Spatial and space-time scan statistics in practice

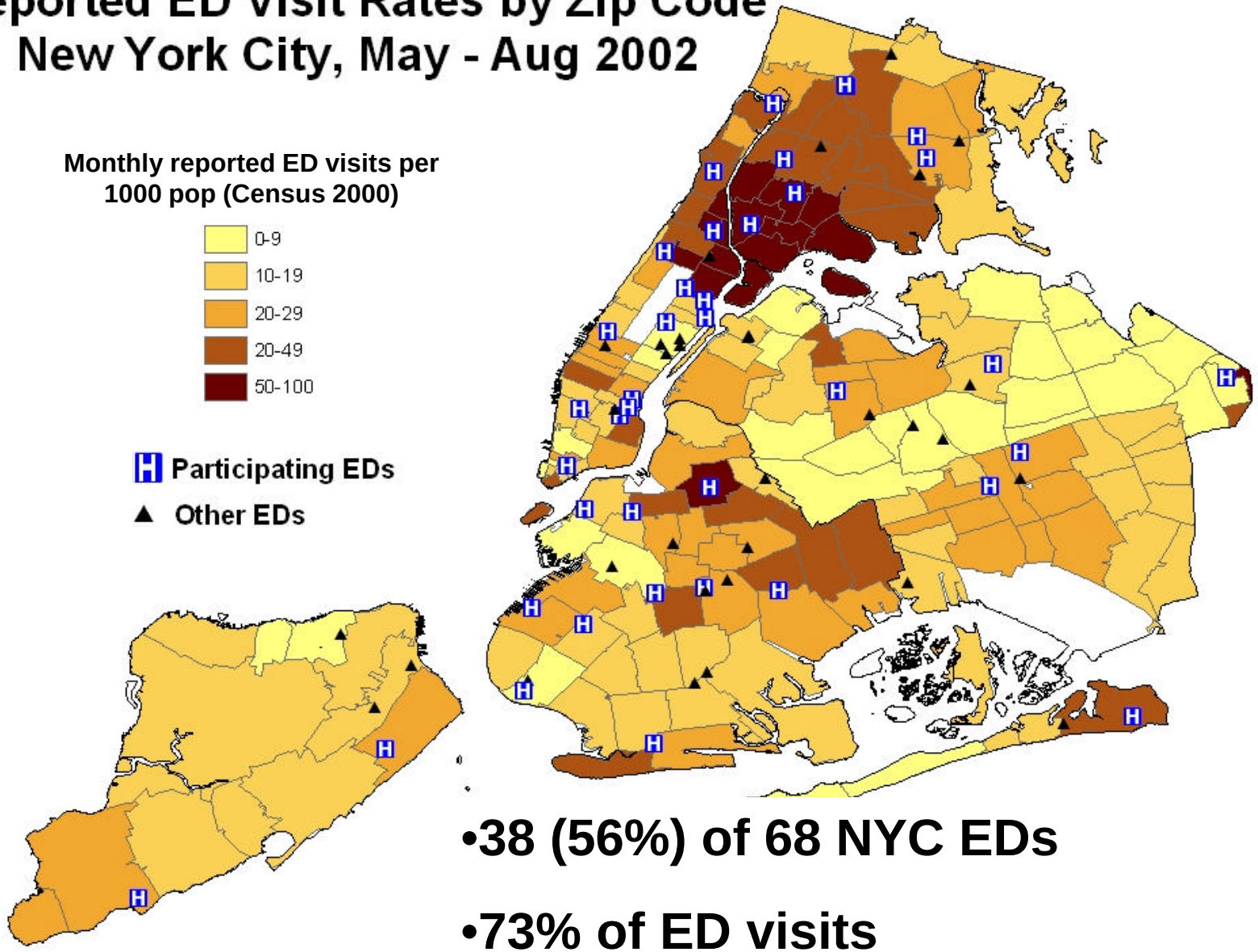
Emergency department chief complaint surveillance, New York City

Rick Heffernan, MPH
Bureau of Communicable Diseases
New York City Department of Health and Mental Hygiene

Objectives

- Briefly describe the NYC emergency department (ED) surveillance system
- Detailed example: respiratory syndrome
  - Source data and chief complaint coding algorithm
  - Purely spatial cluster analysis
  - Space-time cluster analysis
- Gastrointestinal illness in school children

Reported ED Visit Rates by Zip Code
New York City, May - Aug 2002



- 38 (56%) of 68 NYC EDs
- 73% of ED visits

Electronic ED logs

Admission List For 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

- 4% of records have missing or uninformative chief complaint (Eg. ‘See Triage’, ‘Walkout’, ‘N/A’ etc.)

