



Published in final edited form as:

Soc Sci Med. 2025 August ; 378: 118156. doi:10.1016/j.socscimed.2025.118156.

Considerations for the epidemiological evaluation of hyperlocal interventions: A case study of the New York City overdose prevention centers

Bennett Allen^{1,*}, Brandi Moore^{1,2,*}, Victoria A. Jent¹, William C. Goedel³, Khadija Israel^{1,4}, Alexandra B. Collins^{3,5}, Brandon D. L. Marshall⁴, Magdalena Cerdá¹

¹Center for Opioid Epidemiology and Policy, Department of Population Health, NYU Grossman School of Medicine

²Department of Epidemiology, NYU School of Global Public Health

³Department of Epidemiology, Brown University School of Public Health

⁴NYU Silver School of Social Work

⁵Department of Community Health, Tufts University

Abstract

To meet the needs of diverse communities, public health authorities are increasingly reliant on hyperlocal interventions targeting specific health issues and distinct populations. To facilitate epidemiological evaluation of hyperlocal interventions on community-level outcomes, we developed a framework of six practice-based considerations for researchers: spatial zone of impact, temporal resolution of impact, outcome of interest, definition of a plausible comparison group, micro vs. macro impacts, and practitioner engagement. We applied this framework to a case study of an impact evaluation of the New York City (NYC) overdose prevention centers (OPCs) on neighborhood-level drug-related arrests. We used drug arrest data from NYC from January 1, 2014, to September 30, 2023 and US Census data to conduct synthetic control modeling, comparing pre- and post-OPC arrests in the neighborhoods surrounding the two NYC OPCs (East Harlem and Washington Heights). We conducted sensitivity analyses to validate our results and compare our findings with those from a prior published study. Our findings indicate no significant change in drug-related arrests following the OPC openings. The mean absolute differences in daily drug-related arrests between the OPCs and their synthetic controls were 0.63 ($p=0.19$) in

Corresponding Author: Magdalena Cerdá, DrPH, MPH, Professor and Director, Division of Epidemiology, Director, Center for Opioid Epidemiology, Department of Population Health, NYU Grossman School of Medicine, 180 Madison Avenue, 5th Floor, New York, NY 10016, magdalena.cerda@nyulangone.org.

*Denotes joint first authorship

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.

Declaration of Interests

The authors declare no conflicts of interest.

Ethics approval statement

This study was deemed exempt by the NYU Langone Health Institutional Review Board because it did not involve human subjects.

East Harlem and 0.14 ($p=0.22$) in Washington Heights. Sensitivity analyses corroborated our main results. Overall, findings demonstrate how our framework can be used to guide future epidemiological evaluations of diverse, hyperlocal public health interventions.

Keywords

epidemiology; overdose prevention center; drug; arrest; New York City; program evaluation

Introduction

Promoting health equity remains a guiding principle and pressing challenge for public health researchers and practitioners in the face of growing community health disparities along racial and ethnic,¹ socioeconomic,² and geographical³ lines. To meet the needs of diverse communities and cultivate health equity, health authorities have become increasingly reliant on “hyperlocal” interventions targeting specific health issues and tailored to distinct populations with narrowly defined implementation geographies.⁴ Sometimes described as interventions designed for communities, neighborhoods, or small-areas, such hyperlocal interventions have been implemented across a range of public health domains, including, for example, mobile fruit and vegetable distribution to promote nutrition in food deserts,⁵ barbershop-based outreach to increase primary and preventive healthcare utilization among Black men,⁶ and the use of community mediation outreach (i.e., “violence interrupters”) to prevent gun violence.⁷ By prioritizing the needs and well-being of marginalized communities, such hyperlocal interventions hold promise for addressing health disparities.⁸

Despite the deployment of hyperlocal interventions across these diverse public health domains, their impact evaluation presents significant challenges within the applied and localized public health context.⁹ While substantial attention has been paid to the evaluation of the effect(s) of hyperlocal interventions on the health of people receiving such services,^{10,11} less attention has been paid to the unique methodological challenges that arise when evaluating their impact to estimate their causal effects on their surrounding communities. For example, identification of a catchment area (i.e., the neighborhood or community in which a hyperlocal intervention is implemented) that is small enough for an impact to be detected, but large enough for valid estimation, presents a critical epidemiological challenge.¹² Within a plausible catchment area, the hyperlocal context presents additional difficulties to researchers’ understanding of the core components of epidemiological research—e.g., measuring an exposure, specifying an outcome, and defining a target population—all of which have important implications for the findings of a given study. However, to our knowledge, no guidance about the conduct of area-level impact evaluations of hyperlocal interventions to estimate causal effects using methods for policy evaluation currently exists. In this article, we provide a framework to guide applied epidemiologists and public health practitioners in conducting community-level impact evaluations of hyperlocal interventions.

We define “hyperlocal” program evaluation to mean estimation of the small-area neighborhood- or community-level impact of a given intervention, for which impact is

assumed to differ across sub-municipal contexts. While a broad and deep body of social epidemiological literature has been generated about neighborhood effects on health,^{13–16} program evaluation at the hyperlocal level remains less developed. As such, most policy and program evaluation studies estimate at either the level of individual patients, for program evaluation, or broad population domains (e.g., cities, states, or countries), for policy evaluation.

These approaches present a gap for epidemiologists who wish to estimate the impact of programs and policies for which a geospatial effect is anticipated at the small-area level. When evaluating the effect of hyperlocal public health programs on the surrounding community, epidemiologists must situate the core components of epidemiological study design within that local context. To aid epidemiologists and public health researchers, we present six considerations for the evaluation of the neighborhood or community effects of hyperlocal public health programs:

1. *Spatial zone of impact*, or the plausible intervention catchment area;
2. *Temporal resolution of impact*, or the plausible time period during which an intervention impact could be detected;
3. *Outcome of interest*, or the appropriately interpretable specification of the outcome at the hyperlocal level;
4. *Definition of a plausible comparison group*, and by extension, the identification of confounding structures and their implications in the local context, for which residential segregation and other manifestations of structural racism may be particularly salient;
5. *Micro vs. macro impacts*, or the estimation of individual or joint impacts of multiple, hyperlocal interventions in a single parent jurisdiction; and
6. *Practitioner engagement*, or the inclusion of program operators and secondary data stewards throughout the research process.

These considerations correspond to six guiding questions to which epidemiological researchers and practitioners can jointly respond when estimating the impact of hyperlocal programs (Table 1). We illustrate the utility of our evaluation considerations through a case study of the two New York City (NYC) overdose prevention centers (OPCs; also known as “supervised consumption sites” or “drug consumption rooms”), which are community-based facilities where individuals can consume pre-obtained controlled substances under medical or peer observation.¹⁷

Case study: The New York City overdose prevention centers

Unintentional drug overdose remains a leading cause of accidental death in the United States (US),¹⁸ with marked recent increases following the introduction of illicitly manufactured synthetic opioids into US drug markets and disruptions associated with the COVID-19 pandemic.^{19–21} In response, a range of policy and program interventions to reduce and prevent overdose have been implemented.²² While many of these interventions have been implemented at the national, state, or local levels (e.g., state policies mandating prescriber

use of prescription drug monitoring programs²³ and programs to universally equip local first responders with the opioid overdose antidote naloxone²⁴), other interventions have been implemented hyperlocally to serve specific communities at high risk of overdose mortality (e.g., harm reduction vending machines).²⁵

OPCs are one such hyperlocal intervention. At an OPC, trained staff are equipped to intervene to reverse overdoses that may occur, as well as to refer and connect people to healthcare and social services.²⁶ At the neighborhood level, research has shown that OPCs are associated with reductions in overdose mortality,^{27,28} reductions in emergency service calls for overdose response,²⁹ and decreased or unchanged drug-related crime in the neighborhoods in which OPCs are situated.^{30,31} As a low-threshold harm reduction intervention, OPCs have been implemented globally, including in Europe, Canada, and Australia.³² To date, only two publicly recognized OPCs, however, exist in the US, both in NYC.³³

The NYC OPCs of focus in this analysis are embedded within the syringe service program (SSP) OnPoint NYC, which maintains two storefront locations: OnPoint-East Harlem (OPEH) and OnPoint-Washington Heights (OPWH, see Figure 1).³⁴ Each SSP provides a range of services, including syringe access and disposal, overdose education and naloxone distribution, primary care, HIV and hepatitis C testing and treatment, drop-in center services with access to nutrition and hygiene programs, computer and internet access, holistic health services (e.g., acupuncture, reiki, and meditation), crisis and peer-supported mental health counseling, buprenorphine medication for opioid use disorder and referrals to substance use disorder treatment services, case management and benefits access, and drug checking.³⁴ OnPoint NYC added OPC services to both sites on November 30, 2021.³⁵ The NYC OPCs constitute an additional service within the larger OnPoint NYC program infrastructure.

The NYC OPCs represent a community-based service designed for a highly vulnerable population of people who use drugs. During the first year of OPC operations, 22% of OPC clients were street homeless, with an additional 20% reporting precarious housing conditions.³⁶ The OPCs received a median of 167 daily visits comprised of a median of 108 daily participants.³⁶ Given the volume of visitation and large proportion of vulnerable and street-based individuals using the OPC services, we hypothesized that the community-level impact of the OPC services necessarily will be hyperlocal within the neighborhoods in which they are situated.

The anticipated hyperlocal effect of these interventions presents challenges to impact evaluation. Two prior studies have estimated the neighborhood-level impact of the NYC OPCs.^{31,37} One study estimated an 83% decrease in drug-related arrests, among other outcomes, in the neighborhoods surrounding the OPCs using a difference-in-differences (DID) design³¹; another study estimated a 167% increase in property crimes surrounding the OPCs using a synthetic control design.³⁷ Furthermore, non-peer reviewed reports have suggested that neighborhood patterns of arrest may have changed after the OPCs opened.^{34,38} While generating critical early evidence on the impact of OPCs in the US, these prior empirical studies have important limitations. Specifically, they use intervention catchment areas that do not include areas where OPC clientele are known to congregate

(e.g., a large, adjacent park), as well as buffers that may be too small to account for the surrounding communities' unique characteristics, therefore resulting in residual confounding.

