Use-Inspired AI and Machine Learning for Public and Population Health

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

New York University

E-mail: daniel.neill@nyu.edu
Web: http://www.cs.nyu.edu/~neill
http://wp.nyu.edu/ml4good
The Machine Learning for Good Lab @ NYU

- Event Detection and Situational Awareness
  - Public health and disease surveillance
  - Violence, conflict, unrest, human rights
  - Criminal justice
  - Fairness and Equity in Algorithmic Decision-Making
  - Allocating city services
  - Healthcare best practices
  - Environmental health and prevention
  - Crime prediction and prevention
  - Optimizing, Deploying, and Evaluating Targeted Interventions for Good
  - Opioid and overdose surveillance
- Novel ML Theory and Methods for Pattern Detection and Prediction
  - Causal Inference and Learning
  - Allocating city services
  - Healthcare best practices
  - Environmental health and prevention
- Optimizing, Deploying, and Evaluating Targeted Interventions for Good
How can machine learning improve population health?

Event & pattern detection

Better situational awareness

Earlier & more precisely targeted interventions

Interventions to combat the opioid overdose crisis

Providing a safety net for novel disease outbreaks and emerging public health threats
Drug overdoses are an increasingly serious problem in the United States and worldwide.

- In 2020, more than 93,000 drug overdose deaths occurred in the U.S., more than any year in recorded history.
- Nearly three-quarters of these overdose deaths involved opioids.
- Economic costs of the crisis have been estimated at between $78.5 billion and >$1 trillion 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 2020, more than 93,000 drug overdose deaths occurred in the U.S., more than any year in recorded history.
  - Nearly three-quarters of these overdose deaths involved opioids.
  - Economic costs of the crisis have been estimated at between $78.5 billion and >$1 trillion 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 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.

Case study: Geographic surveillance

- We analyzed aggregate monthly counts of fatal opioid overdoses for six New York counties from 1999-2015.

- We developed a new detection approach* which combines Gaussian processes (to model correlations) and subset scan (to identify the most anomalous space-time regions).

- We compared our new method to typical anomaly detection approaches on real and synthetic datasets.
  - GPSS > GP alone: nearby points matter for subtle anomalies
  - GPSS > SS alone: covariance structure matters for correlated data

Case study: Geographic surveillance

Two statistically significant spikes in overdose cases:

- Mid 2006. Just before naloxone programs.
- End of 2015. Recent surge due to fentanyl.
Case study: Geographic surveillance

Simpler anomaly detection methods fail to capture the relevant trends.
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.

![Subpopulation-level monitoring](image)
Overdoses in Allegheny County, PA

• We analyzed* county medical examiner data for fatal accidental drug overdoses, 2008-2015.
• ~2000 cases: for each overdose victim, we have date, location (zip), age, gender, race, and the set of drugs present in their system.
• Reduced to 30 dimensions (age decile, gender, race, presence/absence of 27 common drugs) plus space and time.
• Clusters discovered by MD-Scan were shared with Allegheny County Dept. of Human Services.

Fentanyl is a dangerous opioid which has been a major cause of the recent spike in overdose deaths. This dataset captured the start of the spike in western PA.

January 10 to February 7, 2015:
Cluster of 11 fentanyl-related deaths, mainly black males over 58 years of age, centered in Pittsburgh’s downtown Hill District.

Very unusual demographic: common dealer / shooting gallery?

March 27 to April 21, 2015:
26 deaths county-wide from fentanyl, heroin only present in 11.

Started in the southeast suburbs of Pittsburgh and spread across the city.

Our method could have detected this pattern on March 29, identifying a cluster of four overdose deaths with strong geographic and demographic similarities.
MD-Scan Overdose Results (2)

Another set of discovered overdose clusters each involved a combination of Methadone and Xanax.

Methadone: an opioid used for chronic pain relief and to treat heroin addiction, but also addictive and risk of OD.

Xanax (alprazolam): a benzodiazepine prescribed for panic and anxiety disorders.

The combination produces a strong high but can be deadly (~30% of methadone fatal ODs).

From 2008-2012: multiple M&X OD clusters, 3-7 cases each, localized in space and time.

From 2013-2015: no M&X overdose clusters; 33% and 47% drops in yearly methadone and M&X deaths respectively.

Why did these deaths cluster, when methadone and methadone + other benzo deaths did not?

