

Data acquisition and processing in the New York City Emergency Department Syndromic Surveillance System

Adam Karpati, MD MPH

New York City Department of Health and Mental Hygiene
and Epidemiology Program Office, CDC

Background

- Post-WTC disaster bioterrorism surveillance in 15 emergency departments citywide
 - Dedicated data collection on paper – clinicians categorized patients into syndromes, EIS officers coordinated (24 hrs/day)
 - Data entry on site, hand delivery or email of separate files to central database
 - Concatenation and analysis
- Early chief complaint-based system
 - Faxed logs from 15 hospital emergency departments
 - Hand coding of chief complaints into syndromes
 - Entry of syndrome totals into central database

Current System -- Setup

- Epi personnel based on-site not sustainable
- Fax-based, hand-coded system highly labor-intensive, cumbersome, not scalable
- Moved towards a robust and sustainable system

System Goals

- Surveillance for prodromes consistent with BT agents (early detection and outbreak monitoring)
- Continue use of existing data
- Electronic
- Timely
- Relatively less labor-intensive
- Flexible for other syndromes/diagnoses
- Built-in response capabilities

System setup

- Pre-existing analytic protocols for syndrome-coded data (EMS system, EISO system)
- Context – post-WTC, anthrax
- No dedicated servers or information system

Design Considerations

- Outreach and relationship-building
 - CEO offices
 - Infection Control Practitioners
 - Emergency department directors
 - IT departments
 - On-site consultation by DOH staff
- Key goals: buy-in and cooperation around data transmission and response (not separate tasks)

Data requirements

- Core variables
 - Hospital name
 - Date of visit
 - Time of visit
 - Age
 - Sex
 - Chief complaint (free text)
 - Home zip code
 - +/- Unique identifier
- Discharge diagnosis not generally available in timely manner
- Need to consider response protocols – patient identification, logistics

Data format and transmission

- Not standardized at facility level
 - Ease of reporting
 - Expediency
 - Formats: spreadsheets, delimited and fixed-width files
 - Transmission:
 - 22 email
 - 16 FTP
 - Automation
 - Security

Daily data transmission

HHC



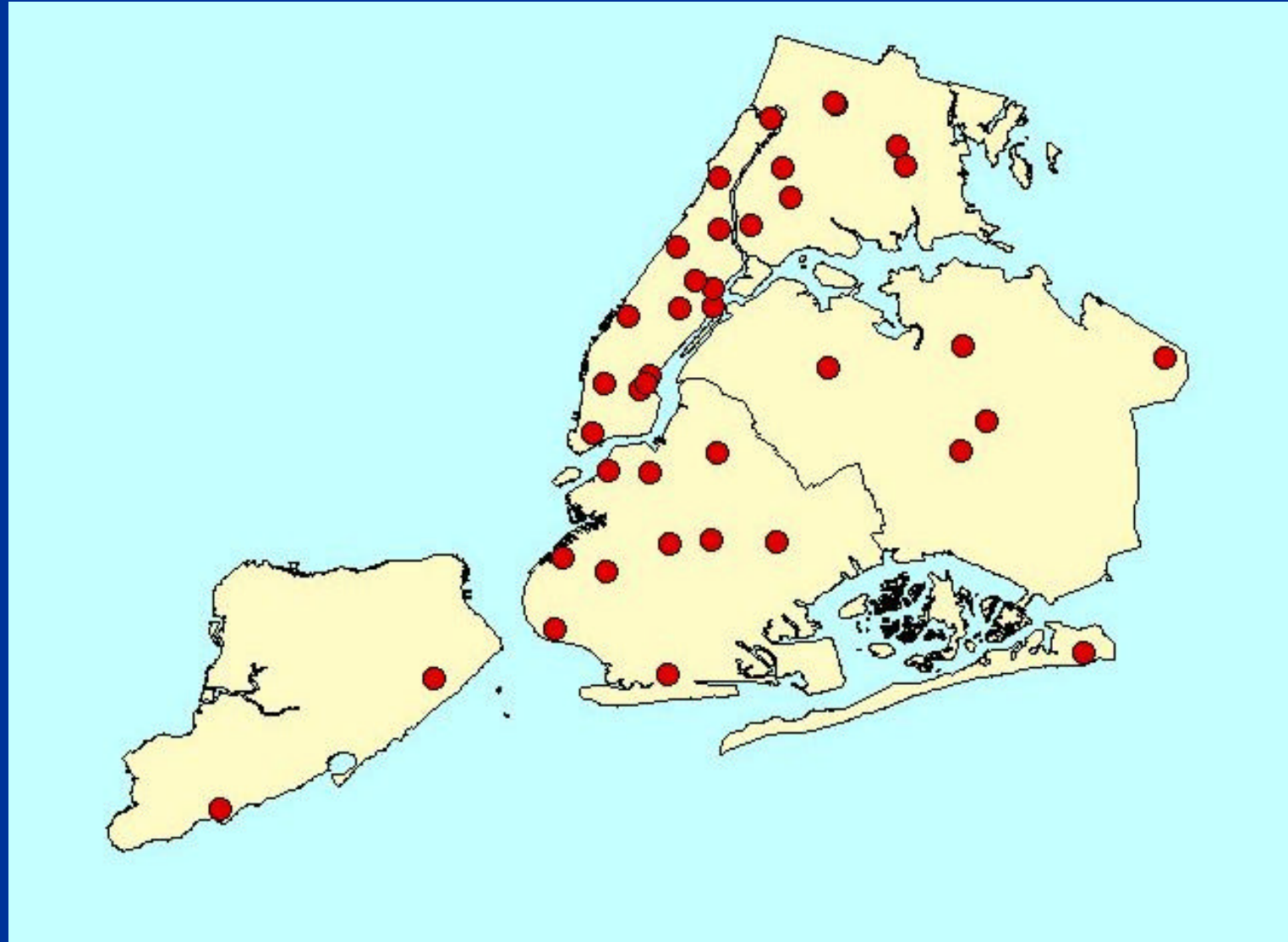
Email server at
DOH (NYC)

