021 October ; 96: 103395. doi:10.1016/j.drugpo.2021.103395.

Machine Learning Takes a Village: Assessing neighbourhood-level vulnerability for an overdose and infectious disease outbreak

Jesse L. Yedinak, MPA¹, Yu Li, MD, PhD, MPH¹, Maxwell S. Krieger, BS¹, Katharine Howe, MPH⁴, Colleen Daley Ndoye⁵, Hyunjoon Lee, MS⁶, Anna M. Civitarese, MPH⁴, Theodore Marak, MPH⁴, Elana Nelson, BS¹, Elizabeth A. Samuels, MD, MPH, MHS^{7,8}, Philip A. Chan, MD, MS^{2,3,4}, Thomas Bertrand, MA, MPH⁴, Brandon D.L. Marshall, PhD¹

¹Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA

²Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, USA

³Department of Medicine, Alpert Medical School of Brown University, Providence, RI, USA

⁴Center for HIV, Hepatitis, STD, and TB Epidemiology, Rhode Island Department of Health, Providence, RI, USA

⁵Project Weber/Renew: Harm Reduction & Recovery Services Provider, Providence, RI, USA

⁶Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

⁷Department of Emergency Medicine, Alpert Medical School of Brown University, Providence, RI, USA

⁸Overdose Prevention Program, Rhode Island Department of Health, Providence, RI, USA

Abstract

***Correspondence Author:** Brandon DL Marshall, PhD, Associate Professor, Department of Epidemiology, Brown University School of Public Health, 121 South Main Street, Box G-S-121-2, Providence, RI, 02912 USA, **Tel:** 401-863-6427, brandon_marshall@brown.edu.

CREDIT AUTHORSHIP STATEMENTS:

JLY was responsible for Conceptualization, Funding acquisition, Supervision, Writing - Original Draft. YL was responsible for Conceptualization, Writing - Original Draft, Methodology, Formal Analysis. MSK was responsible for Investigation, Writing - Original Draft, Visualization, Data Curation. KH was responsible for Project Administration, Funding acquisition, Data Curation. CDN was responsible for Validation, Writing - Review & Editing. HL was responsible for Writing - Review & Editing. AMC was responsible for Data Curation, Writing - Review & Editing. TM was responsible for Data Curation, Writing - Review & Editing. EN was responsible for Investigation, Resources, Writing - Review & Editing. EAS was responsible for Writing - Review & Editing. PAC was responsible for Data Curation, Writing - Review & Editing. TB was responsible for conceptualization, Funding acquisition, Writing - Review & Editing. BDLM was responsible for Conceptualization, Methodology, Writing - Review & Editing.

Declaration of Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Approval

No.

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

BACKGROUND: Multiple areas in the United States of America (USA) are experiencing high rates of overdose and outbreaks of bloodborne infections, including HIV and hepatitis C virus (HCV), due to non-sterile injection drug use. We aimed to identify neighbourhoods at increased vulnerability for overdose and infectious disease outbreaks in Rhode Island, USA. The primary aim was to pilot machine learning methods to identify which neighbourhood-level factors were important for creating “vulnerability assessment scores” across the state. The secondary aim was to engage stakeholders to pilot an interactive mapping tool and visualize the results.

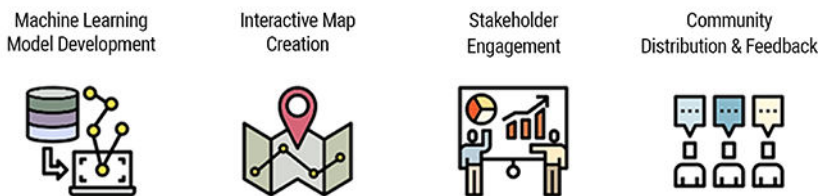
METHODS: From September 2018 to November 2019, we conducted a neighbourhood-level vulnerability assessment and stakeholder engagement process named The VILLAGE Project (Vulnerability Investigation of underlying Local risk And Geographic Events). We developed a predictive analytics model using machine learning methods (LASSO, Elastic Net, and RIDGE) to identify areas with increased vulnerability to an outbreak of overdose, HIV and HCV, using census tract-level counts of overdose deaths as a proxy for injection drug use patterns and related health outcomes. Stakeholders reviewed mapping tools for face validity and community distribution.

RESULTS: Machine learning prediction models were suitable for estimating relative neighbourhood-level vulnerability to an outbreak. Variables of importance in the model included housing cost burden, prior overdose deaths, housing density, and education level. Eighty-nine census tracts (37%) with no prior overdose fatalities were identified as being vulnerable to such an outbreak, and nine of those were identified as having a vulnerability assessment score in the top 25%. Results were disseminated as a vulnerability stratification map and an online interactive mapping tool.

CONCLUSION: Machine learning methods are well suited to predict neighborhoods at higher vulnerability to an outbreak. These methods show promise as a tool to assess structural vulnerabilities and work to prevent outbreaks at the local level.

Graphical Abstract

VILLAGE (Vulnerability Investigation of underLying Local risk And Geographic Events):
Preventing overdose and infectious disease outbreaks at the local level



Keywords

Overdose; HIV; hepatitis C; structural vulnerability; neighbourhood; predictive analytics; machine learning

INTRODUCTION

The United States of America (USA) has faced a growing opioid and overdose crisis for nearly a decade (Rudd et al., 2016). Despite the magnitude of this crisis, many regions of the US are not fully aware of the community-level factors that may make their

region susceptible to an outbreak of drug overdose and infectious disease transmission associated with non-sterile injection drug use (IDU) (Van Handel et al., 2016). Outbreaks of bloodborne infections, particularly HIV and hepatitis C virus (HCV), are increasingly more common among people who use drugs (PWUD) (Crowley & Millett, 2017; Zibbell et al., 2015). Factors such as an adulterated or changing drug supply, restricted access to sterile equipment for injection, and high frequency of injecting behaviors have been found to impact regional susceptibility to a rapid outbreak (Des Jarlais et al., 2016). The sudden outbreak of HIV and HCV in the rural community of Austin in Scott County, Indiana in 2015 served as a wake-up call to the emerging complexities of the opioid crisis (Mercado-Crespo et al., 2014; Rudd et al., 2016; Scholl, 2019).

One of the key interventions for preventing HIV and HCV infections are syringe service programs (SSPs) (Crowley & Millett, 2017); however, due to laws limiting SSP availability (Davis et al., 2019), counties in Indiana were only allowed to implement SSPs after an official public health emergency had been declared (Golding, 2017). Additionally, case investigations reported large networks of people sharing injecting equipment and more frequent injections due to injecting prescription opioids, two factors that contributed to county-level risk (Des Jarlais et al., 2016). Given that the only mechanism to detect an outbreak was through surveillance of HIV testing data, there was a delay in responding to the outbreak (Conrad et al., 2015). It has been estimated that a legal SSP program established prior to the Indiana outbreak could have reduced the scope of the HIV outbreak by up to 90% (Goedel et al., 2020).

In response to this staggering HIV outbreak among PWUD, the Centers for Disease Control and Prevention (CDC) conducted a nationwide vulnerability assessment of US counties at risk of an outbreak of HIV and HCV. The CDC used epidemiological models to forecast which counties may be at risk based on shared characteristics with Scott County, Indiana (Van Handel et al., 2016). Epidemiological models are policy tools used to predict outbreaks of infectious diseases and forecast the course of the outbreak and potential routes for intervention (Metcalf et al., 2015). The CDC's 2016 vulnerability assessment by Van Handel et al. sought to identify potential predictors of a region's susceptibility to an outbreak of HIV and HCV by developing a vulnerability index ranking of US counties. In doing so, states could be better informed and prepared to intervene. As a result of their model, 41 of the highest ranked 220 counties in the US were identified as being in the state of Tennessee (Van Handel et al., 2016), prompting the development of a subsequent statewide epidemiological model. Rickles et al. iterated on the 2016 model by focusing exclusively on counties in the state of Tennessee and, by including additional statewide surveillance data, generating state-specific data to help guide policy making decisions at the county level (Rickles et al., 2018).

The methods used by Van Handel et al. in 2016 also found that few counties in the Northeastern and Northwestern US were vulnerable to an outbreak when using acute HCV infections as a proxy (see Table 1. Glossary) for injection drug use patterns. However, subsequent outbreaks of HIV among PWUD have occurred in several Northeastern cities within Massachusetts, including Boston, Lawrence, and Lowell (Cranston et al., 2019; Freyer, 2019; Schumaker, 2018). The lack of timely data representing factors related to

changes in regional drug patterns may have contributed to this gap in results, as the emergence of the illicit opioid fentanyl is thought to be a major factor in the Massachusetts HIV outbreak, in part due to more frequent daily injections (Cranston et al., 2019). Shared and non sterile injecting equipment can be a large driver of bloodborne disease transmission (Dasgupta et al., 2019; Zibbell et al., 2018). Furthermore, the region was in the midst of an increasing overdose crisis due to the fentanyl-adulterated drug supply and recent rise in HCV infections (Alpren et al., 2020).

The HIV outbreak in Massachusetts was particularly concerning for the neighboring state of Rhode Island, where accidental overdose deaths are a public health crisis (Carroll et al., 2017; Marshall et al., 2017; Mercado-Crespo et al., 2014). Fentanyl overdose deaths in Rhode Island increased by almost 90% from 2011 to 2017, and most occurred as a result of injection drug use (Marshall et al., 2017; Mercado et al., 2018). Finally, Rhode Island and seven other state jurisdictions were excluded from the Van Handel et. al, 2016 analyses because no data on HCV infection existed, leaving room for further examination of available data sources and methods, particularly in the context of an overdose crisis.

