Monitoring and molecular diagnosis of emerging infectious diseases

Raymond Lin, MD
Singapore
Nov 2010
Singapore

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
<th>Land Area</th>
<th>Per Capita GDP</th>
<th>Literacy Rate</th>
<th>Crude Birth Rate</th>
<th>Crude Death Rate</th>
<th>Life Exp at Birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>4.99 Million</td>
<td>710.3 Sq km</td>
<td>S$51,656</td>
<td>96.3%</td>
<td>10.2/10³</td>
<td>4.4/10³</td>
<td>80.9 years</td>
</tr>
</tbody>
</table>

Source: Singapore Department of Statistics
a. zoonotic wildlife b. zoonotic non-wildlife
c. drug-resistant pathogens d. vector-borne

Global distribution of relative risk of an EID event  

Jones 2008 (Nature)

predictors: high population density, richness of wildlife species
• Something new is happening
  – Detect early that something is going on (epidemiology)
  – Find out what it is (microbiology)
clinical epidemiology laboratory
Informs policy and action
Preparedness for (naturally-occurring) infectious disease outbreaks = Preparedness for biothreat events
Using pandemic influenza (H1N1) as an example of threat response
Estimating number infected
Peak of epidemic
3 methods

• Clinical surveillance with lab tests
• Small syndromic sampling
• Serosurveys
Sentinel site monitoring

- General practitioners (mild flu)
- Emergency Departments (moderate)
- Hospitals (severe ICU cases)
- Laboratory confirmation of influenza
- Collection of virus specimens for analysis
When decisions are made on a day-to-day basis, we need to collect data on a daily basis (and must be statistically valid).
Surveillance testing

- Before pandemic: weekly data, 20-50 samples/week
- Pandemic
  - *daily* monitoring
  - Achieve statistical confidence to detect 1% change = 160 samples/day
Use of surveillance

• Decision points used
• Examples
  – 0 to 1% : ? Signal of community spread – transit into mitigation *
  – >15% : justifies empiric use of antivirals in susceptible groups
95% CI for proportion (%) of influenza A/H1N1/2009 among ILI samples from polyclinics and GP clinics

* Figures for E-week 52 are preliminary

Clinical surveillance and lab tests

* Figures for E-week 52 are preliminary
Weekly Influenza Surveillance and Flu A Typing Results

Influenza A positivity from Polyclinics

Influenza A typing results from Polyclinics

* Figures for E-week 52 are preliminary.
When will be the peak of 1st wave?

Peak around at e-week 29/30.
Maximum case number over the first wave of H1N1(2009) pandemic = 133,000 (approx).
Cumulative numbers

Red circles: observed numbers

Fitted by a Richards curve

Maximum case number
From clinical/ lab surveillance, we estimated
5 % infected end of August 2009
9% infected end of March 2010
Serosurvey approach
4 cohorts

• Community 838
• Military 1213
• Hospital staff 558
• Longterm care facilities (nursing home) 300

serosurvey
conclusions

• **Community and military** converted earlier than hospital and LTCF
• >20% conversion in 20-24 yr age group
• Household contact
  – Yes 29%
  – No 12%
conclusions

• Older age protective
  – OR 0.42 per 10 years
Final seroconversion

- Community 13%
- Military 29%
- Hospital staff 7%
Modelling approach
Modeling approach using a small sentinel clinical group
23 GP sentinel clinics without lab test
-June to Aug 2009

Ong 2010 (PLOS One Apr 2010)
From the modeling study

“... the eventual forecast was that 13% of the population had been infected” (9-19%)
Estimated % infected in community in by end-Sep 2009
1. Clinical surveillance 5%
2. Predictive modeling 13%
3. Seroepidemiology 13%

The 3 groups of experts worked independently.
Seroepidemiology is retrospective, accurate and can analyze subgroups (but hard to organize)

Clinical surveillance makes assumptions, has lab confirmation and can be relatively real-time (but need good stats and high volume of testing)

Modeling is predictive, is easy to organize, but needs to be verified by the other methods
Monitoring “virulence”

Primarily clinical and epidemiological
Virological monitoring
In the 2009 pandemic, worldwide, we had unprecedented access to
- Rapid sequencing of many isolates
- Whole genome sequences
- Specialized labs with animal and other biological models

What did those studies tell us?
Monitoring virus mutations

- Antiviral drug resistance
- “Known” virulence markers
- Antigenic changes
- Unknown function
Pathological examination of the lungs of infected cynomolgus macaques.

