

A Systematic Evaluation of Data Streams for Global Disease Surveillance

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

Defense Systems Analysis Division, Los Alamos National Laboratory, Los Alamos, NM, USA

Objective

The overall objective of this project is to provide a robust evaluation of data streams that can be leveraged from existing and developing national and international disease surveillance systems, to create a global disease monitoring system and provide decision makers with timely information to prepare for and mitigate the spread of disease.

Introduction

Living in a closely connected and highly mobile world presents many new mechanisms for rapid disease spread and in recent years, global disease surveillance has become a high priority. In addition, much like the contribution of non-traditional medicine to curing diseases, non-traditional data streams are being considered of value in disease surveillance. Los Alamos National Laboratory (LANL) has been funded by the Defense Threat Reduction Agency to determine the relevance of data streams for an integrated global biosurveillance system through the use of defined metrics and methodologies. Specifically, this project entails the evaluation of data streams either currently in use in surveillance systems or new data streams having the potential to enable early disease detection. An overview of this project will be presented, together with results of data stream evaluation. This project will help gain an understanding of data streams relevant to early warning/monitoring of disease outbreaks.

Methods

Three specific aims were identified to address the overall goal of determining the relevance of data streams for global disease surveillance. First, identify data streams as well as define metrics for the evaluation. Second, evaluate data streams using two different methodologies, decision analysis modeling using a support tool called Logical Decisions® that assigns utility scores to data streams based on weighted metrics and assigned values specific to data stream categories; and a Surveillance Window concept developed at LANL that assigns a window or windows of time specific to a disease within which information coming from various data streams can be determined to have utility. This would obtain a ranked list of useful data streams. Additionally, evaluate data integration algorithms useful for a global disease surveillance system through a review of scientific literature. Finally, validate the top-ranked data streams by application of specific historical outbreaks to determine whether the data streams are capable of providing early warning or detection of the particular disease before it became a large outbreak.

Results

Seventeen categories of data streams were identified that ranged from traditional ones such as clinic/healthcare provider and laboratory records to newly emerging sources of information such as social media and internet search queries. The Logical Decisions® based evaluation of data streams identified 5 data streams that consistently showed utility regardless of the goal of biosurveillance. However, different data streams varied in rank, given different biosurveillance goals, and there is no one top ranked data stream. Surveillance window based evaluation of data streams during disease outbreaks identified data streams that had high utility for early detection and early warning regardless of disease, while others were more disease and operations specific. Additionally, we have built a searchable biosurveillance resource directory that houses information on global disease surveillance systems.

Conclusions

LANL has developed a robust evaluation framework to determine the relevance of various traditional and non-traditional data streams in integrated global disease surveillance. Through the use of defined surveillance goals, metrics and data stream categories, not only have we identified data streams currently in use that have high utility, but also new data streams that could be exploited for the early warning/monitoring of disease outbreaks. Our robust evaluation framework facilitates the identification of a defensible set of options for decision makers to use to prepare for and mitigate the spread of disease.

Keywords

evaluation; disease surveillance; data streams

Acknowledgments

We would like to acknowledge the Defense Threat Reduction Agency (DTRA)-Joint Science and Technology Office (JSTO) for their support and guidance on this project.

References

Dr. Alina Deshpande, Dr. Mac Brown, Ms. Lauren Castro, Dr. William Brent Daniel, Mr. Eric Nicholas Generous, Ms Andrea Hengartner, Dr. Kristen Margevicius, Dr. Kirsten Taylor-McCabe, Los Alamos National Laboratory, Los Alamos, NM 87545.

*Alina Deshpande

E-mail: deshpande_a@lanl.gov

