New Uses of Wastewater Data on High-Risk Substances: Emerging Drug Detection, Overdose Prediction, and Drug Policy Evaluation

Key findings
Based on an analysis of up to 39 weeks of wastewater data on drug analytes and suspected drug overdose events from five counties across four U.S. states, several findings emerged:

- Xylazine concentrations were detectable in wastewater at least one month before two suspected overdoses involving xylazine were reported, confirming that wastewater data can proactively inform harm reduction strategies for emerging drug threats.
- A new Drug-SURGE wastewater alert algorithm correctly flagged 71 to 100% of drug overdoses based on unprecedented patterns in wastewater data, providing early warnings of eight to 11 days that enable officials to act in real time to prevent drug overdoses.
- In the six months following the decriminalization of fentanyl test strips, wastewater concentrations of fentanyl gradually decreased in two Kentucky counties, providing early evidence of the promise of wastewater data to evaluate the effectiveness of new drug policies and programs.
- Wastewater data helped fill gaps in drug overdose data—particularly in rural communities, where data and resources can be sparse—and could shed light on patterns of polysubstance use.

Authors
Aparna Keshaviah, Alex Buben, Jana Epstein, Allison Burrell, and Xindi C. Hu
can further its use, particularly by policymakers. For example, Mathematica’s Covid-SURGE (Signaling Unprecedented Rises in Groupwide Exposure) alert algorithm distinguishes signal from noise in wastewater data to flag community-level disease outbreaks, giving public health officials a timely way to assess the need for heightened pandemic response (Keshaviah et al. 2023).

Building on this prior research, our collaborative team sought to further characterize how wastewater monitoring could help communities adapt to the changing landscape of substance use and prevent overdoses. To do so, Mathematica integrated wastewater data from Biobot on four high-risk substances—fentanyl, methamphetamine, cocaine, and xylazine—with Overdose Detection Mapping Application Program (ODMAP) data from the Washington/Baltimore High Intensity Drug Trafficking Area (W/B HIDTA) program on suspected drug overdoses. We analyzed up to 39 weeks of data on rural, suburban, and urban counties in California, Kentucky, New Jersey, and West Virginia. We selected five counties in these states based on their high prevalence of substance use, availability of longitudinal wastewater data that covered a sizable share of the county’s population, and high-quality reporting of drug overdoses through a statewide ODMAP application programming interface.

Next, we summarize findings from our analysis of wastewater and drug overdose data to assess the ability of wastewater data to detect an emerging drug threat (xylazine), predict suspected drug overdoses, and evaluate the effectiveness of a new drug policy (fentanyl test strip decriminalization).

**Detecting emerging drug threats**

The rising prevalence of xylazine in the illicit drug supply has complicated the public health response to the overdose epidemic. When a person consumes opioids with xylazine, overdose reversal agents such as naloxone have diminished efficacy because they do not counter the sedative effects of xylazine. Further, xylazine use causes skin ulcers and necrotizing wounds, both at the site of injection and elsewhere on the body, and can lead to physical dependence on the drug. Accordingly, having more information on the extent of fentanyl and xylazine co-use could better equip officials to respond to overdoses.

Clinical and wastewater data show xylazine use currently concentrates mainly in eastern parts of the country. A recent study in Kentucky found that 55% of wastewater samples collected from interstate rest areas and weigh stations tested positive for xylazine in the first half of 2021, and by the first half of 2023, the proportion increased to 94%, underscoring the potential for xylazine use to spread rapidly (Delcher et al. 2024). To stem further proliferation of this dangerous drug, officials need to quickly detect when xylazine has entered their community. To this end, we examined the potential of wastewater to serve as an early warning for the presence of xylazine in communities. We looked at trends in wastewater concentrations of xylazine in two counties (one in New Jersey and one in West Virginia) that had a suspected overdose involving xylazine, per ODMAP.

We found that xylazine was detectable in wastewater at least four weeks before the first (and only) xylazine overdose was reported in each county (Exhibit 1). In New Jersey, wastewater concentrations of xylazine increased by 38% three days before the recorded overdose; in West Virginia, the overdose was reported in the middle of a slow but steady rise in xylazine wastewater concentrations, which amounted to a 5 to 10% increase per week over four weeks. Interestingly, across both sites, xylazine concentrations in wastewater fell soon after the report of the drug overdose (though in West Virginia, xylazine levels rebounded and even reached a new high in the subsequent weeks).