What factors could explain the dramatic reduction in M&X overdose clusters?
Another set of discovered overdose clusters each involved a combination of Methadone and Xanax.

**Methadone**: an opioid used for chronic pain relief and to treat heroin addiction, but also addictive and risk of OD.

**Xanax (alprazolam)**: a benzodiazepine prescribed for panic and anxiety disorders.

Increased state oversight of methadone clinics and prescribing physicians after passage of the Methadone Death and Incident Review Act (Oct 2012).

Approval of generic suboxone (buprenorphine + naloxone) in early 2013 lowered cost of suboxone treatment as an alternative to methadone clinics.

Why did these deaths cluster, when methadone and methadone + other benzo deaths did not?

What factors could explain the dramatic reduction in M&X overdose clusters?
Individual-level opioid use monitoring

- Seven years of de-identified data from over 1M individuals provided by Kansas prescription drug monitoring program (PDMP), with unique patient, prescriber, and dispensary identifiers.
- Duration and quantity of prescribed opioids are used to create timelines of morphine milligram equivalents (MME) for individual patients.
- Can we identify early indicators in patient MME timelines which are predictive of later opioid misuse or unsafe prescribing?
Individual-level opioid use monitoring

• Patients are clustered using the *k*-shape algorithm (Paparrizos & Gravano, 2015) to group patients with similar patterns in MME timelines.

• Are some patient clusters associated with higher risk of red flags indicating misuse or unsafe practices?

• For a new patient, can we confidently assess risk of future red flags given a partial MME timeline?
Early individual-level risk assessment by classifying partial trajectories
Early individual-level risk assessment by classifying partial trajectories

Partial trajectory assigned **low risk** based on MME and cluster shape.
Early individual-level risk assessment by classifying partial trajectories

Partial trajectory assigned to high risk cluster #5, and higher MME → high risk
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.
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, & 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.
How can machine learning improve population health?

- Event & pattern detection
- Better situational awareness
- Earlier & more precisely targeted interventions

Interventions to combat the opioid overdose crisis

Providing a safety net for novel disease outbreaks and emerging public health threats

Earlier & more precisely targeted interventions

Better situational awareness

Event & pattern detection
Disease surveillance example

Spike in gastrointestinal (GI) illness near Columbus, Ohio
Disease surveillance example

Spike in gastrointestinal (GI) illness near Columbus, Ohio

“Syndromic surveillance” is a common approach used by public health epidemiologists:
1) Define a set of known disease syndromes: influenza-like illness, respiratory, GI, etc.
2) Classify hospital ED cases, and other indicators like medication sales, to syndromes.
3) Search for emerging clusters of disease cases in space and time.
Detecting rare disease outbreaks with Twitter

Pre-syndromic surveillance

<table>
<thead>
<tr>
<th>Date/time</th>
<th>Hosp.</th>
<th>Age</th>
<th>Complaint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 1 08:00</td>
<td>A</td>
<td>19-24</td>
<td>runny nose</td>
</tr>
<tr>
<td>Jan 1 08:15</td>
<td>B</td>
<td>10-14</td>
<td>fever, chills</td>
</tr>
<tr>
<td>Jan 1 08:16</td>
<td>A</td>
<td>0-1</td>
<td>broken arm</td>
</tr>
<tr>
<td>Jan 2 08:20</td>
<td>C</td>
<td>65+</td>
<td>vomited 3x</td>
</tr>
<tr>
<td>Jan 2 08:22</td>
<td>A</td>
<td>45-64</td>
<td>high temp</td>
</tr>
</tbody>
</table>

Key challenge: A syndrome cannot be created to identify every possible cluster of potential public health significance.

Thus a method is needed to identify relevant clusters of disease cases that do not correspond to existing syndromes.

Use case proposed by NC DOH and NYC DOHMH, solution requirements developed through a public health consultancy at the International Society for Disease Surveillance.
Where do existing methods fail?

The typical syndromic surveillance approach can effectively detect emerging outbreaks with commonly seen, general patterns of symptoms (e.g. ILI).

What happens when something new and scary comes along?
- More specific symptoms (“coughing up blood”)
- Previously unseen symptoms (“nose falls off”)

If we were monitoring these particular symptoms, it would only take a few such cases to realize that an outbreak is occurring!