Due to the political controversy surrounding OPCs in the US,³⁹ rigorous evaluation is critical to inform federal, state, and local policymakers as jurisdictions nationally attempt to open OPC services. With scant and uncertain evidence on their community-level impact in the US, the NYC OPCs are an optimal intervention through which to illustrate the key methodological issues that arise when evaluating the community effects of hyperlocal interventions. Here we apply the hyperlocal evaluation framework detailed above to a quantitative case study of the impact of the NYC OPCs on neighborhood-level drug-related arrests. The study is intended to broadly inform applied epidemiologists working to evaluate hyperlocal interventions, as well as researchers evaluating OPCs in the US.

Methods

Data sources

We used two data sources to evaluate the effect of opening the NYC OPCs on drug-related arrests in the surrounding neighborhoods: NYC Police Department (NYPD) arrest records and US Census/American Community Survey (ACS) data. We obtained NYPD de-identified, individual-level arrest records from January 1, 2014 through September 30, 2023 through NYC Open Data.^{40,41} We selected January 1, 2014 as our start data to align with the change in NYC Mayoral Administration, which corresponded to changes in NYPD community policing policy and practices.⁴² We compared drug-related arrests prior to the opening of OPCs (January 1, 2014–November 30, 2021) and following OPC openings (December 1, 2021 – September 30, 2023). Drug-related arrests were defined as arrests assigned a New York State Penal Code charge beginning with 220 or 178, which correspond to drug possession and sales. We compared drug-related arrests pre- and post-OPC opening in the neighborhoods surrounding the OPCs to comparable neighborhoods with SSPs but without OPC exposure. To identify comparable neighborhoods, we used 2019 ACS estimates on neighborhood demographic, social, and economic characteristics from the US Census (Supplemental Table 1).⁴³ This study was deemed exempt by the NYU Langone Health Institutional Review Board because it did not involve human subjects.

Statistical analysis

We used the synthetic control method (SCM) to compare neighborhood-level changes in post-OPC drug-related arrests surrounding the two OPCs, OPEH and OPWH, and their estimated counterfactual comparison groups, after matching on the pre-OPC period. The SCM is well suited for estimating the impact of hyperlocal interventions with few treated units (here, OPEH and OPWH).⁴⁴ In contrast to DID, which requires an assumption of parallel trends across treated and control units,³ SCM constructs synthetic pre- and post-intervention counterfactual units (here, OPEH and synthetic OPEH and OPWH and synthetic OPWH for the pre- and post-OPC periods).⁴⁵ The synthetic units are derived as weighted averages of the outcomes in the untreated or “donor” units.⁴⁶ In our case, our outcome of interest was drug-related arrest and our donor units were those neighborhoods

with SSPs that did open OPCs. We describe our donor unit selection process in greater detail below.

Since validity of the SCM hinges on the plausibility of the synthetic treated unit as a counterfactual for the observed outcomes of the treated unit, we evaluated pre-intervention fit between the synthetic and observed treated units using the root mean square prediction error (RMSPE) of the synthetic control. The RMSPE is a measure of the lack of fit between the path of the outcome variable for our treated units and their synthetic counterparts. As such, we sought to minimize RMSPE. As the scale of RMSPE is specific to a given outcome variable, there are no established values for acceptable amounts of error, although smaller prediction errors are desirable with a caution toward overfitting (e.g., RMSPE values of 0 or approximately 0); however, we considered RMSPE values below 5 to indicate acceptable fit, consistent with prior work.^{44,47} For all models, we estimated the pre-intervention fit between the synthetic and observed treated units with and without ACS covariates. We performed all analyses in R version 4.2.0 (R Project for Statistical Computing), using the *synth* package for synthetic control modeling.⁴⁸

Considerations for evaluation in the hyperlocal context

Given that our evaluation of the NYC OPCs occurred in the hyperlocal context of the NYC overdose crisis, we applied the SCM to the unique challenge of evaluating the community-level effects of a hyperlocal intervention and outline below our stepwise considerations for demonstration purposes.

Spatial zone of impact: First, we carefully considered and selected a catchment area within which we hypothesized a detectable intervention impact (if present). Specifically, we identified the plausible catchment area as closely linked to the primary outcome (here, drug-related arrests). Prior evaluations of OPCs on health- and safety-related outcomes have used a range of spatial zones of impact—including buffers using 500-meter radii,^{27,49} administratively defined boundaries (e.g., postal codes, public health area),^{29,50,51} and non-administrative approaches (e.g., tessellated hexagonal arrays).³¹ While some studies conduct sensitivity analyses (e.g., varying the buffer radii or hexagonal array size), the shape and structure of the catchment area is most often arbitrarily defined.

We modeled the impacts of OPEH and OPWH using both 500-meter buffers and police precinct boundaries, selecting 500-meter buffers as our primary spatial zone of impact. We used 500-meter buffers because this radius has been used in a number of other spatial OPC evaluation studies^{27,49,52} and, in the neighborhoods around the NYC OPCs, is large enough to capture public spaces where OPC clients are known to congregate.⁵³ In the NYC context, 100 meters approximates 1 street block or 0.5 avenue blocks; thus, in the local context, our 500-meter buffer represents an approximately 5-block radius around the OPCs. We hypothesized that a 500-meter buffer would be small enough to plausibly detect an effect of the OPCs, while large enough to facilitate valid modeling and neighborhood coverage of areas where OPC clients gather. For our primary donor pool, we used 500-meter buffers around all static site and mobile SSPs in NYC, excluding any sites with overlap in the 500-meter buffers. This yielded 26 potential donor units for analysis. We used 500-meter

buffers around only static site SSPs as a donor pool for secondary analyses. As SSPs in NYC are not randomly assigned to neighborhoods—nor were the NYC OPCs randomly assigned—our use of all static site and mobile SSPs ensured a broad and diverse enough donor pool to generate well-fitted synthetic controls while minimizing the risk of overfitting.

We selected precincts as a secondary unit of analysis because these align uniformly with law enforcement jurisdictions and are the units at which drug-related arrests are systematically recorded and managed in NYC. In contrast to distance buffers that may straddle multiple precincts, this choice ensures that our primary outcome measure—drug-related arrests—is accurately captured within a single intervention area of influence. While other neighborhood delineations offer alternative ways to define local contexts, they often do not correspond directly with the operational boundaries of police activities. NYPD precincts are large and complex—with over 117,000 residents in Precinct 34, which contains OPWH, and over 47,000 residents in Precinct 25, which contains OPEH⁵⁴—and, as such, their size may obscure a hyperlocal effect. However, precincts allowed us to leverage existing administrative boundaries that directly reflect the operational and reporting structures of the NYPD as a robustness check to our primary analysis leveraging distance buffers. We excluded Precinct 22, which has no permanent residents and includes only Central Park, and Precinct 33, which straddles the precinct in which OPWH is located to avoid spillover effects, from the possible donor pool. We included all other precincts as potential donors to ensure a diverse donor pool, maximize the size of our potential donor pool, and minimize the risk of overfitting in this secondary analysis.

Temporal resolution of impact: To identify the optimal and plausible temporal resolution for this evaluation, we compared several temporal resolutions within the spatial zones of impact described above: daily, weekly, monthly, quarterly, and 7- and 30-day rolling averages. These were selected to reflect potential detectable short-term changes in police activity (e.g., daily or weekly changes), while including longer time horizons (e.g., monthly and quarterly) to assess for potential longer-term changes and reduced noise.⁵⁵ We used rolling averages to smooth data, despite the necessary limitations to interpretability.⁵⁶ Each resolution was assessed for its ability to capture the intervention's effect on drug-related arrests over the given time horizon. We used RMSPE to evaluate the fit of each temporal resolution, aiming to minimize RMSPE while avoiding overfitting. Ultimately, we selected daily counts of arrests as our temporal resolution of interest, as we were able to achieve a suitable pre-intervention fit while balancing interpretability. Substantively, street-level policing can change from day to day, allowing us to estimate granular changes in community policing activity.

Outcome of interest: We selected counts of drug-related arrests as our outcome specification, rather than arrest rates, for several reasons. First, arrest counts provide a direct measure of police activity and intervention impact without requiring population estimates, which can introduce variability and reduce precision, especially in small, hyperlocal contexts. Second, using counts the interpretation of intervention effects, making it easier to track absolute changes in arrests during the pre- and post-intervention periods. Third, arrests represent an outcome for which an appropriate denominator may not be available, as

we are unable to assume that all individuals arrested in a given catchment area are residents of that catchment area. Finally, counts are less susceptible to fluctuations in population size (which occurred during the COVID-19 pandemic in New York City) than rates, which is important in the context of differential neighborhood size and hyperlocal targeting by police. This approach ensures that our analysis accurately reflects the intervention's direct impact on drug-related law enforcement.

Plausible comparison group: Regardless of the method used, causal epidemiological estimation requires a plausible comparison group. In broad terms, this necessitates an identification of relevant confounding structures and their implications in the hyperlocal context. When considering hyperlocal impact evaluation in the US, it is crucial that researchers consider structural confounding due to racial residential segregation,⁵⁷ as has been discussed in detail in the epidemiological neighborhood effects literature,^{58–60} and other manifestations of structural racism.

In our case, our two treated units represent two distinct NYC neighborhoods: East Harlem and Washington Heights. When considering comparison groups for OPEH and OPWH with respect to drug-related arrests, we faced challenges due to the unequal distribution of drug-related arrests across NYC neighborhoods. As displayed in Figure 2, the NYPD precinct containing East Harlem historically has experienced high levels of drug-related arrests relative to the rest of NYC, with the mean number of daily drug-related arrests the highest in this precinct during our study period (January 1, 2014–September 30, 2023). That one of our treated units is an extreme outlier with respect to our outcome of interest reflects the historical legacy of racialized policing—both generally and drug policing—in East Harlem, a historically Black neighborhood.^{61,62} Inextricable from the history of redlining and racialized urban development in NYC,⁶³ the history of racialized over-policing in East Harlem is well documented,⁶⁴ and, despite changes to racialized policing practices during the study period (e.g., procedural curtailing of the NYPD's stop-and-frisk activities at the start of the de Blasio Administration,⁶⁵ the emergence of the Black Lives Matter movement in 2020,⁶⁶ and a growing popular movement to end racialized policing throughout the years 2014–2021⁶⁷), the legacy of structurally racist policies remains measurable in our outcome and functions as structural confounding to any hyperlocal estimation of changes in policing practices in East Harlem.

