Machine Learning for Opioid and Overdose Surveillance

Daniel B. Neill, Ph.D.

Associate Professor of Computer Science and Public Service
Associate Professor of Urban Analytics, NYU CUSP
Director, Machine Learning for Good (ML4G) Laboratory
Affiliated Faculty, NYU Center for Opioid Epidemiology & Policy

New York University

E-mail: daniel.neill@nyu.edu
Web: http://www.cs.nyu.edu/~neill
http://wp.nyu.edu/ml4good

How can machine learning assist public health practitioners?

Event and pattern detection

→ Better situational awareness

→ Earlier and more precisely targeted interventions

Modeling & mitigating environmental health disparities

Early detection of emerging disease outbreaks

Interventions to combat the opioid crisis

Drug overdoses

- Drug overdoses are an increasingly serious problem in the United States and worldwide.
  - In 2017, more than 72,000 drug overdose deaths occurred in the U.S., more than any year in recorded history.
  - Approximately 68% of these overdose deaths involved opioids.
  - Economic costs of the crisis are estimated at $78.5 billion annually.
- These statistics motivate public health to identify and predict emerging trends in overdoses (geographic, demographic, and behavioral) to better target interventions.
  - **Prevention** of high-risk prescribing and opioid use behaviors
  - **Treatment** of opioid addiction, e.g., medication-assisted therapy
  - **Rescue**, e.g., access to life-saving naloxone
  - **Recovery**, e.g., peer recovery coaches

Drug overdoses

- Drug overdoses are an increasingly serious problem in the United States and worldwide.
  - In 2017, more than 72,000 drug overdose deaths occurred in the U.S., more than any year in recorded history.
  - Approximately 68% of these overdose deaths involved opioids.
  - Economic costs of the crisis are estimated at $78.5 billion annually.
- These statistics motivate public health to identify and predict emerging trends in overdoses (geographic, demographic, and behavioral) to better target interventions.
- Machine learning has potential to **save lives** by detecting subtle, emerging patterns of overdoses in their early stages and targeting an effective public health response at the geographic, subpopulation, individual, and network levels.

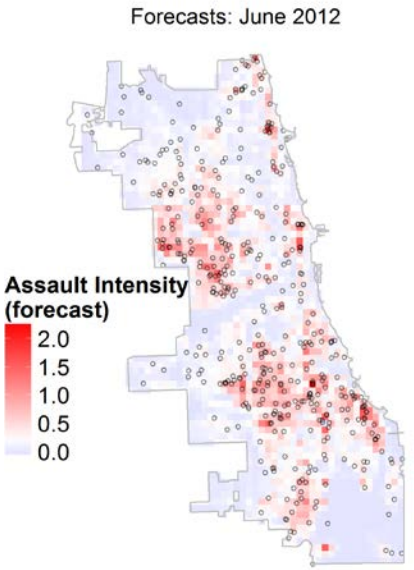
Geographic surveillance

- Answers the question, **where** should I intervene?
- Main goals: estimate predicted overdose trends in space and time; identify anomalous spikes in overdose deaths.

Our recent work* on **scalable Gaussian processes** achieves state-of-the-art accuracy for long-term, small-area forecasting.

Useful predictors include neighborhood characteristics and recent spatio-temporal trends in overdoses and leading indicators.

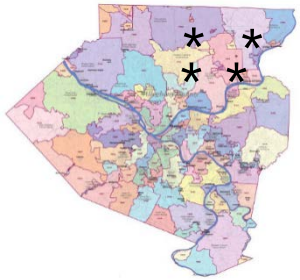
We are currently integrating multiple data sources (ME, EMS, PDMP, census) to **predict overdose risk** and **target interventions** in RI.



*SR Flaxman, AG Wilson, DB Neill, H Nickisch, AJ Smola. Fast Kronecker inference in Gaussian processes with non-Gaussian likelihoods. Proc. 32nd Intl. Conf. on Machine Learning, *PMLR* 37: 607-616, 2015.

Subpopulation-level monitoring

- Answers the question, **for whom** should I intervene?
- Main goal: provide early warning for newly emerging subpopulation-level spikes/clusters of overdose deaths.
- We developed a novel detection method, **multi-dimensional tensor scan**, to detect emerging geographic, demographic, and behavioral patterns.
  - **Earlier detection** of emerging overdose clusters through daily surveillance runs.
  - Better characterization of **where** and **who** is affected.



X white males aged 20-49 X

Subpopulation-level monitoring

- Answers the question, **for whom** should I intervene?
- Main goal: provide early warning for newly emerging subpopulation-level spikes/clusters of overdose deaths.
- We developed a novel detection method, **multi-dimensional tensor scan**, to detect emerging geographic, demographic, and behavioral patterns.
  - **Earlier detection** of emerging overdose clusters through daily surveillance runs.
  - Better characterization of **where** and **who** is affected.
- Analyzed eight years of data from Allegheny County, PA.

