Accounting for the global under-reporting of Emerging Infections Diseases

Ilana Brito

Massachusetts Institute of Technology

CIESIN, Columbia University

Advances in Geospatial Technologies for Health Conference

Santa Fe, September 13, 2011

Motivation

- Number of EIDs is increasing
- Local diseases can spread globally
- Significant morbidity and mortality
- Devastating effects on economies
- **Need to focus surveillance efforts**

But how?

- Step 1:
    - Where is the highest risk?

Emerging Infectious Disease Dataset

- 335 diseases that have emerged between 1940 and 2003
- Jones et al., Nature 2008

Criteria for inclusion:

- Novel disease in human population
- New geographic locales
- Development of drug-resistance
- Dramatic changes in virulence

What causes new diseases to emerge?

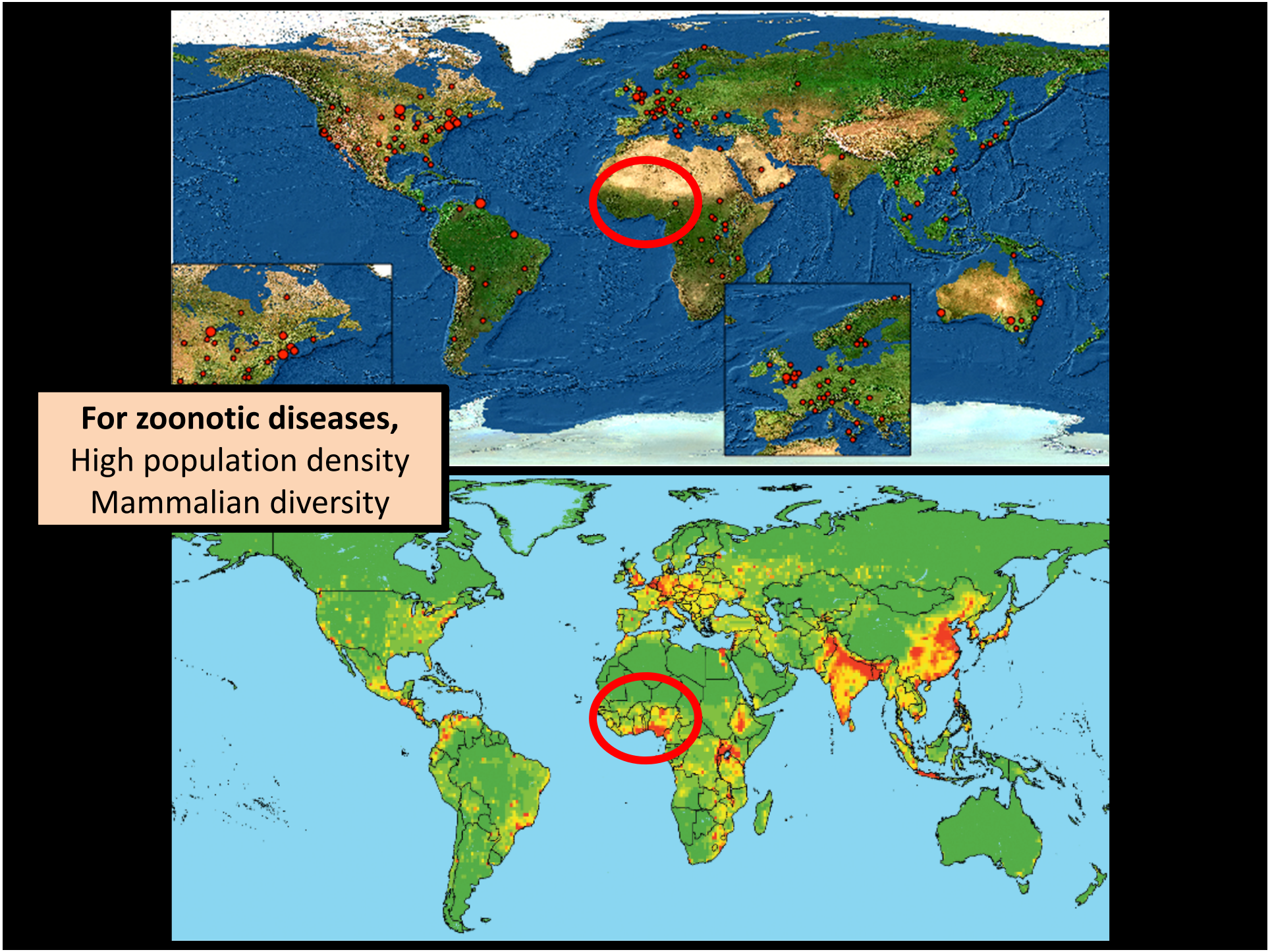
Demographic changes
Age shifts
Population density