The method that we used, SCM, rests on the identification of a plausible comparison group constructed from donor units.⁴⁵ Given that the spatial zone of impact inclusive of our treated unit was located in East Harlem—and, thus, necessarily an outlier relative to other potential donor units across NYC—we relied on a model-based approach to identify a plausible comparison unit, as measured by RMSPE to evaluate model fit. For those models using donor units with acceptably low RMSPE values, we relied on substantive interpretation and expertise in collaboration with study community partners to gauge plausibility based on the computed donor pool and prioritized interpretability of estimates where possible. We considered the breadth of the donor pool to ensure that models were not subject to overfitting.⁶⁸

Micro vs. macro impacts: When evaluating hyperlocal interventions, the anticipated impact necessarily is sub-jurisdictional (e.g., community, neighborhood, precinct, or another spatial zone of impact). Given micro-implementation, multiples of the same intervention would be implemented within a single parent jurisdiction. As such, researchers must decide whether to estimate impacts for multiply implemented interventions individually or jointly, effectively isolating or pooling estimated hyperlocal impacts.

In our case, we estimated the impact of two OPCs in NYC separately, sited in distinct and separate neighborhoods. We elected to estimate the impacts separately for several reasons. OPEH and OPWH operate with different staffing models—medically staffed and peer staffed, respectively. Additionally, the neighborhood surrounding OPEH is an extreme outlier along our outcome, and the East Harlem and Washington Heights neighborhoods have distinct sociodemographic profiles.⁶⁹ Estimating the impacts separately furthers our substantive interest in OPCs as a potential intervention with which to build community health equity. This may warrant different approaches due to distinct neighborhood histories, and we considered that any differential estimates between sites would be meaningful.

Practitioner engagement: Finally, we propose that epidemiological researchers actively seek the active involvement of the people and institutions responsible for creating, managing, and acting upon both the hyperlocal program of interest and the data created and used for any given evaluation. In our case, this included consultation with OPC program operators and public health agencies. Consistent with principles of community-based participatory research (CBPR),⁷⁰ such engagement can support more contextually appropriate study design, clarify the operational meaning of key outcomes, and improve interpretation of results. Incorporating practitioner perspectives throughout the research process may also build trust and facilitate triangulation across studies with divergent findings,⁷¹ such as the current study and prior research on the NYC OPCs.

Sensitivity analyses

Upon selection of a model for our main analyses, we conducted a series of sensitivity analyses to compare our main analyses to those from a prior published study evaluating the impact of the NYC OPCs on drug-related arrests.³¹ These including fitting our model using the same time period as a prior study (January 1, 2019–December 31, 2022), as well as using a 200-meter catchment area around the NYC OPCs designed to approximate the approximately two-block radius enclosed in the tessellated hexagon used by the prior study (Supplemental Figure).³¹ As donor units for our SCM in this sensitivity analysis, we used a 200-meter catchment area surrounding both static-site SSPs and static-site and mobile SSPs.

Results

Drug-related arrests

In the 500 meters surrounding OPEH, the average numbers of daily drug-related arrests were 1.96 in the pre-OPC period (January 1, 2014 – November 30, 2021) and 1.57 in the post-OPC period (December 01, 2021 – September 30, 2023). In the 500 meters surrounding

OPWH, the average numbers of daily drug-related arrests were 0.39 in the pre-OPC period and 0.26 in the post-OPC period.

OnPoint-East Harlem

The East Harlem synthetic control was composed of five donor units with nonzero weights. The pre-OPC daily drug arrest was well fitted between the 500-meter buffer surrounding OPEH and its synthetic control (pre-intervention RMSPE = 2.73). Figure 3 displays the daily counts of drug-related arrests in our primary catchment areas during the study period. Following the opening of OPEH, the mean absolute difference in daily drug-related arrests was 0.63 arrests between the synthetic control and the 500 meters surrounding OPEH. Placebo tests (Figure 4) suggested that these findings were not statistically significant, with a permuted P value of 0.19, indicating that the difference in drug-related arrests in the buffer had a higher probability of being due to chance than to the opening of OPEH.

OnPoint-Washington Heights

The Washington Heights synthetic control was composed of four donor units with nonzero weights. The pre-OPC daily drug arrest was well fitted between the 500-meter buffer surrounding OPWH and its synthetic control (pre-intervention RMSPE = 1.05). Figure 3 displays the daily counts of drug-related arrests during the study period. Following the opening of OPWH, the mean absolute difference in daily drug-related arrests was 0.14 between the synthetic control and the 500 meters surrounding OPWH. Placebo tests (Figure 4) suggested that these findings were not statistically significant, with a permuted P value of 0.22, indicating that the difference in drug-related arrests in the precinct had a higher probability of being due to chance than to the opening of OPWH.

Consistency of findings across temporal resolutions and comparison groups

Consistent with our considerations for hyperlocal evaluation, we modeled a series of temporal resolutions and potential donor comparison groups to check the robustness of our findings (Table 3). In addition to 500-meter buffers surrounding storefront and mobile SSPs, we modeled the following spatial zones of impact as donor pools for our synthetic controls: 500-meter buffers surrounding only storefront SSPs in NYC and NYPD police precincts. For each, we modeled the following temporal resolutions of drug-related arrests: weekly, monthly, quarterly, 7-day rolling average, and 30-day rolling averages. Across all three spatial zones of impact—when modeled with and without covariates—weekly, monthly, and quarterly arrests consistently generated poor model fit for OPEH, and monthly and quarterly arrests generated poor model fit for OPWH, as indicated by the pre-intervention RMSPE values displayed in Table 3. For example, the pre-intervention RMSPE values for the weekly, monthly, and quarterly counts of arrests with covariates for OPEH within a 500-meter buffer of NYC storefront SSPs were, respectively, 10.91, 38.42, and 148.03. For OPWH, the pre-intervention RMSPE values for monthly and quarterly arrests with covariates in that same catchment area were, respectively, 7.40 and 12.00. For combinations of spatial zone and temporal resolution with pre-intervention RMSPE values below 5 (indicating acceptable pre-intervention fit), we present substantive estimates in Supplemental Table 2. Across all scenarios modeled with adequate fit, findings were consistent with our main results.

Sensitivity analyses to compare to prior study

Findings from our sensitivity analyses to approximate the assumptions of a prior published study were consistent with our main analyses and indicated no change in arrest after the OPCs (Supplemental Table 3). Models using a shorter study period (January 1, 2019–December 31–2022) and a 200-meter catchment area around static-site and static-site and mobile SSP donor units were well fitted, with pre-intervention RMSPE values below 5 for all models for both OPEH and OPWH.

Discussion

In this study, we proposed a series of considerations for epidemiologists and public health researchers when evaluating the impact of hyperlocal interventions. Key considerations included the plausible catchment area within which to expect a program impact, the plausible time period during which a program impact may be detected, the appropriate specification of a given outcome, the relevant confounding structures and case-specific implications, and the surrounding jurisdictional context within which a program is established. These considerations are broadly applicable across epidemiological and public health research for the evaluation of community health initiatives.

We started by asking: What is the plausible intervention catchment area within which we would expect to detect program impact? Identification of an appropriate intervention unit is crucial to all public health evaluation research,⁷² but we situated ours in the context of hyperlocal interventions. In our case study, we selected a 500-meter buffer around each of the OPCs as the appropriate spatial zone of impact, with police precinct presented as a secondary catchment area. Our use of a primary distance buffer with a secondary administrative buffer was in contrast to the two prior studies estimating the impact of the NYC OPCs on crime, which used only tessellated hexagonal arrays and distance buffers of 500–2000 feet.^{31,37} Our use of multiple buffers was to avoid the arbitrary exclusion of areas in the surrounding neighborhoods where OPC clients are known to spend time³⁶ Notably, prior work using flexible distance buffers to estimate the impact of the NYC OPCs on crime identified different results depending on the buffer selected,³⁷ underscoring the need for researchers to pre-specify the plausible spatial zone of impact. Likewise, our choice to use a 500-meter distance buffer was determined through ongoing engagement with the program operators, who informed our decisions about how to capture the potential movements of their client population through administrative data. We saw the balance struck between hyperlocal programmatic knowledge and rigorous modeling as working within a tradition of community-engaged epidemiological research.⁷³

After identifying an optimal spatial zone of impact, we turned to the temporal resolution and outcome of interest. Here we sought to balance the applied context within which drug-related arrests occur, interpretability of results for clear communication to practitioner partners, and appropriate model fit. The two prior studies of the NYC OPCs used monthly and weekly counts of arrests.^{31,37} In our case, we selected daily counts, because we were unable to achieve suitable model fit for weekly or monthly counts for both treatment precincts. While we achieved suitable fit using rolling averages—consistent with other literature assessing highly variable outcomes⁷⁴—the challenges to interpretation of

these outcomes for our practitioner partners positioned them as suboptimal for our case. Instead, we considered our rolling average models as confirmatory of our primary findings. Likewise, police behavior, which drives arrests, can shift at a daily unit of time,⁷⁵ imbuing confidence in our use of a daily resolution.

Finally, we considered the relevant, hyperlocal confounding structures and, once identified, assessed how they relate to our choice to estimate separate or joint program impacts. In our case, the primary confounding structure was the history of racialized drug policing in East Harlem, which necessarily positioned the neighborhood as an outlier in NYC for drug-related arrests.^{76,77} While the historical disparities in drug enforcement are well-documented, the East Harlem neighborhood's density of social and health services and frequent citizen complaints related to "quality-of-life" issues may also contribute to heightened police presence and enforcement.^{78,79} Our study does not seek to adjudicate these complex and interrelated drivers of disparities in drug enforcement but rather our evaluation framework highlights the need to account for structural and community-driven influences on law enforcement at the hyperlocal level. Given East Harlem's outlier status, we considered this confounding structure as central to our analysis and ability to evaluate the OPC's impact. Thus, we prioritized model fit for East Harlem, interpretability of our findings, and estimated separate, rather than joint, OPC effects.

Comparison with prior studies

To assess why our findings differed from those of a prior study (Chalfin et al.) that estimated an 83% reduction in drug-related arrests in the neighborhoods surrounding the NYC OPCs, we conducted sensitivity analyses to approximate their study period (January 1, 2019–December 31, 2022, compared with our January 1, 2014–September 30, 2023) and catchment area used to identify comparison groups (a 200-meter buffer around SSPs to approximate their roughly two-block tessellated hexagon). Across these analyses, our findings were consistent with our main results: We estimated no difference in the change in neighborhood-level drug-related arrests before and after the NYC OPCs relative to the synthetic control. While we did observe an absolute decrease in the number of average daily arrests between our pre- and post-OPC periods, we did not estimate these findings to be different from the comparison group. This suggests that the differences between our findings and those of the prior study are unlikely to be driven by the study period selected but rather may be determined by pre-intervention trends in the outcome between treated and control groups, as well as the method selected (i.e., SCM vs. DID).