These early warnings from wastewater data have notable implications for response. Wastewater data provide an opportunity to alert community members, public health officials, and medical personnel to emerging threats weeks before overdose toxicology reports and other established data sources capture them. Further, communities can deploy mobile wound care, adapt training curricula for first responders and community members, and communicate about the emergence of xylazine in real time to people who use drugs. In this manner, communities could use wastewater data to develop proactive intervention strategies that mitigate harms before these novel substances lead to overdoses.

Exhibit 1. Trends in xylazine concentrations before and after reported xylazine overdoses

<table>
<thead>
<tr>
<th>Sampling date</th>
<th>Estimated drug consumption (mg/1,000 people)</th>
<th>NJ: Urban county</th>
<th>WV: Rural county</th>
</tr>
</thead>
<tbody>
<tr>
<td>2023-08-18</td>
<td>0</td>
<td>Xylazine overdose</td>
<td>Xylazine overdose</td>
</tr>
<tr>
<td>2023-08-25</td>
<td>2</td>
<td>Parent drug (xylazine)</td>
<td>Parent drug (xylazine)</td>
</tr>
<tr>
<td>2023-09-01</td>
<td>4</td>
<td>Metabolite (4-hydroxy xylazine)</td>
<td>Metabolite (4-hydroxy xylazine)</td>
</tr>
<tr>
<td>2023-09-15</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The unfilled circles with a value of 0 denote wastewater concentrations that were below the limit of detection.
The benefits of wastewater-based epidemiology are not limited to illicit substances. For example, naloxone concentrations in wastewater could help officials estimate the underreporting of overdoses in data from EMS, law enforcement, and healthcare providers (Endo et al. 2020). A community could use wastewater data on methadone and other medications for substance use to assess treatment adherence, particularly as states provide patients the freedom to take prescriptions outside supervised clinic settings. Ultimately, monitoring wastewater for these compounds in tandem with high-risk substances can yield a more nuanced picture of substance use and the success of interventions.

### Predicting drug overdoses

One of the challenges of using wastewater data for decision-making is that sustained rises in wastewater analyte concentrations can be difficult to distinguish from routine sample-to-sample fluctuations (arising from changes in the flow rate, size of the contributing population, weather, measurement error, and other factors) in real time. For example, our analysis of trends in SARS-CoV-2 wastewater concentrations around major COVID-19 surges in North Carolina showed that none of the commonly reported wastewater metrics (the magnitude of viral concentrations, percentage change in those concentrations, or whether concentrations were detectable), on their own consistently flagged the start of new disease outbreaks (Keshaviah et al. 2023).

To overcome this challenge, Mathematica developed the Covid-SURGE alert algorithm, which logically combines multiple wastewater metrics to reliably flag unprecedented signals in wastewater viral concentrations and empower officials to make better use of wastewater data for decision-making. Based on findings from a multistate validation study, the alert algorithm has better performance in small and large sites alike (Keshaviah et al. 2023).

Building on that work, we sought to develop an analog to the Covid-SURGE alert algorithm, called Drug-SURGE, to flag unprecedented signals in wastewater concentrations of fentanyl, methamphetamine, cocaine, and xylazine in real time. We adapted the logic, thresholds, and lookback period used for Covid-SURGE to suit the patterns we observed in wastewater parent drug and metabolite concentrations leading up to suspected drug overdoses (Exhibit 2). Though not all substance use results in an overdose, we hypothesized that unprecedented rises in drug analyte concentrations in wastewater might signal new substance use patterns and an increased likelihood of one or more overdoses in the community.

### Exhibit 2. Trends in wastewater concentrations of drug analytes and suspected drug overdoses

<table>
<thead>
<tr>
<th>Drug Analyte</th>
<th>County</th>
<th>Report Type</th>
<th>Concentration (mg/1000 people)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methamphetamine</td>
<td>KY: Urban county</td>
<td>Wastewater</td>
<td>12000</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>KY: Rural county</td>
<td>Drug overdose</td>
<td>6000</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>NJ: Urban county</td>
<td>Drug overdose</td>
<td>9000</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>WV: Rural county</td>
<td>Drug overdose</td>
<td>12000</td>
</tr>
</tbody>
</table>