Ecological changes
-Land use changes
- Climate change
-Agriculture
Intensification

Bioterrorism

Changes in drug use
-Personal
-Agriculture
- Factory runoff

Changes in immunity
- HIV status


For zoonotic diseases,
High population density
Mammalian diversity

- Step 1:
    - Where is the highest risk?
- Step 2:
    - Where is surveillance/reporting poor?

Exposure

World Health Organization

CDC
Centers for Disease Control and Prevention

ProMED mail

Novel EID Reported

Exposure

↓

Symptom presentation

↓

Medical visit and examination

↓

Correct diagnosis

↓

Alert authorities

↓

Novel EID Reported

- Cluster of cases
- Access to medical care
- Personal choices
- Medical technologies
- Uniqueness of symptoms
- Medical training
- Communications technology

etc.

Exposure

↓

Symptom presentation

↓

Medical visit and examination

↓

Correct diagnosis

↓

Alert authorities

↓

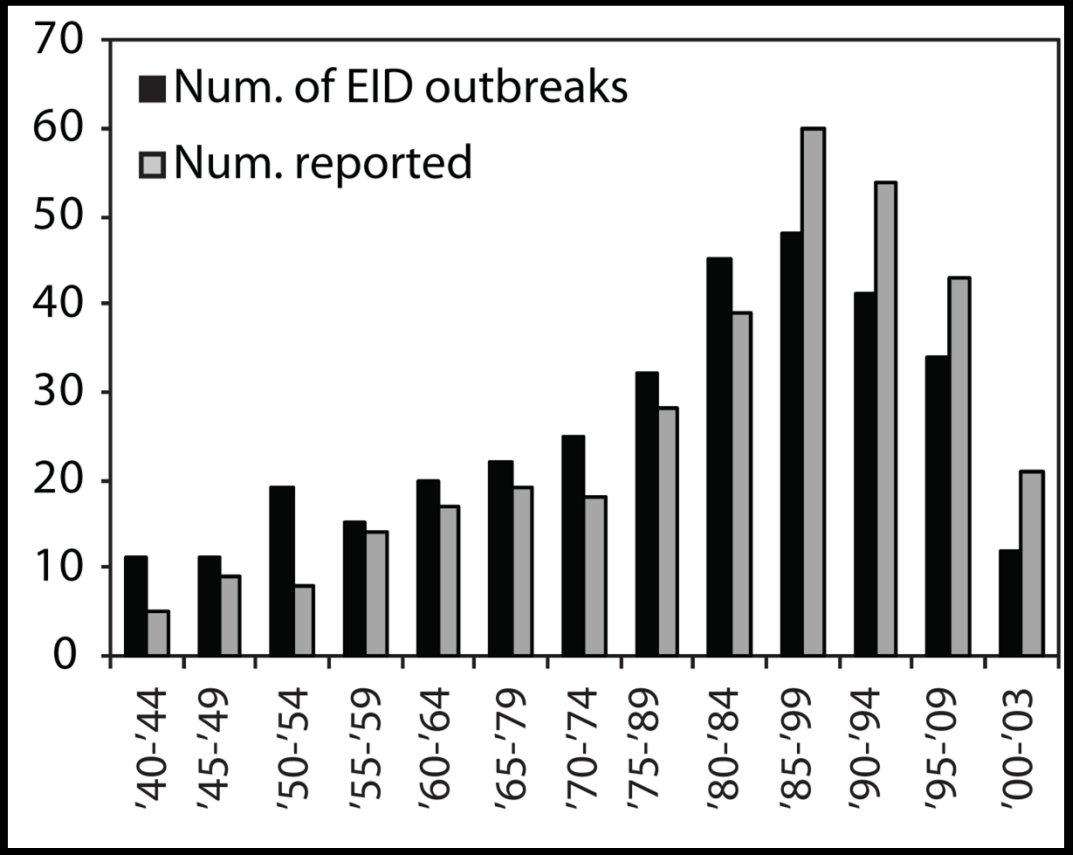
Novel EID Reported

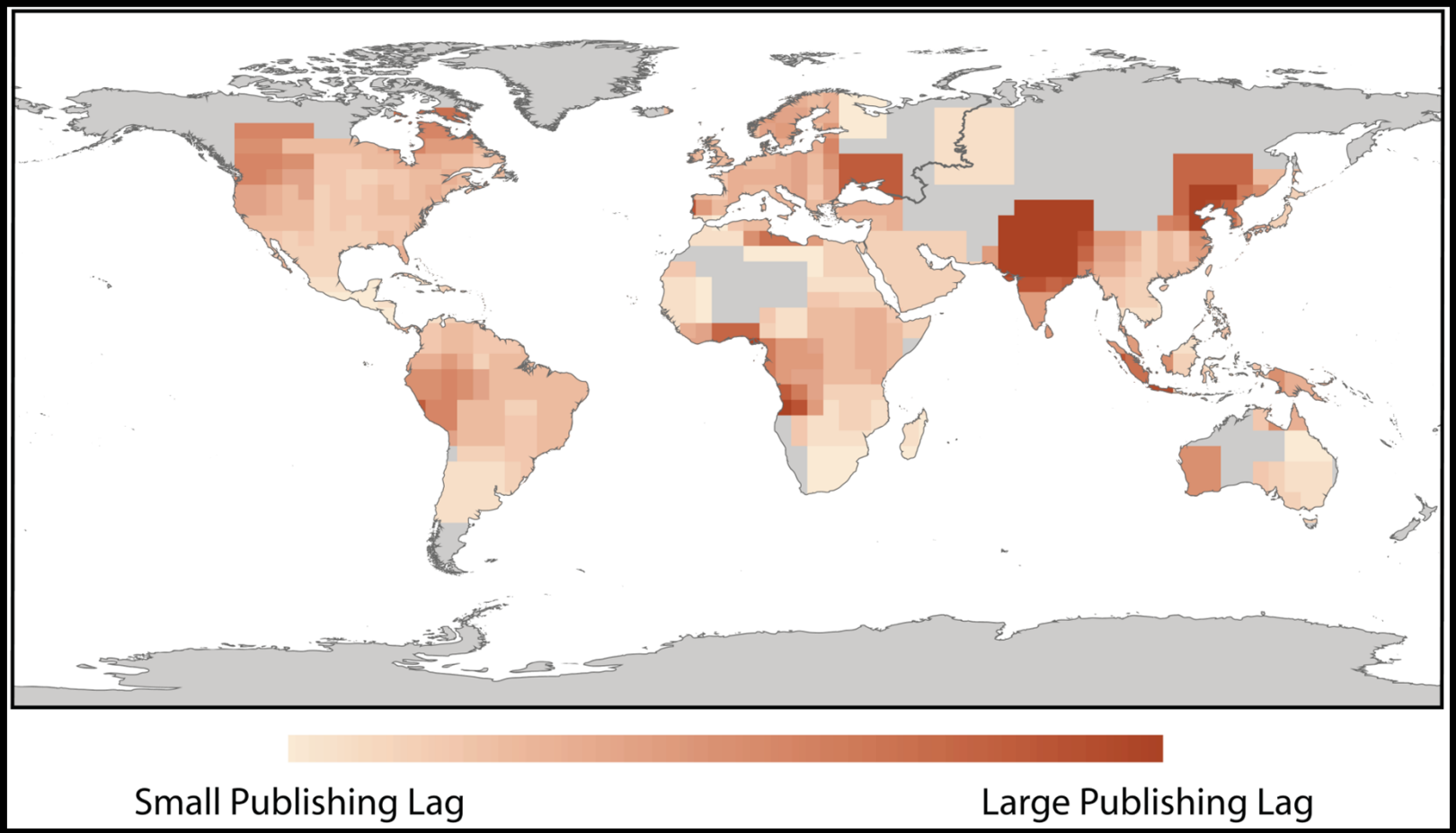
Lag in Reporting

