An ISDS-Based Initiative for Conventions for Biosurveillance Data Analysis Methods: Background and Motivation

Michael Coletta\textsuperscript{1}, Howard Burkom\textsuperscript{2}, Jeffrey Johnson\textsuperscript{3}, Wendy Chapman\textsuperscript{4}

\textsuperscript{1}National Assoc. of County & City Health Officials, Washington, DC.
\textsuperscript{2}Johns Hopkins Applied Physics Laboratory, Laurel, MD
\textsuperscript{3}San Diego County Health & Human Services Agency, San Diego, CA
\textsuperscript{4}University of California San Diego, San Diego, CA

2012 Conference of the International Society for Disease Surveillance Panel Session

San Diego, California December 5, 2012
Motivation

• Early biosurveillance systems were designed by technology developers, not public health users
• Ongoing mismatch between technical methods in academia and PH surveillance needs
• Methods for early alerting, follow-up investigation, threat characterization need to be tuned to health monitor’s purview and needs
• In surveillance system development, IT issues, not algorithms, get priority
Essential Task Interaction in Volatile Data Background

**DATA**

**Medical/Epidemiological**
- filtering/classifying clinical records => syndromes
- interpretation/response to system output
- coding/chief complaint interpretation

**Information Technology**
- surveillance system architecture
- data ingestion/cleaning
- interface between health monitors and system

**Analytical**
- Statistical hypothesis tests
- Data mining/automated learning
- Adaptation of methodology to background data behavior
Essential Task Interaction in Volatile Data Background

**DATA**

- **Medical/Epidemiological**
  - filtering/classifying clinical records => syndromes
  - interpretation/response to system output
  - coding/chief complaint interpretation

- **Information Technology**
  - surveillance system architecture
  - data ingestion/cleaning
  - interface between health monitors and system

- **Analytical**
  - Statistical hypothesis tests
  - Data mining/automated learning
  - Adaptation of methodology to background data behavior
Technical Challenges from the Public Health Practice Community to Statisticians and Analysts

Background:

On May 28th, 2009, the ISDS Research and Public Health Practice Committees hosted a joint panel with the goal of bringing current challenges faced by public health practitioners to the attention of the research community at large. Members of both Committees expressed concern that much current research in disease surveillance has little application for public health practitioners. With an increasing emphasis on health information technology and exchange, public health practitioners need relevant, understandable analytic tools to manage information and make it useful. It was hoped that the invited panelists would be able to discuss concrete technical problems within the current (and future) constraints of practitioners.

This webinar was one of two initiatives to improve the coordination between public health research and practitioners. A special panel session was also held at the 2009 Annual APHA Conference in New Orleans.

A call for abstracts asked submitters to present descriptions of near-term and upcoming projects that address the constraints of epidemiologists doing public health surveillance, such as:

- What analytical tools are needed?
- What objectives and constraints must shape these tools?
- What is the current work environment, and what changes are ahead?

May 2009 ISDS Webinar Presenters

- Julia Gunn
  Boston Public Health Commission
- [Aaron Kite-Powell](#)
  Florida Department of Health
- Michael Coletta
  Virginia Department of Health
- Louise Wilson
  Pandemic Influenza Co-ordination Team
  Edinburgh, Scotland
- Marc Paladini
  New York City Dept. of Health and Mental Hygiene
Detection of an Outbreak of Gastroenteritis in Montgomery County Expedites Public Health Response

Kathy J. Hurt-Mullen, MPH; Christine Lacey, RN, MSN; Sue Wilby, RN; Lynn Frank, MHA, FACHE; Paul Fiumara, RS

Montgomery County Department of Health and Human Services
Example: Detecting Time-of-Arrival Clusters

Emergency Department-based Syndromic Surveillance: Outbreak Detection

Aaron Kite-Powell, Florida Department of Health
International Society of Disease Surveillance Webinar, May 2009

- Opportunity: develop additional analyses that:
  - Look for clustering by the time the patients arrive at the ED.
  - By hospital
  - By patient zip code
  - End result could include detection of smaller outbreaks
Example: Detecting Time-of-Arrival Clusters

A Collaboration to Enhance Detection of Disease Outbreaks Clustered by Time of Patient Arrival

Howard Burkom¹, Wayne Loschen¹, Richard Wojcik¹, Liane Ramac-Thomas¹, Randall Arvizu¹, Chris Lee¹, Aaron Kite-Powell²

¹ Johns Hopkins University Applied Physics Laboratory
National Security Technology Department
² Florida Department of Health, Bureau of Epidemiology