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;
```

NYC Health

Text recognition with SAS

```
IF      index(cc,“PNEUMON")>0
  or    index(cc,“SOB")>0   or   index(cc,“S.O.B.")>0
  or    index(cc,“COUG")>0
  or    (index(cc,“RESP")>0   and  not  index(cc,“RESPO")>0)
  or    index(cc,“786.2")
        etc.
        then RESPIRATORY=1;
```

Key syndrome categories

(1.2 million ED visits reported Nov 2001 – May 2002)

	Age 0-12	Age 13+
Mean # daily visits	1,624	4,762
Key syndromes		
Respiratory	16 %	7 %
Diarrhea	4 %	1 %
Fever/flu-like	21 %	3 %
Vomiting	5 %	2 %
'Other' (not categorized into a syndrome)	39 %	81 %

Respiratory Syndrome, Age 13+

Nov 2001 – May 2002, n = 106,000

Common respiratory complaints

	%
SOB	10.4
COUGH	6.5
DIFFICULTY BREATHING	2.3

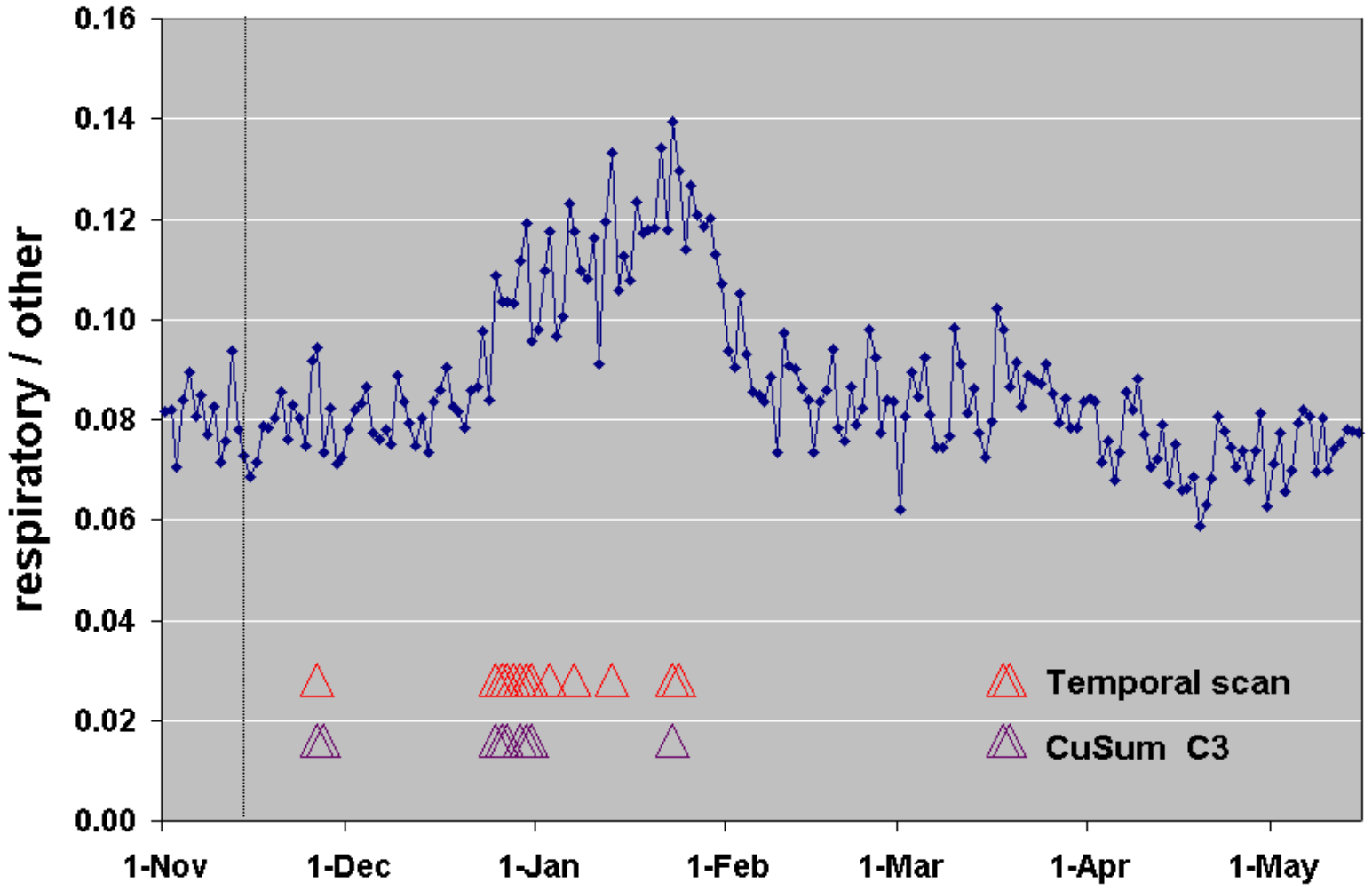
Rare respiratory complaints (occurred once)

PNEUMONIA, ORGANISM UNK

CHEST PAIN AND SOB EKG

ATYPICAL CP ASTHMA BRONCH

Respiratory syndrome, Age 13+
New York City, Nov 2001 – May 2002



Spatial cluster analysis

- Increase sensitivity of signals and timeliness of detection
- Focus epidemiological investigations
- Clustering by hospital address; by patient's home zip code

- **Key Assumption**

  Local exposure ==> localized illness

Spatial scan statistic*

- Can detect clusters of variable size and location
- Probability through Monte Carlo simulations
  - Corrects for multiple comparisons involved in evaluating many possible clusters