H1N1 2009

• *Itoh 2009*
  - Replicate more effectively
  - More severe pathology in lungs of mice, ferrets, primates
  - Transmit among ferrets
• *Maine 2009 Science:* worse disease in ferrets and mice, but lower transmissibility (!)
• Increase affinity for alpha 1,3 receptors (lungs) appeared supported by modeling and glycan studies
H1N1 2009

- *Soundararajan 2009* Nat Biotech
- Theoretical complex: bind mostly alpha 2,6 but also alpha 1,3
Maines 2009
H1N1 2009 binding pattern similar to 1918 virus but lower affinity
• “D222G” (H1) mutation in severe cases
• “Based on currently available virological, epidemiological and clinical information, the **D222G substitution does not appear to pose a major public health issue.**” (WHO)
D222G (H1)

- Found since April 2009
- Kilander 2010 Eurosurveillance
  - 11/61 (18%) severe cases in Norway, 0/205 mild
- Observed elsewhere but lower <10%
- Also found in asymptomatic
- Mutation in receptor binding domain
D222G (H1)

- ?statistical bias
- ?influenced by growth in eggs
- Found in autopsy tissue from Ukraine
- ?selective bias from lung
- Not sure still!
- No strong virological or experimental support
Mutated D222G on reference virus – still maintain predominant affinity for alpha 2,6 receptors
PB2

• Polymerase basic protein 2
  – Adaptation for replication in avian to human
  – Previous pandemic strains had E627K mutation
• Taubenberger 2005 : important for adaptation and pathogenicity
• Mutation absent from H1N1 2009, but what if? ...
PB2

Reconstructed “worst case scenario”

• Jagger 2010 Mbio
• Herfst 2010 J Virol
• Zhu H 2010 Virology
• Jagger 2010 mbio

• Using recombinant/ RG viruses – 627 mutation has no effect or attenuates
Figure 2
Full genome maximum likelihood phylogenetic analysis of pandemic influenza A(H1N1) variants from Singapore, Australia and New Zealand and other non-redundant (>80% identity) strains.

Legend
- A/2009
- Jan-Mar 2010
- Apr-May 2010
- June-Aug 2010

Naturally occurring 2009 pandemic influenza A(H1N1) viral sequences available in GenBank as of 23 August 2010 were downloaded from the NCBI Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). Phylogenetic analysis was conducted on 1,977 strains with full-length nucleotide sequences available for all eight segments. The protein-coding nucleotide sequences for these strains were concatenated such that a single sequence representing a single strain contained nucleotide sequences encoding all 10 proteins. The vaccine strain A/California/07/2009 and recent Singaporean, Thai, New Zealand and Australian strains were set aside and redundancy was removed within the remaining strains with Cci-hit [33] by allowing a maximal sequence identity of 80% to reduce the set to 20 non-redundant strains. This representative set was aligned with the vaccine strain and 16 other strains from Singapore, Thailand, New Zealand and Australia using MAFFT [34] with the FFT-NS-2 option. Sequences flanking the coding region were removed from the alignment. A maximum likelihood tree was created using PhyML [32] with the approximate likelihood ratio test, the HKY85 substitution model and parameters such as for the shape of the gamma distribution (0.298) were estimated by the programme. Substitutions discussed in this analysis were identified and marked in the resulting phylogenetic tree using the MEGA software package [5].

Amino acid changes in the various genes are indicated by name of protein and mutation e.g. PB1 A65V.
2010 viruses are colour-coded by sample date, e.g. 7/10 representing month/year; dec: a virus obtained from a fatal case; vac: a virus obtained from a person vaccinated with pandemic influenza A(H1N1) monovalent vaccine.

The neighbour-joining phylogenetic tree was constructed using the PAUP (v4.0) plugin on Geneious (v5.0.4) and FigTree (v1.3.1) was used to display the summarised and annotated trees with bootstrap values >70 shown on the nodes.
HA: haemagglutinin.
Grey ribbons: influenza A/Brisbane/10/2010(H1N1), sites of all mutations of interest indicated including the positions of the additional ones from the New Zealand strains D111N and V267A.
Green ribbons: a bound antibody in the orientation resulting from superimposition of our HA model with that of the antigenically similar 1918 influenza A(H1) HA in complex with this antibody.
Upper right corner: surface representation of the HA model rotated by 90 degrees with the residues of the Sa epitope coloured yellow.
• Herfst 2010: monitoring should not rely on virulence markers of the past
Other methods of monitoring

GoogleTrends: limited availability
Google Insight: limited by news-driven peaks
Cook 2010 Emerging Infectious Diseases
Twitter and flu trends
-Comparable to Google Flu Trends

Culotta : 2010 Workshop on Social Media Analytics at the Conference on Knowledge Discovery and Data Mining
Influenza Prediction Market to Help Strengthen Early Warning System for U.S. Bird Flu Outbreak

New Robert Wood Johnson Foundation-funded prediction tool will give public health officials better forecasting and response capabilities.