International Society for Disease Surveillance
2010 Annual Conference
Park City, Utah, Dec 2, 2010
Example: Detecting Time-of-Arrival Clusters

Example: Time-of-Arrival Application (2005)

Example: Time-of-Arrival Application (2008-9)


Example: Time-of-Arrival Application (2011)

Finding Time-of-Arrival Clusters of Exposure-Related Visits to Emergency Departments in Contiguous Hospital Groups

Lana Deyneka¹, Zhiheng Xu², Peter Hicks³, Howard Burkom³, Stephen Benoit³, Heather Vaughan-Batten¹, Amy Ising⁴

¹ North Carolina Division of Public Health, ²McKinsey Consulting
³ Centers for Disease Control and Prevention
⁴ University of North Carolina, Chapel Hill
Example: Detecting Time-of-Arrival Clusters

Example: Time-of-Arrival Application (2005)
Example: Time-of-Arrival Application (2008-9)
Example: Time-of-Arrival Application (2011)
Example: Time-of-Arrival Application (2012)

Time of Arrival Analysis in NC DETECT to Find Clusters of Interest from Unclassified Patient Visit Records

Meichun Li¹, Wayne Loschen², Lana Deyneka³, Howard Burkom², Amy Ising¹ and Anna Waller¹
Could the entire development process have been streamlined?

- Problem definition more precise
- More PH involvement in problem specification
- Broader involvement of potential developers
- Broader availability of solution?
Analogy to Cultural Divide in Oncology, ~1970


- “the conference epitomized the … segregation between cancer therapy and cancer science”
- “Chemotherapy and surgery were discussed in one room...carcinogenesis in another”
- “Few scientists...crossed between the two isolated worlds”
- ...”prevailing schizophrenia of the time...”
Establishing Technical Conventions

Vision:

- ISDS Technical Conventions Committee
- Three Core Functions
  1. Problem Specification
  2. Dataset Curation
     - Unrestricted data, and restricted by Data Use Agreements
     - Authentic, “modified” Authentic, Simulated
  3. Solution Validation
- Post, disseminate products of each core function
- An active, broadly inclusive committee devoted to building, sustaining a bridge between problem owners and solution developers
  - A “shell” is insufficient for sustainable communication
  - More than a website or social media tool
Example: Problem Specification

- Problem owner (epidemiologist, department) desiring a technical tool fills out problem template as completely as possible.
- Problem owner sends the template to the conventions group, and a call is scheduled.
- The template is refined and completed during the call, with the user requirements governing the process.
- Provision of datasets for validation, comparison of solutions methods is arranged.
- Completed template is made available to the community of developers in health departments, other government agencies, universities, industry.
Situational Awareness has Many Features, Varying by Institution & Jurisdiction
(from US Medicine Institute meeting, 2007)

Reported benefits include:
• Corroboration of clinical suspicions
• Tracking spread of outbreaks, identifying likely population at risk
• Assessing the likely disease burden of seasonal influenza
• Monitoring of large public events
• Tracking effects of natural disasters
• Explaining away/ruling out/rumor control
e.g., determining that a school-based outbreak is not spreading
• Investigating effects of severe weather
Potential Advantages

• Accessibility of technical issues facing public health to academic, commercial, government research communities

• Strengthen capability to acquire shareable datasets to enable replicable research, thus enable standard methods
  – Ability to answer: “Exactly what do you need our data for?”

• Globalization of disease surveillance research
  – Access to technical issues faced in resource-limited settings
  – Ability to transcend boundaries of hardware/software environments

• Focus on methods, not systems or funded programs
  – Enable technical dialogue outside context of proprietary systems
  – Systems, programs could be enriched by independent methods research

• A way for “biosurveillance” to evolve based on grass roots needs
Reactions from Prominent Academic Researchers and Mentors

• ‘I think it is vital to have these datasets available. We would jump back into the area if they were.’
  – William Woodall, Virginia Tech

• ‘I completely support the functions of...an "interface" committee to help facilitate translation from research to practice.’

• ‘Similar problems arose in the early days of semiconductor chip manufacture and microarray data. The people who owned the data could not share it, for proprietary reasons. Eventually, several consortia organized that published sanitized data for use in testing statistical methods. It was a necessary step forward, and did good for the respective fields.... Yes, I think [this approach] is viable.’
  – David Banks, Duke University
Toward Well-Defined Surveillance Science

To the challenge:

“If syndromic surveillance is the answer, what is the question?”


Biosecur Bioterror. 1(2):77-81.

-- the proposed conventions group would furnish a means to define and collaborate on precise questions of practical importance and thus flesh out the effective role of advanced disease surveillance.