Tectonic shifts in the health information economy: Case studies in mining existing data for public health

John Brownstein, PhD

Children’s Hospital Boston, Harvard Medical School
How to measure population health?
Standard Approaches: RCT and Observational Studies

- Small sample sizes
- Detection of rare events is often impossible
- Rare diseases not usually considered
- Often cost prohibitive
- Difficulty in repurposing data
- Not geared to vulnerable populations
- Data not available in real-time or near real-time
Real-time Population Health Monitoring
Surveillance of citizenry
Patient self-assessment

Information seeking (web clickstream, etc.)
Grocery product sales
Over the counter and prescription medications
School and work absenteeism
Nurse triage telephone calls

Physician office visits
EMS activity
Emergency department visits
Hospitalizations

Orders
Laboratory tests

Preliminary diagnoses
Confirmed diagnoses

Data types

Biological sensors

Incubation Period
Health and healthcare behaviors

Diagnostics
Additional evaluation

Timeliness
Sensitivity?

Specificity
Non-traditional Health care data sources

- Emergency department chief complaints
- International Classification of Disease (ICD) codes
- Text-based notes
- Laboratory data
- Radiological reports
- Physician reports
Non traditional non-healthcare data sources

- Retail sales data
- 911 operators
- Call triage centers
- School absenteeism
- Animal surveillance
- Internet-based reports
- Patient self-assessment
How to measure population health?

Non-traditional healthcare data
Case study: Challenge of Drug Safety

- Premarketing maximize likelihood that approved drugs are safe and effective

- Issues: many small, short term RCT means missing occurrence of rare but serious adverse drug events, drug interactions, late events, pregnancy, different effects in subgroups

- Challenge of early detection relies on US postmarketing system
Goals of pharmacovigilance

- Early and reliable detection of problem drugs
- Characterization of magnitude of impact
Pharmacovigilance and EMR

- Accumulating evidence suggests critical insights can be realized from monitoring large clinical databases.
- Analytical epi studies of large exposed patients clearly outweigh current federal system for collected information.
- Need for automated real-time denominator based methods that is capable of capturing rare events.
- Existing electronic medical record databases means that exposure and outcome data can be linked without much extra cost or invasion of privacy.
Monthly trends in Myocardial Infarction: Sudden Increase

Introduction of COX-2 Inhibitors

18.5% Increase
Signal Detection

First signal:
- 1 year after Celecoxib
- 8 months after Rofecoxib
Effect on patient age

- Negative association between mean age at MI and prescription volume
- Spearman correlation $-0.67$, $P<0.05$
Partners i2b2 Pharmacovigilance Initiative

Methodology Goals

- To develop model that enables retrospective identification of events in defined patient cohort

- Identify a method for identifying patients receiving their care and prescriptions at Partners (to identify event if occurs)

- Provides framework for identifying patients in prospective manner for real-time analysis
**US Population with Diabetes**

<table>
<thead>
<tr>
<th>Race/ethnicity</th>
<th>Total U.S. population</th>
<th>With diagnosed diabetes</th>
<th>With undiagnosed diabetes</th>
<th>Total with diabetes*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population*</td>
<td>301,736</td>
<td>17,486</td>
<td>6,640</td>
<td>24,126</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>199,091</td>
<td>11,403</td>
<td>4,520</td>
<td>15,923</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>37,002</td>
<td>2,775</td>
<td>699</td>
<td>3,474</td>
</tr>
<tr>
<td>Non-Hispanic other</td>
<td>20,101</td>
<td>1,076</td>
<td>317</td>
<td>1,393</td>
</tr>
<tr>
<td>Hispanic</td>
<td>45,541</td>
<td>2,231</td>
<td>1,104</td>
<td>3,335</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>148,744</td>
<td>8,543</td>
<td>3,113</td>
<td>11,656</td>
</tr>
<tr>
<td>Female</td>
<td>152,992</td>
<td>8,943</td>
<td>3,528</td>
<td>12,471</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18</td>
<td>73,878</td>
<td>157</td>
<td>35</td>
<td>192</td>
</tr>
<tr>
<td>18–34</td>
<td>70,373</td>
<td>964</td>
<td>669</td>
<td>1,633</td>
</tr>
<tr>
<td>35–44</td>
<td>43,356</td>
<td>1,686</td>
<td>1,174</td>
<td>2,860</td>
</tr>
<tr>
<td>45–54</td>
<td>43,838</td>
<td>3,443</td>
<td>1,327</td>
<td>4,770</td>
</tr>
<tr>
<td>55–59</td>
<td>18,235</td>
<td>2,307</td>
<td>756</td>
<td>3,063</td>
</tr>
<tr>
<td>60–64</td>
<td>14,323</td>
<td>2,261</td>
<td>775</td>
<td>3,036</td>
</tr>
<tr>
<td>65–69</td>
<td>10,690</td>
<td>1,879</td>
<td>850</td>
<td>2,729</td>
</tr>
<tr>
<td>≥70</td>
<td>27,042</td>
<td>4,788</td>
<td>1,055</td>
<td>5,843</td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>169,886</td>
<td>7,057</td>
<td>3,018</td>
<td>10,075</td>
</tr>
<tr>
<td>Government</td>
<td>91,794</td>
<td>8,997</td>
<td>2,891</td>
<td>11,888</td>
</tr>
<tr>
<td>Uninsured</td>
<td>40,055</td>
<td>1,432</td>
<td>731</td>
<td>2,163</td>
</tr>
</tbody>
</table>