A key difference between our study and a prior evaluation of the impact of the NYC OPCs on drug arrests was our respective selection of comparison groups.³¹ The prior study used a DID approach with buffers surrounding non-OPC SSPs as the primary comparison group, whereas we employed the SCM and selected donor units based on model fit. While DID is a powerful approach when parallel trends can be assumed, its validity depends on the pre-treatment similarity between intervention and comparison groups.⁸⁰ However, SSPs in NYC are not randomly distributed, and many are sited in unique neighborhood contexts with varying levels of drug enforcement, structural policing practices, and community characteristics.⁸¹

The empirical tests of parallel trends that Chalfin et al. presented raise concerns about the validity of this assumption in their study, which the authors acknowledge in the supplement to their published paper. Specifically, they present marked differences in pre-intervention trends for drug-related arrests which reflect, as they state, “considerably higher preintervention levels of drug arrests than comparison sites” (Chalfin et al. Supplement 1, p. 3). Given the higher levels of pre-intervention drug arrests among treated sites, the NYC OPC use case may violate the parallel trends assumption on which the prior study’s DID design hinges, introducing the possibility of residual confounding in their DID estimates.⁸² Our use of the SCM directly addresses this issue by ensuring trend equivalence in the pre-intervention period. This is an especially important consideration when evaluating interventions in neighborhoods with structurally distinct drug enforcement histories, such as East Harlem.

Our use of SCM leveraged a data-driven selection of comparison units, weighting donor units to best approximate pre-intervention trends around the OPCs.⁴⁵ This approach explicitly prioritized comparability in outcome trajectories, reducing confounding bias introduced by unmeasured structural differences across neighborhoods, assuming that neither the OPCs nor other SSPs are randomly assigned to neighborhoods in NYC.⁸³ While SCM in our use case did not impose a strict “apples-to-apples” comparison, it facilitated estimation of the counterfactual by ensuring that synthetic control units shared a similar pre-intervention trajectory in drug-related arrests.⁴⁴ We acknowledge that the prior study’s use of DID aligns with programmatic logic and represents a meaningful alternative approach. However, given the structural disparities between neighborhoods in which OPCs and non-OPC SSPs are located and the challenges in achieving parallel trends across these units, our approach offers an alternative framework that may be particularly useful when policy evaluations are concerned with hyperlocal implementation contexts.⁸³

A further key difference between our approach and that of Chalfin et al. is that we estimated disaggregated impacts for East Harlem and Washington Heights, rather than pooling them into a single treatment group as was done in the prior study. Our choice reflects both substantive and methodological considerations related to the evaluation of micro vs. macro impacts at the hyperlocal level. As we discussed, East Harlem is a structural outlier in drug-related arrests in NYC, with levels exceeding those of all other treated and control neighborhoods in both the present and prior studies. We suggest that pooling the effect of such a structurally distinct site (East Harlem) with another, non-outlier site (Washington Heights) risks obscuring meaningful heterogeneity and potentially introducing bias into estimates of overall intervention impact.⁸⁴ By modeling the two OPCs separately, we sought to preserve internal validity across the two sites and allow for clearer interpretation of hyperlocal effects.

Additionally, it remains possible that our definition of drug-related arrests differed from that the prior study’s; the NYS Penal Codes used to defined drug-related arrests were not reported in the prior publication.³¹ Ultimately however, findings from both this and the prior study estimate *no increase* in drug-related arrests in the time after the NYC OPCs opened. As such, we see our findings as consistent with the prior study’s inasmuch as both studies’ findings offer evidence that the NYC OPCs did not increase drug-related crime—

and concomitantly, drug enforcement measured by arrests—findings that are consistent with studies conducted in other countries.^{32,85}

As epidemiologists expand partnerships with local communities, consideration of the hyperlocal context is an increasingly essential piece of program and policy evaluation. While our framework was applied to a novel drug policy implemented at the neighborhood level in NYC, we see the considerations that we outline as applicable to a broad range of public health programs implemented hyperlocally, including across injury prevention (e.g., community-based violence prevention programs⁸⁶), chronic disease (e.g., community health worker models for care engagement⁸⁷), and infectious disease (e.g., locally implemented COVID-19 responses⁸⁸). In sum, our case study illustrates the need for context, clarity, and practitioner engagement across the full life of the program or policy evaluation to maximize rigor when presented with hyperlocal problems.⁸⁹

Limitations

This study has several limitations. First, although we achieved acceptable preintervention fit between East Harlem and its synthetic control, our East Harlem synthetic control was constructed with a small number of donor units (1-5 donor units), variable by the spatial zone of impact and temporal resolution. However, the unique history of drug policing in East Harlem positioned it as an extreme outlier along our outcome, which presented serious challenges to the identification of a counterfactual in the hyperlocal context. We considered overfitting when assessing models with few donor units, opting not to present models with a single donor unit as our primary result.⁶⁸ However, we acknowledge that this approach does not fully eliminate the risk of overfitting. Future studies should consider incorporating formal validation techniques, such as cross-validation or alternative regularization approaches, to further reduce the risk of overfitting in the context of hyperlocal program evaluation.⁹⁰ Relatedly, while our use of the SCM allowed for the selection of donor units that best approximated preintervention trends, no comparison unit can fully capture the structural and historical factors that shaped enforcement in this neighborhood. Given this, our framework offers a strategy for applied epidemiologists and public health researchers to consider structural confounding in the face of similar challenges to hyperlocal program evaluation. Future research could consider the use of novel methods to outlier detection and synthetic control estimation as strategies to impact evaluation in the presence of structural racism and other structural confounding.⁹¹

Second, the synthetic control method may be subject to residual and unmeasured confounding and, as such, our estimated synthetic control may not provide a good estimation of the counterfactual. However, given that we achieved acceptable pre-intervention fit for both treated sites, it is unlikely that we are subject to large degrees of unmeasured confounding.⁹² Relatedly, the quality of synthetic controls depends on the selection of donor units.⁹³ Given the structural uniqueness of East Harlem as an outlier in drug-related arrests, our pool of comparable donor units was constrained. We mitigated this by ensuring that no single donor unit dominated the synthetic control weights and by conducting sensitivity analyses using alternative donor pools. Likewise, SCM also does not generate traditional confidence intervals but relies on permutation-based inference, which may be sensitive to

the number of available donor units.⁹⁴ To account for this, we used placebo tests to assess statistical significance and evaluated multiple model specifications for consistency. While SCM offers advantages over DID when parallel trends cannot be assumed,⁹⁵ our findings differ from prior DID-based analyses, likely due to differences in pre-intervention trends between treated and comparison units. Despite these limitations, our findings remained stable across multiple sensitivity analyses, reinforcing the robustness of our conclusions.

Third, as a demonstration of our considerations for applied epidemiologists and public health researchers, our case study focused on drug-related arrests and did not assess other policing outcomes relevant to the growing OPC literature (e.g., public order policing, violence-related arrests).³¹ Future evaluation research should consider these and other outcomes (e.g., emergency service calls, community complains for litter) to generate a more comprehensive view of the neighborhood-level impacts of OPCs, and our framework may be useful in the conduct of those studies. Fourth, our outcome was drug-related arrests, which reflect police enforcement against drug-related crimes and serve as an imperfect proxy for drug-related crime. As such, we are unable to interpret our findings as causally related to the impact of OPCs on drug-related crime in neighborhoods following their implementation. Likewise, police maintain considerable discretion related to enforcement during stops,⁹⁶ and we are unable to assess changes in police behavior around the NYC OPCs, as well as policing strategy, discretion, or policy maintained at an institutional level over the study period. As stated above, non-peer-reviewed reports have suggested that neighborhood enforcement practices may have changed after the OPCs opened.³⁸ Likewise, it remains possible that the launch³³ and continued operation of the NYC OPCs may reflect in part an evolving law enforcement posture toward harm reduction in the US,⁹⁷ which may influence the implementation and sustainability of these and other such sites. Future research may consider, where possible, direct indicators of police deployment, discretion, or policy change—including police stops as a prelude to arrest—to further interpret and contextualize the hyperlocal effects of OPCs.

Despite these limitations, given the history of over-policing in NYC—in particular, in East Harlem^{61,62}—and consistently high levels of drug enforcement activity in both neighborhoods during our study period, we posit that it is unlikely that our findings are driven by substantial changes in police activity.

Fifth, our decision to use drug-related arrest counts rather than rates was intended to avoid reliance on potentially unstable denominator estimates, particularly in a hyperlocal context where precise population estimates may be difficult to ascertain. However, this approach assumes a relatively stable underlying population over time,⁹⁸ which we did not explicitly test. As such, it is possible that population shifts in NYC—such as those influenced by COVID-19 or broader neighborhood demographic changes—could have contributed to observed trends.⁹⁹ Future research could explore the impact of population changes on arrest patterns to better contextualize hyperlocal enforcement dynamics.

Sixth, while our study did not detect a statistically significant change in drug-related arrests following OPC implementation, this does not preclude the possibility of a small but meaningful effect. The permuted *P*-values that we estimated (0.19 for OPEH, 0.22 for

OPWH) indicate a potential for Type II error.¹⁰⁰ As such, it is important to caveat that our findings of statistical non-significance do not necessarily imply a definitive absence of effect but rather that any potential change may be smaller than our study was powered to detect, particularly in the context of low arrest counts in our small catchment area.¹⁰⁰ However, our triangulation of findings across donor pools and catchment areas indicates consistency in our results.

Seventh, our selection of our primary and secondary spatial zones of impact were informed by ongoing engagement with program operators to ensure alignment with service delivery and administrative data structures. However, we recognize that this approach does not fully capture the perspectives of OPC practitioners, clients, and other community members who interact with these spaces and their surrounding neighborhoods, nor does it capture the perspectives and input of the creators and stewards of the administrative data that we used (i.e., NYPD). Incorporating insights from people with lived experience—including OPC participants and residents of their surrounding neighborhoods—and relevant data stewards—including NYPD officers and administrators working in the relevant precincts—could further enhance the validity of spatial assumptions made in future crime-related OPC research, as well evaluation research in other hyperlocal contexts, to ensure that studies reflect community-identified realities and priorities.^{101,102} We suggest that future research explicitly engage CBPR principles to avoid a potential “arm’s length approach” to epidemiological research,¹⁰³ ensuring that community-level evaluations are in fact community-based evaluations.