Note: When more than one wastewater treatment plant was sampled in a county, the data from the wastewater treatment plant with the largest service population size was graphed.
We assessed the accuracy of the Drug-SURGE algorithm by comparing the timing of alerts raised with the timing of suspected drug overdoses in each county. We measured sensitivity (true positive rate) by calculating the share of alerts triggered that were followed by a reported overdose event within the next 14 days. We could not assess specificity (true negative rate) because ODMAP does not capture all overdoses (missing, for example, overdoses successfully reversed at home with no resulting EMS call or hospitalization). Without the ability to balance the true positive rate with the false positive rate (keeping the former high and the latter low), the Drug-SURGE algorithm could yield a high sensitivity by simply producing frequent alerts. To avoid this scenario, we calibrated the algorithm to ensure that the overall alert rate—the proportion of days on which we detected unprecedented signals in wastewater—was consistent with an (unobserved) true drug overdose rate, which we approximated using ODMAP data and the literature (Exhibit 3 provides details).

Based on our analyses of five differently sized counties, we found the Drug-SURGE alert algorithm had high sensitivity to predict overdoses due to different drugs (Exhibit 4). Sensitivity was 81% overall and ranged from 71% for cocaine overdoses to 100% for xylazine overdoses (though the latter relied on only two reported overdoses). Among those overdoses that Drug-SURGE flagged correctly, the early warning time provided averaged 8.7 to 11.0 days. Notably, the algorithm had strong performance even though correlations between the wastewater and suspected overdose events were weak to moderate at best (averaging r = 0.36 across sites and drugs), showcasing how advanced analytics can tease out patterns in the data that might not be obvious.

The findings presented here are based on only five counties, and further validation is needed to establish the generalizability of the Drug-SURGE algorithm in other sites and states. However, our results were robust across a range of sensitivity analyses we conducted to examine the impact of analyzing raw versus normalized wastewater concentrations; parent drugs, metabolites, or a combination of the two; and overdoses due to fentanyl versus any opioid. In looking into why the algorithm failed to flag some overdoses, we found many of the overdoses missed had only one wastewater sample collected in the prior 14 days, and thus little opportunity to trigger an alert. When we repeated the analyses

Exhibit 3. Validating the Drug-SURGE algorithm in the absence of a gold standard

ODMAP is one of the timeliest sources of data on suspected overdose events, but it does not represent the full universe of drug overdoses in a community. Research shows that a first responder might never attend to one-third to one-half of opioid overdoses, which thus may not appear in ODMAP (University of Missouri-St. Louis—Missouri Institute of Mental Health n.d.; Pan 2020). These gaps posed challenges because the missing data precluded us from calculating the specificity of the Drug-SURGE algorithm. Consequently, we could not be sure that the raising of an alert when there was no suspected overdose recorded in ODMAP represented a false positive (versus an unrecorded true positive). Likewise, when no alert was raised and no suspected overdose was reported soon thereafter, we could not be certain the lack of alert represented a true negative.

Instead, we calibrated by the algorithm by balancing the sensitivity with the overall alert rate (which included true and false positives). To benchmark this overall alert rate, we imputed the true (unobserved) drug overdose rate in each county. To do so, we adjusted ODMAP estimates of suspected overdoses using a three-step imputation process to account for (1) missing data, because only 2 to 55% of ODMAP overdoses in the counties analyzed had information on suspected substances involved (per Exhibit 7); (2) underreporting of fentanyl overdoses in ODMAP (for example, if overdoses are successfully reversed at home and no call is placed to law enforcement or EMS); and (3) co-use of fentanyl with cocaine or methamphetamine, drawing on estimates from the literature (O’Donnell et al. 2020, Twillman et al. 2020). We had to adjust for co-use because of potential biases in assessing the reported drug involved when no toxicology report was available. When they lack information on the drug(s) ingested, first responders use the success of overdose reversal agents as a proxy. That is, if naloxone successfully revived the person who overdosed, first responders record that fentanyl or another opioid was involved in the overdose. This approach can produce accurate estimates of opioid involvement, but likely leads to underreporting of drugs used in combination with fentanyl or another opioid (such as cocaine and methamphetamine). After we adjusted ODMAP estimates to account for unreported overdoses, we converted the imputed numbers of overdoses into rates (the proportion of days with an overdose), which we used to benchmark the Drug-SURGE alert rate.