*Numbers do not necessarily sum to totals because of rounding.

<table>
<thead>
<tr>
<th>Medication</th>
<th>Route of Administration</th>
<th>Year of Introduction or FDA Approval</th>
<th>Efficacy as Monotherapy, Measured as a Reduction in the Glycated Hemoglobin Concentration (percentage points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin</td>
<td>Parenteral</td>
<td>1921</td>
<td>≥2.5</td>
</tr>
<tr>
<td>Inhaled insulin</td>
<td>Pulmonary</td>
<td>2006</td>
<td>1.5</td>
</tr>
<tr>
<td>Sulfonylureas</td>
<td>Oral</td>
<td>1946</td>
<td>1.5</td>
</tr>
<tr>
<td>Biguanides</td>
<td>Oral</td>
<td>1957</td>
<td></td>
</tr>
<tr>
<td>Metformin†</td>
<td>Oral</td>
<td>1995</td>
<td>1.5</td>
</tr>
<tr>
<td>Alpha-glycosidase inhibitors</td>
<td>Oral</td>
<td>1995</td>
<td>0.5–0.8</td>
</tr>
<tr>
<td>Thiazolidinediones</td>
<td>Oral</td>
<td></td>
<td>0.8–1.0</td>
</tr>
<tr>
<td>Troglitazone‡</td>
<td>Oral</td>
<td>1997</td>
<td></td>
</tr>
<tr>
<td>Rosiglitazone</td>
<td>Oral</td>
<td>1999</td>
<td></td>
</tr>
<tr>
<td>Pioglitazone</td>
<td>Oral</td>
<td>1999</td>
<td></td>
</tr>
<tr>
<td>Glinides</td>
<td>Oral</td>
<td>1997</td>
<td>1.0–1.5</td>
</tr>
<tr>
<td>GLP analogues</td>
<td>Parenteral</td>
<td>2005</td>
<td>0.6</td>
</tr>
<tr>
<td>Amylin analogues</td>
<td>Parenteral</td>
<td>2005</td>
<td>0.6</td>
</tr>
<tr>
<td>DPP-IV inhibitors</td>
<td>Oral</td>
<td>2006</td>
<td>0.5–0.9</td>
</tr>
</tbody>
</table>

* GLP denotes glucagon-like peptide, and DPP-IV dipeptidyl peptidase IV.
† Metformin has been available in other countries since 1957 but was approved in the United States in 1995.
‡ Troglitazone was approved in 1997 but was withdrawn from the market in 2000 because of hepatotoxicity.
In the rosiglitazone group, as compared with the control group, the odds ratio for myocardial infarction was 1.43 (95% confidence interval [CI], 1.03 to 1.98; P=0.03), and the odds ratio for death from cardiovascular causes was 1.64 (95% CI, 0.98 to 2.74; P=0.06).
Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes

Table 4. Rates of Myocardial Infarction and Death from Cardiovascular Causes.