Conclusions

This study discussed a series of considerations for the epidemiological impact evaluation of hyperlocal programs and demonstrated the application of the considerations to a case study of the NYC OPCs. As an evaluation support tool, our considerations have the potential to inform the generation of epidemiological evidence about hyperlocal health programs, which require particular care from evaluators when compared to policy evaluation or patient-level program evaluation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank OnPoint NYC and the NYC Department of Health and Mental Hygiene for their guidance on and contributions to this study.

Funding

This study was supported by the National Institute on Drug Abuse (R01DA058277) and Centers for Disease Control and Prevention (K01CE003586).

Data Availability

Data used for this study are publicly available through NYC Open Data. Analytic code will be published on GitHub upon article acceptance.

References

1. Chandran M, Schulman KA. Racial disparities in healthcare and health. *Health Serv Res.* Apr 2022;57(2):218–222. [PubMed: 35184275]
2. Braveman PA, Cubbin C, Egerter S, Williams DR, Pamuk E. Socioeconomic disparities in health in the United States: what the patterns tell us. *Am J Public Health.* Apr 1 2010;100 Suppl 1(Suppl 1):S186–96. [PubMed: 20147693]
3. Caniglia EC, Murray EJ. Difference-in-Difference in the Time of Cholera: a Gentle Introduction for Epidemiologists. *Curr Epidemiol Rep.* Dec 2020;7(4):203–211. [PubMed: 33791189]
4. Manns BJ, Thomas S, Farinu O, Woolfork M, Walker CL. Hyperlocal lessons from the COVID-19 pandemic: Toward an equity-centered implementation science approach. *Soc Sci Humanit Open.* 2024;9
5. Farley SM, Sacks R, Dannefer R, et al. Evaluation of the New York City Green Carts program. *AIMS Public Health.* 2015;2(4):906–918. [PubMed: 29546140]
6. Releford BJ, Frencher SK Jr., Yancey AK. Health promotion in barbershops: balancing outreach and research in African American communities. *Ethn Dis.* Spring 2010;20(2):185–8.
7. Whitehill JM, Webster DW, Frattaroli S, Parker EM. Interrupting violence: how the CeaseFire Program prevents imminent gun violence through conflict mediation. *J Urban Health.* Feb 2014;91(1):84–95. [PubMed: 23440488]
8. Gómez CA, Kleinman DV, Pronk N, et al. Addressing Health Equity and Social Determinants of Health Through Healthy People 2030. *J Public Health Manag Pract.* Nov-Dec 01 2021;27(Suppl 6):S249–s257. [PubMed: 33729197]
9. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *Bmj.* Sep 29 2008;337:a1655. [PubMed: 18824488]
10. Egan M, McGill E, Penney T, et al. NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. National Institute for Health Research, School for Public Health Research; 2019.
11. Kelly T. Five simple rules for evaluating complex community initiatives. *Community Investments.* 2010;22(1):19–22.
12. Macharia PM, Ray N, Giorgi E, Okiro EA, Snow RW. Defining service catchment areas in low-resource settings. *BMJ Glob Health.* Jul 2021;6(7)
13. Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health.* Nov 2001;91(11):1783–9. [PubMed: 11684601]
14. Berkman LF, Kawachi I, Glymour MM. *Social Epidemiology.* 2nd ed. Oxford University Press; 2014.
15. Pérez E, Braën C, Boyer G, et al. Neighbourhood community life and health: A systematic review of reviews. *Health Place.* Jan 2020;61:102238. [PubMed: 31735517]
16. Arcaya MC, Tucker-Seeley RD, Kim R, Schnake-Mahl A, So M, Subramanian SV. Research on neighborhood effects on health in the United States: A systematic review of study characteristics. *Soc Sci Med.* Nov 2016;168:16–29. [PubMed: 27637089]
17. Levengood TW, Yoon GH, Davoust MJ, et al. Supervised Injection Facilities as Harm Reduction: A Systematic Review. *American journal of preventive medicine.* Nov 2021;61(5):738–749. [PubMed: 34218964]
18. Hedegaard H, Miniño AM, Spencer MR, Warner M. Drug Overdose Deaths in the United States, 1999-2020. *NCHS Data Brief.* Dec 2021;(426):1–8.

19. Mattson CL, Tanz LJ, Quinn K, Kariisa M, Patel P, Davis NL. Trends and Geographic Patterns in Drug and Synthetic Opioid Overdose Deaths - United States, 2013-2019. *MMWR Morb Mortal Wkly Rep.* Feb 12 2021;70(6):202–207. [PubMed: 33571180]
20. Radfar SR, De Jong CAJ, Farhoudian A, et al. Reorganization of Substance Use Treatment and Harm Reduction Services During the COVID-19 Pandemic: A Global Survey. *Front Psychiatry.* 2021;12:639393. [PubMed: 34025471]
21. Cartus AR, Li Y, Macmadu A, et al. Forecasted and Observed Drug Overdose Deaths in the US During the COVID-19 Pandemic in 2020. *JAMA Netw Open.* Mar 1 2022;5(3):e223418. [PubMed: 35311967]
22. Stephenson J. Biden Administration Unveils Overdose Prevention Strategy. *JAMA Health Forum.* Nov 2 2021;2(11):e214252. [PubMed: 36218839]
23. Martin HD, Modi SS, Feldman SS. Barriers and facilitators to PDMP IS Success in the US: A systematic review. *Drug Alcohol Depend.* Feb 1 2021;219:108460. [PubMed: 33387937]
24. Pourtaher E, Payne ER, Fera N, et al. Naloxone administration by law enforcement officers in New York State (2015-2020). *Harm Reduct J.* Sep 19 2022;19(1):102. [PubMed: 36123614]
25. Russell E, Johnson J, Kosinski Z, et al. A scoping review of implementation considerations for harm reduction vending machines. *Harm Reduct J.* Mar 16 2023;20(1):33. [PubMed: 36927354]
26. Yoon GH, Levensgood TW, Davoust MJ, et al. Implementation and sustainability of safe consumption sites: a qualitative systematic review and thematic synthesis. *Harm Reduct J.* Jul 5 2022;19(1):73. [PubMed: 35790994]
27. Marshall BD, Milloy MJ, Wood E, Montaner JS, Kerr T. Reduction in overdose mortality after the opening of North America's first medically supervised safer injecting facility: a retrospective population-based study. *Lancet.* Apr 23 2011;377(9775):1429–37. [PubMed: 21497898]
28. Rammohan I, Gaines T, Scheim A, Bayoumi A, Werb D. Overdose mortality incidence and supervised consumption services in Toronto, Canada: an ecological study and spatial analysis. *Lancet Public Health.* Feb 2024;9(2):e79–e87. [PubMed: 38307685]
29. Salmon AM, van Beek I, Amin J, Kaldor J, Maher L. The impact of a supervised injecting facility on ambulance call-outs in Sydney, Australia. *Addiction.* Apr 2010;105(4):676–83. [PubMed: 20148794]
30. Kennedy MC, Karamouzian M, Kerr T. Public Health and Public Order Outcomes Associated with Supervised Drug Consumption Facilities: a Systematic Review. *Curr HIV/AIDS Rep.* Oct 2017;14(5):161–183. [PubMed: 28875422]
31. Chalfin A, Del Pozo B, Mitre-Becerril D. Overdose Prevention Centers, Crime, and Disorder in New York City. *JAMA Netw Open.* Nov 1 2023;6(11):e2342228. [PubMed: 37955901]
32. Dow-Fleisner SJ, Lomness A, Woolgar L. Impact of safe consumption facilities on individual and community outcomes: A scoping review of the past decade of research. *Emerging Trends in Drugs, Addictions, and Health.* 2022/01/01/ 2022;2:100046.
33. Giglio RE, Mantha S, Harocopos A, et al. The Nation's First Publicly Recognized Overdose Prevention Centers: Lessons Learned in New York City. *J Urban Health.* Apr 2023;100(2):245–254. [PubMed: 37016269]
34. Gibson B, See K, Vargas Estrella B, Rivera S. OnPoint NYC: A baseline report on the operation of the first recognized overdose prevention centers in the United States. *OnPoint NYC;* 2023.
35. Mays JC, Newman A. Nation's First Supervised Drug-Injection Sites Open in New York. *The New York Times.* November 30, 2021:17. <https://www.nytimes.com/2021/11/30/nyregion/supervised-injection-sites-nyc.html>
36. McAteer JM, Mantha S, Gibson BE, et al. NYC's Overdose Prevention Centers: Data from the First Year of Supervised Consumption Services. *NEJM Catalyst.* 2024;5(5):CAT.23.0341.
37. Hall JJ, Ratcliffe JH. Assessing the impact of safe consumption sites on neighborhood crime in New York City: a synthetic control approach. *Journal of Experimental Criminology.* 2024/07/09 2024;
38. Vital City. A Safe Space for Users: A Conversation With Sam Rivera. *Vital City.* December 13, 2023. <https://www.vitalcitynyc.org/articles/sam-rivera-interview>