In recent years, public health officials have turned to machine learning methods (see Table 1. Glossary) to develop enhanced forecasting models for multiple public health issues, including overdose mortality, firearm violence, and HIV acquisition (Adams et al., 2018; Goin et al., 2018; Reeping et al., 2019; Trinquart et al., 2019). Machine learning methods use computational methods and algorithms to process and analyze large amounts of data, using more complex approaches than traditional epidemiological surveillance methods (Bi et al., 2019). Furthermore, these machine learning algorithms have been shown to create robust predictions that are critical to understanding individual risk factors as well as complex structural factors and pathways for intervention (Lo-Ciganic et al., 2019).

Following this trend, and with funding and technical assistance from the CDC, a team at Brown University School of Public Health (Brown) worked in partnership with the Rhode Island Department of Health (RIDOH) to adapt and build on the methods of Van Handel et al. The primary aim was to pilot machine learning methods to create a statewide predictive model that identified significant factors for “vulnerability assessment scores” (see Table 1. Glossary) across the state (see Figure 1). Vulnerability assessment scores are measures to capture the overall combination of factors that may affect a neighbourhood’s relative risk for an outbreak of overdose, HIV and HCV (see Table 1. Glossary). The secondary aim was to engage with community-based and state agency stakeholders to offer context and validity to the vulnerability assessment scores and pilot an interactive mapping tool that would visualize the results by neighbourhood or town (see Figure 1). Rhode Island was an ideal state to construct the predictive model given its highly developed drug overdose data surveillance and collection systems and policy environment, allowing for robust collaboration between state and academic institutions (Marshall, 2018; Marshall et al., 2017; Waye et al., 2018). We named this work The VILLAGE Project (Vulnerability Investigation of underlying Local risk And Geographic Events).

METHODS

Design

From September 2018 to November 2019, a research team at Brown University conducted a neighbourhood-level vulnerability assessment of Rhode Island using supervised machine learning methods (see Table 1. Glossary). During that same timeframe, in collaboration with RIDOH, Brown engaged a group of local stakeholders once per quarter to review the machine learning process, vulnerability assessment scores, and mapping tools. This project was a contracted public health evaluation (see Table 1. Glossary), defined federally as a systematic investigation conducted on behalf of a state agency and designed for local implementation (Office for Human Research Protections (OHRP), 2016), and did not fall under the definition of human subjects research. Therefore, it did not qualify for the Institutional Review Board (IRB) human subjects review. All data sharing described in the following sections was conducted under a fully executed Data Use Agreement between RIDOH and Brown.

AIM 1: Supervised Machine Learning Model

Data Sources—The vulnerability assessment conducted by the CDC in 2016 (Van Handel et al., 2016) used national datasets with geographical detail at the county-level, including indicators from the CDC's Social Vulnerability Index (SVI) (see Table 1. Glossary). To build on prior work, we focused on the need for even greater geographic granularity, as county-level data are not useful for resource prioritization in Rhode Island, a densely populated state with just five counties in total. Furthermore, even within cities, overdose mortality rates can vary substantially across neighbourhoods (Visconti et al., 2015). Therefore, in addition to SVI indicators, we identified additional statewide surveillance data that were available at the census tract level—smaller units within each county representing approximately 4,000 people per tract and a total of 240 tracts in Rhode Island, roughly the size of a neighbourhood or small town (US Census Bureau, 2019). We grouped these statewide indicators together as “all variables except SVI” for comparison. SVI and all other statewide indicators were geocoded and aggregated to the census-tract level. To include data at the ZIP code level, we used a percentage calculation based on the proportion of census tracts within the ZIP codes (Henry & Boscoe, 2008). For example, if the number of treatment admissions in a specific ZIP code was 500, and the ZIP code straddled two census tracts equally, then the estimated number of treatment admissions was also partitioned equally (i.e., 250 in each census tract). To create a final census tract-level dataset for analysis, we joined state-specific variables to publicly available geographic datasets, such as the US Census and the American Community Survey. All variables representing medians with missing values were dropped from the analysis. For all other predictor variables, missing data were recoded as zero. Finally, as described in the Outcome Variable section below, we used 2017 overdose deaths in Rhode Island as our model outcome, representing a proxy for injection drug patterns that may lead to an overdose and infectious disease outbreak. We chose to use this proxy because of the completeness of the data available at the census tract level and the need for timely data to capture changes in the drug supply (i.e., fentanyl adulteration). Full data source descriptions and a list of variables are available in the Supplemental File.

Predictor Variables—Machine learning algorithms allowed for an analysis of the large amounts of data for this evaluation to find the most relevant set of predictor variables for neighbourhood-level vulnerability. Several methods, described below, were used to train our model to “predict” a specified output (2017 overdose deaths), and recognize which variables had the strongest prediction power. Through these methods, we were able to analyze the variables that the literature had previously identified as possibly predictive for risk of a future outbreak related to injection drug use (Rickles et al., 2018; Van Handel et al., 2016), dependent on their availability at our desired geographic level (i.e., census tract). We included variable domains from Van Handel et al., 2016 that were found to be associated with their proxy for injection drug use (acute HCV infection 2012–2013), including drug overdose mortality, access to care (evidence of use of care or treatment services related to IDU), and sociodemographic characteristics associated with geographic areas with higher IDU prevalence (Van Handel et al., 2016). We were unable to include domains for access to prescription opioids, survey-based data regarding the prevalence of IDU, and drug-related arrests, as they did not include census tract level detail at the time of our analysis. Such factors might have provided additional predictive power for injection drug use patterns if they had been available.

Next, we included the CDC’s 2016 SVI factors, which are characterized by four themes: 1. socioeconomic elements such as education level and income, 2. household makeup, 3. demographics such as race and ethnicity, 4. housing units and access to transportation (CDC Social Vulnerability Index (SVI), 2020). The additional statewide indicators, categorized as “all variables except SVI,” included data on the counts and locations of historical injection and non-injection overdose deaths between 2014 and 2016, treatment availability data (including the number of buprenorphine and methadone providers), and the location of structures and facilities (such as hospitals, recreation areas, and schools). Other indicators specific to injection drug use patterns included methadone treatment admissions whose primary route of drug use was injection (2014–2016).

Ultimately, to determine the utility of SVI variables in our dataset and overall model, we separated the predictors into three sets, and performed a dimension reduction process on each, explained in detail under Statistical Analyses. These separated datasets included: (1) SVI variables only, (2) all the variables except for SVI variables, and (3) all the variables.

Outcome Variable—Because no nationwide surveillance data exists to represent injecting risk behaviors, prior models have used the *Summary of Notifiable Infectious Diseases and Condition—United States* database for acute HCV infections in 2012–2013 (Rickles et al., 2018; Van Handel et al., 2016) as a proxy for county-level risk of an infectious disease outbreak of HIV and HCV. Statewide surveillance data for HCV is unavailable for many states, including Rhode Island, as federally funded HCV reporting and surveillance programs are limited (CDC Viral Hepatitis, 2021). HIV cases related to injection drug use were less than 5 per year in Rhode Island, consistent with the low HIV infection rates found in Indiana prior to the outbreak (Conrad et al., 2015). Rapid testing data for both HIV and HCV in Rhode Island only occurred at several fixed site locations in the capital city of Providence and did not include individual addresses. However, Rhode Island does have a robust and timely surveillance system for accidental drug overdoses (Jiang et al., 2018).

Therefore, in our model, counts of accidental drug overdose deaths in each census tract in 2017 were used as a proxy to infer neighbourhood-level risk for susceptibility to an outbreak of overdose, HIV and HCV, and defined as the model outcome. As described previously (Marshall et al., 2017), cases were considered confirmed accidental drug-related overdose fatalities if: (i) the death was pronounced in Rhode Island; (ii) the final manner of death was deemed an accident by the medical examiner, and (iii) a drug is listed on the death certificate as the primary cause of death or a significant contributing factor. We geocoded overdose deaths at the census tract level based on the place of injury, which was obtained from police, Emergency Medical Services (EMS) ambulance run reports, or other files contained in the autopsy reports.

Statistical Analyses—To implement the machine learning model, we used the Elastic Net method, a dimension reduction method that chooses predictors by shrinking model coefficients toward zero, through manipulation of the alpha term (Hastie & Qian, 2014). The method allows for alpha values between 0 and 1, and is analogous to performing ridge regression when the alpha value is set to 0, and performing LASSO when the alpha value is set to 1. (Osman et al., 2017).

The primary purpose of using the Elastic Net method was to reduce the dimension, or size, of the data. Specifically, Elastic Net is powerful for variable selection when there are many variables but a small number of data records (Zou & Hastie, 2005). We determined the relevant variables using the glmnet package in R (Hastie & Qian, 2014). The package uses $l(y, \eta)$, the negative log-likelihood, and adds an elastic net penalty in the second term. The lambda variable controls the overall strength of the penalty (Hastie & Qian, 2014). Through this method, we determined coefficients for each variable of interest. Coefficients that approach zero indicate that the variable does not have strong predictive value, and therefore the variable could be discarded, a process called “regularization.” The dataset was split into a training set (70%) and test set (30%) by the rule of balancing the splits of data according to the outcome variable and to preserve the overall class distribution of the data, using the “caret” package found in R software (Gholamy et al., 2018; Hastie & Qian, 2014; Kuhn, 2008).

The offset in our model was defined as the log of the total population of each census tract. Estimation of coefficients of the included predictors was done by 10-fold cross-validation on the training set. The number of possible lambda values in the sequence tested had been set as 100 and ratio between the maximum and minimum lambda value as $10^{(-6)}$, while the largest lambda at which the mean square error (MSE) was within one standard of the minimal MSE was used in the final model in each scenario (Krstajic et al., 2014). Predicted values and MSE were calculated on the testing set. In order to choose the best fitting model, several models were created and compared with different values of alpha in increments of 0.1 (the range being 0–1).

