New Technologies for Global Public Health Surveillance

John Brownstein, PhD
Children's Hospital Boston, Harvard Medical School

Global outbreaks, the challenge: late reporting and response

Global outbreaks, the solution: early reporting and response

Guangzhou Baiyunshan Pharmaceutical Co.
(白云山 A, Public Shenzhen: 000522.SZ)

January 16
March 10

Delayed response

Lost opportunity for control/risk of international spread

Rapid response
Potential cases prevented/international spread prevented
Emerging threat of infectious diseases

- Breakdown of public health measures
- Drug and pesticide resistance
- Unsuccessful vaccine development
- Environmental change
- Human demographics and behavior
- International travel and commerce (ie: wildlife)

Drivers of outbreaks

Public Health Surveillance

- Need for evidence-based decisions for implementation and targeting of control activities
- Challenge outside the scope of traditional surveillance systems and reporting mechanisms
**Need for early detection**

- **Phase I**: Initial Symptoms
- **Phase II**: Acute Illness

- **Early Detection**
- **Gain of 2 days**
- **Traditional Disease Detection**

**Timeline**
- Exposure
- Incubation Period
- Incubation
- Health and healthcare behaviors
- Diagnostics
- Additional evaluation
- Surveillance of citizenry
- Patient self-assessment
- Information seeking (web clickstream, etc.)
- Over the counter and prescription medications
- School and work absenteeism
- Nurse triage telephone calls

**Data types**
- Orders
- Laboratory tests
- Physician office visits
- EMS activity
- Emergency department visits
- Hospitalizations

**Healthcare data sources**
- Emergency department chief complaints
- International Classification of Disease (ICD) codes
- Text-based notes
- Laboratory data
- Radiological reports
- Physician reports

**“Non traditional” data sources**
- Pharmacy data
- Retail sales data
- 911 operators
- Call triage centers
- School absenteeism
- Animal surveillance
- Internet-based reports
AEGIS

Figure 1. Geographical distribution of a selected subset of outbreaks confirmed by the World Health Organization (WHO) and reported in the “Disease Outbreak News” reports, 1996-2009. Points mark the reported origin of the outbreak, or if unknown, where there were the highest reported morbidity and mortality rates.

Barriers to Global Surveillance Systems

- Weak public health infrastructure in most parts of the world
- Delays in reporting and identification of outbreaks
- Lack of dissemination and feedback to the local level
- Lack of integration of training and surveillance activities
- Limited evaluation of programs
- Limited cross-border collaboration/communication
- Limited national resources and sustainability

WHO confirmed outbreaks (1996-2009)

Traditional Public Health Reporting

Public

Public health practitioners

Local Officials

Labs

Healthcare workers

Clinicians

Ministry of Health

Informal Surveillance

World Bodies

(UN, WHO, FAO, OIE)
**Global Surveillance Capacity Assessment**

Characterize global spatial-temporal trends in the timeliness of outbreak detection and reporting

![Diagram showing outbreak start, outbreak discovery, and public communication with time intervals Δt1 and Δt2]

**Outbreak Database (1996-2009)**

- **Disease / Location**
- **Date of onset of risk factor**
- **Date of local mass gathering**
- **Date of associated wildlife outbreak**
- **Date of exposure**
- **Date of symptom onset**
- **Date of outbreak start**
- **Date of hospitalization or medical visit**
- **Date of outbreak detection**
- **Date of death**
- Date of laboratory confirmation
- Date of announcement by a local
- Date of any earlier mentioned report
- Date of ProMED, GPHIN, HealthMap reports
- Date of WHO notification
- Date of DON report (official)
- Date of mass immunization campaign
- Date of implementation of vector control
- Date of declaration of an epidemic raised
- Date of declaration of end of epidemic

**Drivers of outbreaks**

<table>
<thead>
<tr>
<th>Breakdown of Public Health Measures</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food &amp; agriculture industry changes</td>
<td>51</td>
<td>[32;36.5]</td>
</tr>
<tr>
<td>International travel &amp; commerce</td>
<td>47</td>
<td>[32;56]</td>
</tr>
<tr>
<td>Climate &amp; weather</td>
<td>43</td>
<td>[32;56]</td>
</tr>
<tr>
<td>Unspecified</td>
<td>26</td>
<td>[18;30]</td>
</tr>
<tr>
<td>Human demographics &amp; behavior</td>
<td>26</td>
<td>[18;30]</td>
</tr>
<tr>
<td>Bushmeat</td>
<td>14</td>
<td>[8;18]</td>
</tr>
<tr>
<td>War &amp; famine</td>
<td>8</td>
<td>[2;14]</td>
</tr>
<tr>
<td>Land use changes</td>
<td>2</td>
<td>[2;2]</td>
</tr>
<tr>
<td>Human susceptibility to infection</td>
<td>2</td>
<td>[2;2]</td>
</tr>
</tbody>
</table>

![Boxplots showing median time (and inter-quartile range) between outbreak start and various outbreak "milestones" for a set of WHO-confirmed outbreaks during 1996-2009.]