Mapping specific chief complaints to a broader symptom category can dilute the outbreak signal, delaying or preventing detection.
Where do existing methods fail?

The typical syndromic surveillance approach can effectively detect emerging outbreaks with commonly seen, general patterns of symptoms (e.g. ILI).

Where do existing methods fail?

Mapping specific chief complaints to a broader symptom category can dilute the outbreak signal, delaying or preventing detection.

What happens when something new and scary comes along?

- More specific symptoms (e.g., coughing up blood)
- Previously unseen symptoms (e.g., nose falls off)

If we were monitoring these particular symptoms, it would only take a few such cases to realize that an outbreak is occurring!

Our solution is to combine text-based (topic modeling) and event detection (multidimensional scan) approaches, to detect emerging patterns of keywords.
The semantic scan statistic

<table>
<thead>
<tr>
<th>Date/time</th>
<th>Hosp.</th>
<th>Age</th>
<th>Complaint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 1 08:00</td>
<td>A</td>
<td>19-24</td>
<td>runny nose</td>
</tr>
<tr>
<td>Jan 1 08:15</td>
<td>B</td>
<td>10-14</td>
<td>fever, chills</td>
</tr>
<tr>
<td>Jan 1 08:16</td>
<td>A</td>
<td>0-1</td>
<td>broken arm</td>
</tr>
<tr>
<td>Jan 2 08:20</td>
<td>C</td>
<td>65+</td>
<td>vomited 3x</td>
</tr>
<tr>
<td>Jan 2 08:22</td>
<td>A</td>
<td>45-64</td>
<td>high temp</td>
</tr>
</tbody>
</table>

Bayesian inference using LDA model

- Case prior: $\alpha$
- Topic prior: $\beta$
- Distribution over topics per case: $\theta_1 \ldots \theta_N$
- Observed words: $w_{ij}$

Topics:
- $\phi_1$: vomiting, nausea, diarrhea, ...
- $\phi_2$: dizzy, lightheaded, weak, ...
- $\phi_3$: cough, throat, sore, ...

Classify cases to topics

Time series of hourly counts for each combination of hospital and age group, for each topic $\phi_j$. Now we can do a multidimensional scan, using the learned topics instead of pre-specified syndromes!
Multidimensional Semantic Scan
Learns Two Sets of Topics

• Static Topics
  • Designed to capture common illnesses like flu.
  • Learned over a large set of historical data using a standard LDA topic model.

• Emerging Topics
  • Designed to capture rare or novel diseases that are not well explained by the static topics.
  • Learned over the most recent set of data using a new variant of LDA.
NYC DOHMH dataset

• New York City’s Department of Health and Mental Hygiene, Bureau of Communicable Disease, provided us with 6 years of data (2010-2016) consisting of ~28M chief complaint cases from 53 hospitals in NYC.

• For each case, we have data on the patient’s chief complaint (free text), date and time of arrival, age group, gender, and discharge ICD-9 code.

• Substantial pre-processing of the chief complaint field was necessary because of the size and messiness of the data (typos, abbreviations, etc.).

Variations of the words “vomit” and “vomiting” that appear > 15 times in data
Evaluation on NYC DOHMH data

- Blinded evaluation by NYC DOHMH public health practitioners, comparing our multidimensional semantic scan approach to a state-of-the-art keyword-based scan approach.

- For each method’s 500 highest scoring clusters, users indicated if the cluster is relevant, meaningful, or not of interest.

<table>
<thead>
<tr>
<th>Relevant Clusters of Interest</th>
<th>Meaningful Clusters of Potential Interest</th>
<th>Clusters Not of Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples: bacterial meningitis, synthetic drugs use</td>
<td>Examples: flu, rashes, motor vehicle accidents</td>
<td>Examples: misspellings, non-specific words (i.e. “left”)</td>
</tr>
<tr>
<td><strong>Multidimensional Semantic Scan</strong></td>
<td>53</td>
<td>267</td>
</tr>
<tr>
<td><strong>Keyword Based Method</strong></td>
<td>47</td>
<td>199</td>
</tr>
</tbody>
</table>
The progression of detected clusters after Hurricane Sandy impacted NYC highlights the variety of strains placed on hospital emergency departments following a natural disaster:

- **Acute cases:** falls, SOB, leg injuries
- **Mental health disturbances:** depression, anxiety
- **Burden on medical infrastructure:** methadone, dialysis