Exhibit 4. Performance of the Drug-SURGE alert algorithm

<table>
<thead>
<tr>
<th>High-risk substance</th>
<th>Suspected overdose events reported</th>
<th>Average true positive rate [range by site]</th>
<th>Average lead time [range by site]</th>
<th>Proportion of days with an alert triggered [range by site]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cocaine</td>
<td>73</td>
<td>71% [65–100%]</td>
<td>8.7d [5.0–10.5]</td>
<td>45% [33–67%]</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>28</td>
<td>86% [63–100%]</td>
<td>10.1d [3.5–11.0]</td>
<td>36% [24–69%]</td>
</tr>
<tr>
<td>Fentanyl</td>
<td>129</td>
<td>86% [56–100%]</td>
<td>9.4d [5.8–10.4]</td>
<td>55% [45–73%]</td>
</tr>
<tr>
<td>Xylazine</td>
<td>2</td>
<td>100% [n.a.]</td>
<td>11d [n.a.]</td>
<td>6% [0–19%]</td>
</tr>
</tbody>
</table>

Note: We calculated the average lead time for the subset of overdoses correctly flagged by the algorithm. d = days; n.a. = not applicable (both xylazine overdose events were correctly flagged, so there was no variability by site).
after restricting the data to suspected overdoses with at least two wastewater samples collected in the prior two weeks, the overall true positive rate increased from 81 to 87%, pointing to the importance of routine monitoring.

Lastly, we noted that the urban counties we analyzed saw one or more drug overdoses every day during the analysis period. In such communities, a combination of upstream sampling and predictive modeling to forecast the number of overdoses (versus whether an overdose will occur, which is what Drug-SURGE yields) could help officials to allocate resources adaptively.

**Evaluating the effectiveness of a new drug policy**

When resources are scarce, having evidence on the relative effectiveness of different harm-reduction strategies—such as syringe service programs, fentanyl and xylazine test strips, and naloxone vending machines—can help communities decide which services to expand and how to effectively use opioid abatement settlement funds.

Quasi-experimental impact evaluations provide a robust framework to measure program effectiveness in a real-world setting. For example, using a difference-in-differences design, we can compare changes in pre- versus post-program trends in outcomes between the intervention community and one or more matched comparison communities (which enables us to isolate the effects of the program or policy of interest from external confounding factors and changes associated with the passage of time). However, acquiring data for impact evaluations can be challenging—collecting primary data at the individual level (such as through population surveys) can be costly, while secondary data can be out of date (inhibiting timely insights) or sparse (precluding the ability to measure subtle changes).

Wastewater data offer a more cost-effective option for impact evaluations, yielding granular and timely information that enables communities to customize evaluation designs (for example, to measure impacts overall, but also by neighborhood and drug type). Because quasi-experimental evaluations have not been conducted using wastewater data, we explored whether we could use such data to evaluate a drug policy change. Specifically, we assessed whether wastewater data could capture changes in fentanyl use after Kentucky decriminalized fentanyl test strips on March 31, 2023. We analyzed trends in wastewater concentrations of fentanyl (parent drug plus metabolite) before and after the policy change in two Kentucky counties. In the urban county, Biobot had tested wastewater for fentanyl for a period of one month before the policy became law and roughly five months after the policy change. In the rural county, Biobot had three months of pre-policy data and six months of post-policy data. To compare pre- versus post-policy trends, we ran piecewise linear regression models for each county, with random intercepts at the wastewater treatment plant level and a single knot on the date of the policy change. We tested alternative model specifications (including nested hierarchical effects, random slopes, and two knots to account for delays in policy impacts) and chose the model specification that yielded the best fit to the data, based on the lowest Akaike information criterion.

In both counties, fentanyl concentrations in wastewater were rising before the policy change (slope = 0.013 in the urban county, 0.002 in the rural county); the increase, was more pronounced (and statistically significant, with \( p = 0.037 \)) in the urban county (Exhibit 5). After the decriminalization of fentanyl test strips, fentanyl concentrations in wastewater decreased (slope = -0.001 in both counties), and in the urban county, the difference in the pre- versus post-policy slope was statistically significant (\( p = 0.033 \)). Sensitivity analyses based on the parent drug or metabolite only (versus the sum of the two) left our findings relatively unchanged.

**Exhibit 5. Trends in fentanyl concentrations in wastewater before and after decriminalization of fentanyl test strips**

Note: For each county shown, the data come from the wastewater treatment plant with the largest service population size.
New Uses of Wastewater Data on High-Risk Substances

Running these same exploratory models using ODMAP data on suspected drug overdose deaths due to fentanyl (rather than wastewater concentrations of fentanyl) yielded some interesting findings. In the urban county, overdose counts rose before test strip decriminalization and fell thereafter, and the change was statistically significant ($p < 0.001$). However, we could not use ODMAP data to assess the effectiveness of test strip decriminalization in the rural county because of sparse data—there was only one suspected overdose due to fentanyl in the three months before the policy change and only two in the seven months after the change. In other words, wastewater data helped fill an evidence gap that often arises in rural counties.

