



SOCIAL MEDIA: A NEW DATA SOURCE FOR PUBLIC HEALTH

Mark Dredze
Johns Hopkins University

Michael Paul, Alex Lamb, David Broniatowski

BIG DATA: SOCIAL MEDIA AND HEALTH

- Tweets: ~500 million a day
 - Health Tweets: at least 1 million a day
- Online forums:
 - HealthBoards.com: 4.6m posts, almost 1m members
 - CancerForums.net: 150k posts
 - Drugs-Forum: 500k posts



WHY SHOULD I CARE?

PUBLIC HEALTH:
IMPROVE HEALTH AND
QUALITY OF LIFE IN A
POPULATION

PUBLIC HEALTH:
IMPROVE HEALTH AND
QUALITY OF LIFE IN A
POPULATION

REQUIRES:
DATA ON THE POPULATION

DATA

- Traditional mechanisms
 - Surveys
 - Clinical visits
- **Social media** reflects the ongoing lives of a massive population
 - Health is a big part of that!



PUBLIC HEALTH CYCLE



Population



Doctors

PUBLIC HEALTH CYCLE

Surveillance



Population



Doctors

PUBLIC HEALTH CYCLE

Surveillance



Population



Doctors



Intervention

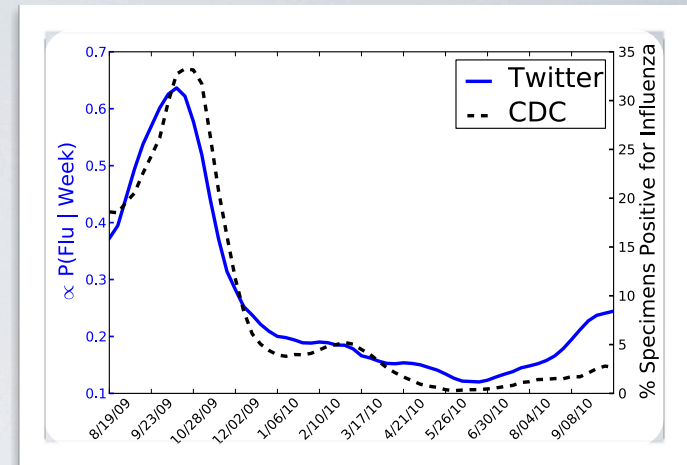
PUBLIC HEALTH CYCLE

- Surveillance
 - Better, faster, cheaper
 - New opportunities
- Intervention
 - Identifying people for communication
 - Tailoring health messages for specific people/populations

