Data Privacy in Low Density Public Health Environments

October 27, 2020

Aaron Katz
Johns Hopkins University Applied Physics Laboratory
The Need for Actionable Public Health Data

A need for open public data standards and sharing in light of COVID-19

The disjointed public health response to the COVID-19 pandemic has demonstrated one clear truth: the value of timely, publicly available, high-quality data on emerging infectious and notifiable diseases in real-time is established. The data should be made available at a spatial and temporal scale that is granular enough to prove useful for planning and modelling purposes. Additionally, a critical component of the proposed system is the democratization of data; all collected information (observing necessary privacy standards) should be made publicly available immediately upon release, in machine-readable formats, and based on open data standards.

“Moving forward, it is imperative that a standardized reporting system for systematically collecting, visualizing, and sharing high-quality data on emerging infectious and notifiable diseases in real-time is established. The data should be made available at a spatial and temporal scale that is granular enough to prove useful for planning and modelling purposes. Additionally, a critical component of the proposed system is the democratization of data; all collected information (observing necessary privacy standards) should be made publicly available immediately upon release, in machine-readable formats, and based on open data standards.”

A need for open public data standards and sharing in light of COVID-19

L. Gardner, J. Ratcliff, E. Dong, A. Katz

August 10, 2020

https://doi.org/10.1016/S1473-3099(20)30635-6
What Data?

Electronic Health Records

Aggregate Data
What Issues?
Numeric Resolution

NSSP BioSense Platform - Virginia User Policy

In the event that Virginia users of the NSSP BioSense Platform wish to share the results of a Permitted Data Use outside their organization, they may request VDH permission to publish the results using the Publication Request Form. The following criteria must be met in order for VDH to approve publication:

- All data with a count greater than zero and less than five are suppressed
- No medical record numbers, visit event identifiers, patient identifiers, insurance IDs, or any other person-specific identifiers are included
- No line-level data are included
- No names of reporting facilities are included (see section on Individual and Entity Identification)

Publication permission shall be sought for finished projects only using the Publication Request Form. The requester will receive a response from VDH within one week regarding the status of their request. A request to publish is not a guarantee of approval. VDH will maintain the ability to deny or halt
How do these changes impact the numbers?

Implemented guidelines to protect privacy and misinterpretation of data means that we are reporting case only for communities with > 1,000 residents. As a result, a couple changes occurred:

2. If communities are <1,000 residents, they are now being reported at the Borough/Census Area level.
   - One case now is displayed in the Yukon-Koyukuk area instead of being combined in the Fairbanks counts.
My Questions

How do we balance legitimate privacy concerns against the need for open data to support response efforts in a public health crisis?

How do we strike this balance responsibly in low density environments (e.g. during the early phases of an outbreak, in remote and rural settings)?