- **Public Communication (n=280)**: 23 days (95% C.I. [18;30])
- **Laboratory Confirmation (n=223)**: 32 days (95% C.I. [28;38.5])
- **WHO Report (n=281)**: 35 days (95% C.I. [32;47])
- **MEDIAN**: 48 days (95% C.I. [40;56])

Figure 2. Boxplots of the median time (and inter-quartile range) between outbreak start and various outbreak "milestones" for a set of WHO-confirmed outbreaks during 1996-2009. Public communication refers to the earliest date of any form of communication (official or informal, written or verbal) to the public about the existence of cases. WHO report refers to the date of the World Health Organization’s (WHO) Disease Outbreak News report about the outbreak.
Time from outbreak start to outbreak discovery

Year of Outbreak Start

Source of outbreak news verified by WHO

Adapted from Heymann 2001

Explosion of epidemic intelligence

- Argus
- BioCaster
- EpiSpider
- GPHIN
- PULS
- HealthMap
- MedISys

Potential Informal Electronic Data Sources

- SMS Messaging
- Emailing
- Social networking
- Blogging
- Video/radio news reporting
- Micro blogging
- Internet searching
- Internet chatting
- Online news reporting
- Health expert reporting
Limitations of many current surveillance systems

- Abundance of disparate electronic resources but none comprehensive
- Each has geographic, expertise, population gaps
- Lack of integration between tools and information sources
- No synthesized view of the current state of global health

HealthMap Overview

- Goal to provide comprehensive overview of global state of infectious diseases

  - Freely available, automated system (24/7)
  - Integrates outbreak data from ‘official’ and ‘unofficial’ sources
  - Aggregates data by disease & location
  - Displays alerts from 29 sources in 7 languages (auto translated)
  - Resource for both public health professionals and general population

HealthMap Article Processing

1 Acquisition
   Over 20,000 sites, every hour, 24/7

2 Extraction
   1800 disease patterns, 5000 locations
   Acholera outbreak ravaging through two regions in northern Cameroon since May…

3 Categorization
   6 million phrases, 91% accuracy
   Breaking News
   Warning
   Old News
   Context
   Not Disease Related
Aggregation
Text matching, Similarity Score, Significance Rating

Breaking News

Warning

HealthMap Article Processing

HHS Command Center  ECDC Command Center  Liberty Science Museum, NYC

HealthMap Users - Over 1 million users per year

Top Visitors:
- CDC
- WHO
- National, state, local public health depts
- NGOs
- Museums
- Clinicians
- Travelers

HealthMap Reports

- > 600 alerts per day from 22 sources from over 20,000 websites
- Alerts in 201 countries & territories and 175 disease categories
- Seven languages – English, French, Spanish, Russian, Chinese, Portuguese, Arabic

Disease covered (n=179)
Median of 12 [95% CI, 9—18] days between suspected and confirmed reports.

Countries with a high GDP tended to have shorter lags (Pearson correlation -0.4; 95% CI, -0.6 — -0.2)

Epidemic Curve - Salmonella Saintpaul

New Mexico Report - 5/31
CDC Report - 6/2

FIFA World Cup Surveillance

Salmonella Saintpaul

June 3
June 6
June 9
June 18
Combining informal data with traditional surveillance

CDC Yellow Book: Dengue

2011 Yellow Book

2008 Yellow Book Risk Areas
2011 Yellow Book Risk Areas
**Model Performance**

<table>
<thead>
<tr>
<th>Region</th>
<th>Latin America</th>
<th>Asia</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.866</td>
<td>0.646</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.895</td>
<td>0.522</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.667</td>
<td>0.600</td>
</tr>
</tbody>
</table>

**Professional networks to support disease reporting**

- Human networks have proven value
- Professional networks play key roles in discovery and validation
- 2 Examples:
  - ProMED ISID
  - GeoSentinel ISTM