CDC Centers for Disease Control and Prevention

World Health Organization

Lag in disease reporting




Small Publishing Lag
Large Publishing Lag

Components of Reporting Lag

Infrastructure

Healthcare spending

Number of doctors, hospitals

Communications

Technology

Biological

Disease etiology

Disease symptoms

Pathogenicity

Disease Traits

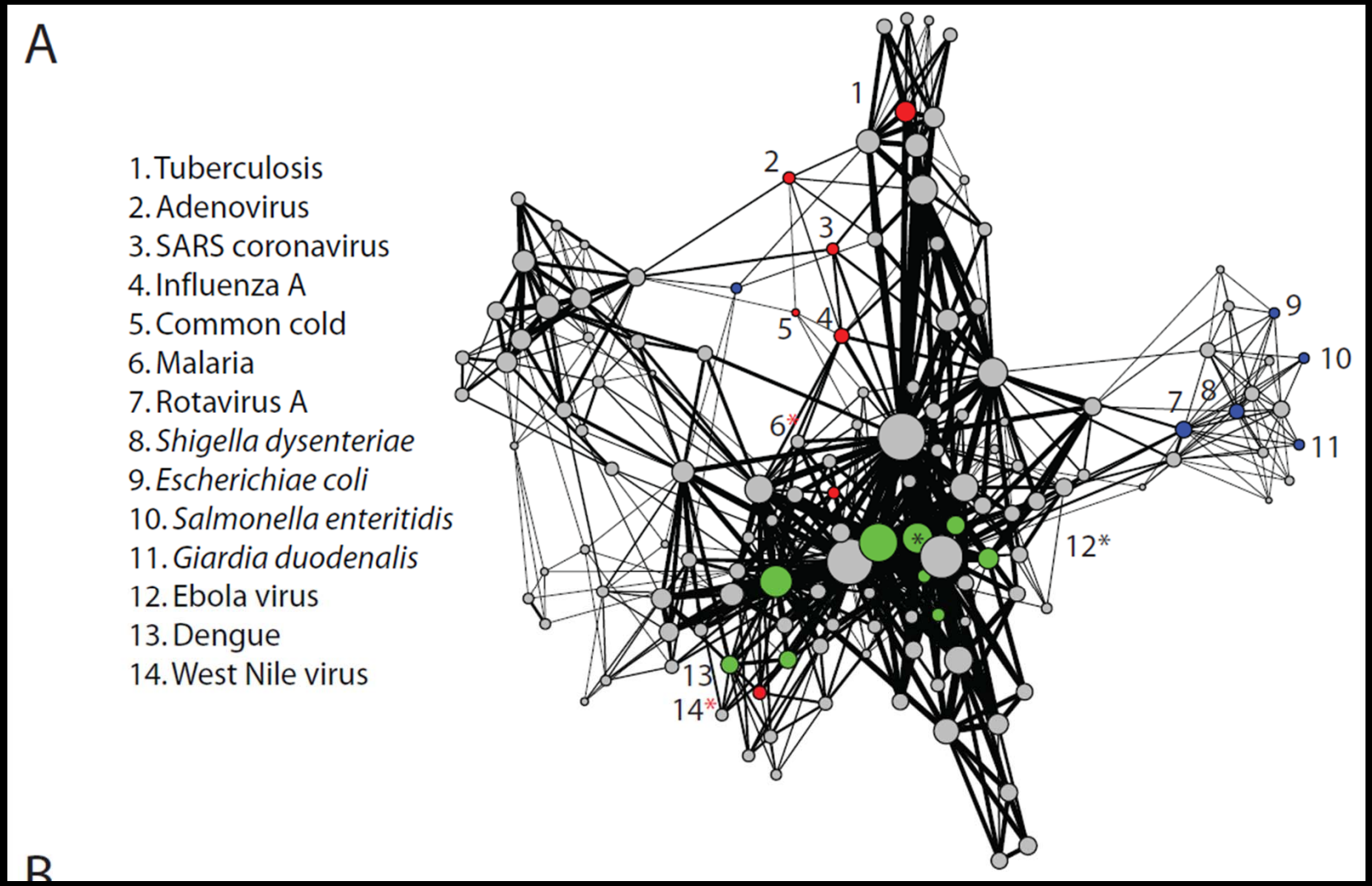
More Likely to Report

- High transmission rate → Large clusters
- High disease severity
- Symptom abnormality