The exploratory analyses described here are just the first step in assessing how exactly to harness wastewater data for program evaluations. A follow-on study to establish what types of wastewater metrics to use to match intervention to control communities can further extend the use of this novel public health data source.

Data sources

Wastewater data

We analyzed 23 to 39 weeks of longitudinal wastewater data from five counties of different sizes across four U.S. states (Exhibit 6). Biobot partnered with utilities to collect untreated wastewater from centralized wastewater treatment plants that serve a portion of the county’s residents. These partners collected samples once or twice weekly using 24-hour composite sampling, after grit removal and bar screen processes.

With wastewater monitoring for substance use, the sample collection day can affect data interpretation. For example, samples collected Tuesday through Friday might reflect more habitual use, whereas those collected Saturday through Monday might reflect more recreational use. The data included in the current analysis came from samples typically collected Tuesday through Friday (with the specific days chosen based on program needs); however, in the New Jersey and California counties, some samples were collected on Mondays, potentially reflecting higher levels of recreational substance use.

After utility partners shipped the wastewater samples to the lab, Biobot used sensitive laboratory assays and mass spectrometry to measure the concentrations of high-risk substances (and their most common metabolites)—including fentanyl (norfentanyl), methamphetamine (amphetamine), cocaine (benzoylcegonine), and xylazine (4-hydroxyxylazine). To account for sample-to-sample fluctuations in the contributing population size and improve the comparability of data across sites, wastewater concentrations were population- and flow-normalized. Multiplying the measured concentrations by the sampling day’s flow rate and dividing the results by the wastewater treatment plant’s service population size yields estimated consumption levels in the units of milligrams per 1,000 people per day, which is what we analyzed.

ODMAP data

ODMAP is a free, web-based, mobile-friendly platform first developed by W/B HIDTA in 2016. The platform enables near-real-time reporting and surveillance of suspected fatal and nonfatal overdose events by licensed first responders, public safety agencies (such as medical examiners and coroners), emergency departments, and public health agencies. ODMAP displays overdose data within and across jurisdictions, helping agencies identify spikes and clusters of suspected overdose events in their community, neighboring communities, and across the country. As of March 2024, ODMAP includes nearly 33,400 users representing more than 5,000 agencies across 50 states, the District of Columbia, and Puerto Rico.

ODMAP relies on contributing agencies to confirm the validity of each overdose entered and requires agencies to submit four pieces of information for each suspected overdose event: date and time, location, fatality status, and naloxone administration status. ODMAP does not require agencies to submit information on the suspected drug(s) involved, but if they do, agencies use drop-down menus listing common substances of abuse, with the option to enter free-text descriptions into an “Other” field. In mid-2023, ODMAP added xylazine to the drop-down menu of substances.

Across the five counties we analyzed, an average of 3.5 overdoses per 1,000 people were reported in ODMAP during the analysis period (with a range of 0.9 to 4.8 overdoses per 1,000 people by county). Between 45 and 98% of overdoses had missing information on the suspected drug involved; among overdoses with drug information reported, the most common substance involved was fentanyl or other opioid(s) (Exhibit 7).
Conclusion

The use of wastewater data for policymaking is relatively new, and the analyses we describe here clarify the types of insights that wastewater data can provide, particularly when combined with other data and advanced analytics. Early warnings from routine monitoring can help officials get ahead of emerging drug threats and shifting patterns in substance use, and wastewater analytics like the Drug-SURGE alert algorithm could help officials predict overdoses and act to avoid them. Using wastewater data for impact evaluations could strengthen the evidence base by identifying the most promising substance use interventions.

As wastewater monitoring expands to include new health targets and geographies, integrating data, analytics, and reporting systems will be critical to gaining insights that inform action. For example, Biobot currently partners with the National Institute on Drug Abuse to test wastewater from 70 locations across the country. Integrating analytics such as Mathematica’s Drug-SURGE alert algorithm into the next phase of work could provide further validation of the algorithm and help communities better interpret and use wastewater data.

Learning models could help overcome data gaps and produce more granular results. Lastly, integrating high-risk substance wastewater data and advanced analytics into existing platforms such as ODMAP and the Centers for Disease Control and Prevention’s National Wastewater Surveillance System could help communities with and without wastewater monitoring use the data to save lives.

References