FTP
server

Stand-alone PC
at DOH

Data available for analysis

— manual
- - - automatic



38 facilities, ~ 6,500 daily visits
90-95% daily completeness

Data tasks – Participating facilities

- Daily file generation from existing information system
- Transmission to health department
- Troubleshooting
- Integration into daily routine

Data Processing

- Quality control (including callbacks)
 - Full file
 - Correct date
 - All fields present
 - Filter example
- Conversion of individual files to standard (SAS) format
- Concatenation and addition to archive

Electronic coding of chief complaints into clinical syndromes

- Performed in SAS
- Text-string recognition
- Mutually exclusive vs. overlapping
- Hierarchy of coding
- Iterative refinement of syndrome definition
- Entire dataset can be recoded easily – allows for changes in definition and addition of new syndromes

Electronic ED logs

Admission List For 01/28/2002 to 01/28/2002

AGE	SEX	TIME	CHIEF COMPLAINT	ZIP
15	M	01:04	ASSAULTED YESTERDAY, RT EYE REDDENED.	11691
1	M	01:17	FEVER 104 AS PER MOTHER.	11455
42	F	03:20		11220
4	F	01:45	FEVER, COUGH, LABORED BREATHING.	11507
62	F	22:51	ASTHMA ATTACK.	10013
48	M	13:04	SOB AT HOME.	10027
26	M	06:02	C/O DIFFICULTY BREATHING.	
66	M	17:01	PT. MOTTLED AND CYANOTIC.	10031

Text recognition with SAS

```
IF    index(cc,"FEV")>0  
     or  index(cc,"HIGH TEMP")>0  
     or  index(cc,"NIGHT SWEAT")>0  
     or  (index(cc,"CHILL")>0 and index(cc,"ACHILLES")=0)  
     or  index(cc,"780.6")  
     etc.  
     then FEVER=1;
```

Data quality and implications for evaluation

- Interhospital differences
 - Different staff types recording chief complaints
 - Different information systems (e.g. administrative vs. clinical)
- Intrahospital differences
 - Descriptive style
 - Lay vs. technical language

Costs – data tasks only

- Setup
- Maintenance
 - 7 days per week operation
 - 1 data manager per day
 - 2-3 hours per day
 - File retrieval and troubleshooting (e.g. callbacks)
 - Conversion to standard formats, concatenation, archiving, coding (all with verifications)

Additional costs

Time

- 0.5 – 1 hr for analysis and review
- 0 to several hours for response to signals

Personnel

- 1 medical epidemiologist (+ consulting epidemiologists as needed)
- Field surveillance/response staff as needed
- System director (overall supervision, improvements, expansion)

ED system only . . .

Goals

- 100% automated FTP
- Secure transmission protocols
 - Secure FTP
 - Secure Socket Link
 - PGP encryption
- Pre-transmittal standardization
- Automated quality control
- Automated processing

Goals, cont.

- More variables (e.g. disposition, tests ordered, ddx)
- Integration of different systems (ED, EMS, Pharmacy, etc.)
- Automated, interactive feedback to hospitals
- Same-day log transmission

Lessons

- Early emphasis on rapid, broad system setup and implementation → later focus on long-term optimization
- Context matters
- MIS and clinical staff partnerships are essential and need to be maintained
- Labor-intensive (economy of scale)
- Different facilities have different capabilities – need to be flexible and consider non-inclusion

Data in context

A large network of reporting facilities, timely and automated transmission of data, sensitive syndrome coding, and sophisticated statistical analysis are all preludes to the real challenge – response.

Acknowledgements

Sharon Balter

Katherine Bornschlegel

Bryan Cherry

Debjani Das

Annie Fine

Rick Heffernan

Marci Layton

Jennifer Leng

Karen Levin

Hadi Makki

Beth Maldin

Farzad Mostashari

Sudha Reddy

Polly Thomas

Don Weiss

Sheryl Young