- Developed for cancer epidemiology

  Numerator = cancer cases

  Denominator = census population
- Adapted for infectious disease surveillance

  Numerator = Cases (Respiratory ED visits)

  Denominator = ???

*Kulldorff, 1997

Modified spatial scan statistic*

Calculation of zip code denominator

ED patients in zip code during 14-day baseline		ED patients in zip code yesterday
$\frac{respiratory: \ 80}{total\ visits: 1300} = 0.07$	1-day lag	$total\ visits: 90$

*Denominator 'pop' = 0.07 * 90 * 1000 = 6,300*

*Mostashari and Kulldorff

NYC Health

How SatScan Calculates Expected Cases in Zip Code)

$$Citywide\ rate = \frac{12 + 7 + 22}{1210 + 950 + 1130} = 0.0125$$

$$Zip\ code\ Expected = 1130 \times 0.0125 = 14$$

NYC Health

Modified spatial scan statistic*

Calculation of expected in each zip code

Day of analysis

$$\frac{283\ respiratory\ city}{5{,}428\ total\ city} = 0.05 \quad \times \quad 50\ total\ in\ zip\ code \quad = \quad 2.5$$

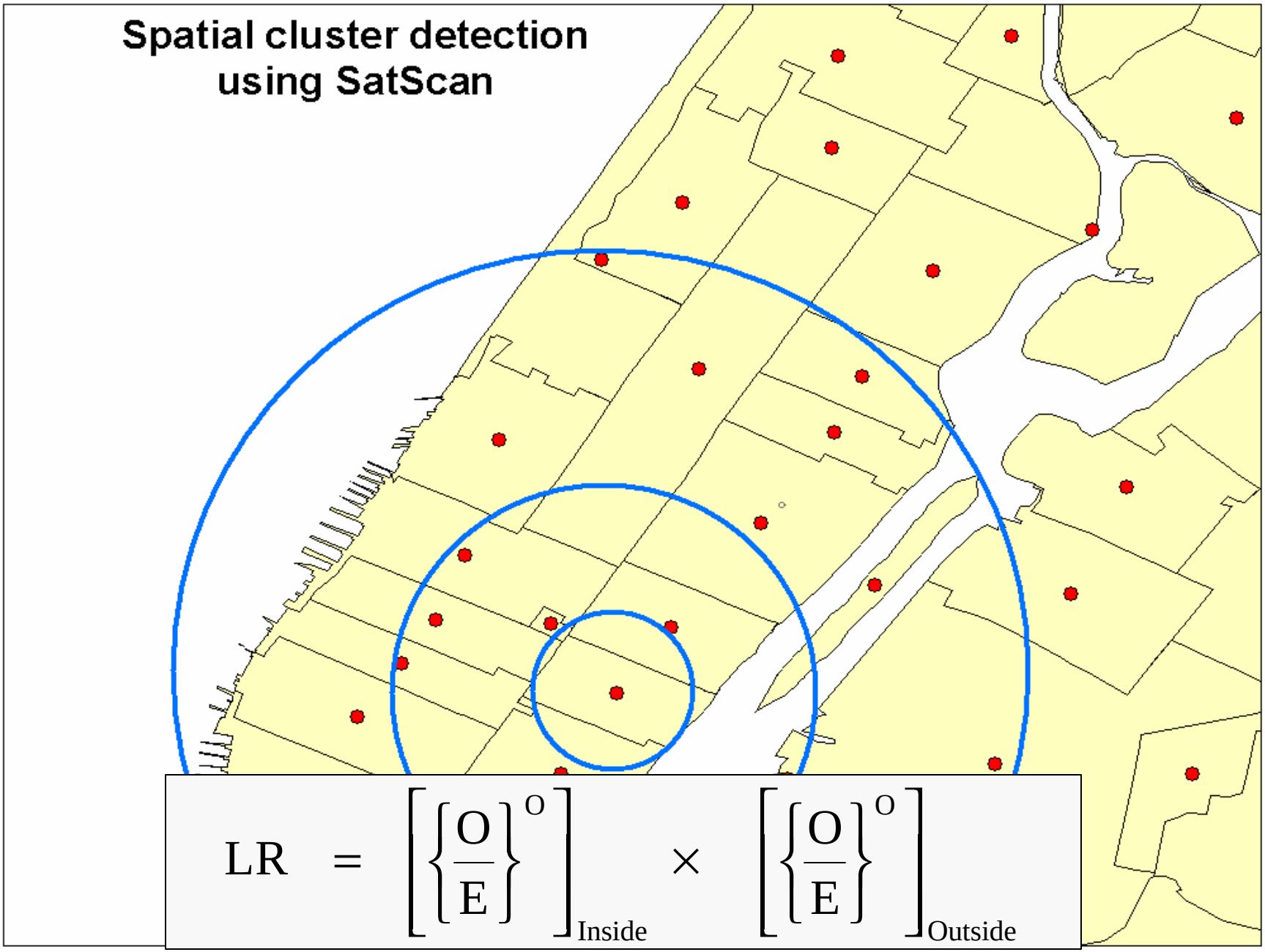
14-day baseline

$$\frac{\%\ resp\ in\ zip \quad = \quad 7\%}{\%\ resp\ outside\ zip \quad = \quad 5\%} \quad \times \quad 2.5 \quad = \quad 3.5 \quad Expected$$

*Mostashari and Kulldorff

NYC Health

Spatial cluster detection using SatScan



$$LR = \left[\left\{\frac{O}{E}\right\}^{O}\right]_{Inside} \times \left[\left\{\frac{O}{E}\right\}^{O}\right]_{Outside}$$

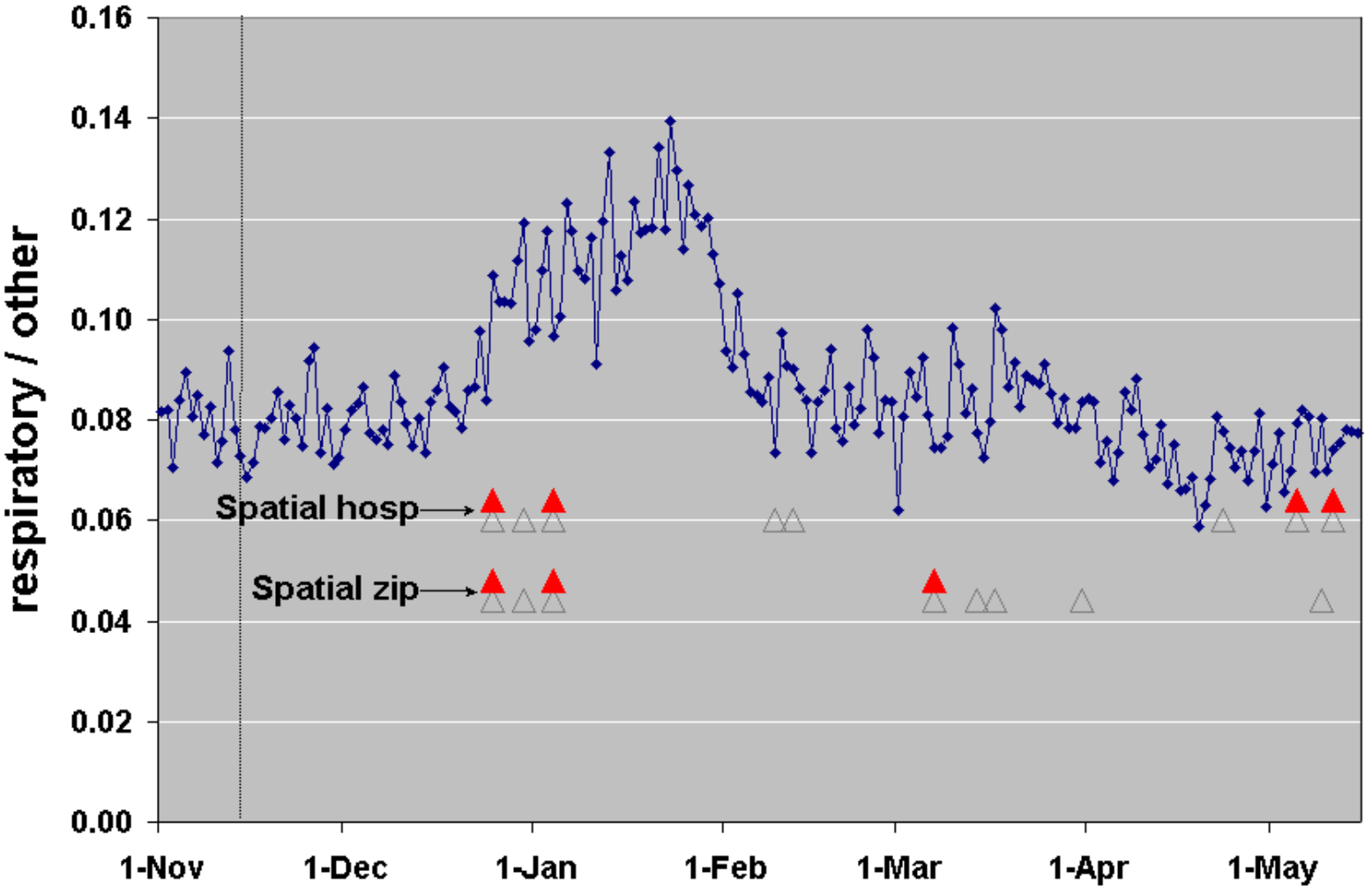
Current spatial cluster analysis

- What question are we answering?