In addition to the standard Poisson model, negative binomial, zero-inflated poisson, and zero-inflated negative binomial models were also attempted but these failed to converge. All statistical analyses were conducted in R software version 3.5.2 (R Core Team, 2014). All of the stated machine learning models with a Poisson distribution were performed using

the “glmnet” package, while both the negative binomial and zero-inflated models were performed using the “mpath” package (Hastie & Qian, 2014; Tibshirani et al., 2012; Wang, 2019).

Finally, our elastic-net regularized Poisson regression estimated the number of overdose deaths for each census tract in Rhode Island. We used the estimated number of 2017 overdose deaths in each census tract as the vulnerability score, a proxy for the risk of future overdoses and drug-related harms including overdose, HIV and HCV infection. Based on the vulnerability assessment scores from our model, a forecasting heatmap was created using ArcGIS software Version 10.4 (Environmental Systems Research Institute (ESRI), Release Date: February 18, 2016). The variables of importance are statistically generated predictors with non-zero coefficients after regularization. We define “importance” as the magnitude of mean change in vulnerability assessment scores given a one-unit increase in the predictor.

AIM 2: Stakeholder Engagement & Mapping Tool

Evaluation Design—Stakeholders were recruited through statewide email listservs and focused outreach with community-based organizations that provided harm reduction and prevention services in collaboration with RIDOH. Ultimately, a total of 13 individuals agreed to participate in the advisory group for one year. Members represented local expertise on community-based street outreach, syringe exchange, naloxone distribution, and HIV and HCV testing and care services.

Data Collection—The stakeholders met four times between November 2018 and November 2019 and were offered US\$50 per meeting. Each meeting was structured to increase understanding of the data sources and the machine learning model process. Meetings included a brief presentation on the datasets used in the predictive model and any preliminary results. Stakeholders offered feedback on the machine learning results, the terminology used relating to drug user health, and contextualization of results (Yedinak et al., 2019). The meetings also included discussions on creating mapping tools for a wider audience of service providers. Stakeholders reviewed the vulnerability score results and helped determine how to present the maps to the statewide community. The decision was made by stakeholders not to show individual census tract ranks or predicted overdose counts on the maps. Stakeholders wanted to avoid negative perceptions of areas with higher vulnerability assessment scores and mitigate the perception of model precision if predicted vs. actual overdose counts did not align.

RESULTS

For Aim 1, we applied the model to the state of Rhode Island which included a population of 1.1 million people across 240 census tracts. The corresponding MSE values for each model were: (1) 1.11 for SVI Variables only, (2) 1.37 for all variables except SVI, and (3) 1.44 for all variables. The final model used option (2), all the variables except for SVI variables, to build the map. We chose this model because it was important for our stakeholders’ understanding of our model, and to build on previous work by incorporating timely and local surveillance data. Our final machine learning model was an elastic-net regularized Poisson

regression ($\alpha = 0.4$) which selected group variables and an alpha value that minimized the MSE.

The model predicted fifteen variables of highest importance based on standardized coefficient magnitudes (see Figure 2.), representing protective and risk factors that contributed to the final statewide vulnerability assessment scores. Variables that lowered the census tract-level vulnerability assessment scores included the percentage of (1) households with income from interest, dividends, or net rent, (2) males aged 25 years or older with a bachelor's degree, and (3) housing units with complete kitchen services, while variables that increased the score included (1) housing cost burden, (2) overdose deaths from 2014–2016, (3) percentage of residents aged 15 years or older who are divorced, and (4) number of injection-related overdose deaths from 2014–2016.

The machine learning prediction models were suitable for estimating relative vulnerability assessment scores at the neighbourhood-level for the state of Rhode Island, using 2017 overdose fatality counts as the proxy outcome (see Supplement File). The final forecasting heat map ranked the census tracts predicted to be at higher risk for experiencing future outbreaks of overdose, HIV and HCV, relative to other census tracts in the state (See Figure 3). The census tract with the highest predicted overdose count (5.4) captured the highest “Risk Level,” while the census tract with the lowest count (0.1) captured the lowest “Risk Level” on a continuous color scale. Eighty-nine census tracts (37%) with no prior overdose fatalities were identified as vulnerable to an outbreak, and nine of those had a vulnerability assessment score in the top 25% (See Figure 4).

For the second aim, the stakeholders finalized both a community report of the vulnerability stratification map and an online interactive version of the mapping tool. They then helped facilitate a statewide workshop to share the VILLAGE Project findings—a half-day event held in November 2019. A total of 65 participants attended the free workshop and breakout sessions. The final results of the work were printed and presented in a plain-language white paper format for community distribution, “The VILLAGE Prevention Plan: Improving prevention efforts and health outcomes for people who use drugs” (Yedinak et al., 2019). The Plan was posted online with interactive Google Maps for use by community organizations on the statewide data dashboard website, www.PreventOverdoseRI.org, and presented at subsequent statewide meetings.

The vulnerability assessment results were transformed into a risk stratification map (Figure 3), for use in a request for proposals for syringe service programs. Applicants were asked to provide services to the neighbourhoods identified as having higher vulnerability assessment scores while preparing their proposals. RIDOH then used the map to ensure optimal coverage for the highest risk areas through their contracting process.

Three neighbourhood-based intervention approaches were recommended to help address risks for injection-related outbreaks of overdose, HIV, and HCV. These were based on neighbourhood-level interventions and were named 1) Universal Precautions, 2) Scale Up, and 3) Take it Further. First, we recommended a “Universal Precautions” campaign to extend access to existing low-barrier interventions to all neighbourhoods—for example,

syringe service programs, naloxone, fentanyl test strip distribution (Krieger et al., 2018) and free hepatitis vaccinations for PWUD (Bowman et al., 2014; Page et al., 2013; Spada et al., 2005). Next, we recommended a “Scale Up” approach for communities with medium or high-risk vulnerability assessment scores to increase resources. Suggestions included a neighbourhood-level needs assessment survey and increased the use of advanced practice practitioners to improve engagement in healthcare (Brooklyn & Sigmon, 2017; Sylla et al., 2007). Finally, we recommended a “Take it Further” campaign to pilot low-barrier practices in the communities with the highest risk scores. Examples included a Medication First model (Winograd et al., 2019) for prescribing medications for opioid use disorder (MOUD) and expanded settings for HCV treatment using nurse practitioners (Olson et al., 2015) through co-located treatment services (Rich et al., 2018). Additional solutions gaining traction across the US include overdose prevention programs, also known as supervised consumption sites (Roth et al., 2019).

DISCUSSION

The results of this study demonstrate that machine learning methods are a suitable technique for identifying neighborhoods at high risk of overdose and infectious disease outbreaks, but that engagement with public health professionals, community-based harm reduction organizations, and other key stakeholders throughout the model development, validation, and dissemination process is critical for uptake. This work also suggests that census tracts may be at risk for an outbreak even with no prior history of overdose fatalities, as predicted by the model. The final variables of importance, however, are consistent with the types of factors identified in previous studies in that they focus heavily on socioeconomic status, household characteristics and composition, and housing types (Linton et al., 2017; Nandi et al., 2006; Perlman & Jordan, 2018). The variables for *total housing cost burden* and *more than 1 housing occupancy per room* were identified as important predictions, indicating that structural vulnerability and housing density variables may help us further understand the neighbourhood-level factors that may lead to an outbreak. The 2016 Van Handel study found a total of six variables that were significant for predicting county-level risk, which primarily reflected the previous knowledge of opioid drug use in rural regions of the US (Suryaprasad et al., 2014), including factors such as prescription opioid sales, unemployment, and white race and non-Hispanic ethnicity (Van Handel et al., 2016). We included unemployment, white race, and non-Hispanic ethnicity in our models but they were not found to be as important as other predictors. As the regional dynamics of the overdose crisis and persistence of fentanyl continue to evolve, local structural characteristics, race and ethnicity, and timely surveillance data offer important considerations for understanding and addressing neighbourhood-level factors that may contribute to injection drug use patterns (Goedel et al., 2019; James & Jordan, 2018; Perlman & Jordan, 2018).

Prior research (in addition to analyses of historical surveillance data in the state of Rhode Island) has shown that neighbourhood-level overdose burden can change rapidly in both space and time (Li et al., 2019). For this reason, resource allocation decisions based on surveillance data alone may not be responsive to emerging regional risks and needs, particularly given the delays and lack of available data for testing and diagnosis of HIV and HCV. Prediction methods—such as those demonstrated here—hold promise for informing

more efficient and equitable deployment of resources to prevent overdose and infectious disease outbreaks before they occur.

In Rhode Island, the resulting vulnerability assessment scores provided localized data to help inform resource allocation as well as funding and programmatic decision-making based on neighbourhood factors. The inclusion of a stakeholder engagement process for the second aim was designed to create more buy-in from the organizations that may use the results of the vulnerability stratification map to inform service delivery. By offering a community-wide workshop, the stakeholder group helped raise awareness of the vulnerability score mapping and increase community acceptability for using machine learning predictions to adapt local service delivery. Stakeholders led multiple educational breakout sessions on harm reduction interventions, stigma and data-driven service delivery. Ultimately, this evaluation highlighted the potential for implementation studies to better understand the role of community organizations in the successful uptake of machine learning-derived products into statewide practice.

