

# Detecting Unanticipated Increases in Emergency Department Chief Complaint Keywords

Ramona Lall\*, Alison Levin-Rector, Robert Mathes and Don Weiss

New York City Department of Health and Mental Hygiene, Queens, NY, USA

## Objective

To detect sudden increases in word frequency in the Emergency Department (ED) syndromic chief complaint (CC) text field.

## Introduction

The CC text field is a rich source of information, but its current use for syndromic surveillance is limited to a fixed set of syndromes that are routine, suspected, expected, or discovered by chance. In addition to syndromes that are routinely monitored by the NYC Department of Health and Mental Hygiene (e.g., diarrhea, respiratory), additional syndromes are occasionally monitored when requested by outside sources or when expected to increase during emergencies. During Hurricane Sandy, we discovered by manual inspection of data for a few EDs an increase in certain words in the CC field (e.g., "METHADONE", "DIALYSIS", and "OXYGEN") that led to the creation of a "needs medication" syndrome. Current syndromic surveillance systems cannot detect unanticipated events that are not defined a priori by keywords. We describe a simple data-driven method that routinely scans the CC field for increases in word frequency that might trigger further investigation and/or temporary monitoring.

## Methods

This method compares the daily citywide frequency of each unique word within the CC field against the average daily frequency of the word during a baseline period. A ratio of the current frequency of a word to the average frequency over a baseline period was computed, and words with high ratios (i.e., ratios  $\geq 4$ ) were identified. A manual review of the "top words" was necessary to decide whether any word warranted further investigation into the nature of the complaints associated with the word. This method was applied daily using a 1-year baseline to ED syndromic data between June and August 2013. To evaluate the usefulness of the method during emergencies, the method was applied to data using a 14-day baseline following two extreme weather-related events: Hurricane Sandy in 2012 and a blizzard in December 2010.

## Results

During the 1-year baseline considered, there were ~59,000 unique words in the CC field, including misspellings, typing errors, and uninformative (e.g., "IS", "AND") and indecipherable words. Daily monitoring of ED CC during June-August 2013 identified words consistent with the summer season, e.g., "BIKE", "MOTORCYCLE", "TRAFFIC", "BEE", "MOSQUITO", "INSECT", "BITES", and "BITTEN". In one instance, there was a sudden and sharp increase in the words "MENINGITIS" (ratio=33) and "EXPOSE" (ratio=32). On further investigation, this was found to be due to a large group of patients visiting a single ED to receive prophylaxis after exposure to a patient with meningitis. Retrospective analyses of the two weather-related events were found to have increases in a similar set of words: "METHADONE", "MEDICATION", "DIALYSIS", "OXYGEN", and "HYPOTHERMIA" (Table). Additionally, there were more frequent mentions of hospital transfers and evacuations during Sandy and of drug-related complaints during the blizzard.

## Conclusions

This method provides a simple approach to routine situational awareness that could prove especially useful during an emergency. Future work will address the following limitations to the current version of this method: (1) some words that are not meaningful (e.g., repeated typos) can have a high ratio; (2) increases in unusual words (e.g., "HYPOTHERMIA") were more likely to be identified than increases in high-frequency words (e.g., "PAIN"), due to the use of ratios; (3) this method used single words (e.g., "EXPOSURE") rather than word combinations or phrases (e.g., "MENINGITIS EXPOSURE") that might be more meaningful. Nevertheless, this method allows health departments to recognize patterns of illness in their communities that might otherwise be overlooked.

Top words identified for two weather-related events

Emergency	No. of words with ratio $\geq 4$	Words of interest among top words
Hurricane Sandy Oct 29-30, 2012	90	METHADONE, NEEDED, O2, OXYGEN, HURRICANE, EVAC, EVACUATION, HYPOTHERMIA, DIALYSIS, SHELTER, INHALATION, REFILLS, TRANSFER, POWER, HYPOXEMIA, PRESCRIPTIONS
Blizzard Dec 27-30, 2010	127	METHADONE, DIALYSIS, HYPOTHERMIA, OPIOID, MEDICATION, SNOW, FROST, NEEDED, ICE, SEPTICEMIA, MARIJUANA

## Keywords

chief complaint; word frequency; situational awareness

\*Ramona Lall

E-mail: rlall@health.nyc.gov

