OJPHI

ISDS
INTERNATIONAL SOCIETY
for DISEASE SURVEILLANCE

Multidimensional Tensor Scan for Drug Overdose Surveillance

Daniel B. Neill*

H.J. Heinz III College, Carnegie Mellon University, Pittsburgh, PA, USA

Objective

We present the multidimensional tensor scan (MDTS), a new method for identifying emerging patterns in multidimensional spatio-temporal data, and demonstrate the utility of this approach for discovering emerging geographic, demographic, and behavioral trends in fatal drug overdoses.

Introduction

Drug overdoses are an increasingly serious problem in the United States and worldwide. The CDC estimates that 47,055 drug overdose deaths occurred in the United States in 2014, 61% of which involved opioids (including heroin, pain relievers such as oxycodone, and synthetics).[1] Overdose deaths involving opioids increased 3-fold from 2000 to 2014.[1] These statistics motivate public health to identify emerging trends in overdoses, including geographic, demographic, and behavioral patterns (e.g., which combinations of drugs are involved). Early detection can inform prevention and response efforts, as well as quantifying the effects of drug legislation and other policy changes.

The fast subset scan[2] detects significant spatial patterns of disease by efficiently maximizing a log-likelihood ratio statistic over subsets of data points, and has recently been extended to multidimensional data (MD-Scan).[3] While MD-Scan is a potentially useful tool for drug overdose surveillance, the high dimensionality and sparsity of the data requires a new approach to estimate and represent baselines (expected counts), maintaining both accuracy and efficient computation when searching over subsets.

Methods

The multidimensional tensor scan (MDTS) is a new approach to subset scanning in multidimensional data. In addition to detecting the spatial area (subset of locations) and time window affected by an emerging outbreak, MDTS can also identify the affected subset of values for each observed attribute. For example, given the drug overdose surveillance data described below, MDTS can identify the affected genders, races, age ranges, and which drugs were involved. MDTS finds subsets of the attribute space with higher than expected case counts, first using a novel tensor decomposition approach to estimate the expected counts. MDTS then iteratively applies a conditional optimization step, optimizing over all subsets of values for each attribute conditional on the current subsets of values for all other attributes[3], and using the linear-time subset scanning property[2] to make each conditional optimization step computationally efficient. The resulting approach has high power to detect and characterize emerging trends which may only affect a subset of the monitored population (e.g., specific ages, genders, neighborhoods, or users of particular combinations of drugs).

Results

We used MDTS to analyze publicly available data from the Allegheny County, PA medical examiner's office and to detect emerging overdose patterns and trends. The dataset consists of ~2000 fatal accidental drug overdoses between 2008 and 2015. For each overdose victim, we have date, location (zip code), age decile, gender, race, and the presence/absence of 27 commonly abused drugs in their system. The highest-scoring clusters discovered by MDTS were shared with Allegheny County's Dept. of Human Services and their feedback obtained.

One set of potentially relevant findings from our analysis involved fentanyl, a dangerous and potent opioid which has been a serious problem in western PA. In addition to identifying two well-known, large clusters of overdoses—14 deaths in January 2014 and 26 deaths in March-April 2015—MDTS was able to provide additional information about each cluster. For example, the first cluster was likely due to fentanyl-laced heroin, while the second was more likely due to fentanyl disguised as heroin (only 11 victims had heroin in their system). Moreover, the second cluster was initially confined to the Pittsburgh suburb of McKeesport and a typical demographic (white males ages 20-49), before spreading across the county. Our analysis demonstrated that prospective surveillance using MDTS would have identified the cluster as early as March 29th, enabling targeted prevention efforts. MDTS also discovered a previously unidentified, highly localized cluster of fentanyl-related overdoses affecting an unusual and underserved demographic (elderly black males near downtown Pittsburgh). This cluster occurred in January-February 2015, and may have been related to the larger cluster of fentanyl-related overdoses that occurred two months later. Finally, we identified multiple overdose clusters involving combinations of methadone and Xanax between 2008 and 2012, and observed dramatic reductions in these clusters corresponding to the passage of the Methadone Death and Incident Review Act (October 2012), which increased state oversight of methadone clinics and prescribing physicians.

Conclusions

Retrospective analysis of Allegheny County overdose data suggests high potential utility for a prospective overdose surveillance system, which would enable public health users to identify emerging patterns of overdoses in their early stages and facilitate targeted and effective health interventions. The MDTS approach can also be used for other multidimensional public health surveillance tasks, such as STI surveillance, where the patterns or outbreaks of interest may have demographic, geographic, and behavioral components.

Keywords

event detection; outbreak detection; subset scan; drug overdose surveillance

References

[1] Rudd RA, Aleshire N, Zibbell JE, Gladden RM. Increases in drug and opioid overdose deaths: United States, 2000–2014. MMWR 2016; 64(40): 1378-1382.
[2] Neill DB. Fast subset scan for spatial pattern detection. J Royal Stat Soc B 2012; 74(2): 337-360.
[3] Neill DB, Kumar T. Fast multidimensional subset scan for outbreak detection and characterization. Online J Pub Health Inform 2013; 5(1): e156.

***Daniel B. Neill**
E-mail: neill@cs.cmu.edu