Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Gesundheit BAG

Automated Spatio-Temporal Outbreak Detection in Low-Count Settings

University of Zurich^UZH

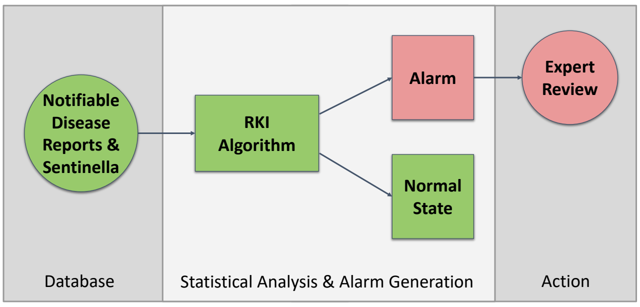
Kelly Reeve[1], Ekkehardt Altpeter[2], and Leonhard Held[1]

[1]University of Zurich, [2]Federal Office of Public Health (FOPH)

Introduction

Public health surveillance: *the ongoing, systematic collection, analysis, and interpretation of health data necessary for timely planning, implementation, and evaluation of public health practice* (Thacker & Berkelman, 1988)

- Modern health surveillance systems consist of both automated and manual processes, including database management, statistical analyses, alarm generation, and actions informed by professional interpretation of the alarms (Hulth et al, 2010)



- Many publications refer to analyses of large geographic regions, highly populated cities or simulations under these conditions (e.g., Enki et al, 2016; Salmon et al, 2016; Mathes et al, 2017)
- However, the population and number of reported disease cases per geographic unit in some European countries can be quite small.

Which spatio-temporal methods are best for public health use in Switzerland, i.e., with low counts?

Algorithms

Farrington algorithm

- Simple regression method that can be applied to many diseases with minimal fine tuning
- Not a truly spatio-temporal method, but can be pragmatically applied to many regions in parallel to mimic one

Steps:

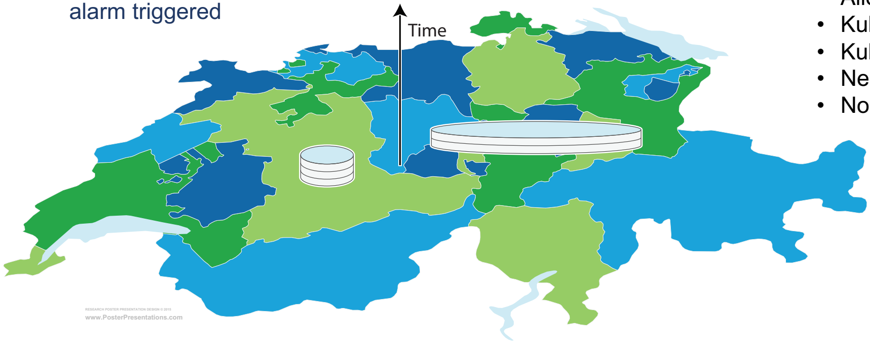
1. Fit a count data model (with seasonality and temporal trend)
2. Compute prediction interval for the expected number of cases
3. If observed number of cases > prediction interval → alarm triggered

Scan statistics

- Identify possible outbreak clusters made up of R regions during the T most recent timepoints.

Steps:

1. Scan over all combinations of contiguous regions and consecutive weeks back
2. Identify space-time clusters with the highest computed statistics (maximum likelihood ratios)
3. If observed statistic in cluster > quantile of MC replicate statistics → alarm triggered



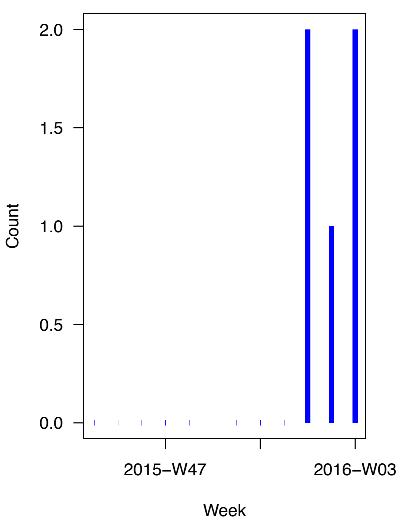
Data

- Database for notifiable diseases and the Sentinella system for common infectious diseases run by the FOPH in Bern, Switzerland
- 3 commonly-reported short-incubation pathogens (salmonella, measles, and meningococcal) and 2 infrequently-reported pathogens (listeria and tularemia)
- Variables: date of diagnosis, location of residence (community, canton) and subtype (if relevant)
- January 1, 1999 to December 31, 2017

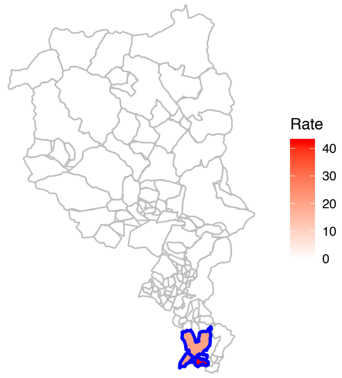
Findings

- Community-level borders change at least yearly → conversion function
- Extremely low counts at the community level → change models used
- No software to aid in scan statistic interpretation → table/map functions
- Alarms when 0 counts observed → automatically suppress them
- Higher computing power and time → prepare users
- Different pathogens behave differently → pathogen-specific settings

Cases in Cluster 1 in Past 12 Weeks



3 week period ending on 2016–W03 (Cluster 1 in blue)



Conclusions

- Scan statistics help identify more precise outbreak locations, even in low count settings, even some missed by older methods
- One algorithm (including parameter settings) does not fit all pathogens!
- Working in R allows for flexible extension in functionality to meet specific needs

References

- Allevius B. & Hoehle, M. (2019) *Scand. J. Stat.* **46**(1).
- Kullforff, M. (2001) *J. Royal Stat. Soc. A* **164.**
- Kulldorff, M., et al. (2005) *PloS Medicine* **2**(3).
- Neill, D. B. (2009) *Int. J. Forecast*. **25**(3).
- Noufaily, A., et al. (2012) *Stat. Med.* **32**(7).

Contact

Kelly is currently a PhD student of Biostatistics at the EBPI at University of Zurich. With funding from the project PrecisionMS, she will pursue topics related to causal inference in clinical trials and observational studies with the aim of improving research and clinical practice related to multiple sclerosis.

kelly.reeve@uzh.ch