A new statistical method to detect disease outbreaks from hospital ED data

Jin Yoon & Justin Boyle | 9 July 2023

Australian e-Health Research Centre

Outline

Why is early detection of disease outbreak important?

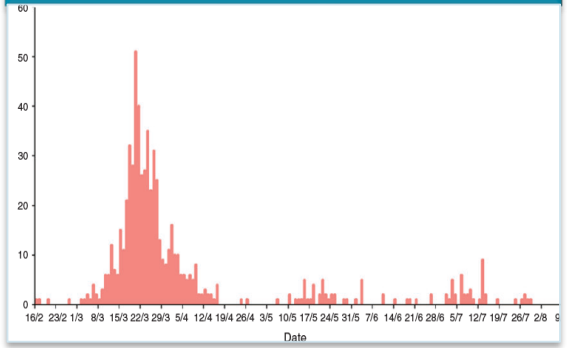
Traditional approach to disease surveillance

Novel data-driven methods for disease surveillance

Validation and conclusion – real data analysis

Disease surveillance: Introduction

2020 was a devastating year for global health



Countries must strengthen preparedness for future pandemics and other emergencies

- COVID-19 pandemic
  - dramatic loss of human life worldwide
- Unprecedented challenge to public health

Early detection of disease outbreak:

- Resilience in health system
- Efficient resource planning
  - to save lives
  - reduce healthcare costs

Graph of confirmed COVID-19 cases in Western Australia (graph data from https://experience.arcgis.com/experience/359bca83a1264e3fb8d3b6f0a028d768, accessed 13 July 2020)

Disease Surveillance: Traditional Approach

Monitor anomalies in reported outbreak cases

(e.g., community-reported ILI from FluTracking and Healthdirect)

Time-consuming data capture and handling:

gather → process → report → release surveillance data through WHO

Statistical methods to improve surveillance systems:

- Improves outbreak detection accuracy
- Minimises false alarms

Disease Surveillance: Novel Method

A new index measure combining concepts from finance and TBE

Data source:

routinely collected data from public hospital emergency departments

Leveraging the dataset and the developed index measure

Disease Surveillance: Novel Index Definition

T_i: the daily median TBE (open, close, low, and high TBE for day i)

V_{ik}: logarithm difference, [log(T_{i+j}) – log(T_i)] from j=1 to k

Index (signal and outbreak)

- Variability index measure S_i, $S_i = \sum_v \{v \in V_{ik} : v < -h\}$
- The Q_n estimator, $Q_n = D_n \cdot \{|V_i - V_j|; i < j\}_{(k)}$
- A new outbreak index, O_i

Determine whether it is outbreak or not

- The minimum consecutive days, d
- The total sum of O_i from day i to day $i + k$

Validation: Emergency Department Data

Emergency Department data from 4 Hospitals and Health Services (HHS)

(Queensland, Australia)

Ground truth: the commencement of the first wave of COVID-19

Different models by varying parameters:

number of periods, target margin, minimum consecutive days, threshold

Results

K	I	Threshold(h, C) Or C*	Hospital and Health Service (HHS)			
			Gold Coast	Metro South	Metro North	Central Queensland
Ground truth (COVID-19)			09/03/2020	11/03/2020	10/03/2020	02/03/2020
7	S_i	(0.3, -2.1)	24/02/2020	18/01/2020	17/01/2020	05/03/2020
		(0.1, -0.7)	02/01/2020	08/01/2020	14/01/2020	05/01/2020
		(0.2, -1.4)	02/01/2020	18/01/2020	14/01/2020	21/02/2020
		C^*	06/03/2020	18/01/2020	17/01/2020	05/03/2020
	Q_n	(0.3, 2.1)	-	-	-	-
		(0.1, 0.7)	23/01/2020	02/01/2020	15/01/2020	06/01/2020
		(0.2, 1.4)	07/03/2020	-	-	05/08/2020
		C^*	25/02/2020	04/05/2020	20/01/2020	21/05/2020

Results

K	I	Threshold(h, C) Or C*	Hospital and Health Service (HHS)			
			Gold Coast	Metro South	Metro North	Central Queensland
Ground truth (COVID-19)			09/03/2020	11/03/2020	10/03/2020	02/03/2020
10	S_i	(0.3, -2.1)	21/02/2020	18/01/2020	14/01/2020	05/03/2020
		(0.1, -0.7)	13/01/2020	08/01/2020	13/01/2020	02/01/2020
		(0.2, -1.4)	02/01/2020	18/01/2020	13/01/2020	05/01/2020
		C*	03/03/2020	18/01/2020	10/01/2020	21/02/2020
	Q_n	(0.3, 2.1)	-	-	-	-
		(0.1, 0.7)	02/01/2020	02/01/2020	09/01/2020	06/01/2020
		(0.2, 1.4)	06/03/2020	-	20/01/2020	04/08/2020
		C*	24/02/2020	23/01/2020	16/01/2020	22/02/2020

Results

K	I	Threshold(h, C) Or C*	Hospital and Health Service (HHS)			
			Gold Coast	Metro South	Metro North	Central Queensland
Ground truth (COVID-19)			09/03/2020	11/03/2020	10/03/2020	02/03/2020
15	S_i	(0.3, -2.1)	13/01/2020	12/01/2020	08/01/2020	05/01/2020
		(0.1, -0.7)	13/01/2020	08/01/2020	08/01/2020	02/01/2020
		(0.2, -1.4)	02/01/2020	10/01/2020	08/01/2020	02/01/2020
		C*	26/02/2020	24/02/2020	14/01/2020	26/02/2020
	Q_n	(0.3, 2.1)	-	-	-	-
		(0.1, 0.7)	14/01/2020	02/01/2020	10/01/2020	07/01/2020
		(0.2, 1.4)	02/03/2020	-	-	08/08/2020
		C*	27/02/2020	19/02/2020	12/01/2020	13/05/2020

Conclusion & Next steps

Aim:

- Develop new variability indices and propose data-driven methodology

Application:

- Method **successfully signalled outbreak a few days before pandemic in two HHSs**
- potential for early warning to health agencies.

Limitation:

- method also flagged outbreaks well in advance of COVID-19 cases in two other HHSs
- possibly due to demographic factors or poor ILI-COVID-19 correlation

Further research:

- application of novel method in additional settings to **build more resilient health systems**

Health Intelligence @ CSIRO AEHRC

Health System Productivity & Efficiency
Operational & Clinical Decision Support
Evidence-driven Policy & Healthcare Delivery

Thank you

Health & Biosecurity
Jin Yoon
Senior Research Scientist

+61 2 9662 7159
Jin.yoon@csiro.au
aehrc.csiro.au

Health & Biosecurity
Justin Boyle
Senior Principal Research Scientist

+61 7 3253 3606
Justin.boyle@csiro.au
aehrc.csiro.au

Australian e-Health Research Centre