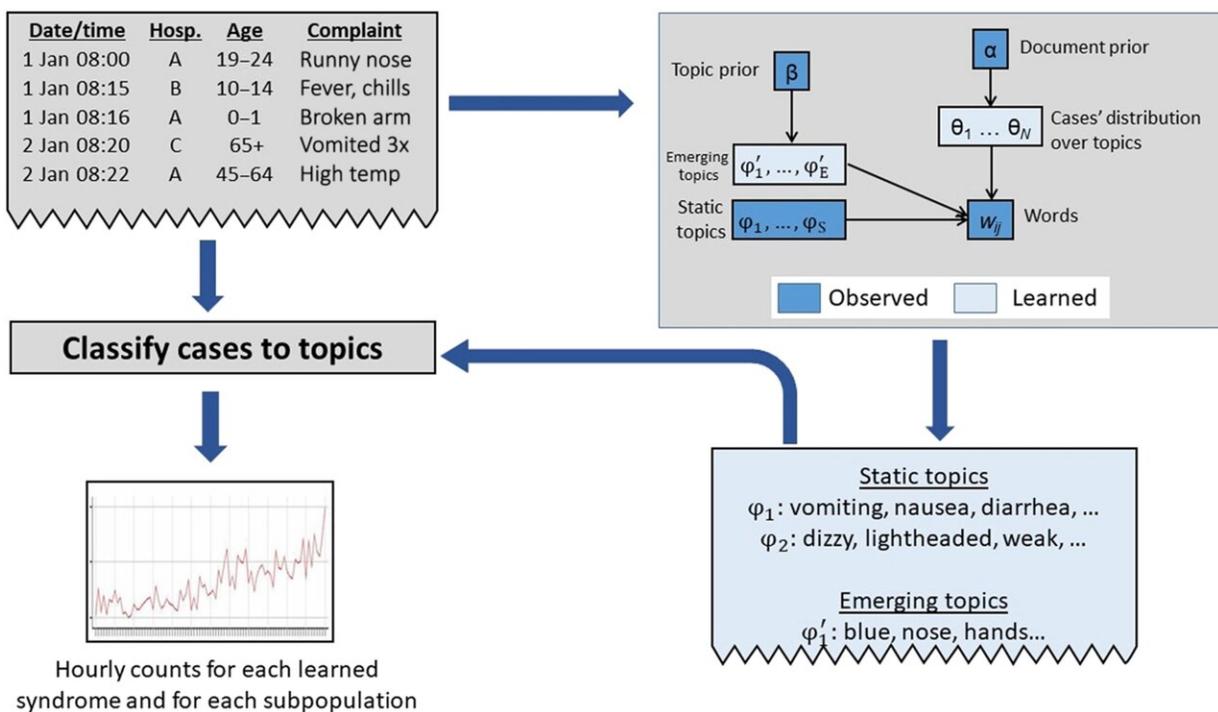


Data-driven, automated machine-learning system for detecting emerging public health threats

November 14 2022, by Robert Polner



Overview of MUSES. MUSES is an innovative approach to presyndromic surveillance that learns newly emerging syndromes directly from free-text chief complaint data from hospital EDs, and detects statistically significant increases in cases related to these syndromes, including case clusters that may be limited to or differentially affect specific subpopulations. Credit: *Science Advances* (2022). DOI: 10.1126/sciadv.abm4920

A dire threat to public health can emerge from a huge variety of sources—for example, infectious diseases, a spate of drug overdoses, or exposures to toxic chemicals. Federal, state, and local health departments must respond rapidly to disease outbreaks and other emerging bio-threats. While the current automated systems for "syndromic surveillance" can help by monitoring health data and detecting disease clusters, they are not able to detect clusters with rare or previously unseen symptomology.

A new study from New York University's Machine Learning for Good Laboratory (ML4G Lab), with colleagues from Carnegie Mellon University and the New York City Department of Health and Mental Hygiene (NYC DOHMH), addresses this critical gap in public [health](#) practice by presenting a new machine-learning approach for "pre-syndromic" surveillance.

The method is incorporated in an automated system that can enable public health practitioners to respond more quickly and effectively in the future to fast-emerging threats, including those that are unusual or novel.

"Existing systems are good at detecting outbreaks of diseases that we already know about and are actively looking for, like flu or COVID," comments NYU professor Daniel B. Neill, the senior author of the study and director of the ML4G Lab. "But what happens when something new and scary comes along? Pre-syndromic surveillance provides a safety net to identify emerging threats that other systems would fail to detect."

The study was published in *Science Advances*.

The authors' approach to disease surveillance is known as pre-syndromic surveillance because it relies on digitally communicated textual data on all patient conditions, rather than classifying case data under existing

disease syndromes (such as "influenza-like" or "gastro-intestinal" illness). The new system enables rapid identification of newly emerging syndromes that health departments are not yet aware of.

To accomplish this, the machine learning technology uses anonymized "chief complaint" data from hospital Emergency Department (ED) visits. A chief complaint is usually provided by the patient in their own words (for example, "I've had a bad headache for the last three days and now my ear hurts") and is recorded by an ED triage nurse.

The method is capable of identifying trends and patterns in the words and phrases of the chief complaints, enabling detection of a localized case cluster. It can incorporate practitioners' feedback in the service of automatically distinguishing between relevant and irrelevant case clusters. It gives personalized and actionable decision support to hospitals and local and state health departments.

Blinded evaluations and case studies by the city health department of the new system—which the researchers dubbed MUSES, or Multidimensional Semantic Scan, after designing, developing and testing it—demonstrate that the pre-syndromic monitoring identifies more events of [public health](#) interest and achieves a lower false positive rate in comparison to traditional methods, according to the study authors.

MUSES, then, offers three significant methodological advances to hospitals and local and state health departments nationally, as it:

- Eliminates the need for pre-defined syndrome categories.
- Identifies localized case clusters through multi-dimensional scan statistics, enabling detection of emerging bio-threats that may affect certain spatial areas or demographic groups of patients.
- Uses a "practitioner in the loop" approach to incorporate user feedback, hone in on relevant patterns, reduce [false positives](#),

and provide local users with actionable insights based on their own criteria for what is, and is not, relevant.

More information: Mallory Nobles et al, Presyndromic surveillance for improved detection of emerging public health threats, *Science Advances* (2022). [DOI: 10.1126/sciadv.abm4920](https://doi.org/10.1126/sciadv.abm4920)

Provided by New York University

Citation: Data-driven, automated machine-learning system for detecting emerging public health threats (2022, November 14) retrieved 9 April 2023 from <https://medicalxpress.com/news/2022-11-data-driven-automated-machine-learning-emerging-health.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.