LIMITATIONS

While important and impactful, this evaluation had several methodological and practical limitations, and results should be interpreted with caution. First, for the machine learning model in aim 1, no statewide measure of injection drug use exists, and we relied on overdose deaths at the census tract level as a single proxy for injection drug use patterns. The Van Handel et al. 2016 study and the Rickles et al. 2018 study used acute HCV incidence 2012–2013 as a proxy for injection drug use (Rickles et al., 2018); however, data for acute HCV infections in Rhode Island are unavailable, and we were unable to incorporate those measures into the machine learning model. HIV case data with prior injection drug use was available at the census tract; however, the yearly totals for confirmed cases related to injection drug use were less than five per year and had little impact on model performance. HIV case data were excluded from final analyses. Additionally, not all data sources we surveyed included sufficient geographic information to use at the scale of our analysis (i.e., census tract level).

Second, given the accelerating overdose crisis in Rhode Island (Hallowell et al., 2020), more attention should be paid to novel and dynamic datasets that could inform factors related to a regional outbreak, including emerging trends in overdose patterns or changes in the drug supply. For example, several studies have evaluated the feasibility of using internet searches or online forum posts to predict state- or county-level trends on overdose morbidity and mortality (Bowen et al., 2019; Campo et al., 2020; Young et al., 2018).

Third, for aim 2 we did not formally evaluate the stakeholder engagement sessions and workshop to examine whether the format increased stakeholder confidence and understanding of the machine learning model and map tools. Future work should also focus on further testing these methods for reproducibility across regions, while examining the uptake of machine learning methods into statewide practice. Finally, the results are not widely generalizable, as they were designed with local input and data sources in mind.

CONCLUSION

Machine learning methods offer the potential for generating new insights to enhance existing surveillance efforts and offer resources for early intervention and innovation within the neighbourhoods identified as at-risk for an outbreak. The process of applying new, novel methods to emerging and overlapping infectious disease and overdose epidemics offers the opportunity to increase the complexity of the data used by state health departments to drive resources to vulnerable neighbourhoods. Recognizing the implications of vulnerability assessment scores, local policy leaders have the opportunity to deploy innovative and forward-thinking methods for geographically prioritizing harm reduction and overdose prevention interventions, which are particularly impactful when designed in collaboration with key stakeholders and service-delivery organizations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We extend our appreciation and thanks to our dedicated community stakeholders for their critical insights throughout this process, the many community based organizations who support our work, and our state agency collaborators in Rhode Island. We would also like to thank our colleagues at Brown University, Jacqueline Goldman, Abdullah Shihpar, and Sam Bessey for their analytic oversight and community engagement support on this project. We would also like to thank Gayle Fraser and the COBRE on Opioids for supporting our community engagement and dissemination efforts.

FUNDING:

The Vulnerability Assessment and Jurisdictional Plan were funded by the Centers for Disease Control & Prevention (CDC-RFA-TP18-1802). Additional evaluation and dissemination activities were funded by the Center of Biomedical Research Excellence (COBRE) on Opioids and Overdose, National Institute of General Medical Sciences of the National Institutes of Health (P20GM125507).

HUMAN SUBJECTS/DATA USE AGREEMENT:

This was a contracted evaluation defined and funded by the CDC, the intent was to provide information for and about a specific program. It did not fall under the definition of human subjects research. All data sharing was conducted under an executed Data Use Agreement (DUA) between Rhode Island Department of Health and Brown University.

List all funding sources for the work

The Vulnerability Assessment and Jurisdictional Plan were funded by the Centers for Disease Control & Prevention (CDC-RFA-TP18-1802). Additional evaluation and dissemination activities were funded by the Center of Biomedical Research Excellence (COBRE) on Opioids and Overdose, National Institute of General Medical Sciences of the National Institutes of Health (P20GM125507).

REFERENCES

- ADAMS JW, LURIE MN, KING MRF, BRADY KA, GALEA S, FRIEDMAN SR, KHAN MR, & MARSHALL BDL (2018). POTENTIAL DRIVERS OF HIV ACQUISITION IN AFRICAN-AMERICAN WOMEN RELATED TO MASS INCARCERATION: AN AGENT-BASED MODELLING STUDY. BMC PUBLIC HEALTH, 18(1), 1387. 10.1186/S12889-018-6304-X [PubMed: 30563496]
- ALPREN C, DAWSON EL, JOHN B, CRANSTON K, PANNEER N, FUKUDA HD, ROOSEVELT K, KLEVENS RM, BRYANT J, PETERS PJ, LYSS SB, SWITZER WM, BURRAGE A,

MURRAY A, AGNEW-BRUNE C, STILES T, MCCLUNG P, CAMPBELL EM, BREEN C, ... BUCHACZ K (2020). OPIOID USE FUELING HIV TRANSMISSION IN AN URBAN SETTING: AN OUTBREAK OF HIV INFECTION AMONG PEOPLE WHO INJECT DRUGS- MASSACHUSETTS, 2015–2018. AMERICAN JOURNAL OF PUBLIC HEALTH, 110(1), 37–44. 10.2105/AJPH.2019.305366 [PubMed: 31725317]

BI Q, GOODMAN KE, KAMINSKY J, & LESSLER J (2019). WHAT IS MACHINE LEARNING? A PRIMER FOR THE EPIDEMIOLOGIST. AMERICAN JOURNAL OF EPIDEMIOLOGY, 188(12), 2222–2239. 10.1093/AJE/KWZ189 [PubMed: 31509183]

BOWEN DA, O'DONNELL J, & SUMNER SA (2019). INCREASES IN ONLINE POSTS ABOUT SYNTHETIC OPIOIDS PRECEDING INCREASES IN SYNTHETIC OPIOID DEATH RATES: A RETROSPECTIVE OBSERVATIONAL STUDY. JOURNAL OF GENERAL INTERNAL MEDICINE, 34(12), 2702–2704. 10.1007/S11606-019-05255-5 [PubMed: 31468343]

BOWMAN S, GRAU LE, SINGER M, SCOTT G, & HEIMER R (2014). FACTORS ASSOCIATED WITH HEPATITIS B VACCINE SERIES COMPLETION IN A RANDOMIZED TRIAL FOR INJECTION DRUG USERS REACHED THROUGH SYRINGE EXCHANGE PROGRAMS IN THREE US CITIES. BMC PUBLIC HEALTH, 14, 820. 10.1186/1471-2458-14-820 [PubMed: 25107530]

BROOKLYN JR, & SIGMON SC (2017). VERMONT HUB-AND-SPOKE MODEL OF CARE FOR OPIOID USE DISORDER: DEVELOPMENT, IMPLEMENTATION, AND IMPACT. JOURNAL OF ADDICTION MEDICINE, 11(4), 286–292. 10.1097/ADM.0000000000000310 [PubMed: 28379862]

CAMPO DS, GUSSLER JW, SUE A, SKUMS P, & KHUDYAKOV Y (2020). ACCURATE SPATIOTEMPORAL MAPPING OF DRUG OVERDOSE DEATHS BY MACHINE LEARNING OF DRUG-RELATED WEB-SEARCHES. PLOS ONE, 15(12), E0243622. 10.1371/JOURNAL.PONE.0243622 [PubMed: 33284864]

CARROLL JJ, MARSHALL BDL, RICH JD, & GREEN TC (2017). EXPOSURE TO FENTANYL-CONTAMINATED HEROIN AND OVERDOSE RISK AMONG ILLICIT OPIOID USERS IN RHODE ISLAND: A MIXED METHODS STUDY. THE INTERNATIONAL JOURNAL ON DRUG POLICY, 46, 136–145. 10.1016/J.DRUGPO.2017.05.023 [PubMed: 28578864]

CDC SOCIAL VULNERABILITY INDEX (SVI). (2020, 9 11). U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES. [HTTPS://WWW.ATSDR.CDC.GOV/PLACEANDHEALTH/SVI/AT-A-GLANCE_SVI.HTML](https://www.atsdr.cdc.gov/placeandhealth/svi/at-a-glance_svi.html)

CDC'S SOCIAL VULNERABILITY INDEX (SVI). (2021, 1 19). [HTTPS://WWW.ATSDR.CDC.GOV/PLACEANDHEALTH/SVI/INDEX.HTML](https://www.atsdr.cdc.gov/placeandhealth/svi/index.html)

CDC VIRAL HEPATITIS. (2021, 2 10). VIRAL HEPATITIS SURVEILLANCE AND CASE MANAGEMENT. [HTTPS://WWW.CDC.GOV/HEPATITIS/STATISTICS/2017SURVEILLANCE/TABLESFIGURES-HEPC.HTM](https://www.cdc.gov/hepatitis/statistics/2017surveillance/tablesfigures-hepc.htm)

CONRAD C, BRADLEY HM, BROZ D, BUDDHA S, CHAPMAN EL, GALANG RR, HILLMAN D, HON J, HOOVER KW, PATEL MR, PEREZ A, PETERS PJ, PONTONES P, ROSEBERRY JC, SANDOVAL M, SHIELDS J, WALTHALL J, WATERHOUSE D, WEIDLE PJ, ... CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). (2015). COMMUNITY OUTBREAK OF HIV INFECTION LINKED TO INJECTION DRUG USE OF OXYMORPHONE--INDIANA, 2015. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 64(16), 443–444. [HTTPS://WWW.NCBI.NLM.NIH.GOV/PUBMED/25928470](https://www.ncbi.nlm.nih.gov/pubmed/25928470) [PubMed: 25928470]

CRANSTON K, ALPREN C, JOHN B, DAWSON E, ROOSEVELT K, BURRAGE A, BRYANT J, SWITZER WM, BREEN C, PETERS PJ, STILES T, MURRAY A, FUKUDA HD, ADIH W, GOLDMAN L, PANNEER N, CALLIS B, CAMPBELL EM, RANDALL L, ... DEMARIA A JR. (2019). NOTES FROM THE FIELD: HIV DIAGNOSES AMONG PERSONS WHO INJECT DRUGS - NORTHEASTERN MASSACHUSETTS, 2015–2018. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 68(10), 253–254. 10.15585/MMWR.MM6810A6 [PubMed: 30870405]