<table>
<thead>
<tr>
<th>Study</th>
<th>Rosiglitazone Group</th>
<th>Control Group</th>
<th>Odds Ratio (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. of events/total no. (N)</td>
<td>no. of events/total no. (N)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Myocardial infarction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small trials combined</td>
<td>44/10,285 (0.43)</td>
<td>22/6106 (0.36)</td>
<td>1.45 (0.88–2.39)</td>
<td>0.15</td>
</tr>
<tr>
<td>DREAM</td>
<td>15/2,635 (0.57)</td>
<td>9/2634 (0.34)</td>
<td>1.65 (0.74–3.68)</td>
<td>0.22</td>
</tr>
<tr>
<td>ADOPT</td>
<td>27/1,456 (1.85)</td>
<td>41/2895 (1.42)</td>
<td>1.33 (0.80–2.21)</td>
<td>0.27</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td>1.43 (1.03–1.98)</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Death from cardiovascular causes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small trials combined</td>
<td>25/6,845 (0.36)</td>
<td>7/3980 (0.18)</td>
<td>2.40 (1.17–4.91)</td>
<td>0.02</td>
</tr>
<tr>
<td>DREAM</td>
<td>12/2,635 (0.46)</td>
<td>10/2634 (0.38)</td>
<td>1.20 (0.52–2.78)</td>
<td>0.67</td>
</tr>
<tr>
<td>ADOPT</td>
<td>2/1,456 (0.14)</td>
<td>5/2895 (0.17)</td>
<td>0.80 (0.17–3.86)</td>
<td>0.78</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td>1.64 (0.98–2.74)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

EMR Pharmacoepi Study

- Inclusion criteria: Patient is on a diabetes medication

- Analyzing patients Outpatient Encounter or Inpatient Encounter over 6 month intervals

- Characterize patients
  - Age, Gender, Prior cardiovascular disease, use of hypertensive Meds, use of hyperlipidemic Meds, Outpatient insulin, Charlson score, Creatinine, Hba1c, Race, insurance

- For each 6 month period
  - Patient must have inpatient/outpatient visit
  - Determine whether prescribed:
    - Rosiglitazone, Metformin, Sulfonylurea, Pioglitazone

- Follow patients longitudinally
  - Earliest time period that pt meets cohort definition
  - Analyze patient observation periods until: Death, Myocardial Infarction (IC9-410.x), End of Study or gap in care
Drug Exposure: Preliminary Results

- Total T2D patient- 6 month periods between 1/1/2000 – 12/31/2006:
  - 159,604 periods (from 34,278 patients)

- Total patient-6 month periods on Rosiglitazone:
  - 8,707 periods (4,274 patients)
  - 2,888 mono-therapy (1,909 patients)
  - 1,936 with metformin alone
  - 1,709 with sulfonylurea alone
  - 2,174 with all three

- Total patient-6 month periods on Metformin:
  - 48,713 periods (18,162 patients)
  - 28,630 mono-therapy (12,774 patients)

- Total patient-6 month periods on Sulfonylurea:
  - 40,694 periods (17,157 patients)
  - 20,838 mono-therapy (11,435 patients)
Pharmacovigilance Rosiglitazone Study
Event Identification in relation to Drug Exposure

- **Total events**
  - 5,108 MI events (64 per 1000 patient-years)

- **Total events associated with Rosiglitazone:**
  - 477 MI events (110 per 1000 patient-years)
  - 203 MI events on mono-therapy (141 per 1000 patient-years)

- **Total events associated with Metformin**
  - 1,544 MI events (63.4 per 1000 patients-years)
  - 721 MI events on mono-therapy (50 per 1000 patients-years)

- **Total events associated with Sulfonylurea:**
  - 2,405 MI (118 per 1000 patients-years)
  - 1512 MI events on mono-therapy (145 per 1000 patient-years)
Diabetes Mellitus Avandia Study
Modeling Strategy

- Fit generalized linear models (GLM)
- Poisson distribution
- Exposure time as offset
- Accounts for overdispersion

Model:
- Univariate
- Full Coverage Model (Only variables with data on each individual)
- Comprehensive Model (All 11 variables)
Rosiglitazone vs Metformin:
- Crude RR: 3.0 (95% CI 2.7-3.3)
- Comprehensive Adjusted RR: 2.5 (95% CI 2.2-3.0)

Rosiglitazone vs Sulfonylurea:
- Crude RR: 1.0 (95% CI 0.9-1.2)
- Comprehensive Adjusted RR: 1.2 (95% CI 1.0-1.5)

Rosiglitazone vs Pio:
- Crude RR: 1.5 (95% CI 1.2-1.9)
- Comprehensive Adjusted RR: 1.6 (95% CI 1.2-2.3)
Prospective Surveillance

Initial Model

Full Coverage Model

![Graph showing relative risk of myocardial infarction over time comparing Rosiglitazone vs Metformin.](image1)