Less Likely to Report

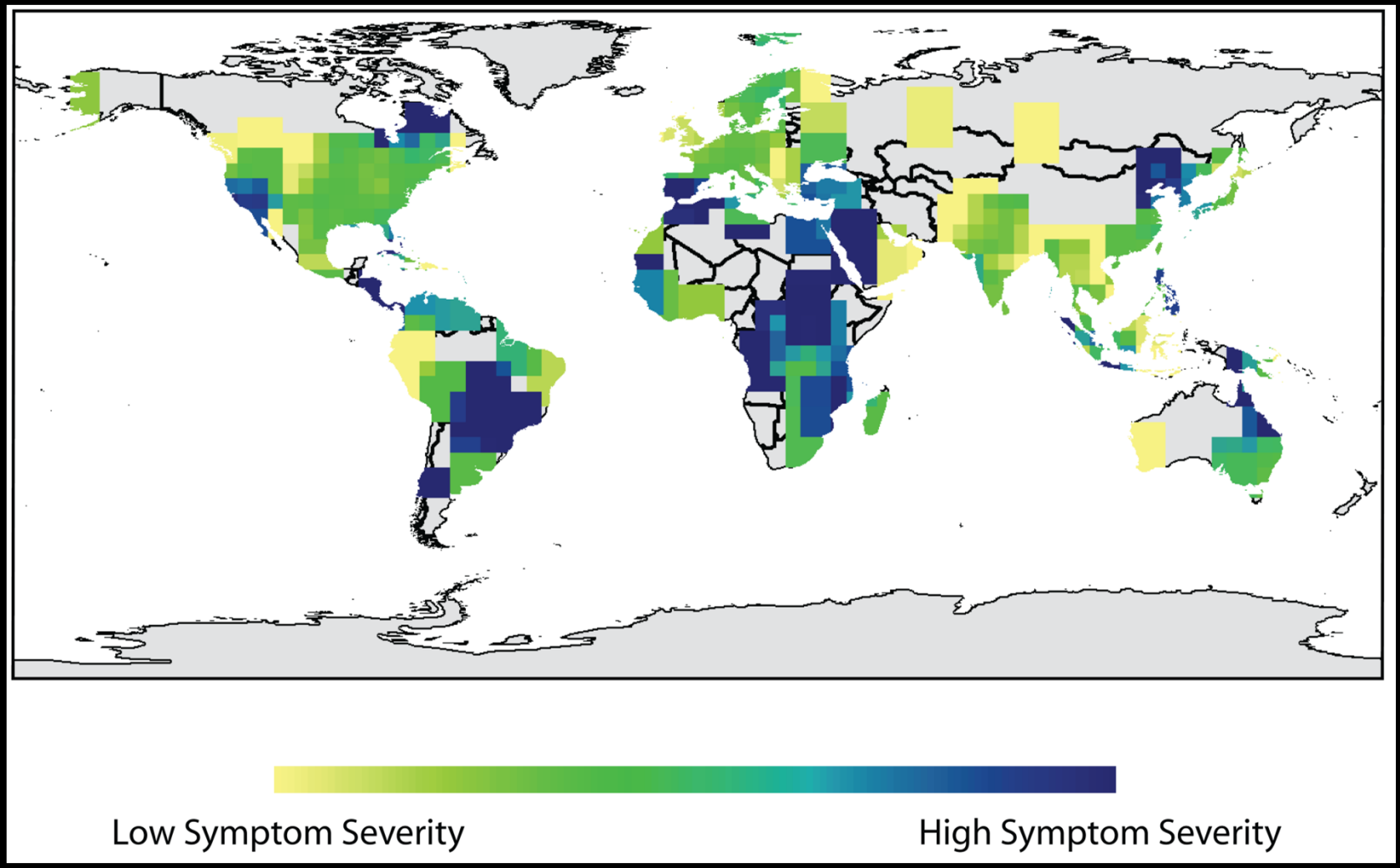
- Fast recovery times
- Overlapping symptoms with commonplace diseases