EXAMPLES

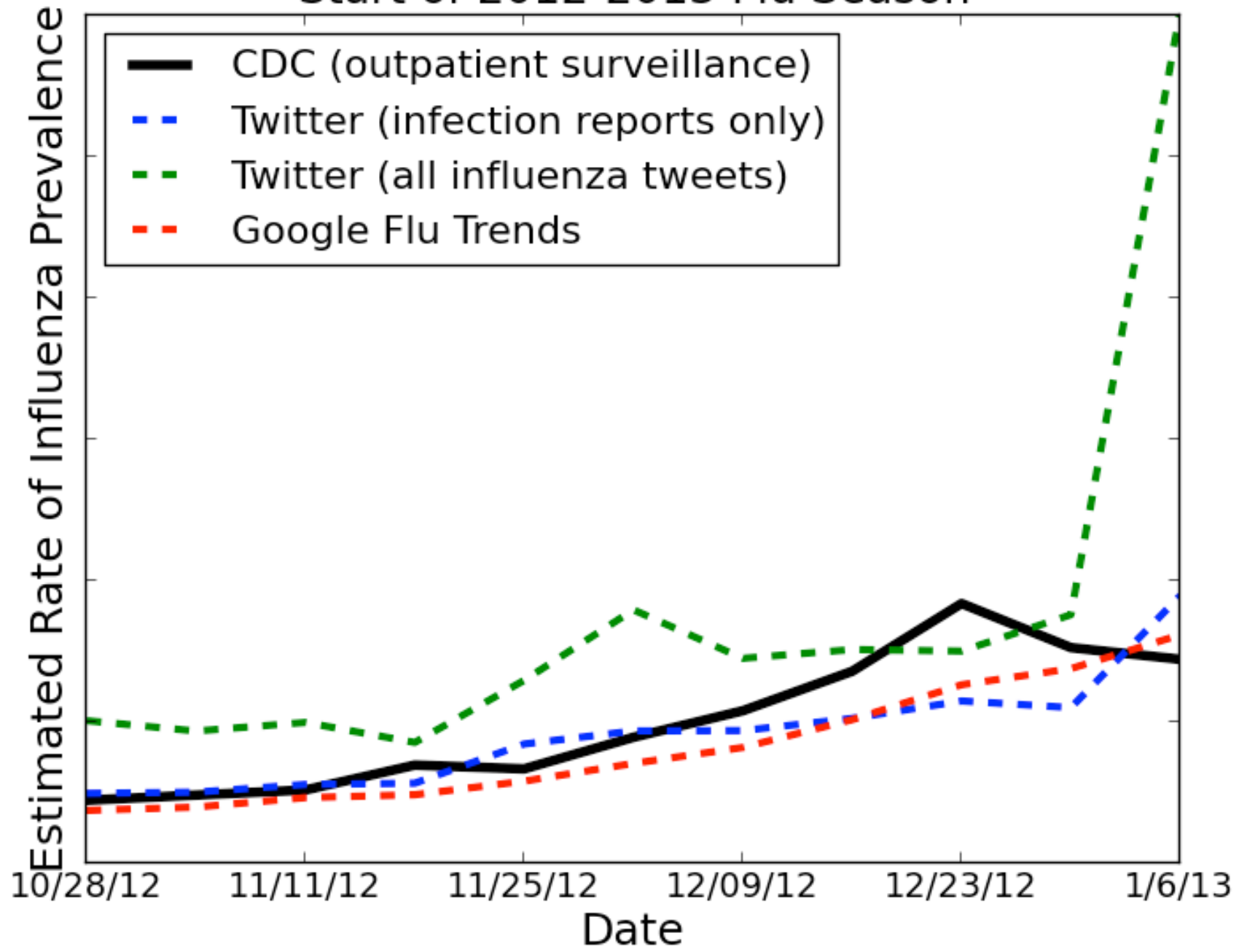
From my computer science research group

DISEASE TRACKING

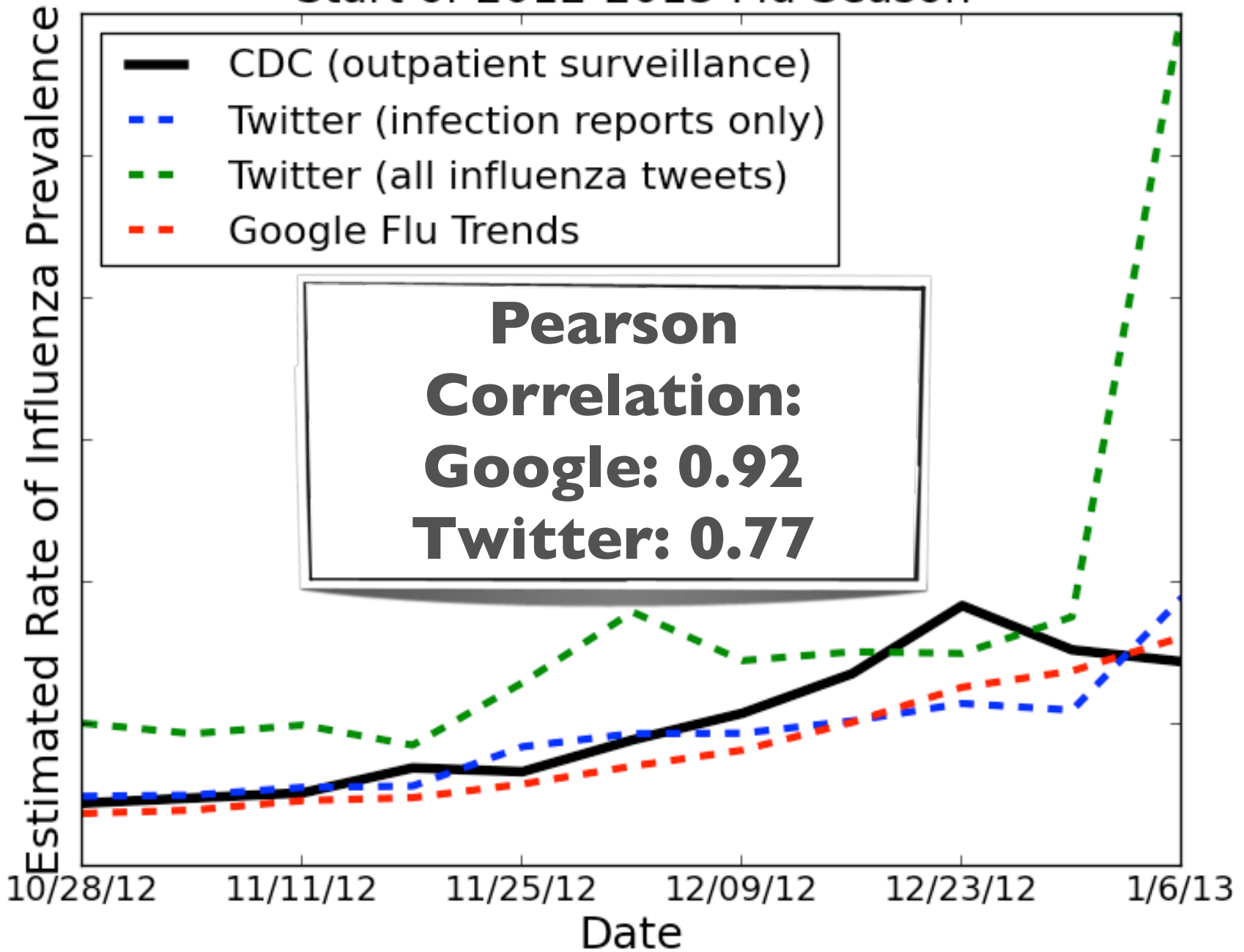


- Disease tracking is slow and resource limited
- Twitter: mine infection reports
- Users tweet about symptoms **4 days** before seeing a doctor; data available **two weeks** before traditional data

Start of 2012-2013 Flu Season



Start of 2012-2013 Flu Season



Predict direction of above average changes

National ILI %, 11 weeks from 2012-13

Predict direction of above average changes

National ILI %, 11 weeks from 2012-13

100% accuracy

TRACKING DRUG TRENDS



- Constantly emerging recreational drugs: more than 1 new drug **a week**
- Drug studies are costly and have slow turnaround time (9 months for annual survey)
- Mine drug user narratives from forums to answer questions
 - How do users smoke Salvia?
 - What are typical dosages for Salvia?
 - What populations use Salvia?

How are people **using** Salvia?

Salvia
Smoking
Usage

"Best way is to use a torch lighter, bong or pipe (bong recommended) and hold in each hit 20-40 seconds."

"A dose of Salvia leaves is 2 grams. A dose of Salvia 5X extract is 0.4 grams (400mg) A dose of salvia 10X extract is 0.2 grams (200mg) A dose of Salvia 25X extract is 0.08 grams (80mg) A dose of Salvia 50X extract is 0.04 grams (40mg) A dose of Salvia 250X extract is impossible."

Salvia
Oral
Usage



What are the **effects** of Salvia?



Salvia Smoking Effects

"He then took one large hit and held it in and laid back and began to feel his body getting heavy and his vision started to get this dim orange brownish light to it and he closed his eyes and moved his body around and it shook both of the feelings off."

Salvia Oral Effects







"When chewed, the first effects are felt after about 15 minutes. After about 30 minutes, the full effects should be realized. Typical Salvia experiences last 5 to 10 minutes on average, with noticeable after-effects lasting up to 1/2 hour."

CDC'S BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM

- Measure correlation between risk factor and Tweets about an ailment by US state

CDC'S BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM

- Measure correlation between risk factor and Tweets about an ailment by US state

CDC Risk Factor	Twitter topics	Correlation	
Asthma	Allergies	0.241	
Exercise	Obesity	-0.201	
Health Care Coverage	All Ailments	-0.253	
Heart Attack	Obesity	0.244	
Obesity	Exercise	-0.267	
Tobacco Use	Cancer	0.648	

MEDICATION USE



- Discover how people self-medicate common illnesses from self reports on Twitter
- Finding:
 - People use Benadryl to treat insomnia
 - Antibiotics incorrectly used to treat viral infections

PATIENT SAFETY



- Medical mistakes are a leading cause of death, possibly killing 250,000 people in the US a year
- Candid patient reports about safety events difficult to obtain
- Tweets about patient safety
 - One of my residents was admitted to die because a surgeon messed up on her back surgery and she's only 43.
 - REALLY?! my doctor screwed up my prescription... I've been taking the wrong dosage for almost a year. HMM. MAYBE THAT EXPLAINS SOMETHING.
 - lets recap. car wreck. hospital. blood clots. best friend died. other best friend had surgery that went wrong. 2nd degree sunburn

PATIENT SAFETY

- Most reports are by patient (80%) or relative (8%)
- Procedural (36%) and medication errors are the most common (22%)
- Anger most common reaction (22%), humor second (11%)



THANK YOU



mdredze@cs.jhu.edu

www.dredze.com

Contact me with your ideas!

Michael J. Paul, Mark Dredze. **You are what you Tweet: Analyzing Twitter for Public Health.** *International Conference on Weblogs and Social Media (ICWSM)*, 2011.

Ralph Passarella, Atul Nakhasi, Sarah Bell, Michael J. Paul, Peter Pronovost, Mark Dredze. **Twitter as a Source for Learning about Patient Safety Events.** *Annual Symposium of the American Medical Informatics Association (AMIA)*, 2012

Alex Lamb, Michael J. Paul, Mark Dredze. **Separating Fact from Fear: Tracking Flu Infections on Twitter.** North American Chapter of the Association for Computational Linguistics (NAACL), 2013.

Michael J. Paul, Mark Dredze. **Drug Extraction from the Web: Summarizing Drug Experiences with Multi-Dimensional Topic Models.** North American Chapter of the Association for Computational Linguistics (NAACL), 2013.