39. Yang YT, Beletsky L. United States vs Safehouse: The implications of the Philadelphia supervised consumption facility ruling for law and social stigma. *Prev Med.* Jun 2020;135:106070. [PubMed: 32243940]
40. New York City Police Department. Data from: NYPD Arrest Data (Year to Date). 2023.
41. New York City Police Department. Data from: NYPD Arrests Data (Historic). 2023.
42. Inaugural Address of Mayor Bill de Blasio: Progress for New York. 2014. <https://www.nyc.gov/office-of-the-mayor/news/005-14/inaugural-address-mayor-bill-de-blasio-progress-new-york#/0>
43. US Census. Data from: American Community Survey. 2023.
44. Bonander C. Compared with what? Estimating the effects of injury prevention policies using the synthetic control method. *Inj Prev.* Jun 2018;24(Suppl 1):i60–i66. [PubMed: 29127114]
45. Bonander C, Humphreys D, Degli Esposti M. Synthetic Control Methods for the Evaluation of Single-Unit Interventions in Epidemiology: A Tutorial. *Am J Epidemiol.* Dec 1 2021;190(12):2700–2711. [PubMed: 34343240]
46. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association.* 2010/06/01 2010;105(490):493–505.
47. Pickett RM, Hill JL, Cowan SK. The Myths of Synthetic Control: Recommendations for Practice. Working Paper. 2022;
48. synth: Synthetic Control Group Method for Comparative Case Studies. Version 1.1-8. R; 2023.
49. Davidson PJ, Lambdin BH, Browne EN, Wenger LD, Kral AH. Impact of an unsanctioned safe consumption site on criminal activity, 2010-2019. *Drug Alcohol Depend.* Mar 1 2021;220:108521. [PubMed: 33485010]
50. Panagiotoglou D. Evaluating the population-level effects of overdose prevention sites and supervised consumption sites in British Columbia, Canada: Controlled interrupted time series. *PLoS One.* 2022;17(3):e0265665. [PubMed: 35316284]
51. Panagiotoglou D, Lim J. Using synthetic controls to estimate the population-level effects of Ontario's recently implemented overdose prevention sites and consumption and treatment services. *Int J Drug Policy.* Dec 2022;110:103881. [PubMed: 36274565]
52. León C, Cardoso LJP, Johnston S, Mackin S, Bock B, Gaeta JM. Changes in public order after the opening of an overdose monitoring facility for people who inject drugs. *Int J Drug Policy.* Mar 2018;53:90–95. [PubMed: 29294417]
53. Interlandi J. One Year Inside a Radical New Approach to America's Overdose Crisis. *The New York Times.* 22 February 2023. <https://www.nytimes.com/2023/02/22/opinion/drug-crisis-addiction-harm-reduction.html>
54. Prison Policy Initiative. Number of people in prison in 2010 from each NYC police precinct, using precinct boundaries as of 2018. Accessed February 12, 2025. https://www.prisonpolicy.org/origin/ny/police_precincts.html
55. Ashby MPJ. Changes in Police Calls for Service During the Early Months of the 2020 Coronavirus Pandemic. *Policing: A Journal of Policy and Practice.* Jun 25 2020;25(10)
56. Raudys A, Lenčiauskas V, Malčius E. Moving Averages for Financial Data Smoothing. Springer Berlin Heidelberg; 2013:34–45.
57. Kramer MR, Hogue CR. Is Segregation Bad for Your Health? *Epidemiologic reviews.* 2009;31(1):178–194. [PubMed: 19465747]
58. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet.* Apr 8 2017;389(10077):1453–1463. [PubMed: 28402827]
59. Riley AR. Neighborhood Disadvantage, Residential Segregation, and Beyond-Lessons for Studying Structural Racism and Health. *J Racial Ethn Health Disparities.* Apr 2018;5(2):357–365. [PubMed: 28573643]
60. Dyer Z, Alcusky MJ, Galea S, Ash A. Measuring The Enduring Imprint Of Structural Racism On American Neighborhoods. *Health Aff (Millwood).* Oct 2023;42(10):1374–1382. [PubMed: 37782878]
61. Thusi IL. The racialized history of vice policing. *UCLA Law Review.* 2023;69:1576.

62. Williams MB. How the Rockefeller Laws Hit the Streets: Drug Policing and the Politics of State Competence in New York City, 1973–1989. *Modern American History*. 2021;4(1):67–90.
63. Kwate NO, Loh JM, White K, Saldana N. Retail redlining in New York City: racialized access to day-to-day retail resources. *J Urban Health*. Aug 2013;90(4):632–52. [PubMed: 22777683]
64. Vitale AS. *The end of policing*. Verso; 2017.
65. González J. *Reclaiming Gotham: Bill de Blasio and the Movement to End America's Tale of Two Cities*. The New Press; 2017.
66. Taylor K-Y. *From #BlackLivesMatter to Black liberation*. Haymarket Books; 2016.
67. Sklansky DA. POLICE REFORM IN DIVIDED TIMES. *American Journal of Law and Equality*. 2022;2:3–35.
68. Abadie A. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*. 2021;59(2):391–425.
69. NYU Furman Center for Real Estate and Urban Policy. *New York Neighborhood Data Profiles*. NYU Robert F. Wagner School of Public Service, NYU School of Law,. Updated 2024. <https://furmancenter.org/neighborhoods>
70. Leung MW, Yen IH, Minkler M. Community based participatory research: a promising approach for increasing epidemiology's relevance in the 21st century. *International Journal of Epidemiology*. 2004;33(3):499–506. [PubMed: 15155709]
71. Christopher S, Watts V, McCormick AK, Young S. Building and maintaining trust in a community-based participatory research partnership. *Am J Public Health*. Aug 2008;98(8):1398–406. [PubMed: 18556605]
72. Elliott P, Wartenberg D. Spatial Epidemiology: Current Approaches and Future Challenges. *Environmental Health Perspectives*. 2004;112(9):998–1006. [PubMed: 15198920]
73. Sapienza JN, Corbie-Smith G, Keim S, Fleischman AR. Community engagement in epidemiological research. *Ambul Pediatr*. May-Jun 2007;7(3):247–52. [PubMed: 17512886]
74. Reis BY, Pagano M, Mandl KD. Using temporal context to improve biosurveillance. *Proc Natl Acad Sci U S A*. Feb 18 2003;100(4):1961–5. [PubMed: 12574522]
75. Spano R. How does reactivity affect police behavior? Describing and quantifying the impact of reactivity as behavioral change in a large-scale observational study of police. *Journal of Criminal Justice*. 2007/07/01/ 2007;35(4):453–465.
76. Stenkamp A, Rempel M. *Racial and Neighborhood Disparities in New York City Criminal Summons Practices*. Data Collaborative for Justice, John Jay College of Criminal Justice, City University of New York; 2024.
77. Guariglia M. *Police and the Empire City: Race and the Origins of Modern Policing in New York*. Duke University Press; 2023.
78. Kang JX, Levanon Seligson A, Dragan KL. Identifying New York City Neighborhoods at Risk of Being Overlooked for Interventions. *Prev Chronic Dis*. Apr 23 2020;17:E32. [PubMed: 32324532]
79. Hinterland K, Naidoo M, King L, et al. *Manhattan Community District 11: East Harlem*. Vol. 11. 2018. *Community Health Profiles* 2018.
80. O'Neill S, Kreif N, Grieve R, Sutton M, Sekhon JS. Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health services & outcomes research methodology*. 2016;16:1–21. [PubMed: 27340369]
81. McAteer JM, Welch A, Tuazon E, Paone D. *Syringe service programs in New York City*. Vol. 110. New York City Department of Health and Mental Hygiene; 2019. *Epi Data Brief*.
82. Tchetgen Tchetgen EJ, Park C, Richardson DB. Universal Difference-in-Differences for Causal Inference in Epidemiology. *Epidemiology*. Jan 1 2024;35(1):16–22. [PubMed: 38032801]
83. Bottmer L, Imbens GW, Spiess J, Warnick M. A Design-Based Perspective on Synthetic Control Methods. *Journal of Business & Economic Statistics*. 2024/04/02 2024;42(2):762–773.
84. Cintron DW, Gottlieb LM, Hagan E, et al. A quantitative assessment of the frequency and magnitude of heterogeneous treatment effects in studies of the health effects of social policies. *SSM Popul Health*. Jun 2023;22:101352. [PubMed: 36873266]

85. Wood E, Tyndall MW, Lai C, Montaner JS, Kerr T. Impact of a medically supervised safer injecting facility on drug dealing and other drug-related crime. *Subst Abuse Treat Prev Policy*. May 8 2006;1:13. [PubMed: 16722550]
86. Patel J, Leach-Kemon K, Curry G, Naghavi M, Sridhar D. Firearm injury-a preventable public health issue. *Lancet Public Health*. Nov 2022;7(11):e976–e982. [PubMed: 36334611]
87. Crespo R, Christiansen M, Tieman K, Wittberg R. An Emerging Model for Community Health Worker-Based Chronic Care Management for Patients With High Health Care Costs in Rural Appalachia. *Prev Chronic Dis*. Feb 13 2020;17:E13. [PubMed: 32053481]
88. Gopalakrishnan V, Pethe S, Kefayati S, et al. Globally local: Hyper-local modeling for accurate forecast of COVID-19. *Epidemics*. Dec 2021;37:100510. [PubMed: 34688165]
89. Haldane V, Chuah FLH, Srivastava A, et al. Community participation in health services development, implementation, and evaluation: A systematic review of empowerment, health, community, and process outcomes. *PLoS One*. 2019;14(5):e0216112. [PubMed: 31075120]
90. Friedrich S, Groll A, Ickstadt K, et al. Regularization approaches in clinical biostatistics: A review of methods and their applications. *Stat Methods Med Res*. Feb 2023;32(2):425–440. [PubMed: 36384320]
91. Mühlbach NS, Nielsen MS. Tree-based Synthetic Control Methods: Consequences of moving the US Embassy. *arXiv*. 2021;1909.03968v3
92. Abadie A, Diamond A, Hainmueller J. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*. 2015;59(2):495–510.
93. Bouttell J, Craig P, Lewsey J, Robinson M, Popham F. Synthetic control methodology as a tool for evaluating population-level health interventions. *Journal of epidemiology and community health*. Aug 2018;72(8):673–678. [PubMed: 29653993]
94. Firpo S, Possebom V. Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets. *Journal of Causal Inference*. 2018;6(2)
95. Nianogo RA, Benmarhnia T, O'Neill S. A comparison of quasi-experimental methods with data before and after an intervention: an introduction for epidemiologists and a simulation study. *Int J Epidemiol*. Oct 5 2023;52(5):1522–1533. [PubMed: 37023467]
96. Del Pozo B, Sights E, Goulka J, et al. Police discretion in encounters with people who use drugs: operationalizing the theory of planned behavior. *Harm Reduct J*. Dec 16 2021;18(1):132. [PubMed: 34915910]
97. del Pozo B, Rouhani S, Bailey A, et al. The effects of message framing on US police chiefs' support for interventions for opioid use disorder: a randomized survey experiment. *Health & Justice*. 2024/12/19 2024;12(1):50. [PubMed: 39699824]
98. Morrison CN, Rundle AG, Branas CC, Chihuri S, Mehranbod C, Li G. The unknown denominator problem in population studies of disease frequency. *Spat Spatiotemporal Epidemiol*. Nov 2020;35:100361. [PubMed: 33138954]
99. Wong WX, Yadavalli A. NYC's Shifting Population: The Latest Statistics. Office of the New York State Comptroller; 2023.
100. Rothman KJ. Curbing type I and type II errors. *Eur J Epidemiol*. Apr 2010;25(4):223–4. [PubMed: 20232112]
101. Michener L, Cook J, Ahmed SM, Yonas MA, Coyne-Beasley T, Aguilar-Gaxiola S. Aligning the goals of community-engaged research: why and how academic health centers can successfully engage with communities to improve health. *Acad Med*. Mar 2012;87(3):285–91. [PubMed: 22373619]
102. Brothers S, Simon C, Vincent L. Community-Driven Research with People Who Use Drugs: A Virtual Project During Multiple Epidemics. *Sociological Methodology*. 2025;55(1):155–181.
103. Bach M, Jordan S, Hartung S, Santos-Hövenner C, Wright MT. Participatory epidemiology: the contribution of participatory research to epidemiology. *Emerg Themes Epidemiol*. 2017;14:2. [PubMed: 28203262]

Highlights

- We present an epidemiologic framework to evaluate hyperlocal health interventions.
- We applied our framework to evaluate New York City's overdose prevention centers.
- We estimated the centers' impacts on drug arrests using a synthetic approach.
- We found no neighborhood change in drug arrests post-implementation.
- Our framework can guide future evaluations of hyperlocal interventions.