CROWLEY JS, & MILLETT GA (2017). PREVENTING HIV AND HEPATITIS INFECTIONS AMONG PEOPLE WHO INJECT DRUGS: LEVERAGING AN INDIANA OUTBREAK RESPONSE TO BREAK THE IMPASSE. AIDS AND BEHAVIOR, 21(4), 968–972. 10.1007/S10461-017-1731-8 [PubMed: 28220312]

- DASGUPTA S, TIE Y, LEMONS A, WU K, BURNETT J, & SHOUSE RL (2019). INJECTION PRACTICES AND SEXUAL BEHAVIORS AMONG PERSONS WITH DIAGNOSED HIV INFECTION WHO INJECT DRUGS - UNITED STATES, 2015–2017. *MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT*, 68(30), 653–657. 10.15585/MMWR.MM6830A1 [PubMed: 31369525]
- DAVIS CS, CARR DH, & SAMUELS EA (2019). PARAPHERNALIA LAWS, CRIMINALIZING POSSESSION AND DISTRIBUTION OF ITEMS USED TO CONSUME ILLICIT DRUGS, AND INJECTION-RELATED HARM. *AMERICAN JOURNAL OF PUBLIC HEALTH*, 109(11), 1564–1567. 10.2105/AJPH.2019.305268 [PubMed: 31536408]
- DES JARLAIS DC, KERR T, CARRIERI P, FEELEMYER J, & ARASTEH K (2016). HIV INFECTION AMONG PERSONS WHO INJECT DRUGS: ENDING OLD EPIDEMICS AND ADDRESSING NEW OUTBREAKS. *AIDS*, 30(6), 815–826. 10.1097/QAD.0000000000001039 [PubMed: 26836787]
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (ESRI). (RELEASE DATE: 2 18, 2016). ARCGIS RELEASE 10.4 (ARCGIS FOR DESKTOP 10.4 FINAL) [BUILD 5524]. [HTTPS://WWW.ESRI.COM/EN-US/ARCGIS/PRODUCTS/ARCGIS-PRO/OVERVIEW](https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview)
- FREYER FJ (2019, 1 29). BOSTON SEES OUTBREAK OF HIV AMONG BOSTON DRUG USERS - THE BOSTON GLOBE. *BOSTONGLOBE.COM*. [HTTPS://WWW.BOSTONGLOBE.COM/METRO/2019/01/29/BOSTON-SEES-OUTBREAK-HIV-AMONG-BOSTON-DRUG-USERS/JTKQOFNIDJRR1PIAZGCOJ/STORY.HTML](https://www.bostonglobe.com/metro/2019/01/29/boston-sees-outbreak-hiv-among-boston-drug-users/jtkqofnidjrr1piazgcoj/story.html)
- GHOLAMY A, KREINOVICH V, & KOSHELEVA O (2018). WHY 70/30 OR 80/20 RELATION BETWEEN TRAINING AND TESTING SETS: A PEDAGOGICAL EXPLANATION. [HTTPS://SCHOLARWORKS.UTEP.EDU/CS_TECHREP/1209/](https://scholarworks.utep.edu/cs_techrep/1209/)
- GOEDEL WC, GREEN TC, VINER-BROWN S, RICH JD, & MARSHALL BDL (2019). INCREASED OVERDOSE MORTALITY DURING THE FIRST WEEK OF THE MONTH: REVISITING THE “CHECK EFFECT” THROUGH A SPATIAL LENS. *DRUG AND ALCOHOL DEPENDENCE*, 197, 49–55. 10.1016/J.DRUGALCDEP.2018.12.024 [PubMed: 30776571]
- GOEDEL WC, KING MRF, LURIE MN, GALEA S, TOWNSEND JP, GALVANI AP, FRIEDMAN SR, & MARSHALL BDL (2020). IMPLEMENTATION OF SYRINGE SERVICES PROGRAMS TO PREVENT RAPID HUMAN IMMUNODEFICIENCY VIRUS TRANSMISSION IN RURAL COUNTIES IN THE UNITED STATES: A MODELING STUDY. *CLINICAL INFECTIOUS DISEASES: AN OFFICIAL PUBLICATION OF THE INFECTIOUS DISEASES SOCIETY OF AMERICA*, 70(6), 1096–1102. 10.1093/CID/CIZ321 [PubMed: 31143944]
- GOIN DE, RUDOLPH KE, & AHERN J (2018). PREDICTORS OF FIREARM VIOLENCE IN URBAN COMMUNITIES: A MACHINE-LEARNING APPROACH. *HEALTH & PLACE*, 51, 61–67. 10.1016/J.HEALTHPLACE.2018.02.013 [PubMed: 29549755]
- GOLDING NJ (2017). THE NEEDLE AND THE DAMAGE DONE: INDIANA’S RESPONSE TO THE 2015 HIV EPIDEMIC AND THE NEED TO CHANGE STATE AND FEDERAL POLICIES REGARDING NEEDLE EXCHANGES AND INTRAVENOUS DRUG USERS. *INDIANA HEALTH LAW REVIEW*, 14, 173–220. [HTTPS://HEINONLINE.ORG/HOL/P?H=HEIN.JOURNALS/INHEALR14&I=387](https://heinonline.org/hol/p?h=hein.journals/inhealr14&i=387)
- HALLOWELL BD, WEIDLE HR, & SCAGOS RP (2020). ACCIDENTAL DRUG OVERDOSE DEATHS IN RHODE ISLAND: JANUARY 1, 2016–JULY 31, 2020. *RI MEDICAL JOURNAL*, 103(10), 62–65. [HTTP://WWW.RIMED.ORG/RIMEDICALJOURNAL/2020/12/2020-12-62-HEALTH-HALLOWELL.PDF](http://www.rimed.org/rimedicaljournal/2020/12/2020-12-62-health-hallowell.pdf)
- HASTIE T, & QIAN J (2014, 6 26). GLMNET VIGNETTE. [HTTPS://WEB.STANFORD.EDU/~HASTIE/GLMNET/GLMNET_ALPHA.HTML](https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html)
- HENRY KA, & BOSCOE FP (2008). ESTIMATING THE ACCURACY OF GEOGRAPHICAL IMPUTATION. *INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS*, 7, 3. 10.1186/1476-072X-7-3 [PubMed: 18215308]
- JAMES K, & JORDAN A (2018). THE OPIOID CRISIS IN BLACK COMMUNITIES. *THE JOURNAL OF LAW, MEDICINE & ETHICS: A JOURNAL OF THE AMERICAN SOCIETY OF LAW, MEDICINE & ETHICS*, 46(2), 404–421. 10.1177/1073110518782949

- JIANG Y, MCDONALD JV, WILSON ME, KOZIOL J, GOLDSCHMIDT A, KING E, VINER-BROWN S, POWELL SM, & ALEXANDER-SCOTT N (2018). RHODE ISLAND UNINTENTIONAL DRUG OVERDOSE DEATH TRENDS AND RANKING - OFFICE OF THE STATE MEDICAL EXAMINERS DATABASE. RHODE ISLAND MEDICAL JOURNAL, 101(1), 33–36. [HTTPS://WWW.NCBI.NLM.NIH.GOV/PUBMED/29393310](https://www.ncbi.nlm.nih.gov/pubmed/29393310) [PubMed: 29393310]
- KRIEGER MS, YEDINAK JL, BUXTON JA, LYSYSHYN M, BERNSTEIN E, RICH JD, GREEN TC, HADLAND SE, & MARSHALL BDL (2018). HIGH WILLINGNESS TO USE RAPID FENTANYL TEST STRIPS AMONG YOUNG ADULTS WHO USE DRUGS. HARM REDUCTION JOURNAL, 15(1), 7. 10.1186/S12954-018-0213-2 [PubMed: 29422052]
- KRSTAJIC D, BUTUROVIC LJ, LEAHY DE, & THOMAS S (2014). CROSS-VALIDATION PITFALLS WHEN SELECTING AND ASSESSING REGRESSION AND CLASSIFICATION MODELS. JOURNAL OF CHEMINFORMATICS, 6(1), 10. 10.1186/1758-2946-6-10 [PubMed: 24678909]
- KUHN M (2008). BUILDING PREDICTIVE MODELS IN R USING THE CARET PACKAGE. JOURNAL OF STATISTICAL SOFTWARE, 28(5), 1–26. 10.18637/JSS.V028.I05 [PubMed: 27774042]
- LINTON SL, COOPER HL, KELLEY ME, KARNES CC, ROSS Z, WOLFE ME, FRIEDMAN SR, JARLAIS DD, SEMAAN S, TEMPALSKI B, SIONEAN C, DINENNO E, WEJNERT C, PAZ-BAILEY G, & NATIONAL HIV BEHAVIORAL SURVEILLANCE STUDY GROUP. (2017). CROSS-SECTIONAL ASSOCIATION BETWEEN ZIP CODE-LEVEL GENTRIFICATION AND HOMELESSNESS AMONG A LARGE COMMUNITY-BASED SAMPLE OF PEOPLE WHO INJECT DRUGS IN 19 US CITIES. BMJ OPEN, 7(6), E013823. 10.1136/BMJOPEN-2016-013823
- LI ZR, XIE E, CRAWFORD FW, WARREN JL, MCCONNELL K, COPPLE JT, JOHNSON T, & GONSALVES GS (2019). SUSPECTED HEROIN-RELATED OVERDOSES INCIDENTS IN CINCINNATI, OHIO: A SPATIOTEMPORAL ANALYSIS. PLOS MEDICINE, 16(11), E1002956. 10.1371/JOURNAL.PMED.1002956 [PubMed: 31714940]
- LO-CIGANIC W-H, HUANG JL, ZHANG HH, WEISS JC, WU Y, KWOK CK, DONOHUE JM, COCHRAN G, GORDON AJ, MALONE DC, KUZA CC, & GELLAD WF (2019). EVALUATION OF MACHINE-LEARNING ALGORITHMS FOR PREDICTING OPIOID OVERDOSE RISK AMONG MEDICARE BENEFICIARIES WITH OPIOID PRESCRIPTIONS. JAMA NETWORK OPEN, 2(3), E190968. 10.1001/JAMANETWORKOPEN.2019.0968 [PubMed: 30901048]
- MARSHALL BDL (2018). USING DATA TO GUIDE AND EVALUATE RESPONSES TO THE OPIOID CRISIS: RHODE ISLAND'S DRUG OVERDOSE DASHBOARD. [HTTPS://ESCHOLARSHIP.UMASSMED.EDU/CGI/VIEWCONTENT.CGI?ARTICLE=1054&CONTEXT=NER](https://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1054&context=ner)
- MARSHALL BDL, KRIEGER MS, YEDINAK JL, OGERA P, BANERJEE P, ALEXANDER-SCOTT NE, RICH JD, & GREEN TC (2017). EPIDEMIOLOGY OF FENTANYL-INVOLVED DRUG OVERDOSE DEATHS: A GEOSPATIAL RETROSPECTIVE STUDY IN RHODE ISLAND, USA. THE INTERNATIONAL JOURNAL ON DRUG POLICY, 46, 130–135. 10.1016/J.DRUGPO.2017.05.029 [PubMed: 28601512]
- MERCADO-CRESPO MC, SUMNER SA, SPELKE MB, SUGERMAN DE, STANLEY C, & EIS OFFICER CDC. (2014). NOTES FROM THE FIELD: INCREASE IN FENTANYL-RELATED OVERDOSE DEATHS - RHODE ISLAND, NOVEMBER 2013-MARCH 2014. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 63(24), 531. [HTTPS://WWW.NCBI.NLM.NIH.GOV/PUBMED/24941333](https://www.ncbi.nlm.nih.gov/pubmed/24941333) [PubMed: 24941333]
- MERCADO MC, SUMNER SA, SPELKE MB, BOHM MK, SUGERMAN DE, & STANLEY C (2018). INCREASE IN DRUG OVERDOSE DEATHS INVOLVING FENTANYL-RHODE ISLAND, JANUARY 2012-MARCH 2014. PAIN MEDICINE, 19(3), 511–523. 10.1093/PM/PNX015 [PubMed: 28340233]
- METCALF CJE, EDMUNDS WJ, & LESSLER J (2015). SIX CHALLENGES IN MODELLING FOR PUBLIC HEALTH POLICY. EPIDEMICS, 10, 93–96. 10.1016/J.EPIDEM.2014.08.008 [PubMed: 25843392]

