# Data, models and science

Jonathan Dushoff, McMaster University

DAIDD 2020

# Outline

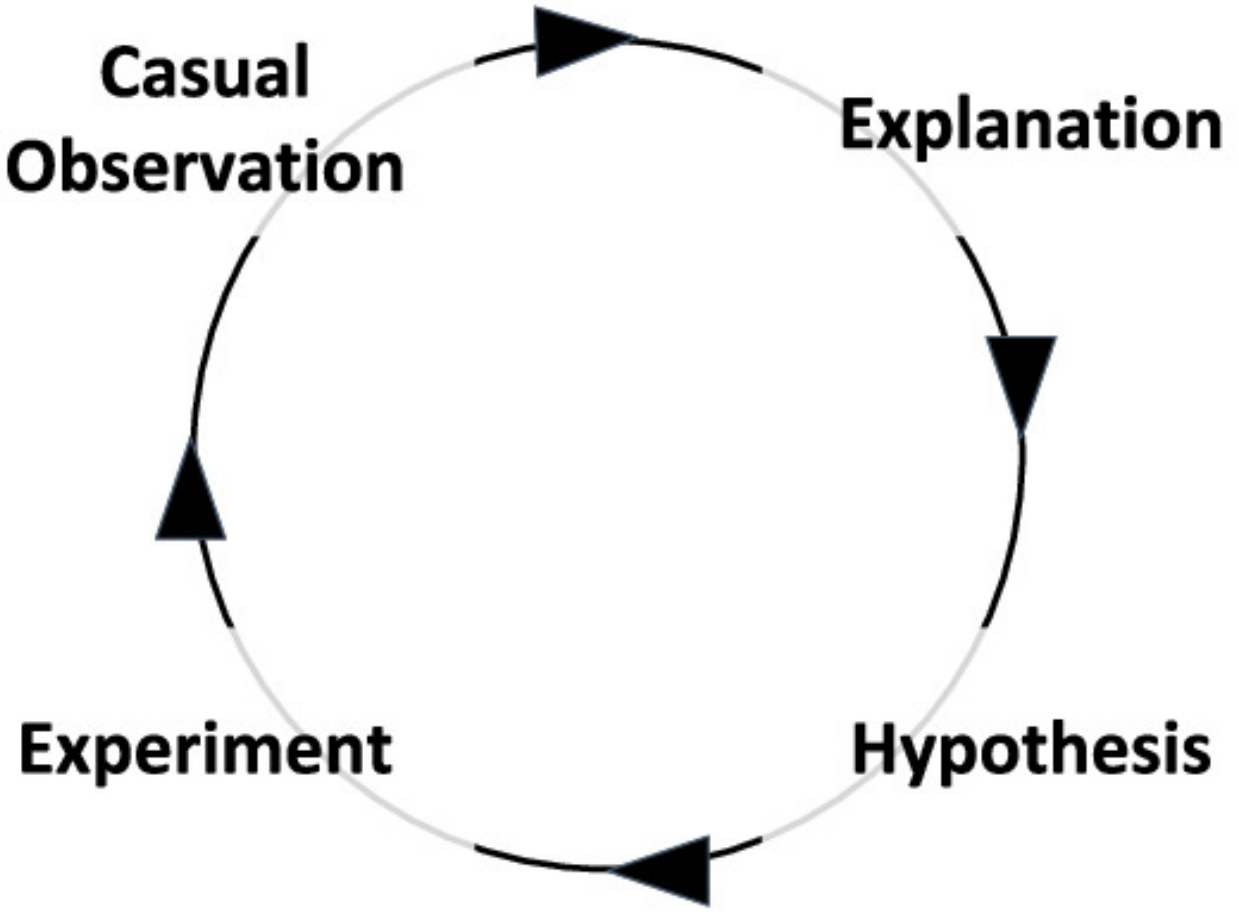
# Goals

- Process of science
- How science informs public health
  - Specific examples
- Approaches to epidemiology