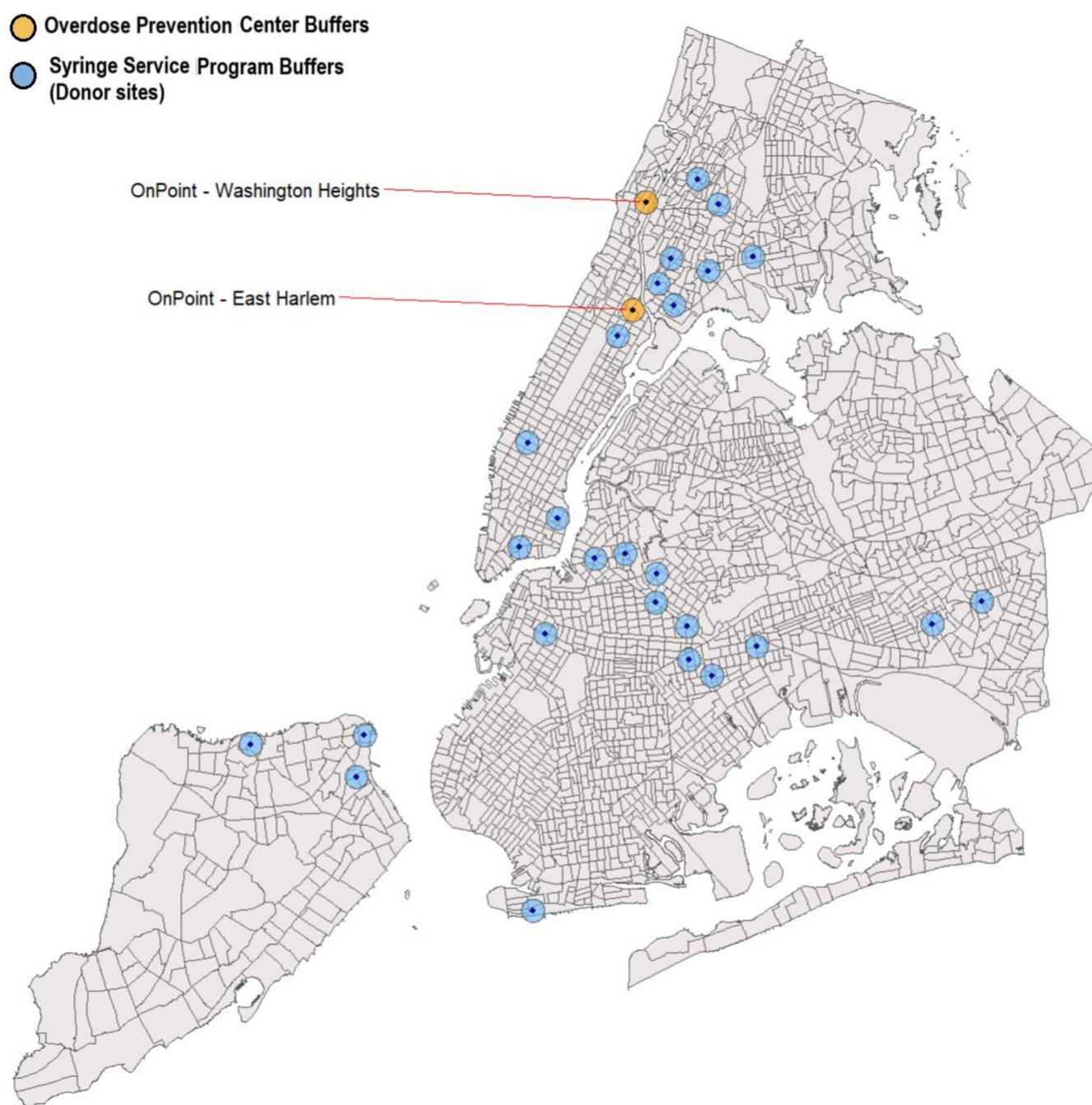


Figure 1.
Location of overdose prevention centers and syringe service programs with 500-meter buffers

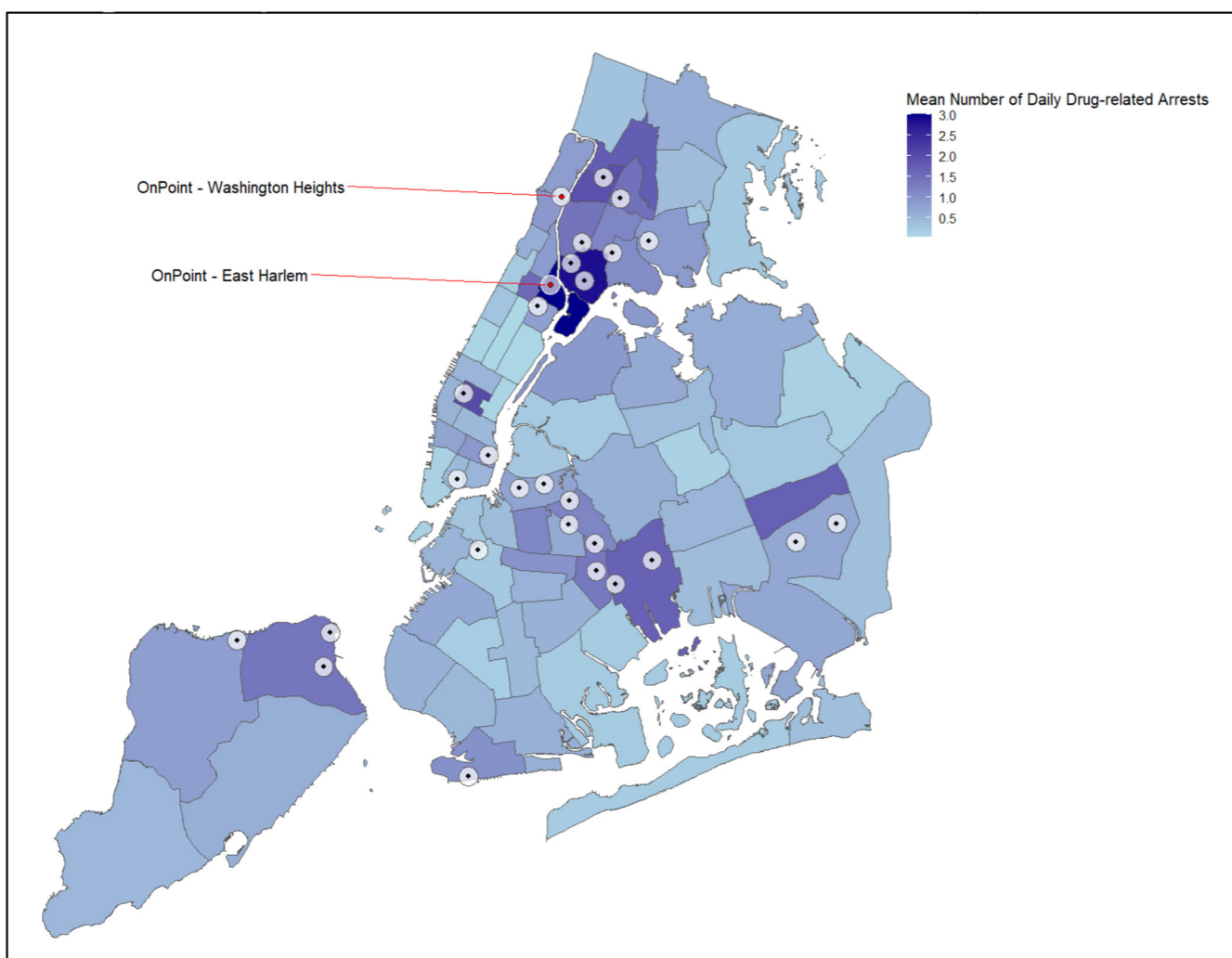


Figure 2. Mean number of daily drug-related arrests by New York City police precincts, January 1, 2014-September 30, 2023

Note: Boundaries denote New York City police precincts; points reflect donor sites with 500-meter buffers