Changing demographics of risk

Cluster of 11 fentanyl-related deaths in 2015, elderly black males in downtown Pittsburgh.

Impacts of policy

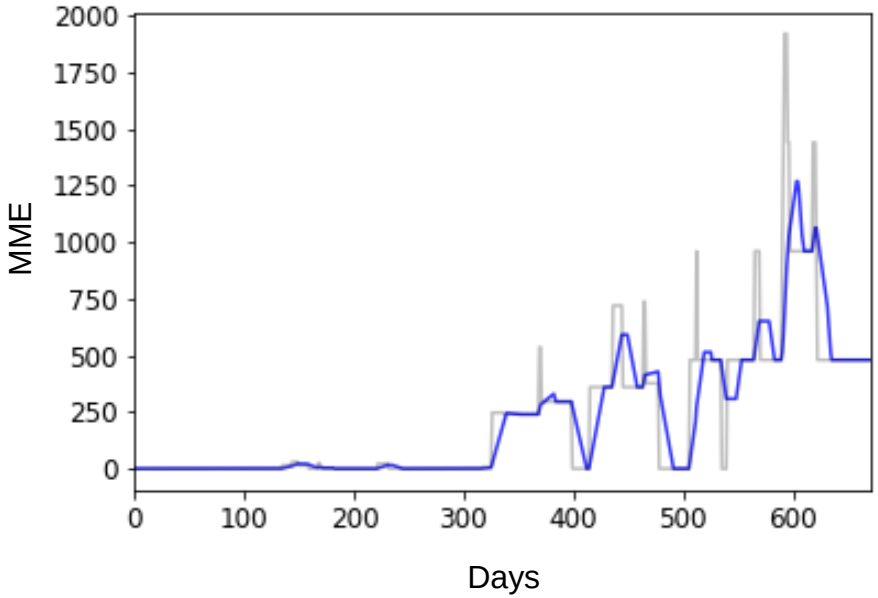
"Methadone + Xanax" overdose clusters were reduced by PA's passage of the Methadone Death & Incident Review Act.

Individual-level opioid use monitoring

(Joint work with Dylan Fitzpatrick)

- Seven years of de-identified data from over 1M individuals provided by Kansas prescription drug monitoring program (PDMP).
- Duration and quantity of prescribed opioids are used to create timelines of morphine milligram equivalents (MME) for individual patients.
- We were able to identify **early indicators** in patient MME timelines which were highly predictive of later opioid misuse.

Smoothed MME Timeline for a Single Patient



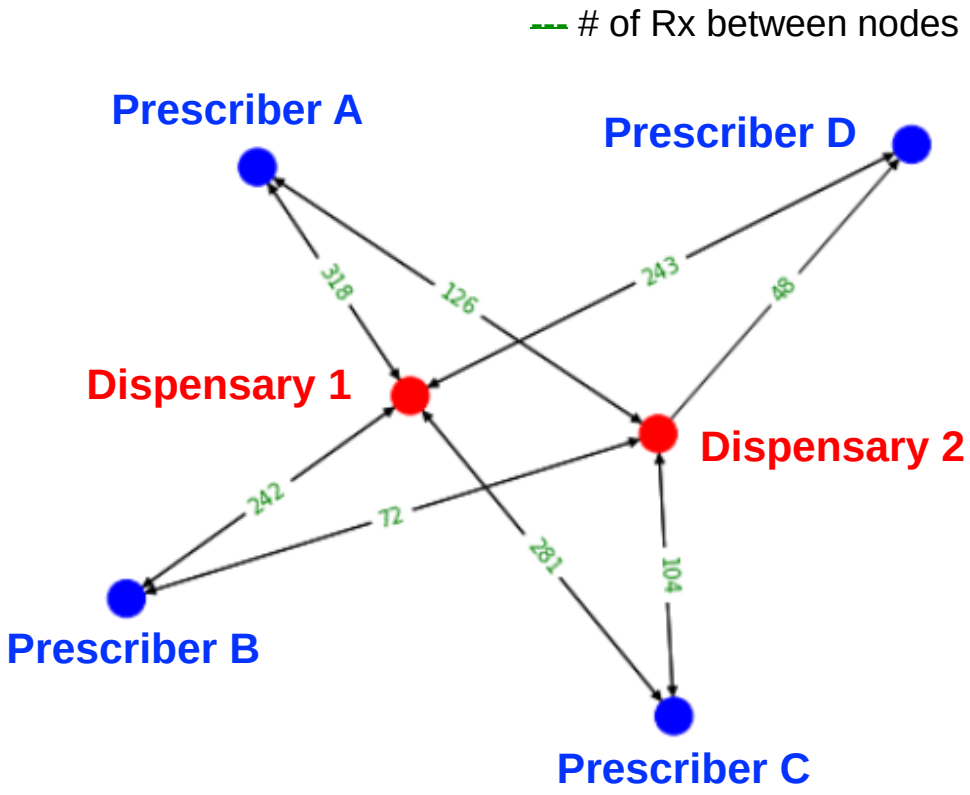
Monitoring networks of prescribers

(Joint work with Katie Rosman)