Many other events of public health interest were identified:

- **Accidents**
  - Motor vehicle
  - Ferry
  - School bus
  - Elevator

- **Contagious Diseases**
  - Meningitis
  - Scabies
  - Ringworm
  - Hepatitis

- **Other**
  - Drug overdoses
  - Smoke inhalation
  - Carbon monoxide poisoning
  - Crime related, e.g., pepper spray attacks
Example of a detected cluster

<table>
<thead>
<tr>
<th>Arrival Date</th>
<th>Arrival Time</th>
<th>Hospital ID</th>
<th>Chief Complaint</th>
<th>Patient Sex</th>
<th>Patient Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/28/2014</td>
<td>7:52:00</td>
<td>HOSP5</td>
<td>EVALUATION, DRANK COFFEE WITH CRUS</td>
<td>M</td>
<td>45-49</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>7:53:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>M</td>
<td>65-69</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>7:57:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>F</td>
<td>20-24</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>7:59:00</td>
<td>HOSP5</td>
<td>INGESTED TAINTED COFFEE</td>
<td>M</td>
<td>35-39</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>8:01:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>M</td>
<td>45-49</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>8:03:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>M</td>
<td>40-44</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>8:04:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>M</td>
<td>30-34</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>8:06:00</td>
<td>HOSP5</td>
<td>DRANK TAINTED COFFEE</td>
<td>M</td>
<td>35-39</td>
</tr>
<tr>
<td>11/28/2014</td>
<td>8:09:00</td>
<td>HOSP5</td>
<td>INGESTED TAINTED COFFEE</td>
<td>M</td>
<td>25-29</td>
</tr>
</tbody>
</table>

This detected cluster represents 9 patients complaining of ingesting tainted coffee, and demonstrates Semantic Scan’s ability to detect rare and novel events.
Incorporating user feedback

- Our system enables continual improvement of performance by including public health practitioners in the loop and incorporating their feedback.

- Users can add new syndromes and specify if they would like the system to monitor or ignore them in the future.

- Blinded user studies show that this Practitioner in the Loop approach enables the system to report more relevant clusters and to avoid overwhelming the user with irrelevant findings.
Pre-syndromic surveillance is a safety net that can supplement existing ED syndromic surveillance systems by alerting public health to unusual or newly emerging threats.

Our recently proposed semantic scan can accurately and automatically discover pre-syndromic case clusters corresponding to novel outbreaks and other patterns of interest.
How can machine learning improve population health?

ML can address many other aspects of pandemic preparedness and response, for example:

- Contact tracing
- Epidemic modeling
- Estimating the causal impacts of various public health interventions

How can machine learning improve population health?

Interventions to combat the opioid overdose crisis

Earlier & more precisely targeted interventions

Better situational awareness

Event & pattern detection

Providing a safety net for novel disease outbreaks and emerging public health threats
How can machine learning improve population health?

- Event & pattern detection
- Better situational awareness
- Earlier & more precisely targeted interventions

Interventions to combat the opioid overdose crisis

And there are many other ways ML can improve population health, such as:

- **Causal inference** methods to assess impacts of environmental exposures, such as poor-quality housing, on health.
- **Algorithmic fairness** to allocate resources and reduce health disparities.

Providing a safety net for novel disease outbreaks and emerging public health threats
Identifying **causal** effects of environmental exposures

We are using Medicaid data linked to detailed building characteristics in order to identify impacts of poor-quality housing on chronic health.

---

**Key idea**: treatment effects may be **heterogeneous**; use multidimensional scan to identify most affected subpopulations.

---

“Which housing conditions impact which health conditions, for which subpopulations, to what extent?”

---

Must adjust for known confounders, selection into treatment (exposure).

**Step 1: Predictive model at building level**

\[X = 65 \text{ diagnoses } \times \{\text{adult, child}\}\]

\[Y = \text{building on landlord watch list?}\]

- Adult asthma and COPD
- Mental health (ADHD, adjust. disorder)
- Injuries (children and adults)

**Step 2: Heterogeneous treatment effect scan**

- “**Crowded housing** is associated with increased respiratory conditions & injuries among Asians living in Manhattan.”

---

Must account for multiple hypothesis testing to bound false positive rate.

We have also developed an alternative scan-based approach to causal inference, based on automated discovery of natural experiments.
Thanks for listening!

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

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