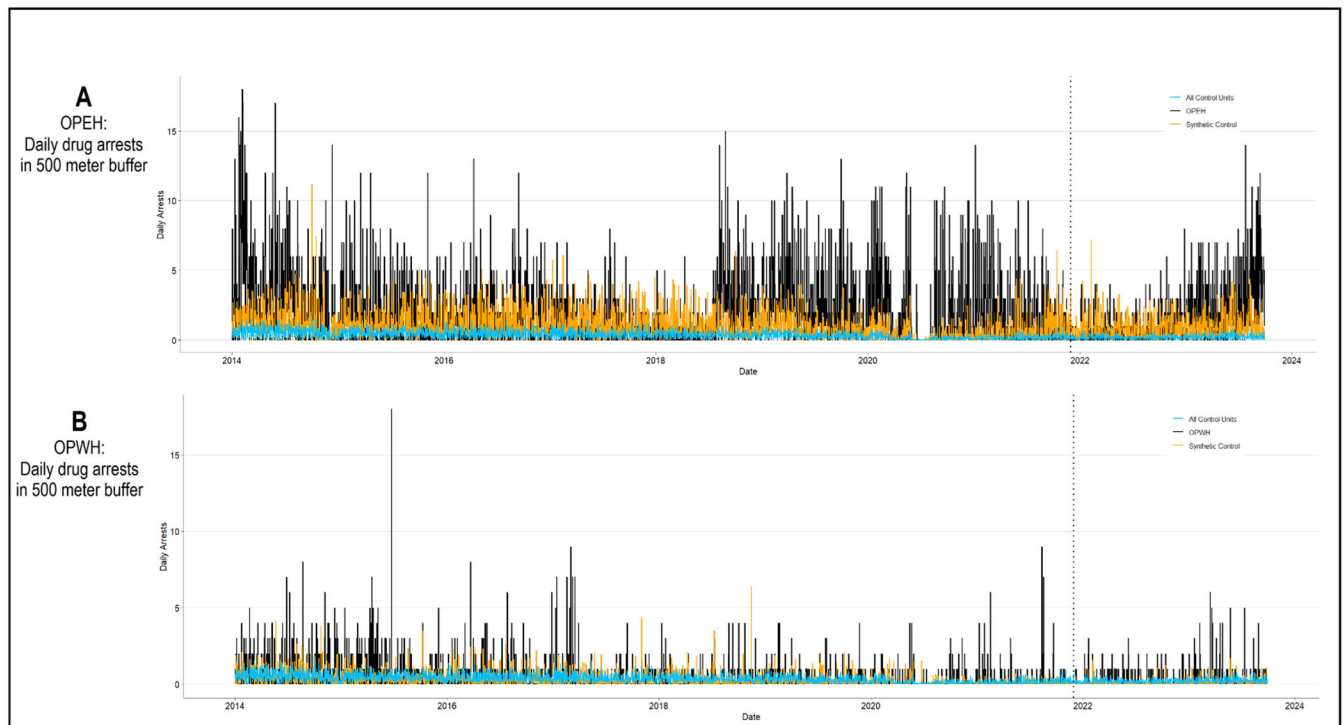


Figure 3. Daily counts of drug-related arrests, 500-meter buffers centered around East Harlem and Washington Heights overdose prevention centers, January 1, 2014-September 30, 2023
Note: OPEH vs. synthetic OPEH (A) and OPWH vs. synthetic OPWH (B), January 1, 2014-September 30, 2023. The vertical lines indicate the opening of the overdose prevention center services. All control units reflect an unweighted average of daily drug arrest counts over the study period, inclusive of all control buffers surrounding syringe service programs.
Abbreviations: OPEH, OnPoint-East Harlem; OPWH, OnPoint-Washington Heights

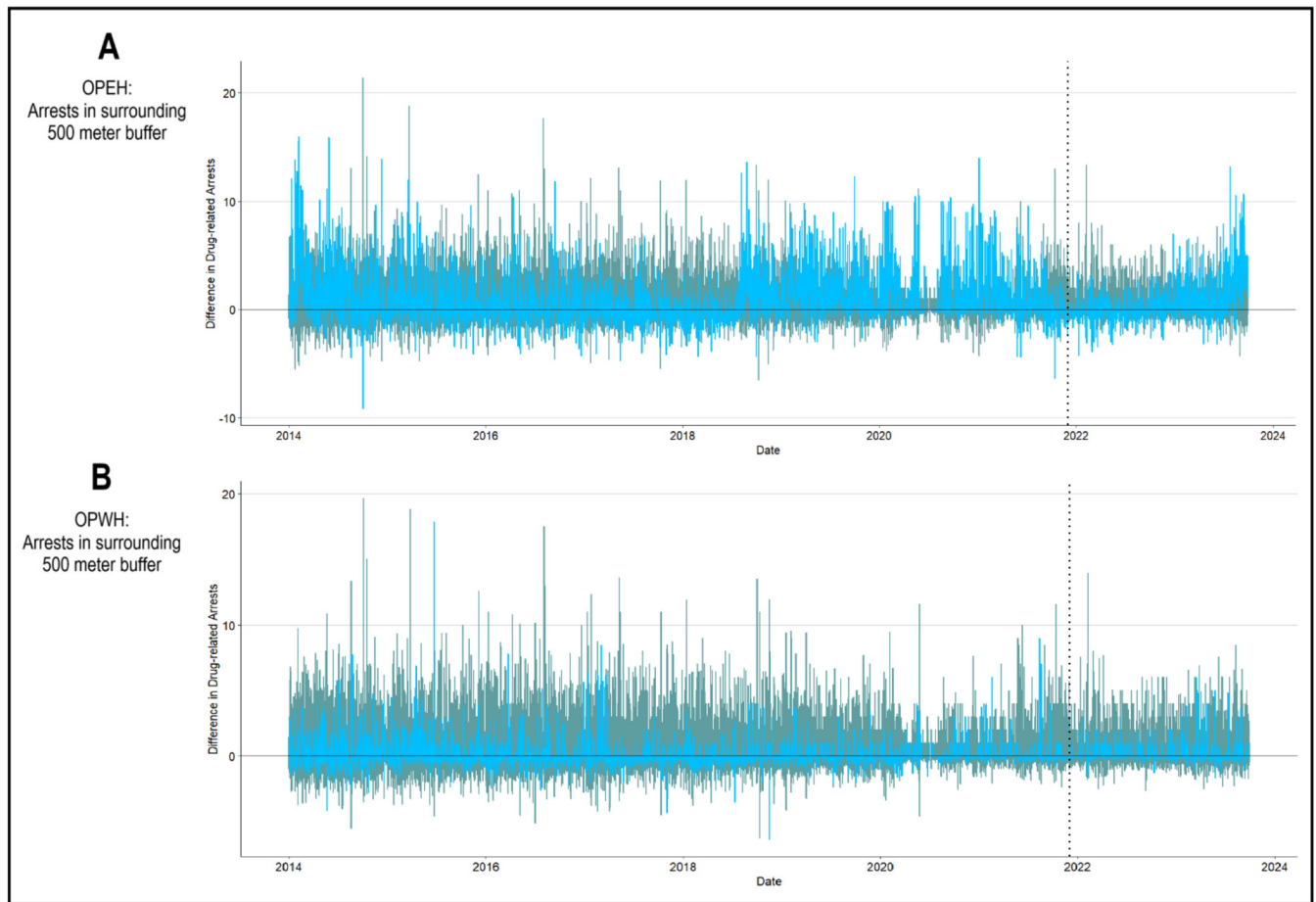


Figure 4. Placebo tests of daily drug-related arrests, 500-meter buffers centered around East Harlem and Washington Heights overdose prevention centers, January 1, 2014-September 30, 2023

Note: Each panel displays the difference between the observed and estimated daily counts of drug-related arrests within 500-meter, circular buffers centered around surrounding OnPoint-East Harlem (A) and OnPoint-Washington Heights (B). The vertical lines indicate the opening of the overdose prevention center services.

Abbreviations: OPEH, OnPoint-East Harlem; OPWH, OnPoint-Washington Heights

Table 1.

Considerations for the evaluation of hyperlocal public health programs using secondary data

Consideration	Guiding Question	Case Application
Spatial zone of impact	What is the plausible intervention catchment area within which we would expect to detect program impact?	Considered distance-based and administrative catchment areas
Temporal resolution of impact	What is the plausible time period during which an intervention impact could be detected?	Considered daily, weekly, monthly, quarterly counts and 7- and 30-day rolling averages
Outcome of interest	What is the appropriate specification of the outcome?	Considered counts versus rates of drug-related arrests
Plausible comparison group	What are the relevant confounding structures and their implications in the hyperlocal context?	Considered neighborhood historical context as related to racialized policing as a structural confounder
Micro vs. macro impacts	What is the surrounding jurisdictional context within which we could estimate individual or joint impacts of multiple, hyperlocal interventions?	Estimated intervention impacts separately due to differences between program operations and surrounding neighborhoods
Practitioner engagement	How have the individuals or institutions responsible for implementing the intervention and generating the relevant data been engaged in the design or interpretation of the evaluation?	Ensure collaborative research design and interpretation with program operators (i.e., OnPoint NYC) and creators and stewards of relevant administrative data (i.e., NYPD)

Table 2.

Difference in daily drug-related arrests between East Harlem and Washington Heights 500-meter buffers and synthetic controls in the postimplementation period, November 30, 2021-September 30, 2023

Variable	East Harlem	Washington Heights
Mean post-OPC difference	0.63	0.14
Pre-RMSPE	2.73	1.05
Post-RMSPE	2.20	0.74
Post/pre-RMSPE ratio	0.81	0.71
Permutation ratio (permuted P)	0.19	0.22

Abbreviations: OPC, overdose prevention center; RMSPE, root mean squared prediction error

Note: Synthetic control donors included all stationary and mobile syringe service programs in New York City

Table 3.

Pre-intervention model fit comparisons

	Root Mean Square Prediction Error	
Outcome	East Harlem	Washington Heights
	<i>Synthetic Control Donors: Stationary Syringe Service Programs (500-meter buffer)</i>	
Daily Arrests	2.99	1.02
<i>With covariates</i>	2.72	1.06
7-day Rolling Average of Daily Arrests	1.59	0.41
<i>With covariates</i>	1.55	0.42
30-day Rolling Average of Daily Arrests	1.30	0.25
<i>With covariates</i>	1.29	0.25
Weekly Arrests	11.17	2.79
<i>With covariates</i>	10.91	2.92
Monthly Arrests	38.67	7.44
<i>With covariates</i>	38.42	7.40
Quarterly Arrests	148.03	18.05
<i>With covariates</i>	148.03	12.00
	<i>Synthetic Control Donors: All Syringe Service Programs (Stationary and Mobile; 500-meter buffer)</i>	
Daily Arrests	2.99	0.99
<i>With covariates</i>	2.73	1.05
7-day Rolling Average of Daily Arrests	1.59	0.39
<i>With covariates</i>	1.56	0.40
30-day Rolling Average of Daily Arrests	1.30	0.24
<i>With covariates</i>	1.29	0.23
Weekly Arrests	11.17	2.68
<i>With covariates</i>	11.00	2.80
Monthly Arrests	38.67	7.16
<i>With covariates</i>	38.61	6.68
Quarterly Arrests	148.03	16.51
<i>With covariates</i>	148.03	13.76
	<i>Synthetic Control Donors: New York City Police Precincts</i>	
Daily Arrests	4.67	1.53
<i>With covariates</i>	4.18	1.56
7-day Rolling Average of Daily Arrests	2.39	0.62
<i>With covariates</i>	2.28	0.61
30-day Rolling Average of Daily Arrests	1.48	0.405
<i>With covariates</i>	1.43	0.34
Weekly Arrests	16.94	4.31
<i>With covariates</i>	16.26	4.15
Monthly Arrests	43.19	12.12

	Root Mean Square Prediction Error	
Outcome	East Harlem	Washington Heights
<i>With covariates</i>	41.60	10.16
Quarterly Arrests	126.14	26.33
<i>With covariates</i>	119.28	14.40

Note: Excluding Precinct 22 (Central Park) and Precinct 33 (precinct Washington Heights location straddles)