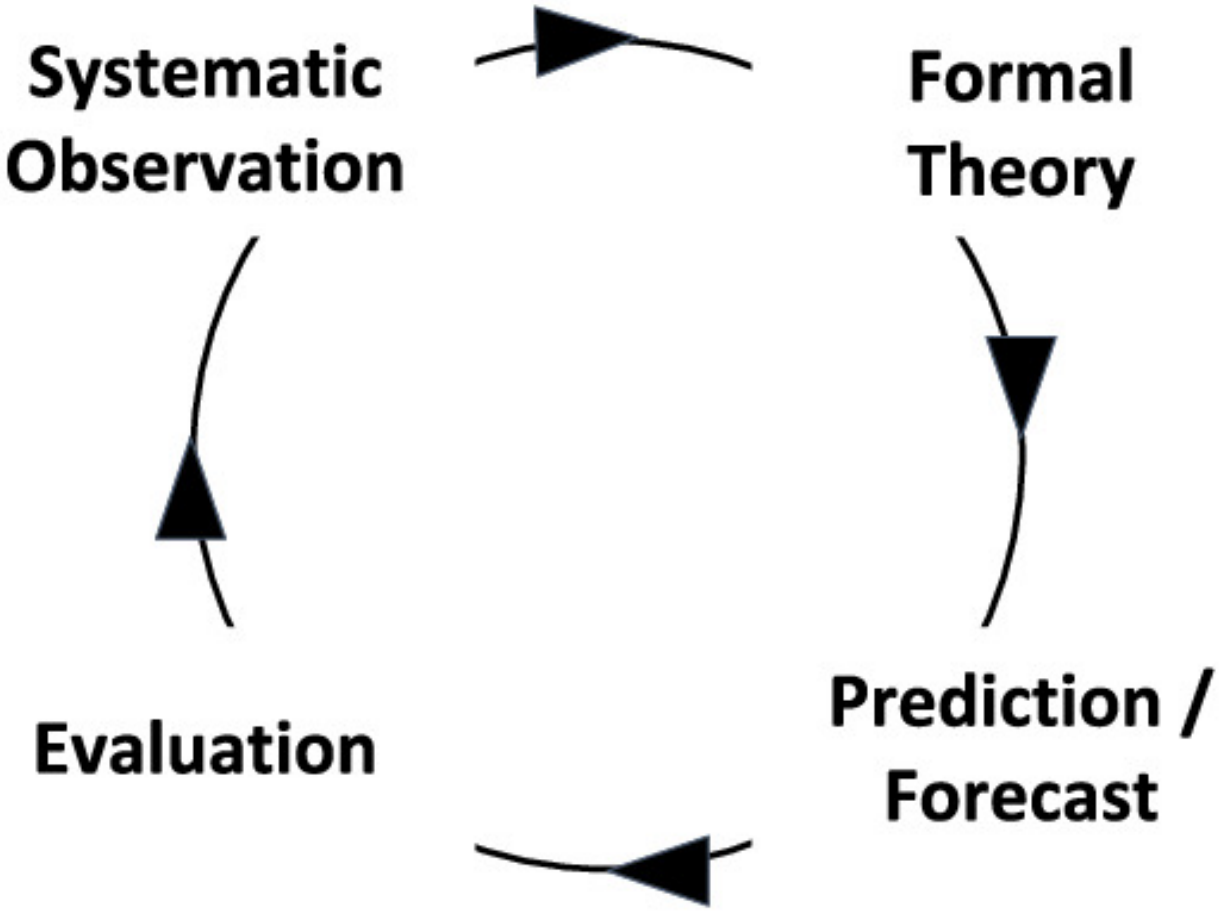
# Science is a *process*

- Observe and experiment with reality to *discover* and *challenge* ideas about how it works
- A key to science is that everything is open to question
  - Science is the belief in the ignorance of experts – *Feynman*

