OJPHI ONLINE JOURNAL OF PUBLIC HEALTH INFORMATICS

Detection of Outbreak Signals Using R

Steve E. Rigdon,[*,1] George Turabelidze,[2] Richard DeClue,[1] Sarah Patrick,[1] Ehsan Jahanpour,[2] Rong He,[2] and Jennifer Lloyd[2]

Author information ► Article notes ► Copyright and License information ►

Objective

Go to: ☑

Develop a statistically rigorous automated process for weekly communicable disease report analysis to improve the speed and accuracy of outbreak detection in Missouri.

Introduction

Go to: ☑

The goal of adequate biosurveillance is to signal that an outbreak may be occurring and through subsequent work is confirmed or refuted. Such a system should be equally able to detect outbreaks of diseases of extremely low reporting frequency, or those with high seasonality.

Methods of detecting increases in notifiable communicable diseases reported to the Missouri Department of Health and Senior Services (MDHSS) were based on quartile comparisons to 5-year historical disease reports for the report week and resulted in frequent detection of statistically significant increases that were, in fact, not indicative of disease outbreaks. Frequently generated alerts led to "alarm fatigue" in epidemiologists.

Methods

Go to: ☑

MDHSS's 10-year communicable diseases database (2002-2011) was used to establish the background frequency of selected diseases: E. coli O157:H7 (a predominantly foodborne disease with strong outbreak potential, EC), erhlichiosis (a highly seasonal tickborne infection, ERH), Q fever (low incidence disease, Q), and pertussis (high incidence disease, P).

Seasonality was modeled by fitting a first- or second-order trigonometric model1 to the average reported cases in each week of the year. Assuming a Poisson distribution for the number of cases per week, we can get an estimate of the mean (or expected) count for each week. Since the mean and variance of the Poisson are equal, we are able to estimate the variance, and therefore the standard deviation, of the count from this

Temporal and Spatio-Temporal Modeling and Monitoring of Epidemic Phenomena

Statistical methods for the modeling and monitoring of time series of counts, proportions and categorical data, as well as for the modeling of continuous-time point processes of epidemic phenomena. The monitoring methods focus on aberration detection in count data time series from public health surveillance of communicable diseases, but applications could just as well originate from environmetrics, reliability engineering, econometrics, or social sciences. The package implements many typical outbreak detection procedures such as the (improved) Farrington algorithm, or the negative binomial GLR-CUSUM method of Höhle and Paul (2008) <doi:10.1016/j.csda.2008.02.015>. A novel CUSUM approach combining logistic and multinomial logistic modeling is also included. The package contains several real-world data sets, the ability to simulate outbreak data, and to visualize the results of the monitoring in a temporal, spatial or spatio-temporal fashion. A recent overview of the available monitoring procedures is given by Salmon et al. (2016) <doi:10.18637/jss.v070.i10>. For the retrospective analysis of epidemic spread, the package provides three endemic-epidemic modeling frameworks with tools for visualization, likelihood inference, and simulation. 'hhh4' estimates models for (multivariate) count time series following Paul and Held (2011) <doi:10.1002/sim.4177> and Meyer and Held (2014) <doi:10.1214/14-AOAS743>. 'twinSIR' models the susceptible-infectious-recovered (SIR) event history of a fixed population, e.g. epidemics across farms or networks, as a multivariate point process as proposed by Höhle (2009) <doi:10.1002/bimj.200900050>. 'twinstim' estimates self-exciting point process models for a spatio-temporal point pattern of infective events, e.g., time-stamped geo-referenced surveillance data, as proposed by Meyer et al. (2012) <doi:10.1111/j.1541-0420.2011.01684.x>. A recent overview of the implemented space-time modeling frameworks for epidemic phenomena is given by Meyer et al. (2017) <doi:10.18637/jss.v077.i11>.

Maintainer: Sebastian Meyer <seb.meyer@fau.de>

Author(s): Michael Höhle*, Sebastian Meyer*, Michaela Paul*, Leonhard Held*, Howard Burkom*, Thais Correa*, Mathias Hofmann*, Christian Lang*, Juliane Manitz*, Andrea Riebler*, Daniel Sabanés Bové*, Maëlle Salmon*, Dirk Schumacher*, Stefan Steiner*, Mikko Virtanen*, Wei Wei*, Valentin Wimmer*, R Core Team* (A few code segments are modified versions of code from base R)

Install package and any missing dependencies by running this line in your R console:

```
install.packages("surveillance")
```

Documentation

Manual: surveillance.pdf
Vignettes:

- twinSIR: Individual-level epidemic modeling for a fixed population with known distances
- hhh4 (spatio-temporal): Endemic-epidemic modeling of areal count time series
- Monitoring count time series in R: Aberration detection in public health surveillance
- hhh4: An endemic-epidemic modelling framework for infectious disease counts
- Getting started with outbreak detection
- algo.glrnb: Count data regression charts using the generalized likelihood ratio statistic
- twinstim: An endemic-epidemic modeling framework for spatio-temporal point patterns

Dependencies Table	Dependencies Graph	Reverse Table

Depends	xtable, polyCub, sp
Imports	spatstat, Rcpp, nlme, MASS, Matrix
Suggests	fanplot, intervals, rgeos, colorspace, maptools, gpclib, splancs, xts, lattice, runjags, ggplot2, scales, testthat, hhh4contacts, spdep, gridExtra, msm, animation, MGLM, quadprog, gsl, numDeriv, knitr, spc, coda, polyclip, maxLik, gamlss, memoise
Enhances	
Linking to	