Published: Mar 01, 2007  | Iowa City, Iowa

To prepare for a bird flu pandemic, the University of Iowa today unveiled a new tool to help public health officials better predict when the disease will strike and plan ways to stem its effects. Through a unique collaboration among the University's Colleges of Business and Medicine, the Iowa Health Prediction Market is launching the 'Avian Flu Market (AFM),' an information trading and aggregation system to help public health officials around the world collect and analyze information to forecast the timing and extent of a human-to-human bird flu outbreak. The project is supported by a $245,685 grant from the Robert Wood Johnson Foundation (RWJF).
the challenge: massive digitalized data

the tools: data mining in its many forms

the objective: patterns and clusters and alerts
Care Quest
Infection Surveillance and Management (ISaM)
MRSA/Other Organisms

Powered by: Global Mobility Laboratory
Infection Surveillance & Management (ISaM)
CareQuest: RT bioinformatics data warehouse
CQ/ISaM Framework

Care Quest
- Bio-Informatic DW Stores
  - patient caches
  - movements tracking
  - KPIs
  - reporting, trending
  - UMLS coding standards
    - SNOMED CT
    - LOINC, ICD9 & 10

Patient Tracker
- patient caches
- pre-admission watch
- movements watch
- post discharge watch

Surveillance Engine
- Definitions & Knowledge rules
  - triggers
  - case typing
  - attributions
  - compliance cohorting

HMS
- Real-time
- Stakeholders Alerting
- Close-loop mobile comm.
  - ICN
  - BMU
  - Wards
  - Doctor-of-care

CQ/ISaM Framework

Care Quest

Patient Tracker

Surveillance Engine

HMS

Strictly Confidential, Copyright 2010 - Global Mobility Laboratory Pte Ltd
Surveillance

- MRSA
- MRSA Program
- MRSA Bacteraemia
- Clostridium Difficile
- VRE
- AFB Smear
- MDRO
  - Acinetobacter baumanii
  - Klebsiella pneumoniae
  - Pseudomonas aeruginosa
- Alert Organism
  - Neisseria meningitides
  - Salmonella typhi
  - Salmonella paratyphi A
  - Vibrio cholerae
- Surgical Site Infection
- Sharps/Mucous Membrane Exposure
- NDM-1 Superbug
  - Carbapenem resistant Klebsiella
  - Carbapenem resistant E. coli
MRSA/Other Organisms

Old Workflow (manual process)

Patient Information (SAP)
- Admission
- Discharge
- Transfer

Lab Results (LIS)

Combine, compare & categorize
- Case type and attribution
- Follow a set of definitions

Statistics
- Prevalence, transmission, burden, etc
- Graph

Great amount of time is spent on collating, arranging and presenting of the information.
MRSA/Other Organisms

New Workflow (automated in real-time)

Lab Results (LIS)

Data Management
- Case type and attribution
- Follow a set of definition

Patient Information (SAP)
- Admission
- Discharge
- Transfer

Care Quest
Infection Surveillance

Statistics
- Prevalence, transmission, burden, etc
- Graph

Infected patients are now identified through Care Quest.
MRSA/Other Organisms

Statistics on surveillance performance
MRSA – Cohortung Compliance

Old Workflow (many manual steps)

The MRSA positive patient is picked up during the daily laboratory download.

ICN will pass the daily download of MRSA positive patients to the lab staff.

ICN will then check through the SAP system to check the time in which these MRSA patients are cohorting according to the definition.

ICN will collect the number of patients that are cohorting according to the definition based on the cut off time daily.
MRSA – Cohorting Compliance

New Workflow (automated, empowered staff validation policy)

The MRSA positive patient is identified through Care Quest.

The cohorting compliance data is identified through the acknowledgement of the staff at the ward.

Cohort compliance is tabulated in Care Quest.

ICN will collect the number of patient that is cohorted according to the definition based on the cut off time.

Refers to HMS Surveillance
“syndromic surveillance” has many forms and its value in detecting new pathogens and biothreat events is unproven
the alert individual physician functioning within a responsive network is vital and is opposed to systematic syndromic data collection

Examples of alert individuals
-AIDS in 1981
-SARS 2003
-And many at local level
Laboratory methods for diagnosis

- Train many labs with PCR
- System to centralize pathogen culture, genetic and biological studies
- New technologies under study
  - MALDI-TOF
  - Roche GX Junior
  - Agilent MassTag
  - Pathogen chips and rapid devices
  - Bead micro-arrays
  - IBIS (Abbott T5000)