**Professional networks and epidemic intelligence**

**Collaborations**

- Automated Feed
- Automated Extraction
- Community Input
- Added Value
  1. Discovery of new content
  2. Provides two-way contextualization
  3. Validation studies

- Visualization
- Alerting
- Curation
- Verification
Limitations of current approaches

- Can’t mine all possible sources
- Can’t mine in all languages
- Can’t mine all data types
- Delay required for searching, curating and processing
- Resource limited process (both machine and human)
Ushahidi

- Began in response to post-election violence in Kenya in 2007
- The system collects individual reports from users through SMS, Web and email
- Provides tools for translating, classifying and georeferencing the reports
- Presents aggregated information on a map-based Web interface
Mechanical Turk

- Human Intelligence Tasks (HITs)
  - Individual tasks that require human intelligence to complete
  - Example: Categorize the objects in these pictures.
- Workers
  - Flexible work schedule and additional income
- Requesters
  - Access to a global, on-demand workforce
  - Pay when satisfied with quality of work

Incentive 2: Social Facebook

- Social Networking site with 2B registered users
- People spend over 500,000,000,000 minutes per month on Facebook
- Over 70% of users engage with applications
Engaging Public Social Networks: Facebook

Get Well Soon!

I want to tell others about my symptoms!

- Runny Nose
- Headache
- Fever

See what symptoms are going around in your network

Provides social incentives for self-reporting of flu data

10,000+ users in worldwide pilot

Healthysocial.org

Get Well Soon!

Please fill out the following anonymous questionnaire [enter link]

When did the symptoms start?
- 3 days ago

Did you have a fever?
- Yes
- No

How high was your fever?

When did your fever start?
- 3 days ago

Did your fever start abruptly (within 48 hrs)?
- Yes
- No
- Don’t know

Did you have to alter your daily routine?
- Yes, I stayed at home.
- Yes, I had to work/school.
- No, I did everything as usual.

International launch in 10 countries next month through partnership with European Epiwork project.

Social Networking and Patient Data

Submit a disease report:

- Headline:
- Disease: Please select a disease
- Location:
- Description:

Upload a related photo (optional): Choose File

Email: 

By submitting, you agree to the Terms of Submission.

Submit

Real time group and individual measurement of outcomes, treatment side effects, adverse events
Report Your Health Status

Thank you for registering for the Flu Near You project! Now it's time to start your first survey.

What symptoms have you experienced in the past week?
- Fever
- Headache
- Sore throat
- Runny nose
- Muscle aches
- Chills

What was your highest temperature measured?
- Less than or equal to 99.5°C
- More than 99.5°C
- Not measured

What day did you start feeling ill?

Did you receive the flu vaccine after July 31, 2011?
- Yes
- No
- Don't know

Did you receive the flu vaccine last year (between July 31, 2010 - July 31, 2011)?
- Yes
- No
- Don't know

Submit

Flu Activity in Your Area

User submitted flu reports in Springfield, Massachusetts, United States in the past month.

APHA flu near you Challenge

Individual Top Awards

1st $25,000
2nd $15,000
3rd $10,000

Awarded to top three individuals with highest total number person-forms

Minimum Threshold eligibility: 10,000 person-forms
(one form allowed per person per week)

Group Top Awards

1st $50,000
2nd $25,000
3rd $15,000

Awarded to top three groups with highest total number person-forms divided by number of persons in group

Minimum Threshold eligibility: 100,000 person-forms
(one form allowed per person per week)

Explore flu trends across the U.S.

We've found that certain search terms are good indicators of flu activity. Google Flu Trends uses data up to two weeks later than traditional systems. Read more →

United States flu activity:

2003 - 2009: Past years

2009 - 2010: Current season

Explore flu trends across the U.S.
Twitter as a predictor of influenza-like illness

Alessio Signorini. University of Iowa

Adapted from Dzenowagis, WHO, 2005 and T. Kass-Hout, 2009

Growth of Mobile Technologies

Thailand

(per 1,000 people)

<table>
<thead>
<tr>
<th>years</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tel. mainlines in largest city</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>50</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
</tr>
<tr>
<td>Telephone mainlines</td>
<td>150</td>
<td>200</td>
<td>250</td>
<td>300</td>
<td>350</td>
</tr>
<tr>
<td>Personal computers</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td>600</td>
</tr>
<tr>
<td>Telephone mainlines, waiting list</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
</tr>
<tr>
<td>Internet host</td>
<td>50</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
</tr>
</tbody>
</table>

Adapted from Dzenowagis, WHO, 2005 and T. Kass-Hout, 2009

Is a Swine Flu Outbreak Coming? Ask Your iPhone

By Bryan Walsh

Wednesday, Sep 10, 2009

TIME IN ACTION

Current Location

1. 8-year-old falls victim to H1N1 in Ills.
6/23/09 8:59 PM
A health care worker died in the U.S. with swine flu, making her the first American to die from the disease.

