



Public Health
England

Protecting and improving the nation's health

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



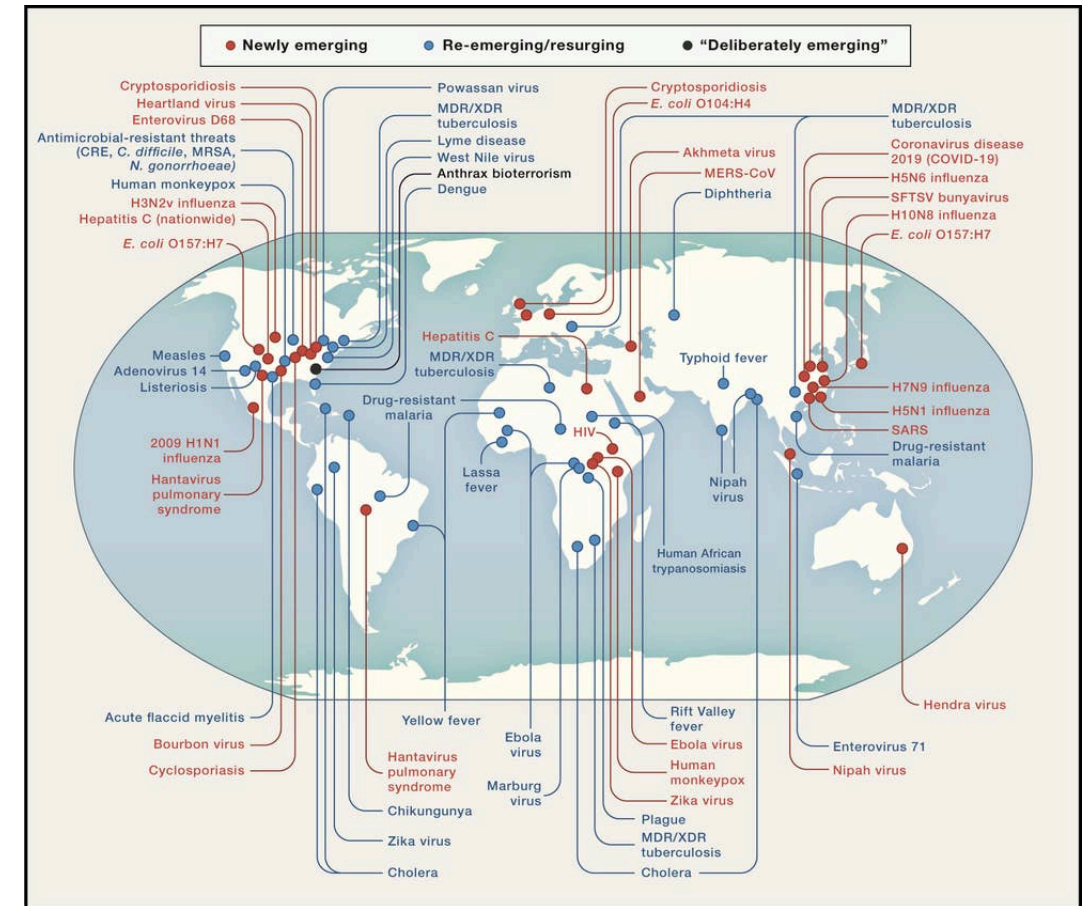
Data Science in the Field

David Kennedy,
Data Scientist, UK-PHRST and Research Fellow, LSHTM

The role of data during an outbreak

We can use data to generate evidence for answering the key epidemiological questions:

- **Who?**
 - Who is the population at risk
- **What?**
 - What is happening to them?
- **Where?**
 - Where are the affected people?
- **When?**
 - When did this begin?



Morens DM, Fauci AS. Emerging Pandemic Diseases: How We Got to COVID-19 [published online ahead of print, 2020 Aug 15]. Cell. 2020;S0092-8674(20)31012-6. doi:10.1016/j.cell.2020.08.021

Lots to consider!

The image shows a handwritten data table on a grid, organized into columns for dates: 12/08, 13/08, 14/08, and 15/12/2018. Each date column has sub-columns labeled 'REP', 'NORM', 'A', 'VIS', and 'SORTS'. The data is written in red and black ink. The table is divided into two main sections by a horizontal line. The top section contains data for the first four dates, while the bottom section contains data for the same dates but with different values. The data is organized into rows, with some rows having multiple entries for the same date. The table is a complex grid of numbers and dates, with some cells containing multiple values or symbols.

