Innovations in Public Health Surveillance: The National Bioterrorism Syndromic Surveillance Demonstration Project

Ross Lazarus MB, BS MPH
Director of Bioinformatics,
Channing Laboratory
Brigham and Women’s Hospital and Harvard Medical School

June 10, 2003
# Innovations in Public Health Surveillance: The National Bioterrorism Syndromic Surveillance Demonstration Project

---

**1. REPORT DATE**
10 JUN 2003

**2. REPORT TYPE**
N/A

**3. DATES COVERED**
-

**4. TITLE AND SUBTITLE**
Innovations in Public Health Surveillance: The National Bioterrorism Syndromic Surveillance Demonstration Project

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**6. AUTHOR(S)**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Channing Laboratory; Brigham and Womens Hospital; Harvard Medical School

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

**10. SPONSOR/MONITOR’S ACRONYM(S)**

**11. SPONSOR/MONITOR’S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release, distribution unlimited

**13. SUPPLEMENTARY NOTES**
See also ADM001576., The original document contains color images.

**14. ABSTRACT**

**15. SUBJECT TERMS**

**16. SECURITY CLASSIFICATION OF:**

<table>
<thead>
<tr>
<th>a. REPORT</th>
<th>b. ABSTRACT</th>
<th>c. THIS PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>unclassified</td>
<td>unclassified</td>
<td>unclassified</td>
</tr>
</tbody>
</table>

**17. LIMITATION OF ABSTRACT**
UU

**18. NUMBER OF PAGES**
23

**19a. NAME OF RESPONSIBLE PERSON**

---

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18
Outline

♦ Electronic records for surveillance
♦ Syndromic surveillance for BT
♦ HPHC/MDPH System
♦ National Demonstration Project
♦ Issues and challenges
♦ Lessons
Electronic records for surveillance

♦ BT - timeliness is crucial
♦ EMR potentially “real time” with Dx
♦ Increasingly available
  ♦ Large group practices, HMOs
♦ Low marginal cost
  ♦ Already in a database somewhere
♦ “Value add” - eg BT surveillance
Syndromic Surveillance for BT

- First signs - non-specific prodrome
- ICD highly specific/granular
  - Clinician coding styles
  - Systematic coding influences
- Aggregate ICD into broad syndromes
- Currently 12 - eg “Respiratory Infection”
- HPHC proof of concept
Public web site with access to publications and information about the project

Suggestions or inquiries may be directed to Ross Lazarus, Infection Disease and Epidemiology related questions to John Paltt, statistical questions to Ken Kleinman, questions about report programming and Epicare interface to Inna Zashinsky.

If you have been assigned a UserID and password for this site, you may click on the "View the reports" link above, then type these into the authentication dialogue box to identify yourself to gain access. Please use the link at the bottom of this page to contact the Webmaster if you have not been assigned a UserID and password but are experiencing technical difficulties accessing the site. We regret that technical support may not be available outside normal business hours (excluding weekends) and not available outside normal business hours (excluding weekends).

Please direct any requests for access to the material to the project Administrative Officer.

Please note that your IP address (179.233.258.139) and all your activity on this website is logged. If this is not acceptable, please disconnect immediately.
Boston Bioterrorism Surveillance Daily Reports

Massachusetts Department of Public Health
Daily Public Surveillance Report of Office Visits
With Diagnoses Corresponding to Infection Syndromes
Previous 30 Days Only. Click here to see the entire period

Click on a date (e.g. Monday, 21 April, 2003) to see a Summary report.
Click on an individual syndrome colored table bar for a detailed syndrome report and map.
UR = Upper Respiratory Infections, LR = Lower Respiratory Infections, LG = Gastrointestinal Infections

http://btsurveillance.org

Regular daily reports since October 31 2001
(Authorized users only)

© Copyright 2001 Bioterrorism Lab All Rights Reserved.
Iri syndrome by census tract

http://btsurveillance.org

Next day daily results from Generalised Linear Mixed Models - P values for major syndromes

(Authorized users only)

<table>
<thead>
<tr>
<th>Tract</th>
<th>Code</th>
<th>Cases</th>
<th>Pop</th>
<th>Known</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swampscott</td>
<td>250002021</td>
<td>2</td>
<td>335</td>
<td>0</td>
</tr>
<tr>
<td>Arlington</td>
<td>250173563</td>
<td>2</td>
<td>498</td>
<td>0</td>
</tr>
<tr>
<td>Wellesley</td>
<td>250214044</td>
<td>2</td>
<td>650</td>
<td>0</td>
</tr>
</tbody>
</table>

* The 5 most extreme tracts are shown, plus all with counts not expected to occur more than once per month.

* Estimated number of years between daily counts this extreme in
**http://btsurveillance.org**

