Syndromic Surveillance

David Buckeridge, MD PhD FRCPC
Canada Research Chair in Public Health Informatics
The Surveillance Lab, McGill Clinical and Health Informatics
Department of Epidemiology and Biostatistics, McGill University
Agence de la santé et des services sociaux de Montréal, Direction de santé publique

2013 Environmental Health Surveillance Workshop
NCCEH and BCCDC
Ottawa, Canada, February 26, 2013
Syndromic Surveillance in Context

**Indicator-based surveillance**

- Identified risks
  - Mandatory notification
  - Laboratory surveillance
- Emerging risks
  - Syndromic surveillance
- Non-healthcare based
  - Poison centres
  - Behavioral surveillance
  - Environmental surveillance
  - Veterinary surveillance
  - Food safety / Water supply
  - Drug post-licensing monitoring

**Event-based surveillance**

- Data
  - Collect
  - Analyze
  - Interpret
- Event monitoring
  - Capture
  - Filter
  - Verify
- Signal
  - Assess
  - Investigate

**Surveillance** systems

- Public health
- Alert
- Control measures

**Event-based surveillance**

- Disseminate
  - Confidential / Limited Dissemination
  - Public Dissemination

**Domestic**
- Media review
- EI focal points

**International**
- Infor scanning tools
- Distribution lists
- International agencies

**Panel: Definition of syndromic surveillance**

A real-time (or near real-time) collection, analysis, interpretation, and dissemination of health-related data to enable the early identification of the impact (or absence of impact) of potential human or veterinary public health threats that require effective public health action.

Syndromic surveillance is based not on the laboratory-confirmed diagnosis of a disease but on non-specific health indicators including clinical signs, symptoms as well as measures (e.g., absenteeism, drug sales, animal production collapse) that lead to a provisional diagnosis (or “syndrome”).

The data are usually collected for purposes other than surveillance and, where possible, are automatically generated so as not to impose an additional burden on the data providers.

This surveillance tends to be non-specific yet sensitive and rapid, and can augment and complement the information provided by traditional test-based surveillance systems.

**Characteristics**
1. Automation
2. Syndromes

**Implications**
1. Real Time
2. No Reporting Burden

**Applications**
1. Identify Impact
2. Augment Systems

---

The US CDC BioSense 2.0 System
The UK HPA Syndromic Systems

Syndromic Systems and Bulletin Archive

Below are links to information on each syndromic surveillance system used at the HPA, illustrating the key characteristics of each system.

Archived copies of previous surveillance bulletins are also available through these links.

- HPA/NHS Direct Syndromic Surveillance System
- HPA/QSurvveillance National Syndromic Surveillance System
- GP Out-of-Hours/Unscheduled Care Surveillance System
- Emergency Department Syndromic Surveillance System (EDSSS)
- Royal College of General Practitioners Weekly Returns Service (RCGP WRS)
- NHS24 syndromic surveillance system

Related Information

- Norovirus
- Real-time Syndromic Surveillance Team References
- Seasonal Influenza
- The London 2012 Olympic and Paralympic Games
- Extreme weather events and natural disasters

External Links

- NHS Direct
- QSurvillance®
- RCGP Research & Surveillance Centre
- Triple-S - Syndromic Surveillance Systems in Europe
Syndromic Surveillance Process

1. Automatic Capture, Transmission
2. Syndrome detection
3. Routine Analysis
4. Ad hoc Analysis
DATA CAPTURE
Common Settings for Data Capture

