A BRIEF ON METHODOLOGY

USING PROXIMITY ANALYSIS TO STUDY THE IMPACT OF SUBSTANCE USE SERVICES ON LOCAL NEIGHBOURHOODS
Purpose
This series provides plain language introductions to an array of methods that are used in research on substance use and health, in order to support policy makers, community members and researchers in critically assessing evidence.

Objectives
- To promote the use of strong methodological approaches that promote equity and fairness in generating evidence on substance use and health.
- To showcase examples from research that illustrate good practices and common pitfalls.

About Co/Lab
Co/Lab is a collaborative network for research and knowledge exchange that aims to promote health and health equity for people who use drugs (including alcohol, other licit, and illicit drugs). Co/Lab activities are guided by collaborations with people who use drugs, families, health care providers, researchers and policy makers, and are focused on generating practical evidence that can be used to enhance substance use services and supporting policies.

Co/Lab is funded by Health Canada’s Substance Use and Addictions Program. The views expressed in this brief are solely those of the authors.

Co/Lab Core Team (listed alphabetically)
Kiffer G. Card, PhD
Brittany Graham, MPH
Lacey Mesley, MPH
Bernie Pauly, RN, PhD
Timothy Stockwell, PhD
Heather Strosher, MA
Karen Urbanoski, PhD

Suggested Citation
Background

Local governments are frequently tasked with making decisions about funding and placement of essential health and social services. These decisions are complicated when the benefits of a service have to be balanced with impacts on the surrounding community. For example, opening a new emergency room at a local hospital could significantly increase the traffic in a neighborhood. Yet, the need for some health services and social supports cannot be overstated due to the demonstrated benefits they have for population health and equity. As such, evaluations of these programs must be equity-oriented and solutions focused to ensure harms are minimized and benefits are maximized. Indeed, even essential and life-saving services and supports — such as hospitals, housing, shelters, and harm reduction services (e.g., safe consumption sites) — can contribute to both positive and negative changes in crime rates, property values, social order, and equity. Local governments should have a clear understanding of these impacts so they can implement optimal policies and protections for those who use services and the surrounding community without reducing access to needed supports. One technique that is commonly used in community impact studies is called **proximity analysis**.

Example Indicators Used in Proximity Analysis

- Average property values for private homes in a neighbourhood over a specified period of time.
- Count of crime events reported to police within 200m of a service, per unit of time.
- Count of improperly discarded needles found in a neighbourhood, per unit of time.
- Count of emergency response calls made by residents of a neighbourhood during a specified period of time.
- Count of overdose events recorded within 500m of a service, per unit of time.

Proximity analysis refers to a set of tools that can be used to assess the geographic impact of selected outcomes. It is often used by health geographers to determine the relationship between distance from a feature (a health or social service) and one or more outcomes (behaviours and events of health or social significance). Because these tools are designed to work specifically with spatial data, they are particularly useful as part of studies that aim to determine the effects of features and outcomes in relation to their physical location.

Proximity analysis is just one of many approaches that can be usefully incorporated into a community impact study. In practice, we recommend that community impact studies include a combination of different approaches, as any single method or technique used on its own will
provide only a part of the picture. With this understanding, we focus here on how to conduct a rigorous proximity analysis, covering general principles, guidance on key analytical decisions, and limitations. In doing so, we will address five fundamental questions that analysts must ask themselves when conducting a proximity analysis:

1. “What outcomes are appropriate to analyze?”
2. “What unit of analysis should be used?”
3. “What time periods should be compared?”
4. “What comparison groups should be considered?”
5. “What statistical methods should be used?”

Following the responses to each of these questions, several case examples are presented to provide some examples of how proximity analyses have been used in prior studies. We conclude with recommendations for undertaking or interpreting results from a proximity analysis, along with some suggestions for further reading.

**What outcomes are appropriate to analyze?**

As in any study, the outcomes examined in a proximity analysis should be important to the communities that are affected by the service under study. This includes potential service users, service providers, community members, and decision makers. Often times, outcomes in a proximity analysis are taken from administrative data on health (e.g., hospitalizations, emergency room visits, overdoses, access to harm reduction supplies, paramedic calls, other health events), law enforcement (e.g., reported and suspected crime events, criminal charges, police calls), or other relevant data sources (e.g., property values, housing prices). While these types of data are not collected specifically for research purposes, they can nonetheless be useful – especially because they allow analysts to compare outcomes across regions. In the remainder of this section, we consider the characteristics that outcomes must share to be usable in a proximity analysis and end with a consideration on the selection and combination of different types of outcomes.