12/08					13/08					14/08					15/12/2018				
REP	NORM	A	VIS	SORTS	REP	NORM	A	VIS	SORTS	REP	NORM	A	VIS	SORTS	REP	NORM	A	VIS	SORTS
0	19	19	0	19	0	19	19	1	18	0	18	18							
0	16	16	0	16	0	16	16	0	16	0	16	16							
0	75	71	0	75	0	75	75	0	75	57	132	132							
2	120	122	0	122	7	122	122	0	122	0	122	122							
0	74	74	0	74	0	74	74	0	74	0	74	74							
0	8	8	0	8	0	8	8	0	8	0	8	8							
0	2	2	0	2	0	2	2	0	2	0	2	2							
0	70	70	0	70	0	70	70	0	70	0	70	70							
0	3	3	0	3	0	3	3	0	3	0	3	3							
0	13	13	0	13	0	13	13	0	13	0	13	13							
0	4	4	0	4	0	4	4	0	4	0	4	4							
0	5	5	0	5	0	5	5	0	5	0	5	5							
0	3	3	0	3	0	3	3	0	3	0	3	3							
0	3	3	0	3	0	3	3	0	3	0	3	3							
12	20	20	0	20	0	20	20	0	20	12	32	32							
0	0	0	0	0	0	0	0	0	0	0	0	0							
499	4	503	40	503	7	510	520	0	509	141	650	650							
80	0	80	80	80	0	80	80	0	80	0	80	80							
73	0	73	73	73	0	73	73	0	73	0	73	73							
21	0	21	21	21	0	21	21	0	21	0	21	21							
174	10	174	174	174	0	174	174	0	174	174	168	168							

- Is there an existing surveillance system in place? Is it functional?
- What type of event? What do we need to monitor? Is the tool used for the right purpose?

Lots to consider!



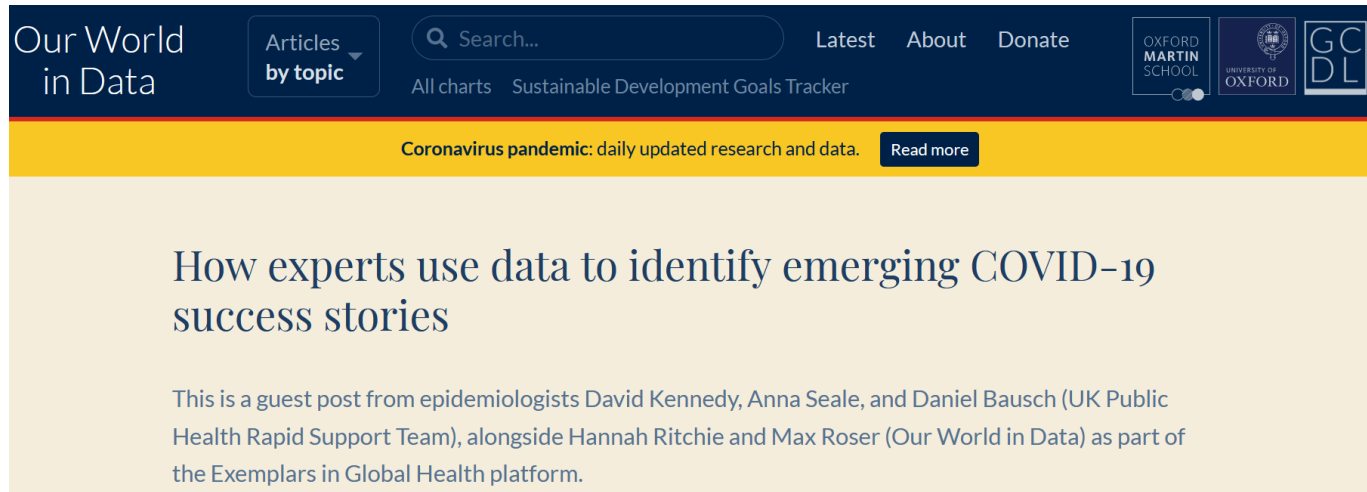
- What are the resource implications (costs, human resources) for set up and training?
- How quickly can it be rolled out?
- Is there technical capacity to implement and maintain the system/tool?
- Can it be easily adapted? Does it use open source software?

Lots to consider!



- What are the barriers to getting good data?
- Data security: Where is the data stored?
Where is the data sent to?
- Data visualisation: How can the interpretation of the data be communicated to the target audience?