Automated Capture and Transmission
# Standards for Data Capture

**Final Recommendation:**
Core Processes and EHR Requirements for Health Syndromic Surveillance

*International Society for Disease Surveillance (ISDS)*
Meaningful Use Workgroup

January 31, 2011

<table>
<thead>
<tr>
<th>Data Element Name</th>
<th>Description of Field</th>
<th>Inpatient</th>
<th>Ambulatory</th>
<th>Emergency Department / Urgent Care</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Patient age at time of visit</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Age units</td>
<td>Unit corresponding to numeric value of patient age</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Gender</td>
<td>Stated gender of patient</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Race</td>
<td>Race of patient</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Ethnicity of patient</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Patient City / Town</td>
<td>City or town of patient residence</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Patient ZIP Code</td>
<td>ZIP Code of patient residence</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Patient County</td>
<td>County of patient residence</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Patient State</td>
<td>State of patient residence</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Patient Country</td>
<td>Country of patient residence</td>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td><strong>Visit Information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chief Complaint / Reason for Visit</td>
<td>Patient’s self reported chief complaint or reason for visit</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Admit or Encounter Reason</td>
<td>Provider’s reason for a patient admission or encounter</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Admission or Encounter Date/Time</td>
<td>Date and time of patient admission or encounter</td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Date of Onset</td>
<td>Date that patient began having symptoms of condition being reported</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Patient Class</td>
<td>Patient classification within facility</td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Hospital Unit</td>
<td>Hospital unit where patient is treated</td>
<td></td>
<td>RE</td>
<td>-</td>
</tr>
<tr>
<td><strong>Diagnostic and Predictive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis Type</td>
<td>Qualifier for Diagnosis / Injury Code specifying type of diagnosis</td>
<td></td>
<td>RE</td>
<td>-</td>
</tr>
<tr>
<td>Primary Diagnosis</td>
<td>Primary diagnosis of the patient's condition</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
<tr>
<td>Additional Diagnosis</td>
<td>Additional diagnosis of the patient's condition</td>
<td></td>
<td>RE</td>
<td>RE</td>
</tr>
</tbody>
</table>
SYNDROME DETECTION
Detecting Syndrome Cases

• Each visit is classified into a syndrome (e.g., respiratory, influenza-like-illness, ...)

• Classification uses information from a clinical information system
  – **Code Based**: Search for specific codes; or
  – **Natural Language**: 1) Build a statistical model of the word distribution in true as opposed to false cases, or 2) Use text ‘templates’ to search for words or expressions
# Standards for Syndrome Definition

## Clinical condition: vomiting

<table>
<thead>
<tr>
<th>Concept (relation to condition)</th>
<th>Keywords, regular expressions representing keywords and UMLS CUI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vomiting</strong> (Condition name)</td>
<td>vmt, n/v, v/n, v/d, d/v, v+d, n+v, v+n, v + d, d + v, d &amp; v, dry heaves, emesis, f v, f v, f v, n v, n &amp; v, n v d, n&amp;v, n+v, nv, nvd, retching, v d, v f, v&amp;d, v+d, vomiting, viomiting, vmt, vo, voimitting, voiting, vom, vometing, vomi, vomiktng, vomintg, vomintng, vomit, vomited, vomiti, vomitibg, vomitin, vomitine, vomiting, vominingg, vomiting, vomitt, vomitting, vomitting, vomitus, vommitng, vomoting, vomting, vomting, vomting, vomting, vomting, puke, n v, retching, v d, vomit, vomicking, cant hold any food down, threw up, throw up, throwing up, emesis, emisis</td>
</tr>
<tr>
<td><strong>Spitting up</strong> (Synonym)</td>
<td>bringing up, spitting up</td>
</tr>
<tr>
<td><strong>Hematemesis</strong> (Related concept)</td>
<td>hematemesis, coffee ground emesis, throw* up blood, vomit* blood</td>
</tr>
</tbody>
</table>

Regular expressions (3 of 73 shown below) and UMLS CUI:

- \bd\s&\sv\b (Vomiting): C0042963
- \bd\dry\heaves\b (Retching): C0232602
- \b\vomitus\b (Vomitus): C0042965
- \b\bringing\sup\b (Spitting up): C0042963
- \b\spitting\sup\b (Spitting up): C0042963
- \b\hematemesis\b (Hematemesis): C00189261
- \b\coffee\sground\semesis\b (Coffee ground vomiting): C1510416
- \b\(\text{throw}\w*?\s+up}\vomit\s*\b (Hematemesis): C00189261
- \s*\blood\b (Hematemesis): C00189261