First, as noted above, proximity analysis uses spatial data to assess the geographic distribution of outcomes, relative to a given feature. This means that we need to be able to pinpoint the outcomes to specific physical locations, or geographic points on a map. For example, a crime can be said to occur at a specific address or at an intersection of two streets; a property value can be attributed to a specific home or business; or a paramedic call can be located to a particular address. In some circumstances, it is also possible to estimate or impute a geographic location for an analysis. Continuing our example, a crime that is recorded to have occurred somewhere on the “300 block of Main Street” may be randomly assigned to a more specific location – perhaps at “345 Main Street.” Randomly assigning a location within pre-specified area can help to minimize the bias that may arise.
from purposely selecting a given or known location within that area. Assessing outcomes across multiple geographic areas will allow analysts to make appropriate comparisons.

Second, the data that are used to generate outcomes must be collected in the same way over time. This is important so that the data collected prior to the opening of the service (at baseline) can be legitimately compared to data collected after the service is up and running. Even so, caution must always be taken to ensure that other factors that are correlated with the opening of a service are taken into account. This requires analysts to be consider what else was happening at the time that may affect the outcomes under study. For example, if you were measuring the number of people stopped by a police officer in a neighbourhood, you would want to control for the amount of time that police officers spent in a given neighborhood looking for people to stop. This is why it is recommended to measure outcomes as a rate (e.g., number of stops / hour of policing) rather than as a count (e.g., number of stops). Otherwise, in this example, you would just be documenting changes in policing not impacts of the service.

It is advisable to include an array of outcomes of different types, so that the effects of a service can be examined across different domains or areas of interest. Focusing in on effects related to crime or health alone will only provide part of the picture, which can hamper decision-making that is in the interest of all community members. Likewise, a single outcome domain can be measured in a number of different ways, and examining a number of different outcomes within a specific domain may yield interesting findings. For example, a rate can be calculated for hospitalizations or for hospitalizations for a specific reason (e.g., overdoses, mental health, or injuries); police calls can be counted overall or separately for specific reasons (e.g., property crimes, noise disturbances, domestic disputes, emergency medicine). It can be the case that a service affects different outcomes in different ways. Outcomes should be selected based on theory and prior research in order to generate a comprehensive picture of the community impacts of a service.

Finally, interpretations must be consistent with the data being analysed. For example, if data show an increase in police calls for a specific neighbourhood, there may be a number reasons underlying this finding: it could reflect an increase in crime, or equally, an increase in the need for emergency services. People contact the police for a wide range of reasons. Accurate interpretation of an increase in police calls requires that the reason for the calls is either ascertained or (if that is not possible) that multiple underlying explanations are considered in the interpretation. Interpreting an increase in police calls as simply an increase in crime is an inaccurate interpretation of the data, and might cause undue alarm. Understanding the nuanced nature of the data you are using will allow you to accurately interpret results from a given analysis.
What unit of analysis should be used?

After you have selected outcomes, you then must decide upon the unit of analysis. In a proximity analysis, the unit of analysis refers to geographic distances or areas of a given size. The appropriate unit of analysis depends on the outcome; specifically, on whether the location of the outcome can vary (e.g., police calls, overdoses) or is fixed in space (e.g., property values).

When the outcome varies in location, a common technique is to define a buffer zone (or radius) of a given size and count the number of events that occur with that zone. For example, if the outcome is a rate of crime, you would identify the location of each crime and then count the number of events within the buffer zone over a period of time. The typical size of a buffer zone ranges from 100 to 500 metres (100 metres is about the length of a city block).

When the outcome is fixed in location, you can calculate the distance between each fixed location and the location of the service under study. The distance can be measured using a straight line (“as the crow flies”, known as Euclidian distance) or as a walking path (known as network distance).