![Graph showing relative risk of myocardial infarction over time comparing Rosiglitazone vs Metformin.](image2)
Prospective Surveillance

Initial Model

Full Coverage Model
The graph shows the relative risk of myocardial infarction over the years 2000 to 2007 for three different comparisons:

- **Rosiglitazone vs Pioglitazone**
- **Rosiglitazone vs Metformin**
- **Rosiglitazone vs Sulfonylurea**

The relative risk is measured on the y-axis, while the years are indicated on the x-axis. The error bars represent the variability or uncertainty in the data.
Study findings

- Despite model assumptions, identified potential CV risk of rosiglitazone based on real-word population
- Confounding by indication → pio results provide confidence
- Risk is consistent with other studies
  - RR against Sulfonylurea of 1.2 → same as NEJM meta-analysis, but higher than others
  - RR against Metformin of 2.5 → higher than other studies
  - Trend is similar to McAffee et al with increased risk when comparing to metformin but results in that study were not significant
  - RR against Pio of 1.6 → consistent with recent Gerritis paper (showed 22% reduction in AMI for PIO)
- EMRs as an innovative approach to rapid safety signal detection
Ongoing Research Challenges

- Much electronic information is in free text
- Medical record data are not complete
- Data quality
- Relationship to RCT data
- Patient Privacy
Next Steps

- NLP approaches to augment drug exposure identification
- Experiment with other drug-event associations
- Relevance networks to look at all comparisons and identify unknown associations
- Implementing methodology and visualization prospective monitoring
How to measure population health?

Non-traditional non-healthcare data
gathering population data outside healthcare
Case Study: 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)
Public Health Surveillance

Need for evidence-based decisions for implementation and targeting of control activities

Challenge outside the scope and budget of traditional surveillance systems
Traditional public health reporting

World Bodies (UN, WHO, FAO, OIE)

Ministry of Health

Local officials

Practitioners

Public

Labs

Ministry of Health

Local officials

Practitioners

Public

Labs

Local officials

Practitioners

Public

Local officials

Practitioners

Public

Practitioners

Public

Practitioners

Public

Practitioners

Public
Informal reporting

World Bodies (UN, WHO, FAO, OIE)

Local Health Officials

Ministries of Health

General Public

Healthcare workers, Clinicians

Laboratories

Public health practitioners

Informal Surveillance
Influenza A (H1N1) Virus, 2009 — Online Monitoring
John S. Brownstein, Ph.D., Clark C. Freifeld, B.S., and Lawrence C. Madoff, M.D.

Veracruz: reporta agente municipal extraño brote epidémico que ha cobrado dos vidas

La funcionaria de La Gloria informó que el raro padecimiento ha afectado a 60 por ciento de sus tres mil habitan tes con infecciones respiratorias.
**HealthMap Article Processing**

**Acquisition**
- >20,000 sites
- Every hour; 24/7

**Extraction**
- 1800 disease patterns
- 5000 location patterns

**Categorization**
- 6 million phrases
- 91% accuracy

**Aggregation**
- Text Matching
- Similarity Score
1,000-150,000 unique visitors/day

>1,000,000 since launch in 9/06

Top visitors:
- CDC.gov
- WHO.int
- DHS.gov
- National, state, local public health depts
- NGOs
- National Conventions
Tool for general population

Researchers Track Disease With Google News, Google.org Money

By Alexis Madrigal July 17, 2008 | 7:00 AM Pacific Time Categories: Disease, Pandemic, Web/Tech

A new tool tracks diseases, contaminants and other threats as they occur worldwide.

By Adam Hartman

What's New Now you can satisfy your morbid curiosity and discover what diseases are plaguing your favorite vacation spot before hopping the next flight. Instead of stumbling through a host of blogs and news sites, visit http://www.healthmap.org, created by Boston epidemiologist Clark Freifeld and research software developer John Brownstein to track disease outbreaks around the world.
HealthMap Stats

- > 200 alerts per day
- >200,000 alerts so far
- Alerts in 201 countries & territories
- 175 disease categories
- Five languages – English, French, Spanish, Russian, Chinese, Portuguese, Arabic
Influenza A (H1N1) Reports

Source:
- Informal Sources (Media)
- Official Sources (e.g., CDC, WHO)

Category:
- Ruled Out
- Suspected Cases
- Suspected Deaths
- Confirmed Cases
- Confirmed Deaths