2. West Nile virus found in Boston
6/11/09 3:47 PM
By Stephen Delaney, Online Staff For the first time this summer, West Nile virus has been detected in mosquitoes in Boston, city health authorities said today. The infected insects were discovered in Jamaica Plain.
Geo-reporting of disease events

Technical Challenges
- Formatting, quality
- Georeferencing while preserving privacy
- Potential for abuse
- Scalability
- Rules and features for automated filtering
  - Distance from phone to event
  - Photo
  - Reputation

Validation Process: Example

*Galvin middle school teacher; teacher diagnosed with bacterial meningitis.*

Additional info
- Coordinates of iphone
- Email for follow-up
- Associated news media

>100k downloads
Validation Process: Example

- Galvin middle school teacher; teacher diagnosed with bacterial meningitis.

Limitations

- Validation
- Lack of details
- Issues of self-diagnosis
- Intentional false alarms

What about aggregation?
**iPhone Submissions vs CDC sentinel surveillance**

- CDC Sentinel Physician Network (%ILI)
- Outbreaks Near Me iPhone app (%H1N1 submissions/Downloads)

\[ R^2 = 0.74 \]

---

**Mobile phones and crisis surveillance**

---

**SMS Based reporting networks: Haiti Example**

1. A Haitian with a need sends an SMS to the 4636 shortcode.
2. The SMS is then routed through the Nantum SMS gateway and onto the Emergency Information Service (EIS).
3. Once at EIS, the information in the SMS is analyzed, tracked, and then forwarded to the crowdflower.com website.
4. A Haitian volunteer or staff member logs onto the website and translates the SMS, adding meta and geospatial information.
5. After translation, the information is turned into a report that goes out to multiple organizations involved in the crisis response and recovery effort.
Cholera Surveillance in Haiti

Estimate Reproductive Number ($R_0$)

Phase 1: Informal sources 1.54-6.89 compared to official sources 1.27-3.72
Phase 2: Informal sources 1.04-1.51 compared to official sources 1.06-1.73

Using social media to build an epidemic curve

Case counts
Hospitalizations
HealthMap
Twitter
Clinical Reporting and

- Improve Reporting from Field
- HTML 5 Application designed for iPad
- Electronically organizes patient information
- Real-time visualization of ID diagnosis and syndromes
- Lightweight and easily updated
- Works offline
- Automated location capture
- Easily deployed on other devices
  - laptops, tablets, mobile phones (iPhone, Google phone)

HealthMap Pilot, Haiti

The Web beyond infectious disease

MedWatcher App

- Submit adverse event to FDA
- See latest drug safety reports
- Post reviews as a patient
- Read drug reviews
>10k medications
Stores personal medication list

• 20k downloads
• Top downloaded medical app
• 1% of users report adverse events
Mining of chat room allows for quantitative comparisons of product-specific discussions and health outcomes.

As of July 30, 2008
Entire database contains:
2,281,572 posts
Unique authors discussing prescription opioids:
27,065 screen names
Using search volume to estimate abortion rates


Internet users seek out information on quitting and cheap cigarettes in response to the SCHIP cigarette tax increase.

Conclusions

- Value in the fusion and visualization of distributed electronic resources (online epidemic intelligence, social networks, mobile technology)

- Novel Internet-based collaborative systems can play an important complementary role in gathering information quickly and improving coverage and accessibility.

- These early efforts at tapping the power of digital tools demonstrate important steps in improving health systems as well as engaging the public as participants in the public health process.

Acknowledgments

HealthMap Team
- Clark Freifeld
- Susan Aman
- Katelynn O’Brien
- Mikaela Keller, PhD
- Rumi Chunara, PhD
- Katia Charland, PhD
- Amy Sonricker, MPH
- Emily Chan, MSc
- Sumiko Mekaru, DVM
- Annie Gatwood, PhD
- Leila Amerling, MBA
- Vina Chhaya, MPH

Children’s Hospital Informatics Program
- Ken Mandl, MD MPH
- Ben Reis, PhD
- Isaac Kohane, MD PhD

Funding