# The process of science



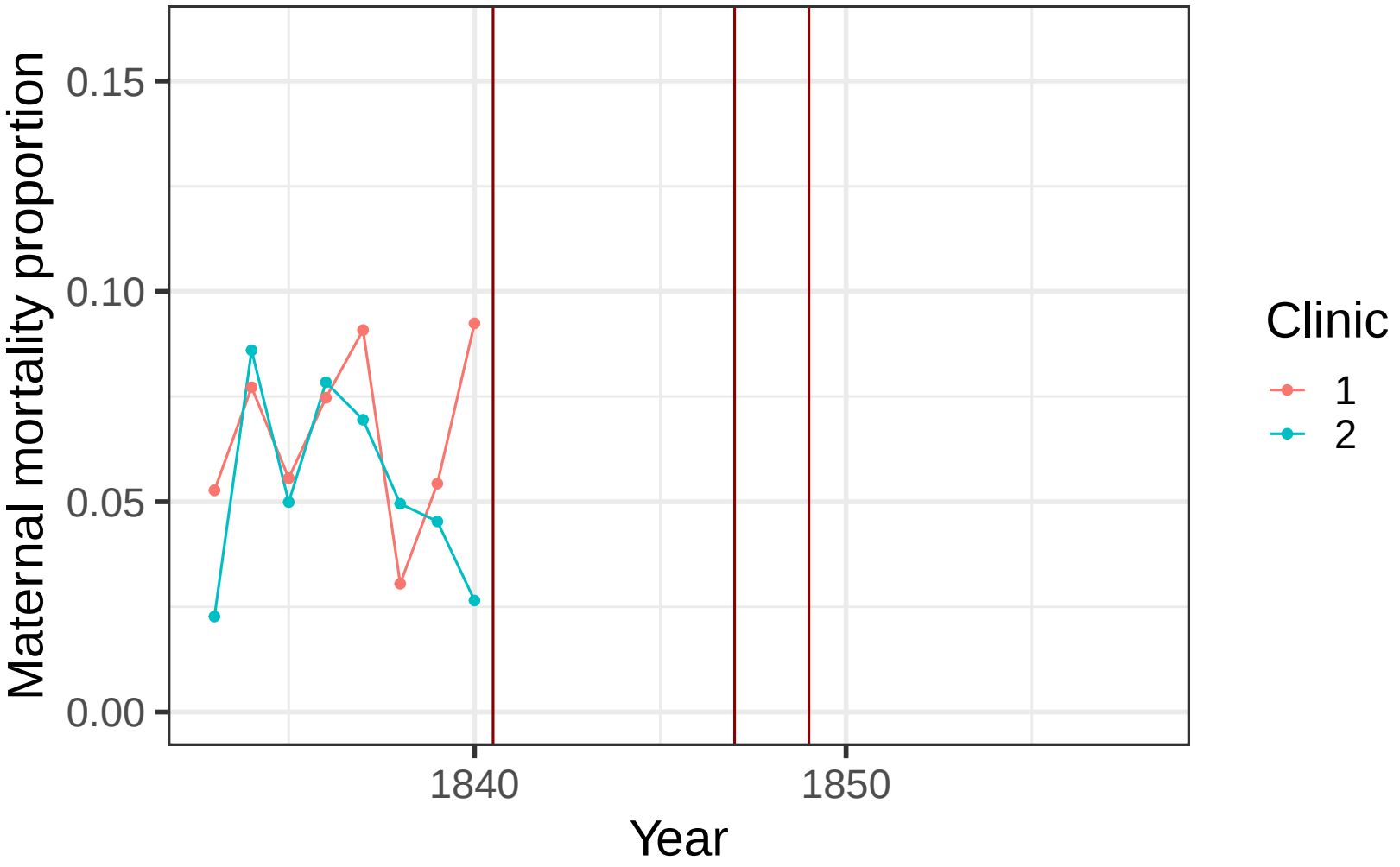
# Science without experiments



# Outline

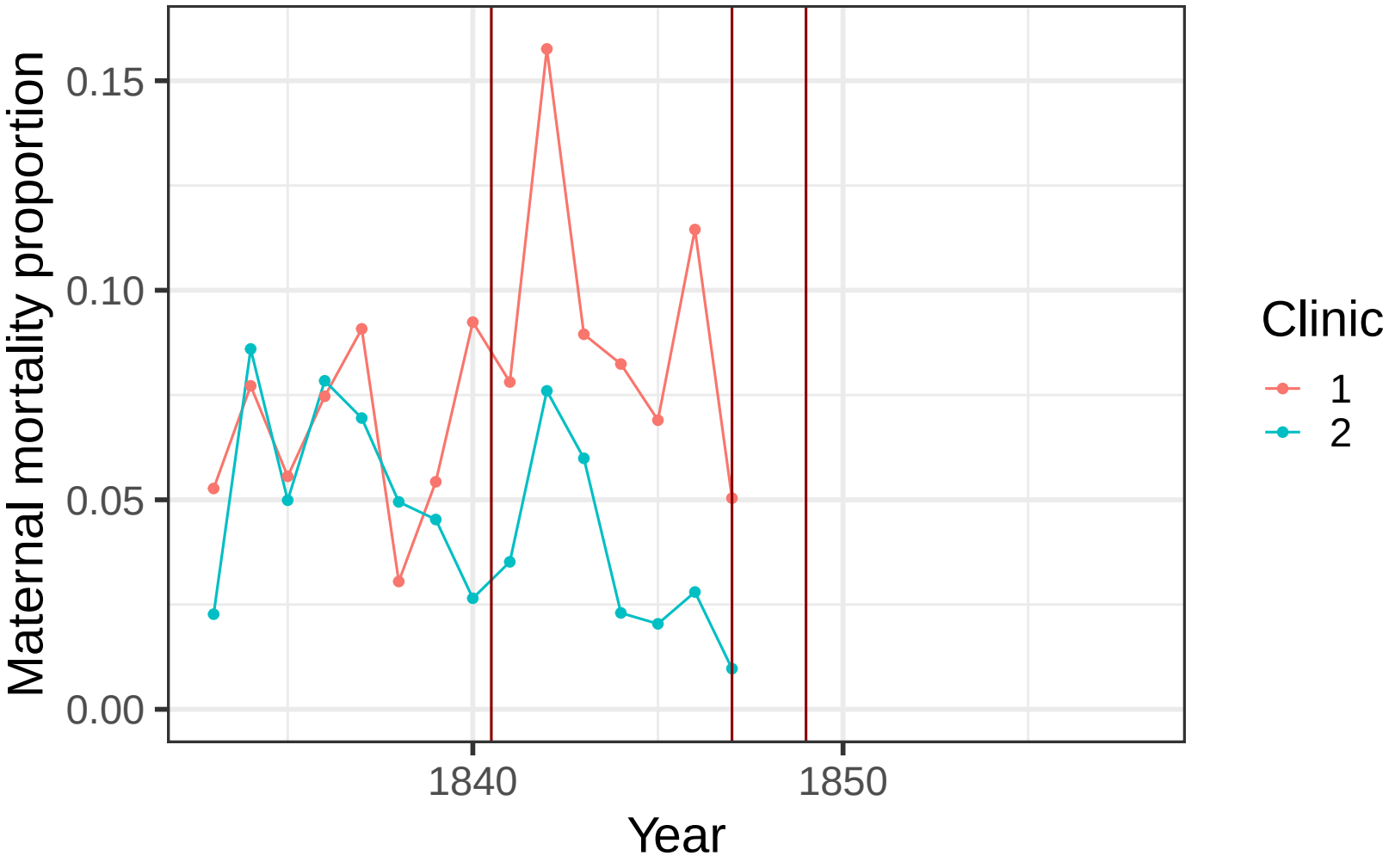