- NANDI A, GALEA S, AHERN J, BUCCIARELLI A, VLAHOV D, & TARDIFF K (2006). WHAT EXPLAINS THE ASSOCIATION BETWEEN NEIGHBORHOOD-LEVEL INCOME INEQUALITY AND THE RISK OF FATAL OVERDOSE IN NEW YORK CITY? SOCIAL SCIENCE & MEDICINE, 63(3), 662–674. 10.1016/J.SOCSCIMED.2006.02.001 [PubMed: 16597478]
- OFFICE FOR HUMAN RESEARCH PROTECTIONS (OHRP). (2016, 2 16). 45 CFR 46. US DEPARTMENT OF HEALTH AND HUMAN SERVICES. [HTTPS://WWW.HHS.GOV/OHRP/REGULATIONS-AND-POLICY/REGULATIONS/45-CFR-46/INDEX.HTML](https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html)
- OLSON MC, GARDENIER D, & JACOBSON IM (2015). THE REVOLUTION OF HEPATITIS C TREATMENTS: REVIEW FOR NURSE PRACTITIONERS. THE JOURNAL FOR NURSE PRACTITIONERS: JNP, 11(1), 116–123. 10.1016/J.NURPRA.2014.10.013
- OSMAN H, GHAFARI M, & NIERSTRASZ O (2017). AUTOMATIC FEATURE SELECTION BY REGULARIZATION TO IMPROVE BUG PREDICTION ACCURACY. 2017 IEEE WORKSHOP ON MACHINE LEARNING TECHNIQUES FOR SOFTWARE QUALITY EVALUATION (MALTESQUE), 27–32. 10.1109/MALTESQUE.2017.7882013
- PAGE K, MORRIS MD, HAHN JA, MAHER L, & PRINS M (2013). INJECTION DRUG USE AND HEPATITIS C VIRUS INFECTION IN YOUNG ADULT INJECTORS: USING EVIDENCE TO INFORM COMPREHENSIVE PREVENTION. CLINICAL INFECTIOUS DISEASES: AN OFFICIAL PUBLICATION OF THE INFECTIOUS DISEASES SOCIETY OF AMERICA, 57 SUPPL 2, S32–S38. 10.1093/CID/CIT300 [PubMed: 23884063]
- PERLMAN DC, & JORDAN AE (2018). THE SYNDROMIC OF OPIOID MISUSE, OVERDOSE, HCV, AND HIV: STRUCTURAL-LEVEL CAUSES AND INTERVENTIONS. CURRENT HIV/AIDS REPORTS, 15(2), 96–112. 10.1007/S11904-018-0390-3 [PubMed: 29460225]
- R CORE TEAM. (2014). R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL COMPUTING. R FOUNDATION FOR STATISTICAL COMPUTING (VERSION 3.5.2) [COMPUTER SOFTWARE]. [HTTPS://WWW.R-PROJECT.ORG/](https://www.r-project.org/)
- REEPING PM, CERDÁ M, KALESAN B, WIEBE DJ, GALEA S, & BRANAS CC (2019). STATE GUN LAWS, GUN OWNERSHIP, AND MASS SHOOTINGS IN THE US: CROSS SECTIONAL TIME SERIES. BMJ, 364, L542. 10.1136/BMJ.L542 [PubMed: 30842105]
- RICH KM, BIA J, ALTICE FL, & FEINBERG J (2018). INTEGRATED MODELS OF CARE FOR INDIVIDUALS WITH OPIOID USE DISORDER: HOW DO WE PREVENT HIV AND HCV? CURRENT HIV/AIDS REPORTS, 15(3), 266–275. 10.1007/S11904-018-0396-X [PubMed: 29774442]
- RICKLES M, REBEIRO PF, SIZEMORE L, JUAREZ P, MUTTER M, WESTER C, & MCPHEETERS M (2018). TENNESSEE'S IN-STATE VULNERABILITY ASSESSMENT FOR A "RAPID DISSEMINATION OF HUMAN IMMUNODEFICIENCY VIRUS OR HEPATITIS C VIRUS INFECTION" EVENT UTILIZING DATA ABOUT THE OPIOID EPIDEMIC. CLINICAL INFECTIOUS DISEASES: AN OFFICIAL PUBLICATION OF THE INFECTIOUS DISEASES SOCIETY OF AMERICA, 66(11), 1722–1732. 10.1093/CID/CIX1079 [PubMed: 29228122]
- ROTH AM, KRAL AH, MITCHELL A, MUKHERJEE R, DAVIDSON P, & LANKENAU SE (2019). OVERDOSE PREVENTION SITE ACCEPTABILITY AMONG RESIDENTS AND BUSINESSES SURROUNDING A PROPOSED SITE IN PHILADELPHIA, USA. JOURNAL OF URBAN HEALTH: BULLETIN OF THE NEW YORK ACADEMY OF MEDICINE, 96(3), 341–352. 10.1007/S11524-019-00364-2 [PubMed: 31168733]
- RUDD RA, ALESHIRE N, ZIBBELL JE, & GLADDEN RM (2016). INCREASES IN DRUG AND OPIOID OVERDOSE DEATHS--UNITED STATES, 2000–2014. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 64(50–51), 1378–1382. 10.15585/MMWR.MM6450A3 [PubMed: 26720857]
- RUDD RA, SETH P, DAVID F, & SCHOLL L (2016). INCREASES IN DRUG AND OPIOID-INVOLVED OVERDOSE DEATHS - UNITED STATES, 2010–2015. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 65(5051), 1445–1452. 10.15585/MMWR.MM655051E1 [PubMed: 28033313]