We are also using the PDMP data for **network analysis**: we identify connected networks of prescribers and dispensaries who are engaging in high-risk and possibly illicit prescribing behaviors.

Step 1: compute the **anomalousness** of each prescriber and dispensary based on Rx and patient-level attributes.

Step 2: Identify the **most anomalous clusters** by maximizing a nonparametric scan statistic over connected subgraphs.



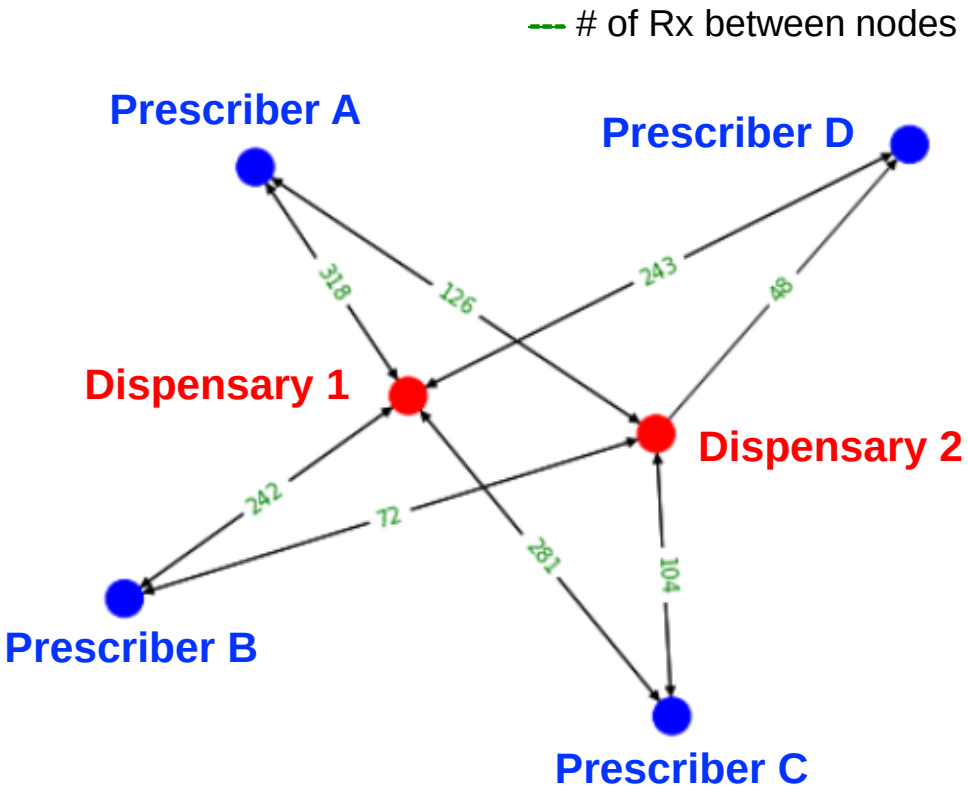
Monitoring networks of prescribers

(Joint work with Katie Rosman)

We are also using the PDMP data for **network analysis**: we identify connected networks of prescribers and dispensaries who are engaging in high-risk and possibly illicit prescribing behaviors.

This detected subgraph of four prescribers and two dispensaries had ~8K prescriptions and ~1,800 patients associated with it.

- 77% of prescriptions were opioids (1.5x expected)
- Average daily dose of opioids per patient was 135 MME (6x expected).
- 30% of prescriptions paid for in cash (3x expected).



Discussion

Here we described several new methods that can be used for **early warning** and **advance forecasting** of overdoses at geographic, subpopulation, individual, and network levels.

Our retrospective analyses of overdose and opioid use data from Pennsylvania, New York, and Kansas suggest high potential utility for **prospective** drug overdose surveillance systems, to facilitate targeted and effective interventions.

We are currently collaborating with an interdisciplinary team of investigators and public health practitioners, with the goals of deploying targeted interventions to prevent overdoses and evaluating their effectiveness through randomized trials.

Thanks for listening!

More details on our web site:
http://wp.nyu.edu/ml4good

Or e-mail me at:
daniel.neill@nyu.edu