# Outline

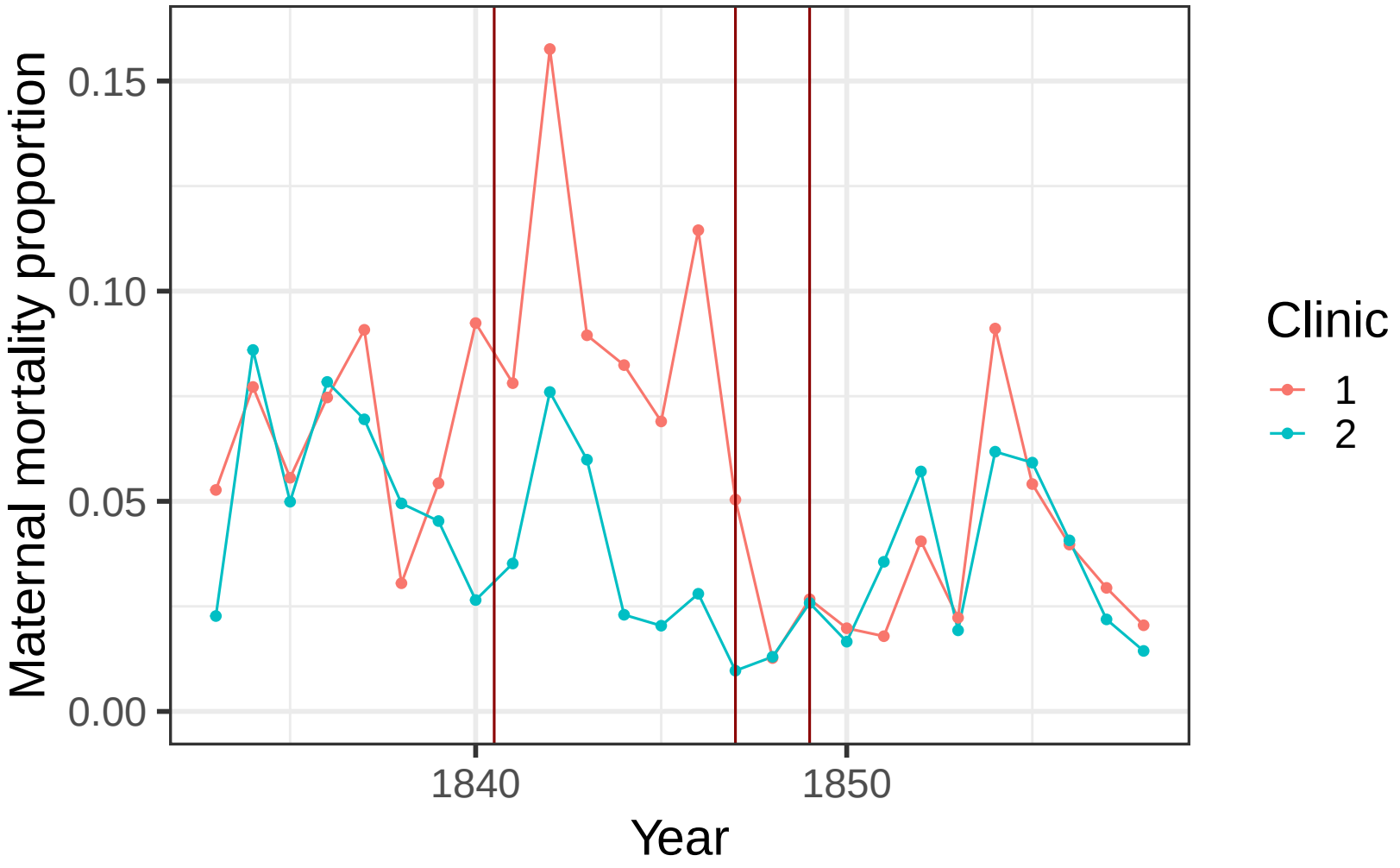
# Maternal mortality

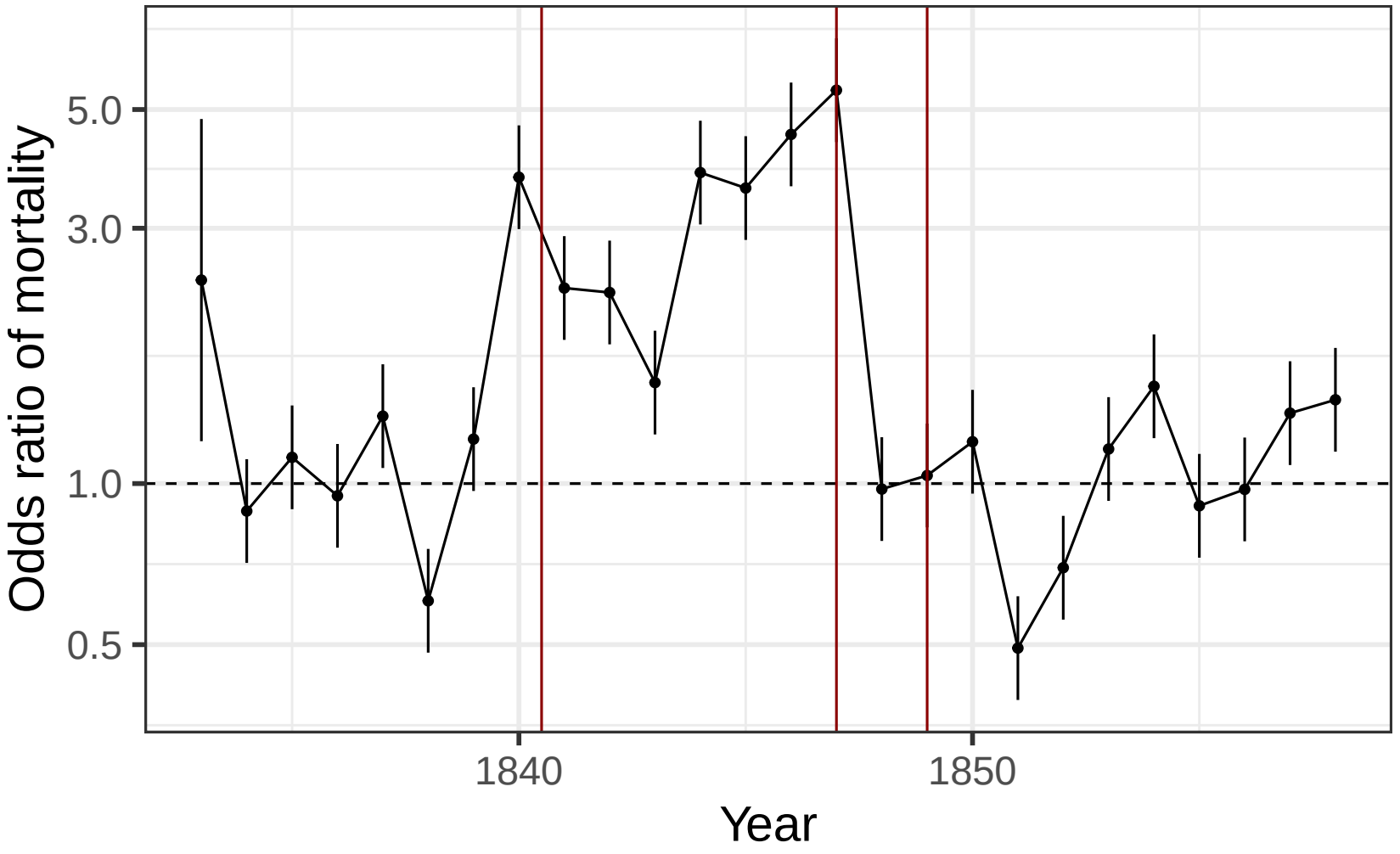

Maternal mortality proportion
0.15
0.10
0.05
0.00
1840
1850
Year
Clinic
1
2

# Observation and action

- In 1840, medical students stopped visiting Clinic 1
- In 1847, a surgeon died from infection following a scalpel injury
  - Igor Semmelweiss made medical students wash their hands