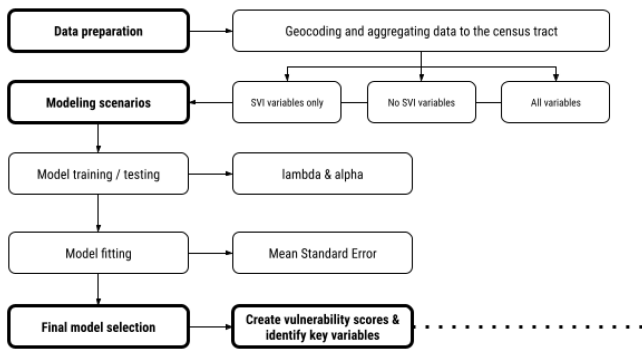
- SCHOLL L (2019). DRUG AND OPIOID-INVOLVED OVERDOSE DEATHS — UNITED STATES, 2013–2017. *MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT*, 67. 10.15585/MMWR.MM6751521E1
- SCHUMAKER E (2018, 8 6). OPIOIDS HAVE SPARKED AN HIV OUTBREAK IN MASSACHUSETTS. HUFFPOST. [HTTPS://WWW.HUFFPOST.COM/ENTRY/MASSACHUSETTS-FENTANYL-OPIOIDS-HIV_N_5B6470BFE4B0B15ABAA2958C](https://www.huffpost.com/entry/massachusetts-fentanyl-opioids-hiv_n_5b6470bfe4b0b15abaa2958c)
- SPADA E, GENOVESE D, TOSTI ME, MARIANO A, CUCCUINI M, PROIETTI L, GIULI CD, LAVAGNA A, CRAPA GE, MORACE G, TAFFON S, MELE A, REZZA G, & RAPICETTA M (2005). AN OUTBREAK OF HEPATITIS A VIRUS INFECTION WITH A HIGH CASE-FATALITY RATE AMONG INJECTING DRUG USERS. *JOURNAL OF HEPATOLOGY*, 43(6), 958–964. 10.1016/J.JHEP.2005.06.012 [PubMed: 16143420]
- SURYAPRASAD AG, WHITE JZ, XU F, EICHLER B-A, HAMILTON J, PATEL A, HAMDOUNIA SB, CHURCH DR, BARTON K, FISHER C, MACOMBER K, STANLEY M, GUILFOYLE SM, SWEET K, LIU S, IQBAL K, TOHME R, SHARAPOV U, KUPRONIS BA, ... HOLMBERG SD (2014). EMERGING EPIDEMIC OF HEPATITIS C VIRUS INFECTIONS AMONG YOUNG NONURBAN PERSONS WHO INJECT DRUGS IN THE UNITED STATES, 2006–2012. *CLINICAL INFECTIOUS DISEASES: AN OFFICIAL PUBLICATION OF THE INFECTIOUS DISEASES SOCIETY OF AMERICA*, 59(10), 1411–1419. 10.1093/CID/CIU643 [PubMed: 25114031]
- SYLLA L, BRUCE RD, KAMARULZAMAN A, & ALTICE FL (2007). INTEGRATION AND COLOCATION OF HIV/AIDS, TUBERCULOSIS AND DRUG TREATMENT SERVICES. *THE INTERNATIONAL JOURNAL ON DRUG POLICY*, 18(4), 306–312. 10.1016/J.DRUGPO.2007.03.001 [PubMed: 17689379]
- TIBSHIRANI R, BIEN J, FRIEDMAN J, HASTIE T, SIMON N, TAYLOR J, & TIBSHIRANI RJ (2012). STRONG RULES FOR DISCARDING PREDICTORS IN LASSO-TYPE PROBLEMS. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY. SERIES B, STATISTICAL METHODOLOGY*, 74(2), 245–266. 10.1111/J.1467-9868.2011.01004.X [PubMed: 25506256]
- TRINQUART L, ERLINGER AL, PETERSEN JM, FOX M, & GALEA S (2019). APPLYING THE EVALUATE TO ASSESS THE ROBUSTNESS OF EPIDEMIOLOGIC FIELDS OF INQUIRY TO UNMEASURED CONFOUNDING. *AMERICAN JOURNAL OF EPIDEMIOLOGY*. 10.1093/AJE/KWZ063
- US CENSUS BUREAU. (2019, 9 16). UNITED STATES CENSUS BUREAU GEOGRAPHY PROGRAM GLOSSARY. US CENSUS BUREAU. [HTTPS://WWW.CENSUS.GOV/PROGRAMS-SURVEYS/GEOGRAPHY/ABOUT/GLOSSARY.HTML#PAR_TEXTIMAGE_13](https://www.census.gov/programs-surveys/geography/about/glossary.html#PAR_TEXTIMAGE_13)
- VAN HANDEL MM, ROSE CE, HALLISEY EJ, KOLLING JL, ZIBBELL JE, LEWIS B, BOHM MK, JONES CM, FLANAGAN BE, SIDDIQI A-E-A, IQBAL K, DENT AL, MERMIN JH, MCCRAY E, WARD JW, & BROOKS JT (2016). COUNTY-LEVEL VULNERABILITY ASSESSMENT FOR RAPID DISSEMINATION OF HIV OR HCV INFECTIONS AMONG PERSONS WHO INJECT DRUGS, UNITED STATES. *JOURNAL OF ACQUIRED IMMUNE DEFICIENCY SYNDROMES*, 73(3), 323–331. 10.1097/QAI.0000000000001098 [PubMed: 27763996]
- VISCONTI AJ, SANTOS G-M, LEMOS NP, BURKE C, & COFFIN PO (2015). OPIOID OVERDOSE DEATHS IN THE CITY AND COUNTY OF SAN FRANCISCO: PREVALENCE, DISTRIBUTION, AND DISPARITIES. *JOURNAL OF URBAN HEALTH: BULLETIN OF THE NEW YORK ACADEMY OF MEDICINE*, 92(4), 758–772. 10.1007/S11524-015-9967-Y [PubMed: 26077643]
- WANG Z (2019). MPATH: REGULARIZED LINEAR MODELS (VERSION VERSION 0.4–2.19) [R]. [HTTPS://CRAN.R-PROJECT.ORG/WEB/PACKAGES/MPATH/INDEX.HTML](https://cran.r-project.org/web/packages/mpath/index.html)
- WAYE KM, YEDINAK JL, KOZIOL J, & MARSHALL BDL (2018). ACTION-FOCUSED, PLAIN LANGUAGE COMMUNICATION FOR OVERDOSE PREVENTION: A QUALITATIVE ANALYSIS OF RHODE ISLAND'S OVERDOSE SURVEILLANCE AND INFORMATION DASHBOARD. *INTERNATIONAL JOURNAL OF DRUG POLICY*, 62, 86–93. 10.1016/J.DRUGPO.2018.08.010 [PubMed: 30384027]
- WINOGRAD RP, PRESNALL N, STRINGFELLOW E, WOOD C, HORN P, DUELLO A, GREEN L, & RUDDER T (2019). THE CASE FOR A MEDICATION FIRST APPROACH TO THE

- TREATMENT OF OPIOID USE DISORDER. THE AMERICAN JOURNAL OF DRUG AND ALCOHOL ABUSE, 45(4), 333–340. 10.1080/00952990.2019.1605372 [PubMed: 31084515]
- YEDINAK JL, KRIEGER MS, LI Y, NELSON E, SHIHIPAR A, GOLDMAN JE, & MARSHALL BDL (2019). THE VILLAGE PREVENTION PLAN: IMPROVING PREVENTION EFFORTS AND HEALTH OUTCOMES FOR PEOPLE WHO USE DRUGS. BROWN UNIVERSITY SCHOOL OF PUBLIC HEALTH. [HTTPS://PREVENTOVERDOSERI.ORG/WP-CONTENT/UPLOADS/2020/01/VILLAGE-PREVENTION-PLAN.PDF](https://preventoverdoseri.org/wp-content/uploads/2020/01/village-prevention-plan.pdf)
- YOUNG SD, ZHENG K, CHU LF, & HUMPHREYS K (2018). INTERNET SEARCHES FOR OPIOIDS PREDICT FUTURE EMERGENCY DEPARTMENT HEROIN ADMISSIONS. DRUG AND ALCOHOL DEPENDENCE, 190, 166–169. 10.1016/J.DRUGALCDEP.2018.05.009 [PubMed: 30036853]
- ZIBBELL JE, ASHER AK, PATEL RC, KUPRONIS B, IQBAL K, WARD JW, & HOLTZMAN D (2018). INCREASES IN ACUTE HEPATITIS C VIRUS INFECTION RELATED TO A GROWING OPIOID EPIDEMIC AND ASSOCIATED INJECTION DRUG USE, UNITED STATES, 2004 TO 2014. AMERICAN JOURNAL OF PUBLIC HEALTH, 108(2), 175–181. 10.2105/AJPH.2017.304132 [PubMed: 29267061]
- ZIBBELL JE, IQBAL K, PATEL RC, SURYAPRASAD A, SANDERS KJ, MOORE-MORAVIAN L, SERRECCHIA J, BLANKENSHIP S, WARD JW, HOLTZMAN D, & CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC). (2015). INCREASES IN HEPATITIS C VIRUS INFECTION RELATED TO INJECTION DRUG USE AMONG PERSONS AGED 30 YEARS - KENTUCKY, TENNESSEE, VIRGINIA, AND WEST VIRGINIA, 2006–2012. MMWR. MORBIDITY AND MORTALITY WEEKLY REPORT, 64(17), 453–458. [HTTPS://WWW.NCBI.NLM.NIH.GOV/PUBMED/25950251](https://www.ncbi.nlm.nih.gov/pubmed/25950251) [PubMed: 25950251]
- ZOU H, & HASTIE T (2005). REGULARIZATION AND VARIABLE SELECTION VIA THE ELASTIC NET. JOURNAL OF THE ROYAL STATISTICAL SOCIETY. SERIES B, STATISTICAL METHODOLOGY, 67(2), 301–320. 10.1111/J.1467-9868.2005.00503.X

HIGHLIGHTS:

- Injection drug use is a driver for emerging outbreaks of HIV and HCV.
- We assessed risk for an outbreak in a state with a high burden of overdose.
- Machine learning methods can identify neighborhood-level factors for risk.
- Identification of neighborhood-level factors can guide interventions.
- Local experts in service delivery provide important context for data interpretation.