Zoom to country:
- Afghanistan
- Albania
- American Samoa (USA)
- Antigua & Barbuda
- Argentina
- Australia
- Austria
- Azerbaijan
- Bahamas
- Bangladesh
- Barbados
- Belarus
- Belgium

Date range: 1 Apr - 1 Aug


Official data is obtained from CDC and WHO daily Influenza A (H1N1) updates. Informal data sources are a subset of reports from the HealthMap database. The case numbers shown are cumulative counts. HealthMap is a public website bringing together disparate data sources to achieve a unified view of the current global state.
Global Spread of H1N1

Surveillance of Clusters Closures
Community transmission Changes in transmission Changes in age distribution

Number of Countries with Confirmed Cases

Date
04/11 04/18 04/25 05/02 05/09 05/16 05/23 05/30 06/06 06/13 06/20 06/27
School closures

- 37 schools with at least one confirmed case in 25 counties in 14 states
- 32 public schools, 5 private;
- 13 schools in urban districts; 16 schools in suburban districts; 6 schools in smaller towns; 2 schools in rural areas

37 schools with cases compared with the other 10,035 schools in the 25 counties containing the outbreak schools
→ Increased risk in urban more affluent areas (increased travel history, healthcare access, reporting bias...)

37 schools
Moving from passive to active surveillance: Participatory Epidemiology
Engaging Users

HealthMap Community

- Web contributors
- Data curators
- Email subscribers
- Alert Partners
- Twitter followers
- Social networks
HealthMap and Participatory Epidemiology

Watch live: New York officials discuss swine flu outbreak

Swine flu outbreak tracked with Twitter

Woman Being Tested in Thailand: http://tinyurl.com/d36t26 #swineflu
about 1 hour ago from web

1st Case Confirmed in Israel: http://tinyurl.com/cp23l8 #swineflu
about 1 hour ago from web

3 confirmed swine flu cases in New Zealand:
http://tinyurl.com/uri... #swineflu
about 3 hours ago from web

Israel confirms first swine flu case: http://tinyurl.com/7m344
#swineflu
about 3 hours ago from web

Ukraine urges calm and caution, not panic for swine flu
http://tinyurl.com/9lu9w #swineflu
about 1 hour ago from web
Iceland reports its first death from swine flu, cases have spiked on island in recent weeks

By THE ASSOCIATED PRESS (CP) – 1 day ago

REYKJAVIK, Iceland — Icelandic health officials reported their country’s first swine flu death Monday amid a spike in cases on the North Atlantic island.

The victim, an 18-year-old disabled woman, died after falling ill about 11 days ago, Iceland’s Directorate of Health said in a statement. The woman had a pre-existing lung condition which contributed to her death, according to Dr. Olafur Baldursson, the director of medicine at Reykjavik’s Landspitali University Hospital, where she died.

“The girl had been multiply disabled for a long period of time and suffered from a chronic pulmonary disease,” Baldursson told journalists Monday.

Iceland, a country of about 300,000 people, has so far seen a total of 479 laboratory-confirmed cases of the virus - many of them occurring in the past two weeks. Until recently the virus was largely confined to the capital, but swine flu has now spread across the country, resulting in "a huge drop" in school attendance in some rural areas, according to the health directorate.

The directorate added that the virus was adding considerably to the work of the small nation's health care service.

Swine flu, first identified in April, is a global epidemic. The World Health Organization says there have been nearly 400,000 laboratory confirmed cases and over 4,700 deaths linked to the illness. But many countries have stopped counting individual cases and the organization says the true totals are likely to be much higher.

Outbreak missing from the map? Help us by adding it:
http://URL-of-Disease-Outbreak-Report
Community input: Digital Disease Detectives

WHEAT RIDGE, Colo. — The first U.S. case of Marburg hemorrhagic fever has been confirmed in Colorado, and authorities say the patient — who contracted the rare illness while traveling in Uganda — has since recovered.

The patient, who traveled to Uganda, visited a pychit cave in the Mantisa-PUC forest in Queens Elizabeth Park and encountered fruit bats, which can carry the Marburg virus. The Ugandan government closed the cave after a tourist from the Netherlands died from Marburg in July.

The patient was treated at Lutheran Medical Center in January 2008 and sought follow-up care in July, after learning of the tourist’s death. The patient recovered and his or her identity wasn’t disclosed.

