OJPHI

ISDS 2013 Conference Abstracts

ISDS INTERNATIONAL SOCIETY FOR DISEASE SURVEILLANCE

Biosurveillance Data Stream Framework: A Novel Approach to Characterization and Evaluation

Kristen Margevicius*, Eric Generous, Kirsten Taylor-McCabe, Mac Brown, W. Brent Daniel, Lauren Castro, Andrea Hengartner and Alina Deshpande

Los Alamos National Laboratory, Los Alamos, NM, USA

Objective

To develop a data stream-centric framework that can be used to systematically categorize data streams useful for biosurveillance systems, supporting comparative analysis

Introduction

Multiple data sources are used in a variety of biosurveillance systems. With the advent of new technologies, globalization, high performance computing, and 'big data' opportunities, there are seemingly unlimited potential data streams that could be useful in biosurveillance. Data streams have not been universally defined in either the literature or by specific biosurveillance systems. The definitions and framework that we have developed enable a characterization methodology that facilitates understanding of data streams and can be universally applicable for use in evaluating and understanding a wide range of biosurveillance activities- filling a gap recognized in both the public health and biosurveillance communities.

Methods

Infectious disease biosurveillance terminology, types of data streams used (past, current or considered for future use) in infectious disease biosurveillance systems, and operational biosurveillance systems were identified through web and literature searches and through information provided by a subject matter expert (SME) panel and others in the biosurveillance community. The SME panel represented experts involved in animal, plant, and human infectious disease biosurveillance from U.S. federal government agencies, national laboratories, and academic institutions. Data streams were identified and common characteristics and attributes were cataloged. As the framework developed, the collected data streams were binned back into the framework classification which informed the development of the data stream categories first by suggesting what the categories should be and second by validating that any data stream was indeed classifiable.

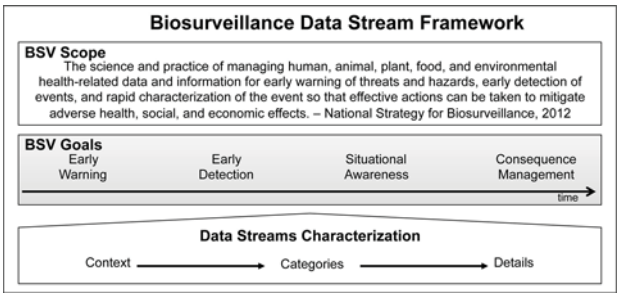
Results

An overview of the data stream framework is shown (see figure). Sixteen broad data stream categories were identified. Population, Type, and Data Stream categories characterize the kind of information that is being collected, and how it could impact biosurveillance. Also associated with each individual data stream are data stream details. All of these descriptors inform the quality and usefulness of the specific data stream. Within the scope of biosurveillance data streams are put into context by specifically defined biosurveillance goals. We were able to successfully apply our characterization schema to both traditional and non-traditional data streams that are currently being considered for use in disease surveillance as identified by the SME panel, found in the literature, and used by surveillance systems.

An advantage of deconstructing data streams is in the ability to discuss data streams in broad terms yet still retain the detailed information that may be very important regarding actual utility of specific data sources. We used this framework in building a relational database (The Biosurveillance Resource Directory, BRD) to categorize information on operational infectious disease biosurveillance systems. From 115 systems cataloged 272 data streams were recorded and tallied. Data streams representing 14 of our 16 categories were used in at least one of these systems. The most commonly used data streams were laboratory records and clinic/health care provider records. As the BRD continues to be updated, using our framework we can monitor the use of different data streams for different diseases, locations and purposes.

Conclusions

The development of the Biosurveillance Data Stream Framework resulted from a systematic and comprehensive analysis of data streams used or potentially useful for biosurveillance. The resulting framework puts data streams into context. The foundation underlying this framework is both novel, because it has not been done before as comprehensively, iteratively, or categorically, and timely, because the many individual data streams that might by considered for biosurveillance can now be systematically categorized to enable comparison and analysis.



Keywords

data stream; biosurveillance; infectious disease

Acknowledgments

This project is supported by the Chemical and Biological Technologies Directorate Joint Science and Technology Office (JSTO), Defense Threat Reduction Agency (DTRA).

*Kristen Margevicius
E-mail: kmargevicius@lanl.gov