Pre-syndromic surveillance for improved detection of emerging public health threats

Daniel B. Neill, Ph.D.

Machine Learning for Good Laboratory
New York University
daniel.neill@nyu.edu

Joint work with Mallory Nobles (CMU), Ramona Lall (NYC DOHMH), and Rob Mathes (NYC DOHMH).
This presentation is based on our recent publication:
M. Nobles, R. Lall, R.W. Mathes, and D.B. Neill*, Science Advances 8(44): eabm4920, 2022.
DOI: 10.1126/sciadv.abm4920 (open access)

Acknowledgements:
This work was partially supported by NSF grants IIS-0916345, IIS-0911032, and IIS-0953330 and the DHS Hidden Signals Challenge.
Thanks to the BCD Syndromic Surveillance Unit at NYC DOHMH for providing retrospective data for this study and participating in blinded evaluations, and to Boyuan Chen and Andy Wei for their work in developing the open-source Python version of MUSES.
Pre-syndromic surveillance

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

**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 ("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!

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

Time series of hourly counts for each combination of hospital and age group, for each topic $\varphi_j$.

Classify cases to topics

Bayesian inference using LDA model

Now we can do a multidimensional scan, using the learned topics instead of pre-specified syndromes!
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 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
The progression of detected clusters after Hurricane Sandy impacted NYC highlights the variety of strains placed on hospital emergency departments following a natural disaster:

Many other events of public health interest were identified:

- **Acute cases:** falls, SOB, leg injuries

- **Mental health disturbances:** depression, anxiety

- **Burden on medical infrastructure:** methadone, dialysis

**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

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

This detected cluster represents 9 patients complaining of ingesting tainted coffee, and demonstrates Semantic Scan’s ability to detect rare and novel events.
“Highly relevant” clusters included bacterial meningitis and synthetic drug use. “Meaningful” but not “highly relevant” clusters included motor vehicle accidents. “Not meaningful” clusters could be due to typos, coincidence, etc.
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.
Second blinded user study

PITL detected 49 highly relevant clusters (53% increase vs. fixed), corresponding to 24 distinct event types (33% increase vs. fixed).

Fig. 4. Results from a blinded user study comparing the fixed and PITL models. Blue lines: PITL model. Red lines: fixed model. (A) Cumulative number of highly relevant clusters detected by each method, after each 2 week time period. The performance gap between the PITL and fixed models increases monotonically as a function of the number of labeled clusters used as training data by the PITL model. (B) Cumulative number of clusters detected by each method that were similar to clusters previously labeled “to ignore” by the user, after each 2-week time period. During the experiment, the fixed model detected 78 irrelevant clusters similar to those labeled “to ignore,” while the PITL model only identified three such clusters. (C) Cumulative number of clusters detected by each method that were similar to clusters previously labeled “to monitor” by the user, after each 2-week time period. The PITL model identified a total of 10 highly relevant clusters that the practitioner had previously expressed interest in monitoring, as compared to 6 for the fixed model.
### COVID results (March-June 2020)

Table 2. Results from MUSES runs on ED chief complaint data from NYC DOHMH during the first wave of the novel coronavirus (COVID-19) pandemic in NYC, 1 March through 30 June 2020. Highest-scoring clusters found with 25 static and 25 emerging topics, scanning over both static and emerging topics. For each cluster, we report the data, de-identified hospital ID, number of cases and number of hours, whether the cluster is COVID-related, the most common chief complaints, and the cluster's log likelihood ratio score. ICD-10 diagnosis codes were noted when used consistently to describe cases in the cluster (9 of 33 clusters). At least 30 of the 33 detected clusters were COVID related. Thirty of 33 clusters occurred during the peak of the pandemic in NYC (17 March through 5 April), and 32 of 33 clusters corresponded to emerging topics rather than static topics.

| Date        | Hosp ID | No. of cases | No. of hours | COVID | Description                                                                 | Score |
|-------------|---------|--------------|--------------|-------|------------------------------------------------------------------------------|-------|
| 27 March    | 31      | 164          | 12           | Y     | "Covid 19 exposure", flu-like symptoms, testing, cough, sob                  | 244   |
| 28 March    | 31      | 152          | 10           | Y     | Testing, exposure, cough, sore throat, syncope                               | 178   |
| 25 March    | 19      | 43           | 5            | Y     | "Coronavirus" [ICD-10: B97.29, cough, fever, headache, sob                   | 75    |
| 29 March    | 31      | 111          | 11           | Y     | Testing, exposure, cough, fever, diarrhea, pneumonia                         | 69    |
| 1 April     | 40      | 26           | 2            | Y     | Influenza-like respiratory [ICD-10: J10.1]                                   | 69    |
| 17 March    | 7       | 42           | 8            | Y     | Smoke inhalation [ICD-10: J90.5], cough                                      | 65    |
| 26 March    | 1       | 14           | 3            | Y     | "Covid", cough, sore throat, body ache, measured C2                          | 58    |
| 2 April     | 52      | 64           | 11           | Y     | Screening for viral disease [ICD-10: Z11.99, cough, fever, sob               | 54    |
| 27 April    | 7       | 19           | 5            | Y     | "Covid 19 screening", cough, fever, sob                                      | 53    |
| 24 March    | 4       | 30           | 6            | Y     | Respiratory, headache                                                        | 53    |
Discussion

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 multidimensional semantic scan (MUSES) can accurately and automatically discover pre-syndromic case clusters corresponding to novel outbreaks and other patterns of interest.
Thanks for listening!

More details and MUSES open-source software on our project page: https://wp.nyu.edu/ml4good/preadsyr/syndromic-surveillance

Check out our MUSES demo later this afternoon! Or e-mail me at: daniel.neill@nyu.edu