Syndromic Case Detection can be Accurate for Broad Categories

Sensitivity Specificity PPV NPV

Respiratory ILI (Broad) ILI (Narrow)

Syndromic Surveillance Summary

• Defining Syndromic Surveillance
  – Characteristics: Automated, syndromes
  – Implications: Real time, no reporting burden
  – Applications: Impact (rule-out), augment systems

• Innovations Applicable to Surveillance
  – Automation of data capture for ‘clinical’ settings
  – Advances in real-time data analysis
  – Integration of data from multiple systems
Pattern Analysis (in Public Health Surveillance)

David Buckeridge, MD PhD FRCPC
Canada Research Chair in Public Health Informatics
The Surveillance Lab, McGill Clinical and Health Informatics
Department of Epidemiology and Biostatistics, McGill University
Agence de la santé et des services sociaux de Montréal, Direction de santé publique

2013 Environmental Health Surveillance Workshop
NCCEH and BCCDC
Ottawa, Canada, February 26, 2013
• Benefits of detection are measured by intervention outcomes: reducing morbidity, mortality and cost
• Intervention strategies are outbreak-specific:
  – anthrax: early medical treatment of infected individuals
  – waterborne *c.parvum*: preventing new infections by limiting exposure
• Intervention outcomes depend not only on timeliness of detection
  – level of compliance (e.g. with boil-water advisory)
  – extend and duration of exposure to pathogen
  – incubation time
  – ...
Public Health Context

• Prospective analysis
  – Repeated routinely
  – Analysis should build on previous results

• Multiple data sources
  – Statistical analysis of one data source rarely provides definitive information for action
  – Integration of analysis results is a difficult problem

• Dynamic decision making
  – Surveillance informs actions
  – Possible actions should drive surveillance
DETECTING PATTERNS
Surveillance Analysis Framework

• Target
  – Omnibus alternative hypothesis
  – Specified alternative hypothesis

• Dimensions
  – Time: Critical aspect of any analysis
  – Place: Often spatial, sometimes space—time
  – Person: Usually stratification

• Integration
  – Combining multiple data sources and systems
  – Linking analysis results with actions, effectiveness
Temporal Detection

Space—Time Detection

• Spatial data model
  – Point based
  – Region based

• Popular approaches
  – Independent monitoring of sub-regions
  – (Bayesian) Spatial regression
  – Scan statistics
Space—Time Cluster Detection

Example TB cluster with no person-to-person transmission

- Genetically Isolated
- Genetic Group 21
- Genetic Group 30

Time Slice: 1997-05-01 to 1997-05-31
EVALUATING DETECTION
Evaluation Framework

• Evaluation Paradigms
  – Statistical Measures
    • Diagnostic test: sensitivity, specificity, timeliness
    • Process control: average run length
  – Health and Economic Outcomes
    • Prevented utilization, morbidity, mortality
    • Cost-Effectiveness very rarely assessed

• Evaluation Approaches
  – Real data: Limited availability, low numbers
  – Simulated data: Alone or with real data, complexity
Statistical Measures, Real Data

1999-03 to 1999-07
Person-to-Person Transmission

2000-01 to 2000-01
NO Person-to-Person Transmission

2002-02 to 2002-02
Statistically significant clusters

2004-01 to 2004-01
LOW Person-to-Person Transmission
Diagnostic Test Evaluation Approach
Simulation for Integrated Evaluation

Evaluating Impact – Infections Averted

Contamination duration (days)

Boil water advisory time (days)

Proportion of Infections Averted
Pattern Analysis Summary

• Context is moving from single time series to rich, multi-dimensional data sets
• Innovations in analytical methods
  – Automated temporal analysis
  – Space—time analysis more complicated
  – High-dimensional methods in other disciplines
• Methods and tools needed for integration
• Training programs needed desperately