Overlapping symptoms

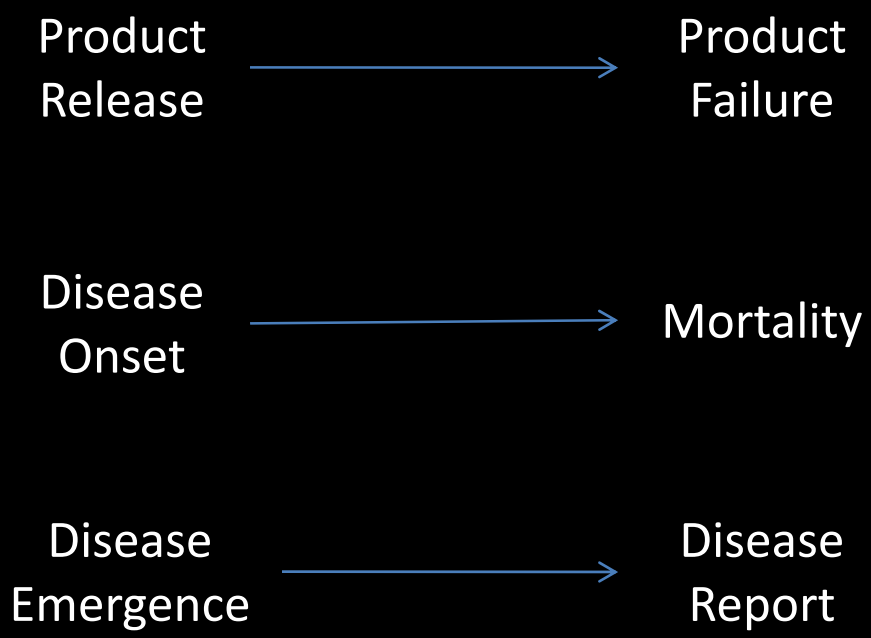


Scoring Symptoms

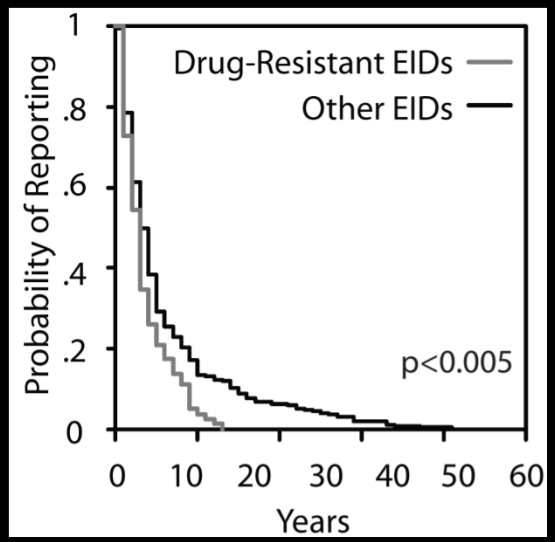
Symptom	Score
Headache, Cough, Abdominal pain, Sneezing, Back pain	1
Fever, Diarrhea, Localized rash, Muscle swelling, Pustules	2
Cardiac or pulmonary pain, Genital lesions, Abdominal mass, Vomitting, Macules, Blood in urine	3
Pneumonia, Gastrointestinal obstruction, Eye worm, Hemorrhagic rash, Seizures, Meningitis	4
Death, Hemmohagic fever, Paralysis, Coma	5

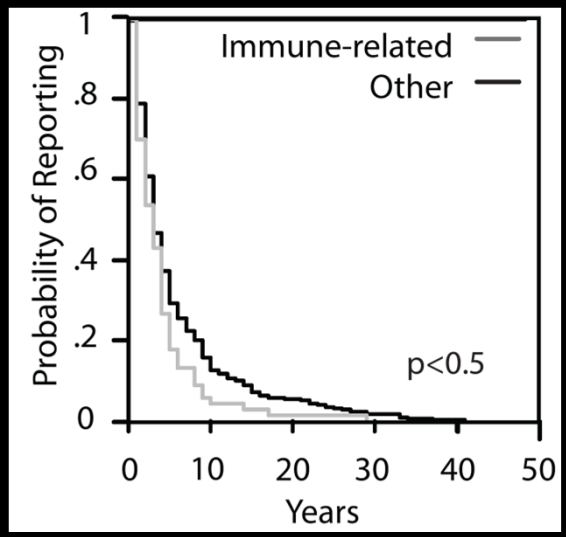

Low Symptom Severity
High Symptom Severity

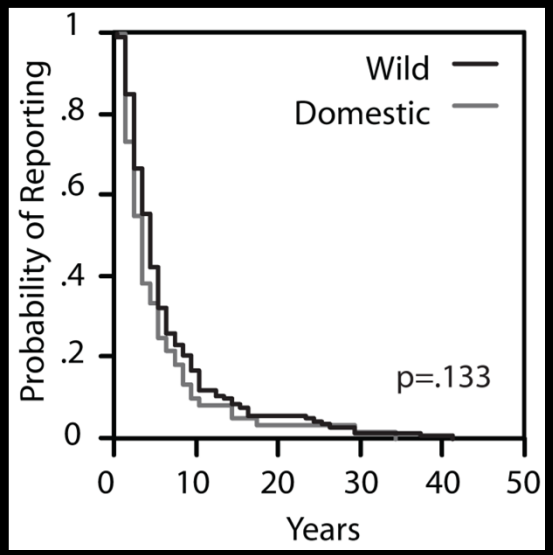
Applying Survival Analysis to Reporting Lag