  After adjusting for purely temporal (citywide) trends and for purely spatial clustering observed during baseline…

  did an excess of respiratory visits occur in any zip code or group of zip codes on the day of interest?

Respiratory syndrome, Age 13+
Spatial signals



Limitations of spatial scan statistic as applied to date

- Circles only
  - Future SatScan will handle ellipses
- One-day purely spatial analysis may miss borderline signals spread over several days
- Separate univariate analyses do not evaluate spatial interaction (eg. Simultaneous fever and respiratory increase in same area)

NYC Health

Space-time scan statistic*

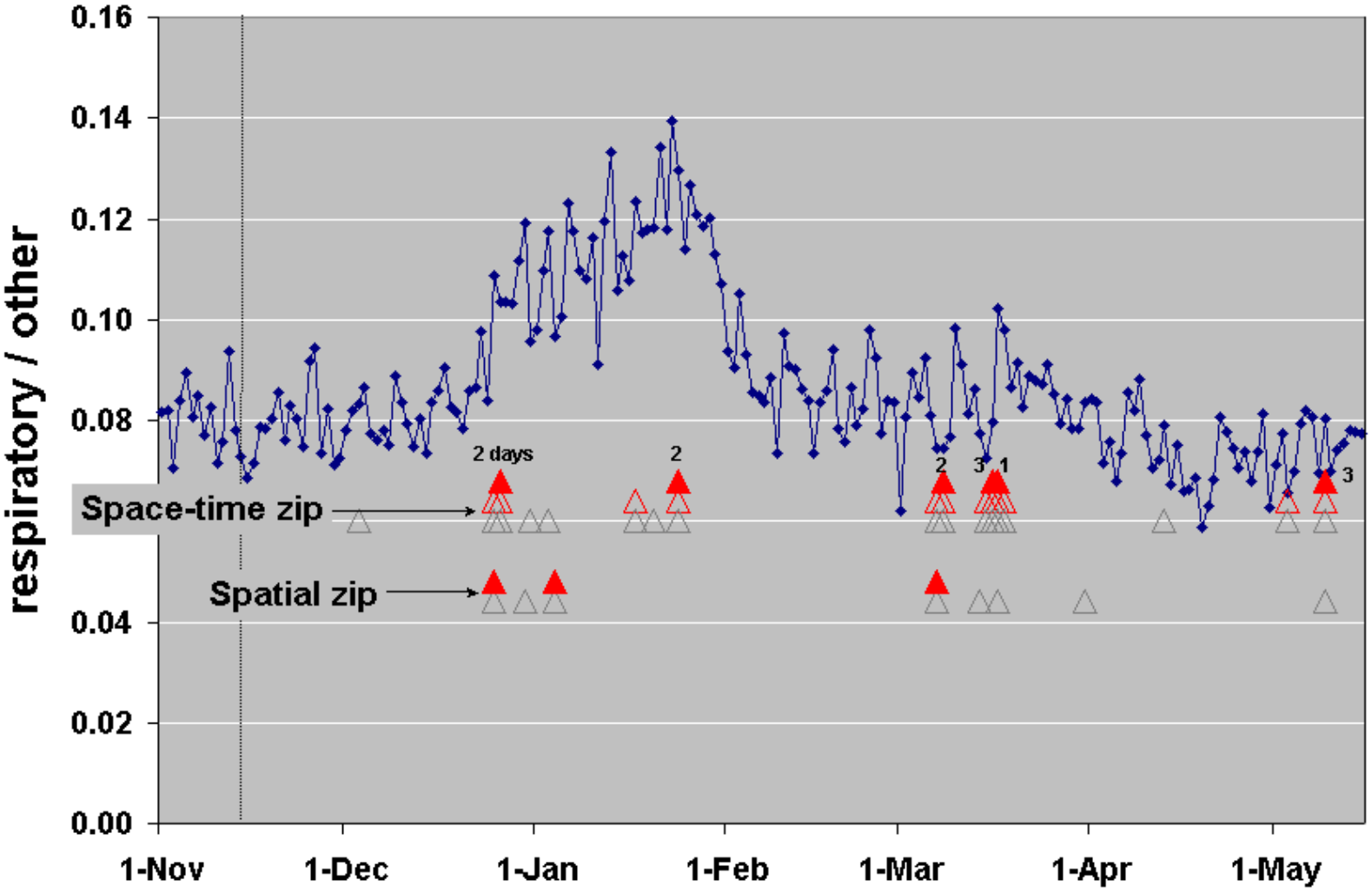
- Flexible window in space and time
- We chose a short 3-day maximum spatial window

  For every unique zip code and group of zip codes, evaluates 1-day, 2-day and 3-day clusters
- SatScan run time increases from 7 to 34 seconds per run (from 5 to 15 minutes for daily analyses)

*Kuldorff, 2001

Respiratory syndrome

Zip code spatial and space-time signals



GI outbreak