A recommended practice is to use sensitivity analyses to assess the impact of the decisions made when selecting a unit of analysis. That is, the analysis should be repeated a few times with buffer zones of different sizes (100m, 200m, 300m, 400m, and 500m) to see the impact that this has on the results. Likewise, you can calculate distances as both a straight line and walking path and compare the results. Doing sensitivity analyses can help to give you a better understanding of how outcomes are distributed geographically, and reduce biases that arise because of the decisions made during analysis.

What time periods should be compared?

Knowing current property values or crime rates within a 200m radius of a health or social service does not actually tell you much about the impact of that facility on its neighbourhood. You need to also compare the outcome from one time period to another. As noted earlier, the time period of interest will span from before the service was opened (baseline period) to some point after the service was opened (follow-up period). As a rule of thumb, a longer follow-up time is better, because some of the initial disruption of a new service will often work itself out over time. It is better to examine the outcome when conditions have reached a state of equilibrium (i.e., when
they are not in the process of changing). If results are needed immediately, it is important to plan to reassess the outcomes with updated data to see if the results change. Also, if your outcome is a rate, such as a count of events per unit of time, then the time periods being compared must match in length in order to be comparable. Finally, it is a good idea to take repeated measures over multiple months or years and calculate the average. This can help to smooth out changes in the outcome, reducing the impact of random error.

What comparison groups should be considered?

Changes over time show only a part of the picture. Indeed, changes in a given outcome can occur for any number of reasons. Take for instance property values: Property values are regularly studied in relation to new services because changes in the quality of a neighborhood are frequently reflected in the price that properties sell for. Therefore, the impact of a new health service can be measured by associated changes on the average sale price for nearby properties. Yet, if you were to compare property values in 2010 to those in 2020, you would almost certainly find that the value of most properties increased regardless of which neighbourhood they were in.

To rule out alternative explanations for changes in an outcome over time, a robust proximity analysis will include a comparison group. Continuing the above example, we can look at whether the observed change in property values over time in one neighbourhood (where the service is located) differs from that in comparable neighbourhoods. This raises the question of how to identify comparable neighborhoods. There are a number of strategies that analysts can use, some of which are reviewed below. In general, though, consider that it is typically the case that we want to compare geographic areas of the same size. This is one reason why proximity analyses commonly use buffer zones to create geographic areas, rather than relying on existing administrative boundaries (e.g., hospital catchment areas, police beats) as these often vary widely in size. Size is only one factor to consider, however, as areas of the same size can still differ in population density, businesses and amenities, and the characteristics of people living in the area – and these differences can the validity of comparing them. For example, it may not be reasonable to expect that the changes seen in an outcome in an affluent neighborhood will be similar to those seen in a lower income neighbourhood. Likewise, a 100m buffer zone in the downtown core of a city will be incomparable to one located in a new suburb on the edge of town. Strategies for
minimizing these sources of bias can be undertaken at the stage of designing a proximity analysis, while others are undertaken at the analysis stage. The latter are discussed in the next section. At the design stage, here are some strategies for improving the comparability of neighbourhoods or other geographic areas.

First, you can compare areas inside and outside of the buffer created around a service. Second, you can sample areas across a city; for example, by randomly selecting points on a map and then sampling data from each of those points or by selecting points at a fixed interval (i.e., creating a grid of sampling units). Third, you can compare an inner and outer buffer to one another. For example, by comparing values of properties within 0 and 200 metres to those between 200 and 400 metres, you might be able to discern whether properties closer to a facility are impacted differently than those that are far away.

With all of these strategies, comparability still rests on whether the characteristics in the area where the service is located are similar to those in other areas (regardless of how those buffer zones are defined and created). For example, if the area where the service is located is not representative of the city overall, then comparing areas inside and outside of the buffer zone created around a service will be misleading. You might observe a significant difference between areas that has little to do with the service under study.

What statistical methods should be used?

