Big Brother Is Watching : Using Digital Surveillance Tools for Near Real-Time Mapping of the Risk of International Infectious Disease Spread

Sangeeta Bhatia, Anne Cori and Pierre Nouvellet

In our increasingly interconnected world, it is crucial to understand the risk of an outbreak originating in one country/region and spreading to the rest of the world. Rapid recognition and response to potential pandemics and emerging diseases have become essential global health priorities. Digital disease surveillance tools such as ProMed and HealthMap have the potential to serve as important early warning systems as well as complement the field surveillance data during an ongoing outbreak. While there are a number of systems that carry out digital disease surveillance, there is as yet a lack of tools that can compile and analyse the generated data to produce easily understood actionable reports. We present a flexible statistical model that uses different streams of data (such as disease surveillance data, mobility data etc.) for short-term incidence trend forecasting.

In validating the model using data collected by ProMED and HealthMap during the 2014-2016 West African Ebola outbreak, we provide a realistic appraisal of the strengths and limitations of such data in incidence forecasting. We infer incidence trends at finer spatial scales from aggregated data. Our work shows how the data from event based surveillance systems (EBS) can complement the data collected from traditional public health infrastructure. During an ongoing crisis, combining data from different sources gives stakeholders a more complete picture.

1   Introduction

other tools that do similar stuff - EpiDMS (Liu et al., 2016) slightly old paper - authors curated news themselves! (Chowell et al., 2016)

2   Results

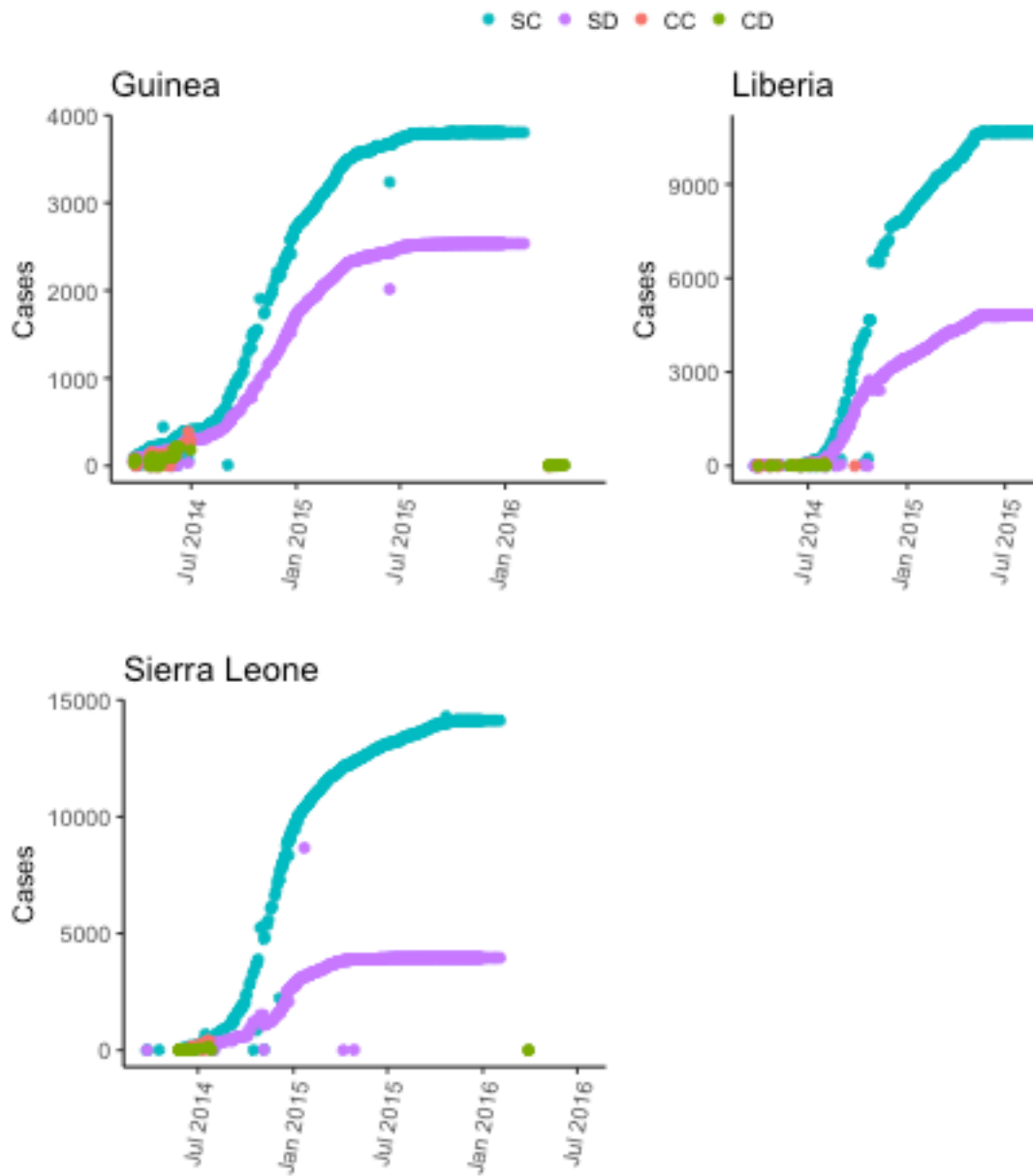The raw data from ProMED is shown in Figure 1

Figure 1: Raw data from HealthMap

2.1 Cleaning ProMED/HealthMap data feeds

The data cleaning consists of the following steps : extracting the total case counts as a sum of desired case categories, merging duplicate alerts, removing data that made the cumulative case counts inconsistent and finally, fill in the missing data using interpolation. Results at each step pre-processing workflow for Sierra Leone are show in Figure 2.