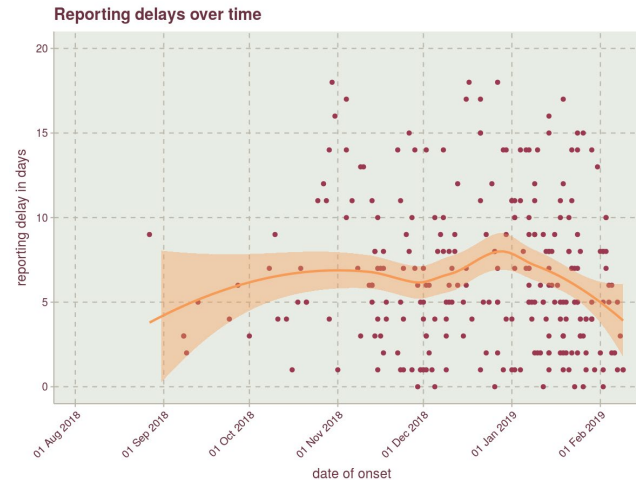
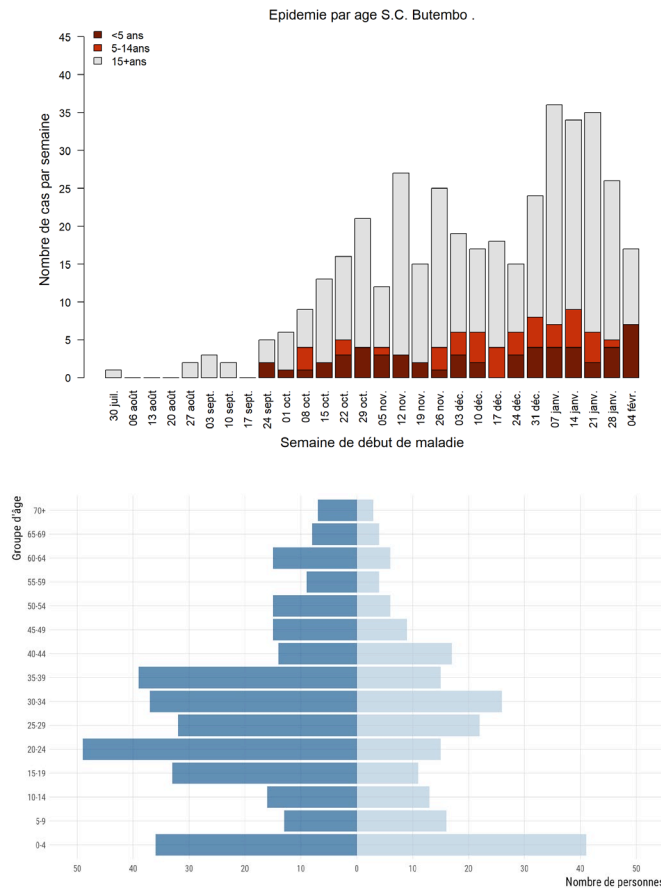
Using data for outbreak research



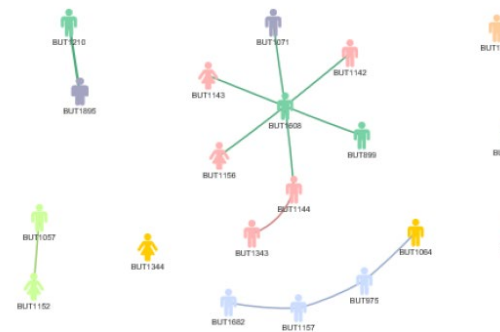
- Work with partners to identify knowledge gaps
- Develop protocols to answer questions
- Analyse a variety of data sources to build up a clear picture

Challenges in data analytics

EVD outbreak, North Kivu, DRC, Aug – Feb 2018



- Excel, STATA, ArcGIS, R, ...
- Often ad-hoc and person-specific
- Sustainability using R remains difficult
- Automation, but do not forget epidemiological interpretation
- Automation: data integration, cleaning



Questions on the roll-out of electronic tools

- Is there an existing surveillance system in place? Is it functional?
- What type of event? What do we need to monitor? Is the tool used for the right purpose?
- What are the resource implications (costs, human resources) for set up and training?
- How quickly can it be rolled out?
- Is there technical capacity to implement and maintain the system/tool?
- Can it be easily adapted? Does it use open source software?
- Can it be easily integrated with other surveillance tools, platforms?
- Data security: Where is the data stored? Where is the data sent to?



Public Health
England

Protecting and improving the nation's health

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Thank you for joining

Data Science in the Field

David Kennedy,
Data Scientist, UK-PHRST and Research Fellow, LSHTM