**Detailed individual records on line for immediate access**

*(Highly restricted authorized users only)*

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Visit ID#</th>
<th>Age</th>
<th>Sex</th>
<th>Temp</th>
<th>Diag at the visit</th>
<th>Diagnosis Description</th>
<th>Enc. Type</th>
<th>Patient ZIP code</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEVER</td>
<td>161</td>
<td>9</td>
<td>F</td>
<td>100.3</td>
<td>034.0</td>
<td>STREP SORE THROAT</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>FEVER</td>
<td>162</td>
<td>39</td>
<td>F</td>
<td>100.1</td>
<td>034.0</td>
<td>STREP SORE THROAT</td>
<td>URGE</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>28</td>
<td>73</td>
<td>F</td>
<td></td>
<td>733.00</td>
<td>OSTEOPOROSIS NOS</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>123</td>
<td>71</td>
<td>F</td>
<td></td>
<td>287.5</td>
<td>THROMBOCYTOPENIA NOS</td>
<td>Tele</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>79</td>
<td>87</td>
<td>F</td>
<td></td>
<td>787.91</td>
<td>DIARRHEA</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>79</td>
<td>87</td>
<td>F</td>
<td></td>
<td>401.9</td>
<td>HYPERTENSION NOS</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>79</td>
<td>87</td>
<td>F</td>
<td></td>
<td>287.2</td>
<td>PURPUR A NOS</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>79</td>
<td>87</td>
<td>F</td>
<td></td>
<td>569.42</td>
<td>ANAL OR RECTAL PAIN</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>79</td>
<td>87</td>
<td>F</td>
<td></td>
<td>715.90</td>
<td>OA NOS-UNSPEC</td>
<td>Visi</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>192</td>
<td>86</td>
<td>F</td>
<td></td>
<td>287.5</td>
<td>THROMBOCYTOPENIA NOS</td>
<td>Tele</td>
<td></td>
</tr>
<tr>
<td>DERM</td>
<td>42</td>
<td>31</td>
<td>M</td>
<td></td>
<td>788.3</td>
<td>SPONTANEOUS ECCHYMOSIS</td>
<td>Visi</td>
<td></td>
</tr>
</tbody>
</table>
National Demonstration Project

- Electronic health records
- Office visits, nurse telephone triage calls
- ~20 million people.
- 50 states.
- Distributed processing
- Centralized reporting
- Syndrome count data only
- HIPAA “deidentified” (statistician’s cert.)
National BT Demo Program

Optum 11/02
- 1-50: Funded
- 51-250: Pending or in discussion
- 251+: Others
National Demonstration Project design features

♦ Scalable, data from many sources.
♦ Health plans put data extracts on a PC they control for processing.
♦ Analysis and reporting programs provided by data center.
♦ Internet-based communication.
♦ Encounter-level data stays with health plan or provider until health department requires it.
♦ Adaptable to other public health uses.
Surveillance strategy

♦ Extract encounters with an ICD9 code of interest.
♦ Initially daily. Potentially more frequently
♦ Group encounters into syndromes.
♦ Identify INITIAL visit for each episode of illness.
♦ Map episodes to the patient’s home ZIP code initially
♦ Geocoding -> census tract when available
♦ Transfer summary data to data center
♦ Identify statistically “extreme” regions each period.
♦ Notify data provider and public health agency.
Identifiable data is all processed locally using software provided by the Datacenter.

Identifiable lists of encounters available locally to be passed on in case of an event.

XML file, counts by zip by period only to Datacenter.
<Period>
   <startTimeStamp>2003-04-10T01:00:00-05:00</startTimeStamp>
   <endTimeStamp>2003-04-11T00:59:59-05:00</endTimeStamp>
   - <geography TYPE="ZIP">
      <geovalue>01463</geovalue>
      - <syndrome TYPE="Respiratory">
         - <payer NAME="PR">
            <count>1</count>
         </payer>
      </syndrome>
   </geography>
   - <geography TYPE="ZIP">
      <geovalue>01527</geovalue>
      - <syndrome TYPE="Hemorrhagic">
         - <payer NAME="PR">
            <count>1</count>
         </payer>
      </syndrome>
   </geography>
   - <geography TYPE="ZIP">
      <geovalue>01581</geovalue>
      - <syndrome TYPE="Fever">
         - <payer NAME="PR">
            <count>1</count>
         </payer>
      </syndrome>
   </geography>
</Period>
Data viewable as time series within minutes of arrival at datacenter

GLMM being run as historical data becomes available.
Outline

♦ Electronic records for surveillance
♦ Syndromic surveillance for BT
♦ HPHC/MDPH System
♦ National Demonstration Project
♦ Issues and challenges
♦ Conclusions
Issues: non-technical

♦ Surveillance data model
  ♦ Traditionally centralized identifiable
  ♦ Alternative - only counts centralized

♦ HIPAA

♦ Privacy/Security

♦ Authentication & administration

♦ Data presentation, visualization

♦ User interface for ad-hoc queries

♦ Notification - who, how, when
Issues: technical

- Noisy data!
- “Optimal” models
- ICD9 granularity $\Rightarrow$ Syndromes
- “Optimal” syndromes & ICD9 mapping
- Standards for data exchange - XML
- Authentication and access control
- Security for internet services - SSL
Overcoming challenges

♦ Shared definitions - CDC led workgroup
♦ Ad hoc queries.
♦ CDC PHIN/NEDSS compliance
♦ Open source infrastructure
♦ Open standards
♦ Open source code
Lessons

- EMR works for timely surveillance
- Distributed processing model works
- Design for additional data sources
- Standardized syndromes
- Standardized data exchange
- Strong encryption
- Best security practices
Collaborators

- Richard Platt
- Ken Klienman
- Inna Dashevsky
- Katherine Yih
- Courtney Adams
- Virginia Rego
Innovations in Public Health Surveillance: The National Bioterrorism Syndromic Surveillance Demonstration Project

Ross Lazarus
Director of Bioinformatics, Channing Laboratory
Brigham and Women’s Hospital and Harvard Medical School

June 10, 2003
Assigns diagnoses during encounter, extracts selected encounters nightly, identifies new episodes, assigns to zip code.

More data as needed

Linelist

Health department

Alert! Cluster type, size, location.

Website: Public areas
Private areas

Data center: aggregates data from different providers, identifies unusual clusters

Instruction to send linelist to health dept.

Count by zip