Pierre Rollin, acting chief of the Special Pathogens Branch of the CDC, said specialized tests of the initial sample taken in January 2008 confirmed the illness in the Colorado patient in December.
Florida

30 Jun Two horses in eastern Leon County euthanized after testing positive ...

30 Jun Clay County Reports EEE in Three Horses | Firstcoastnews.com | Local ... (submitted by: a.kitpowell)
Is a Swine Flu Outbreak Coming? Ask Your iPhone

By Bryan Walsh
2. **26-year-old falls victim to H1N1 in Mass.**  
8/17/09 8:43 PM  
3 hours 39 min 26 sec ago (NECN: Boston, Mass.) - The H1N1 flu is dominating news headlines today. Massachusetts is reporting its 11th swine flu death, as Boston Mayor Tom Menino looks to prepare for a possible resurgence of the virus this fall. Boston

3. **West Nile virus found in Boston**  
8/13/09 10:18 PM  
By Stephen Smith, Globe 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-alerting

Diseases to receive alerts
- Soybean Rust
- Swine Flu H1N1
- Tick-borne disease
- Vaccine Complication

Location
- Current Location

Email

Geo-reporting

Submit Outbreak Report

Anthrax

Location
- Current Location

Upload Related Photo
- Photo Attached

Email
80k downloads in first week
crowd-sourced surveillance

- **Closures** “Outbreak of laboratory-confirmed H1N1 in schools in the provinces of Lucca and Pisa, Tuscany, Central Italy” “Canterbury elementary school closed until 10/23 due to 30% percent of students out with flu.”

- **Clinical** “10 year old boy confirmed test positive. Initial onset came on fast with extreme headache, fatigue, and low/med grade fever. He began tamiflu same day. Is asthmatic. So far is recovering well.”

- **Clusters** “First my 5 year old son got it then my 18 month old daughter got it. Now my wife and I both have it.”
Part II: Conclusions

- Value in the fusion and visualization of distributed electronic resources

- Complements traditional healthcare data sources by providing customized real-time intelligence for the broad scope of international public health activities

- Importance of multi-lingual, collaborative approach that minimizes information overload and engages users
Key questions

- What are the regulatory obstacles impacting your work?
- What are the resource needs required to replicate your work at other institutions?
- What are the priority short term "translational" questions in your fields that would represent the most rapid payoff on investment?
Regulatory Obstacles

Non-traditional healthcare data
- IRB
- Data security
- Anonymization of data (especially when considering rate events)
- Data standards
- Adoption of EMRs is low (only 33% of physician offices use)
- Lack of universal patient identifier
- Data often silo’ed by clinical specialty, lab service, etc

Non-traditional non-healthcare data
- Much of the non-traditional data falls outside regulatory structures
Resources Needed

Non-traditional healthcare data

- NLP approaches (eg: augment drug exposure identification)
- Database storage capacity
- IT infrastructure
- Signal processing technology
- Implementing methodology and visualization
- Prospective monitoring (human analysts)
- Validation efforts
- Common data “clearinghouses” – allow parties to publish data while retaining control
  - Patient locator/identifier/anonymization services
  - Patient consent
  - Billing code resolution
  - Drug metadata
Resources Needed

Non-traditional non-healthcare data

- Validation efforts
- Population representativeness
- Extrapolation of results
- Appropriate denominators
- Historical baseline
- Self-report biases
- Sampling frame
- Signal to noise
- Validity of empirical surveillance data for epidemiologic studies
- Value to make regulatory decisions
Priority short term question

- Mining EMRs for discovery (safety, benefit, drug interactions)
- Mining search usage data for discovery (safety and benefit and drug interactions)
- Use of social networking for clinical trials recruiting, safety and discovery
- Mobile devices and EMR, to reach underserved populations US & worldwide
- Beyond drug safety: could we have discovered virus - CFS association from EMR data?
Public Health 2.0 and EMRs

Community input

Post to public

Review

Notify contributor

Aggregate
Social Networking and EMRs

Real time group and individual measurement of outcomes, treatment side effects, adverse events
Mobile Devices and EMRs
HealthMap Team
- John Brownstein, PhD
- Clark Freifeld
- Mikaela Keller, PhD
- Susan Aman
- Amy Sonricker, MPH
- Emily Chan, MSc

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

Partners
- Susanne Churchill, PhD
- Shawn Murphy, MD PhD

Collaborators
- Larry Madoff (MA-DPH/ProMED)
- Marjorie Pollack (ProMED)
- Tim Brewer (ProMED)