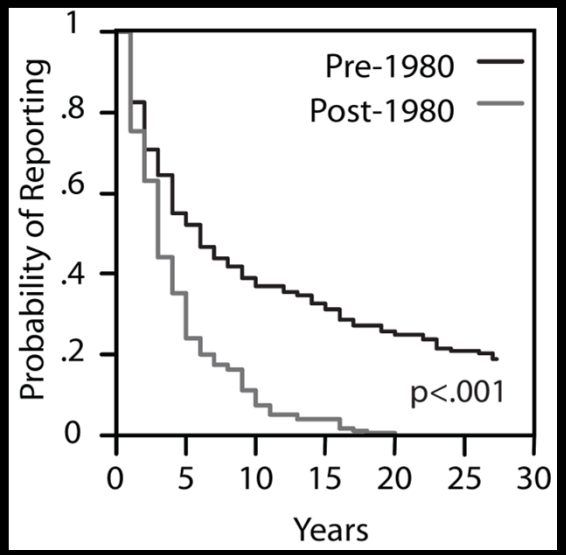


$$h_i(t) = h_0(t)e^{\{\beta_1 X_{i1} + \ldots + \beta_k X_{ik}\}}$$


Probability of Reporting
1
.8
.6
.4
.2
0
Drug-Resistant EIDs
Other EIDs
p<0.005
0 10 20 30 40 50 60
Years


Probability of Reporting
1
.8
.6
.4
.2
0
Immune-related
Other
p<0.5
0 10 20 30 40 50
Years


Probability of Reporting
1
.8
.6
.4
.2
0
Wild
Domestic
p=.133
0 10 20 30 40 50
Years


Probability of Reporting
1
.8
.6
.4
.2
0
Pre-1980
Post-1980
p<.001
0 5 10 15 20 25 30
Years

What factors contribute the lag?

- Type of disease (zoonotic, vector-borne, drug-resistance)
- Pathogen type (virus, bacteria etc.)
- Disease severity
- Population Density → Cluster size
- Disease Burden
- Hospital beds
- Health expenditure (as % of GDP)
- Political rights and Civil liberties
- Number of Physicians (per 1000 people)

Health Infrastructure

- Immunizations
- Civilian coverage (birth and death certificates)
- Internet Users (per 100 people)
- Telephone Lines (per 100 people)
- Cell phone Subscriptions (per 100 people)

Communication Infrastructure

Model to account for reporting lag

Zoonotic disease (n=158)

Variable	β	s.e.(β)	p-value
Disease burden (all causes/10,000 people)	-1.9×10^{-7}	9.3×10^{-8}	*
Disease severity score	.0054	.002	**

Including drug resistant diseases (n=204)

Variable	β	s.e.(β)	p-value
Disease burden (all causes/10,000 people)	-9.6×10^{-8}	8.0×10^{-8}	
Disease severity score	.0018	.0019	

Healthcare (% GDP) also significant

Conclusions

- Spatial pattern to the lag in disease reporting
- Use the lag to understand which diseases are reported in a more timely fashion.
- Lag is associated with disease types, severity of symptoms, and disease burden.
- Overlapping symptoms make novel disease detection more difficult without additional technology.

2 Final Thoughts:

- Knowing where diseases may emerge and where the gaps in disease reporting are suggests where surveillance efforts should be bolstered.

- Diseases may be observed (now or at any later time) or never observed at all. What is the true burden of EIDs?

Thank you!

CIESIN

Marc Levy
Susana Adamo

EcoHealth Alliance

Tiffany Bogich
Carlos Zambrana
Peter Daszak

The Earth Institute
AT COLUMBIA UNIVERSITY

Ruth DeFries

ZSL
LIVING CONSERVATION

Kate Jones
Sebastian Funk

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

USAID
FROM THE AMERICAN PEOPLE

PREDICT

National Science Foundation
WHERE DISCOVERIES BEGIN