As noted in the previous section, strategies to improve the validity of results are undertaken at both the design and analysis stage. In order to conduct a meaningful proximity analysis, researchers must select the appropriate statistical methods that will allow for valid conclusions about changes in selected outcomes over time, and in comparison to appropriately selected neighbourhoods. When possible, matching and multivariable modelling should be used to control for confounding factors.\(^1\) Matching involves identifying characteristics that are directly comparable and using paired statistical tests to conduct the analysis. Multivariable modelling can also be used when data on potential confounders are available. For example, when studying housing prices, it may be useful to control for number of bedrooms, square footage, and the age of the property as well as the conditions of the neighbourhood.\(^2\) Even broader economic trends and policies might impact your result if confounders are not appropriately identified and addressed. In general, the more factors controlled in the analysis, the more useful the results of the analysis will be to helping decision makers understand the effect that the service itself had on
its neighbourhood. A full review of statistical models is beyond the scope of this brief. However, interested readers are directed to difference-in-differences techniques, an economic modeling tool that is commonly used to compare changes over time between neighbourhoods.\(^3,4\)

## Case Studies

Below are a few case studies that highlight some ways that proximity analysis has been used to study the impact of health services and supports on surrounding neighbourhoods. Other case studies also exist.\(^1,2,5-11\) The case studies discussed below were selected because of the important lessons they illustrate, not because they are perfect (or even good) representations of how proximity analysis can be used in community impact assessments.

1. **Doherty et al. (2000)** examined drug related paraphernalia across 32 city blocks randomly sampled over a two-year period before and after the opening of a needle exchange site in Baltimore, Maryland. Each block was sampled 6 times over the two-year period. The authors found that over time the number of needle debris (measured as number of needles per 100 trash items collected) decreased from 2.42 before the program opened to 1.30 2 years later. They found that there was no difference in the number of discarded needles by distance from the program site. This is a strong case example because it compares the prevalence of drug related paraphernalia prior to and after the opening of a needle exchange site, involves multiple periods of data collection, and includes a carefully selected comparison group.

2. **Schilling et al. (2004)** conducted a proximity analysis by recruiting individuals within 10 blocks of a needle exchange program and beyond 10 blocks of a needle exchange program.\(^12\) The main outcome measures of this study were whether participants shared syringes and whether they used condoms. The authors found those within 10 blocks of a needle exchange program were less likely to share needles and more likely to use condoms. This study highlights the potential utility of conducting proximity analyses using person-level data collection rather than administrative data. Novel sampling and methodological strategies such as this, which are increasingly common, may give analysts more control over aspects of the study and avoid potential ecological biases. However, pre-/post-testing could further improve this methodology. For example, Marshall et al. (2011) used cohort data to and compared individuals within 500 meters of a safe consumption site to those more than 500 meters away (but still in the city) for the periods before (2001-2003) and after (2003-2005) the opening of Vancouver’s first safe injection facility. They found that within 500 meters of the facility there was a 35% decrease in overdose deaths compared to 9% in the rest of the city. By using longitudinal data covering the periods before and after the opening of the safe injection site, Marshall’s analysis
provides strong evidence for a decline in overdose. By including a control group, the analysis further shows that the decline is likely, at least in part, attributable to the institution of a safe consumption site.

(3) The Alberta Safe Consumption Site Review Committee evaluated the impact of 7 safe consumption sites in Grand Prairie, Red Deer, Lethbridge, Calgary, and Edmonton. The main outcomes were perceived change in needle debris, number of reported crimes (theft, trespassing, disturbance/nuisance, drugs, and suspicious behaviour), number of emergency response calls, and number of poisoning deaths. Perceived needle debris was measured retrospectively based on subjective perceptions of a convenient sample within three communities. Crime data was reported comparing 2017 to a partially completed year (2018) without any adjustments for changes in policing activities. The comparison units varied from city to city and analysis to analysis. One comparison used a 250-meter buffer and compared it to both the city centre and the rest of the city; one compared a 50 to 500-meter buffer vs. a 500+ meter buffer (inclusive of the rest of the city); a third analysis compared a 500 meter buffer to a 501 – 2000 meter buffer; a fourth analysis used police beats instead of the buffer approach. Only the years (or partial years) immediately prior to or after the opening of these sites were considered. It is unclear why consistent methodologies were not used nor is it apparent that any sound statistical analyses were conducted as part of this review. Given these methodological limitations and the fact these results differ from previous studies, findings from this review are suspect.