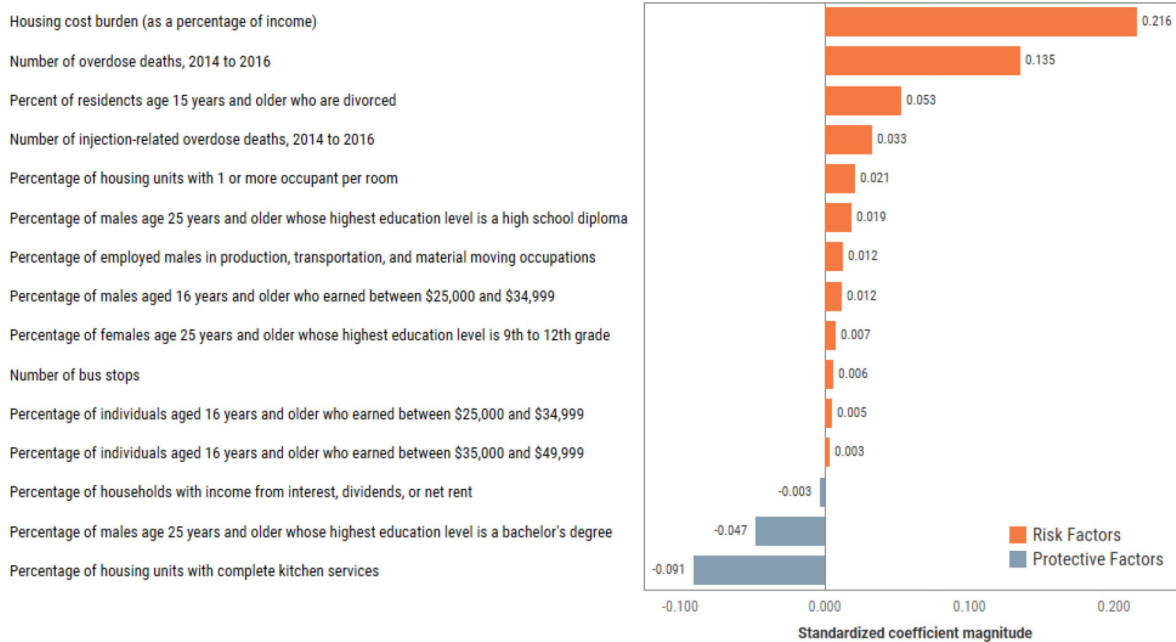
Aim 1. Supervised Machine Learning Model



Aim 2. Stakeholder Engagement & Mapping Tool



Figure 1.
The VILLAGE Project Workflow



The variables of importance are statistically generated predictors with non-zero coefficients after regularization. We define "importance" as the magnitude of mean change in vulnerability scores given a one-unit increase in the predictor.

Figure 2. Protective factors and risk factors contributing to statewide vulnerability assessment scores in Rhode Island, US.

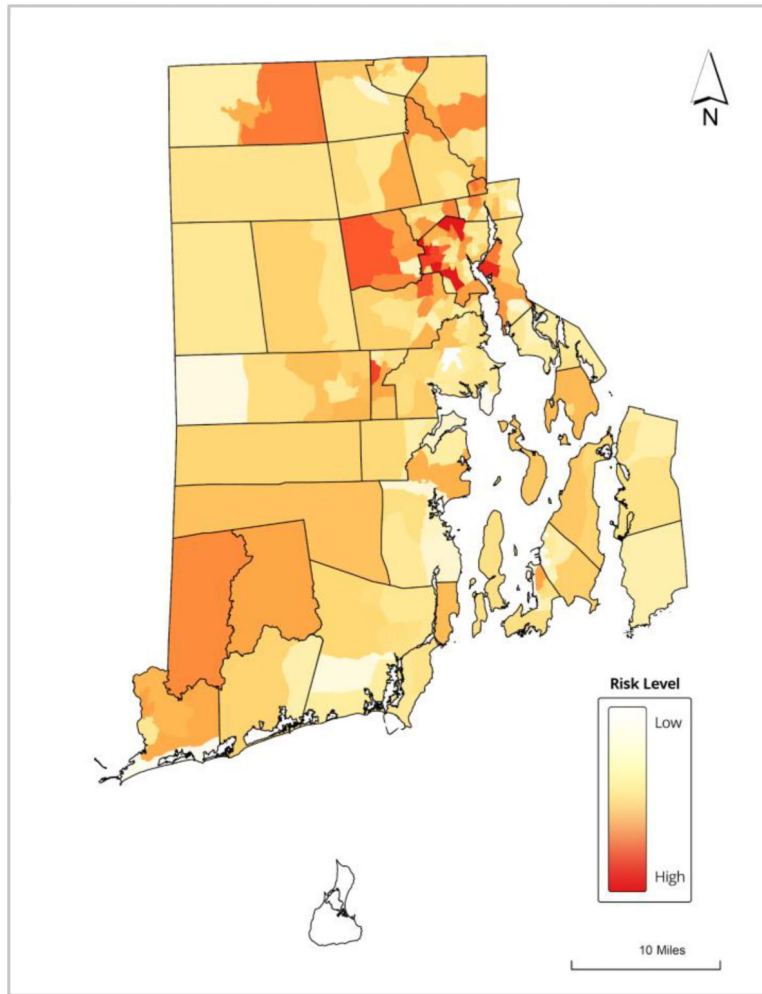
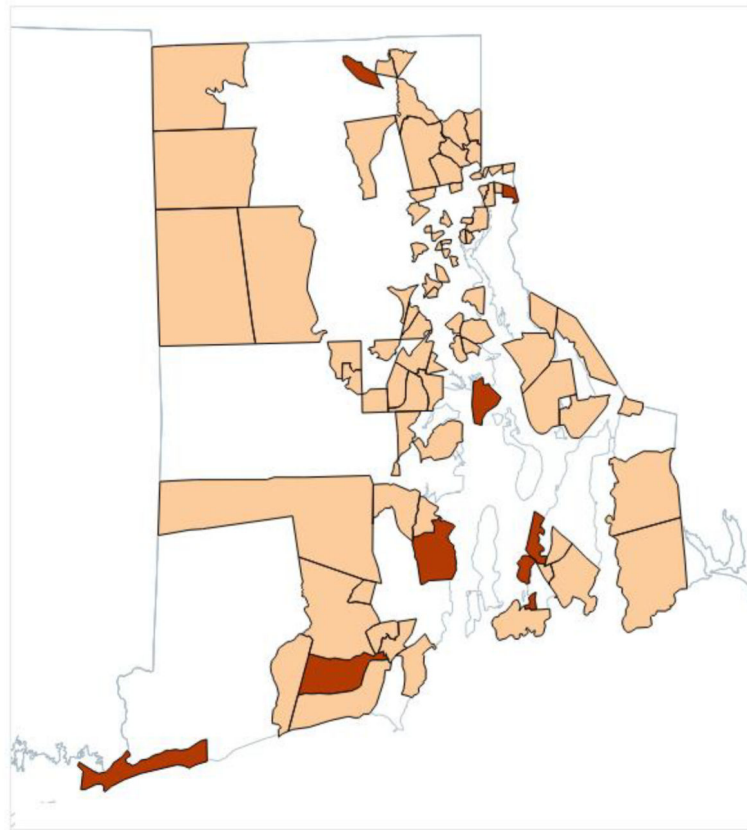


Figure 3. Vulnerability assessment score stratification across census tracts in Rhode Island, US



Eighty-nine census tracts (37%) previously experiencing no overdose deaths were identified as being vulnerable to an outbreak of overdose, HIV and HCV. **Nine** census tracts with no historical overdose deaths were identified as having a vulnerability assessment score in the top 25%.

- Newly identified areas vulnerable to an outbreak
- Newly identified areas vulnerable to an outbreak in Top 25%

Figure 4. Newly predicted vulnerability assessment scores by census tract, Rhode Island, US

Table 1.

Glossary

Term	Definition
Outbreak	We use the term <i>outbreak</i> to refer to a significant increase in the incidence of accidental drug overdose, HIV, or hepatitis C (HCV) transmission in a region. Regional susceptibility to an outbreak can be assessed by identifying structural factors such as changes to the drug supply, availability of syringe service programs and prevention services, and access to affordable housing.
Proxy	We use the term proxy because no clear data exists for what we are hoping to measure, neighbourhood injection drug use patterns and related health outcomes, so we chose a single variable to best represent, or infer, those patterns. In The VILLAGE Project, we used 2017 overdose deaths in Rhode Island as our model outcome, representing a single proxy for IDU-related behaviors that may lead to an outbreak.
Machine Learning	<i>Machine learning</i> is an analytical approach that uses computational methods and algorithms to process and analyze large amounts of data (Bi Q., et al., 2019). They have been shown to create robust prediction models that (in this case) can help identify complex structural factors and pathways for intervention (Lo-Ciganic et al., 2019).
Supervised Machine Learning	<i>Supervised machine learning</i> refers to when the machine learning methods rely on a “training set” to help cross-validate the results and evaluate the accuracy of the predictions. In The VILLAGE Project, we split 70% of the dataset into a “training set”, and the remaining 30% was used as the “test set” consistent with common standards for cross-validation in the field (Gholamy et al., 2018).
Evaluation	The VILLAGE Project was determined to be a public health <i>evaluation</i> , rather than a research study, because the activities were conducted on behalf of a state health department and designed for local implementation. Therefore, it does not meet the US federal definition of human subjects research according to 45 CFR 46 Protection of Human Subjects (Office for Human Research Protections (OHRP), 2016).
Vulnerability Assessment Score	The machine learning methods used helped to identify significant variables in the data, such as community factors, that might increase the likelihood of the model outcome--in this case it was overdose deaths, our proxy for an outbreak. Looking at the combination of significant variables for each neighbourhood, we were able to assign <i>vulnerability assessment scores</i> across the state to represent the areas at higher risk (relative to other areas in the state) for experiencing overdose deaths. This work builds on previous studies of regional vulnerability using risk scoring, with HCV infection as a proxy for an outbreak (Ricklees et al., 2018; Van Handel et al., 2016).
CDC’s Social Vulnerability Index (SVI)	The US Centers for Disease Control and Prevention (CDC) created a <i>Social Vulnerability Index</i> (SVI) using 15 US Census data variables to help identify regions potentially in need of increased resources in the face of natural disasters or public health emergencies (CDC’s Social Vulnerability Index (SVI), 2021).