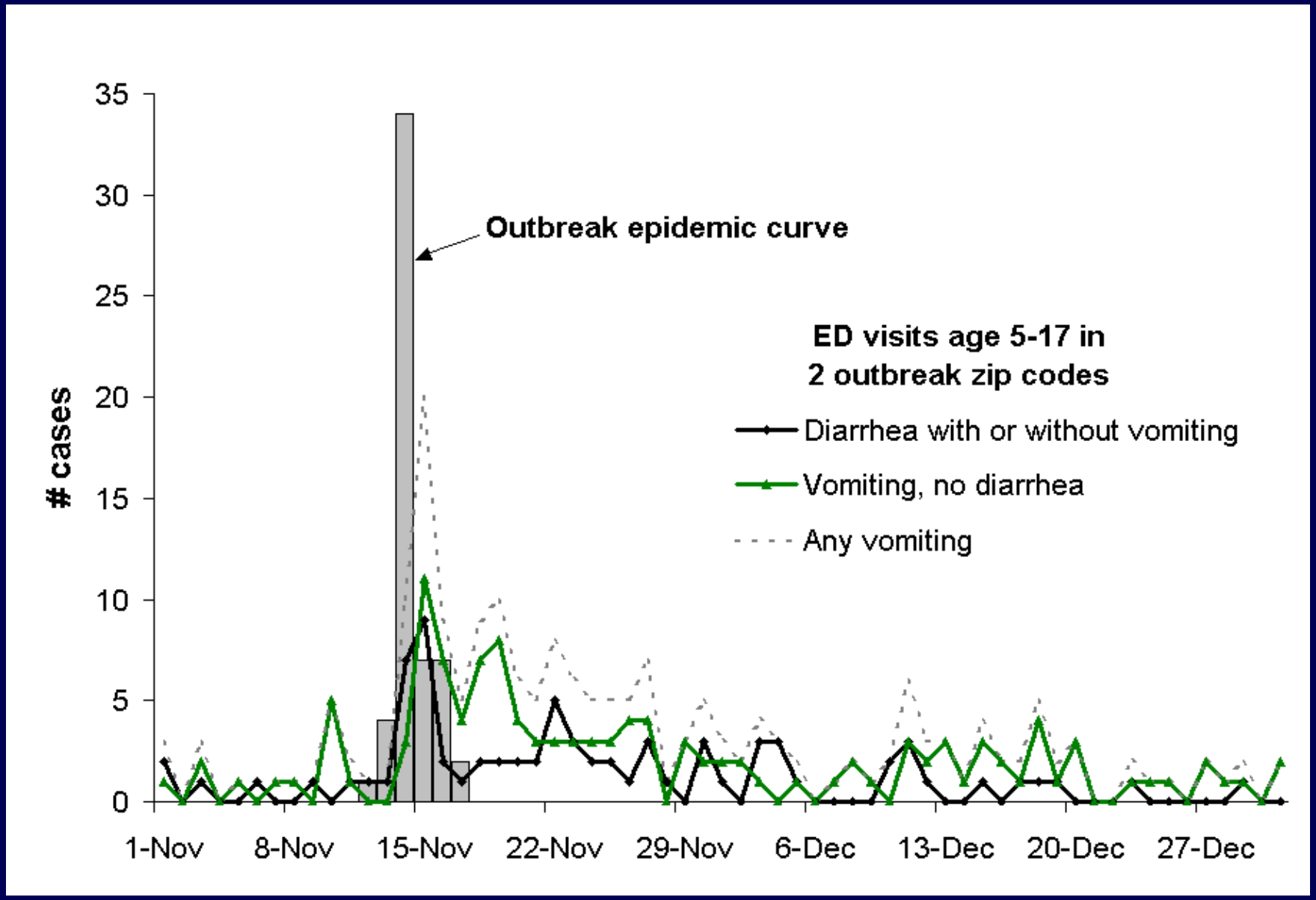
- DOHMH informed of GI outbreak at grade school
  - 19 children “vomiting” after lunch
  - ~150 total cases over course of outbreak
  - DOHMH interviewed 79 cases, 18 (23%) visited ED

    Vomiting 100%

    Diarrhea 75%

    Fever 62%
  - Stools negative for bacterial pathogens
  - No source of exposure was identified
  - One stool culture grew Calicivirus, a Norwalk-like virus

GI illness outbreak



Spatial and Space-time signals in GI outbreak zip codes, by hospital and zip code

			Nov 14	Nov 15
Spatial only	Diarrhea	All ages	ZIP	
		Ages 5-17	ZIP	ZIP
	Vomiting	All ages		
		Ages 5-17		
Space-time	Diarrhea	All ages	ZIP	ZIP
		Ages 5-17	HOSP, ZIP	HOSP, ZIP
	Vomiting	All ages		
		Ages 5-17		HOSP, ZIP

P<0.01, p<0.05

NYC Health

Conclusions

- Spatial cluster analysis of ED chief complaint data can work
  - Both purely spatial and space-time methods detected a signal for this outbreak
- Requirement: EDs involved in outbreak have to be in the surveillance system
- Sensitivity affected by somewhat arbitrary syndrome definitions and age groupings
- Space-time analysis may be more sensitive than spatial only

NYC Health

Future

- Space-time analysis into daily routine
- Next version of SatScan will handle ellipses
- ‘Scan’ other features of data, eg. age
- Space-time interaction between syndromes
- Simulations to better define sensitivity
- Investigate syndromic signal of more real-life outbreaks

NYC Health

Acknowledgements

NYC Department of Health and Mental Hygiene

Analysts
Debjani Das
Hyunok Choi
Sudha Reddy
Bryan Cherry
Katie Bornschlegel
Jessica Hartman
Rich Rosselli

'Cluster Docs'
Don Weiss
Sharon Balter
Jennifer Leng
Polly Thomas
Joel Ackelsberg
Mike Phillips
Elsie Lee
Adam Karpati
Farzad Mostashari
Marci Layton

Field Surveillance
Amanda Adams
Lacretia Jones
Sheryl Young

MIS
Ed Carubis
Hadi Makki
Chris Liang
Jian Liu
Julien Yuen
Shelly Curry

NYC Hospitals: Emergency Departments, MIS and Infection Control staff

Sloan Foundation
Martin Kulldorff (University of Connecticut)

Centers for Disease Control and Prevention (CDC)
Matthew Seaman and Lori Hutwagner

NYC Health