Limitations of Proximity Analysis

As demonstrated in the examples above, proximity analysis has broad applications. When used appropriately and treated with rigour, this method is a powerful tool. However, proximity analysis is not without its limitations. For example, proximity analysis frequently relies on administrative data that is collected for non-research purposes. These data sources thus have limitations of their own, including those arising from inconsistencies in the way these data are generated over space and time. For example, a new service might increase the number of emergency response calls due to both an objective increase in the number of calls and a change in the reporting practices of individuals who make these calls. Recognizing limitations such as these, researchers and analysts must carefully conduct and use proximity analyses. As frequently noted throughout this bulletin, sensitivity analyses should be regularly used to assess the extent to which analytic decisions made by researchers might impact the results of an analysis. When it is apparent that results from a proximity analysis are highly variable, we recommend interpreting results with consideration to this variation. Moreover, authors should take special care to ensure that the limitations of their analyses are appropriately communicated to decision makers. In turn, policy makers and practitioners must also be able
to recognize when limitations such as these might explain the results they are observing. In addition to careful exploration and accurate communication of limitations, other forms of community consultation can help contextualize results from a proximity analysis. In conducting these consultations, it is important to involve people with experience utilizing and operating the service being evaluated. In many cases, these individuals have important insights that are not obvious to analysts. Including these insights will ensure that the analysis sufficiently captures the realities underlying the data being used. Involving these communities is consistent with equity-based approaches that address the social determinants of health.16

Discussion

This brief provides an introductory overview of some key design and analytical decisions that need to be made when planning and conducting a proximity analysis as part of a community impact assessment. When used properly, proximity analysis can serve as a highly useful tool for determining the impact of a new health or social service on the surrounding neighbourhood. As with any research tool, when used improperly, it can provide a misleading or biased picture. Our recommendations for analytical decisions are summarized below. In addition to these analytical decisions, we recommend that researchers undertaking a proximity analysis work collaboratively with local communities – including those who use and operate supports and services – to understand the causes for observed results. This is essential to understanding the impacts that a service might have on their local community. By engaging with communities – particularly those who utilize or administer a specific service – analysts will be able to identify ways to improve the services being offered. When negative impacts on the community are identified, community leaders can help identify solutions that can remediate these negative impacts.

Finally, we also recommend that researchers combine proximity analysis with other tools in conducting a comprehensive community impact assessment. The geographic proximity of a service to neighbourhood outcomes reflects only a sliver of the true benefits and costs of a specific service. Results from a proximity analysis should be considered within the broader framework of a multi-criteria cost-benefit analysis that fully captures the impacts of a service on various stakeholder groups (including those who use the service). Interpreting results from different perspectives will help to ensure that decisions made regarding the provision of essential health services are made prudently. After all, results showing that emergency rooms are associated with increased sound pollution or even greater traffic incidents are not sufficient to justify the closure of a hospital. When proximity analysis is included as part of a broader community impact study, it has the potential to provide essential information and assist decision makers in understanding the impacts of services on local neighbourhoods. When treated rigorously, proximity analyses can help policy makers improve the services they deliver and reduce collateral harms to local communities.
Recommendations

Based on the case studies reviewed in this brief and our collective expertise in epidemiological study design, we recommend proximity analysis as a tool for discerning the impact of health services and supports on local neighbourhoods in communities provided the following expectations are met:

1. Impact should be based on meaningful changes that are important to local communities.
2. Outcomes should be readily measurable. They should not rely on the subjective and retrospective recollections of individual community members.
3. Change should be measured over time and appropriate control groups should be utilized.
4. Buffers, with radii between 0 and 1,000 meters, should be used in lieu of administrative units.
5. Changes in outcomes overtime should be measured by comparing (a) the time prior to any speculation that a new service is to be opened and (b) a period of several years following the initial opening of the facility. This will ensure outcomes are being measured in a state of equilibrium (i.e., no changes are occurring).
6. Control neighbourhoods should be carefully selected with consideration of the unique characteristics of the neighborhoods being studied. This often involves utilizing administrative data and working with local communities to help inform optimal sampling strategies.
7. Observations should capture enough data as to provide information regarding the natural variability of the outcome across time (i.e., year to year changes in rates) and space. Statistical tests should be appropriately utilized to account for this natural variability.
8. Sensitivity analyses should be conducted and reported to verify the analytic choices made by the research team.
9. Interpretations of results should be grounded in the experiences of local community members and service administrators. Doing so will allow study authors to identify alternative causal explanations (e.g., changes in policing behaviour, external factors impacting home prices).
References and Further Readings