2.2 Comparison of incidence trends from different sources

References

Chowell, G., Cleaton, J. M., and Viboud, C. (2016). Elucidating transmission patterns from internet reports: Ebola and middle east respiratory syndrome as case studies. *The Journal of Infectious Diseases*, 214(suppl_4):S421–S426.

Liu, S., Poccia, S., Candan, K. S., Chowell, G., and Sapino, M. L. (2016). epidms: Data management and analytics for decision-making from epidemic spread simulation ensembles. *The Journal of Infectious Diseases*, 214(suppl_4):S427–S432.
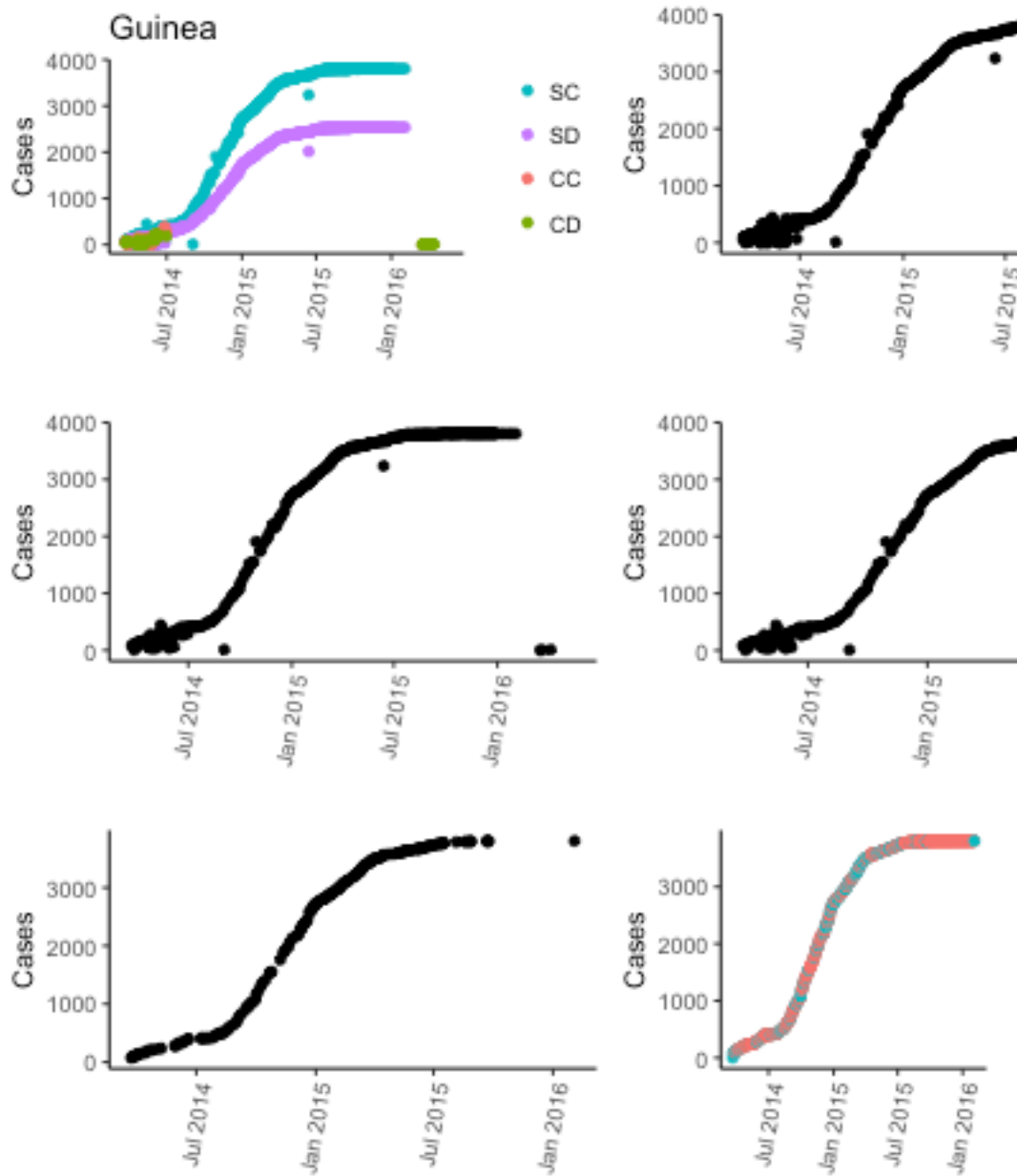
Figure 2: Data clean-up steps