Maternal mortality proportion
0.15
0.10
0.05
0.00
1840
1850
Year
Clinic
1
2


Maternal mortality proportion
0.15
0.10
0.05
0.00
1840
1850
Year
Clinic
1
2

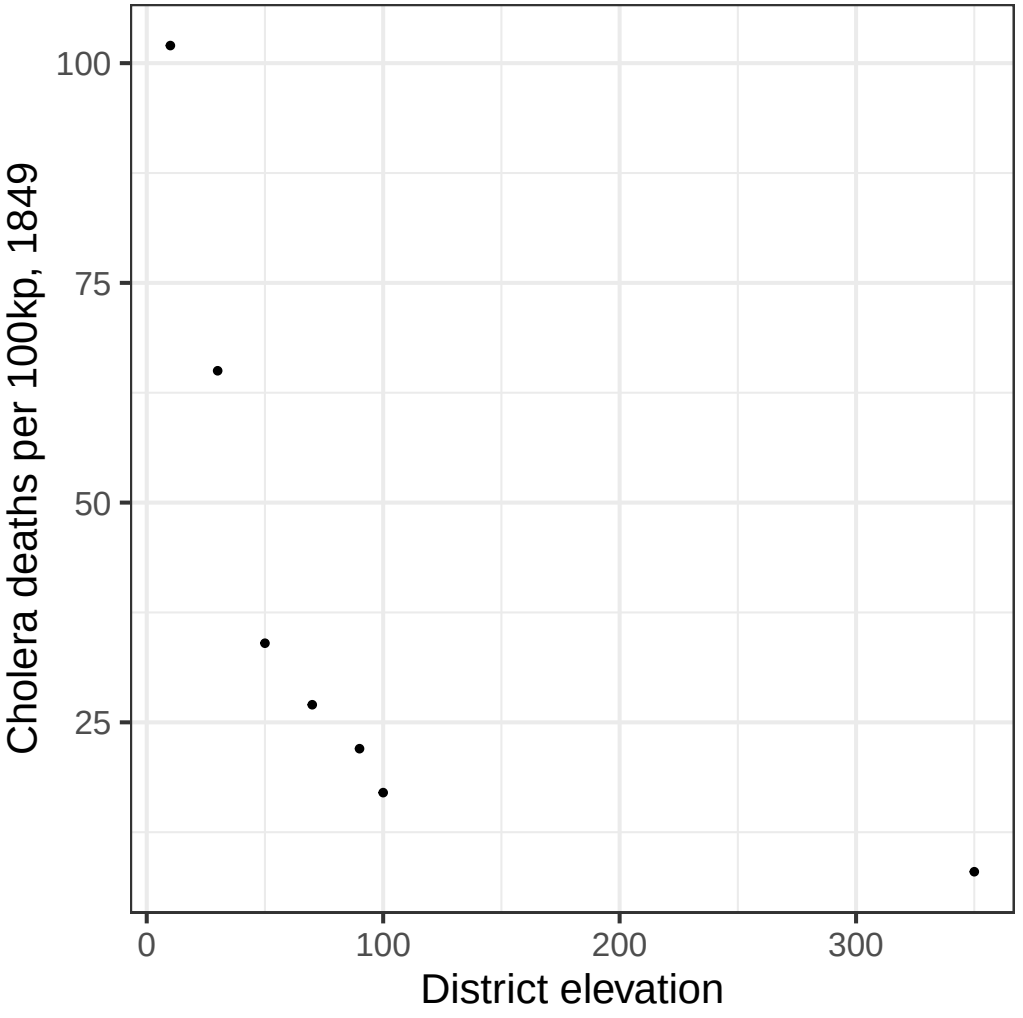

Odds ratio of mortality
5.0
3.0
1.0
0.5
1840
1850
Year

# Outline

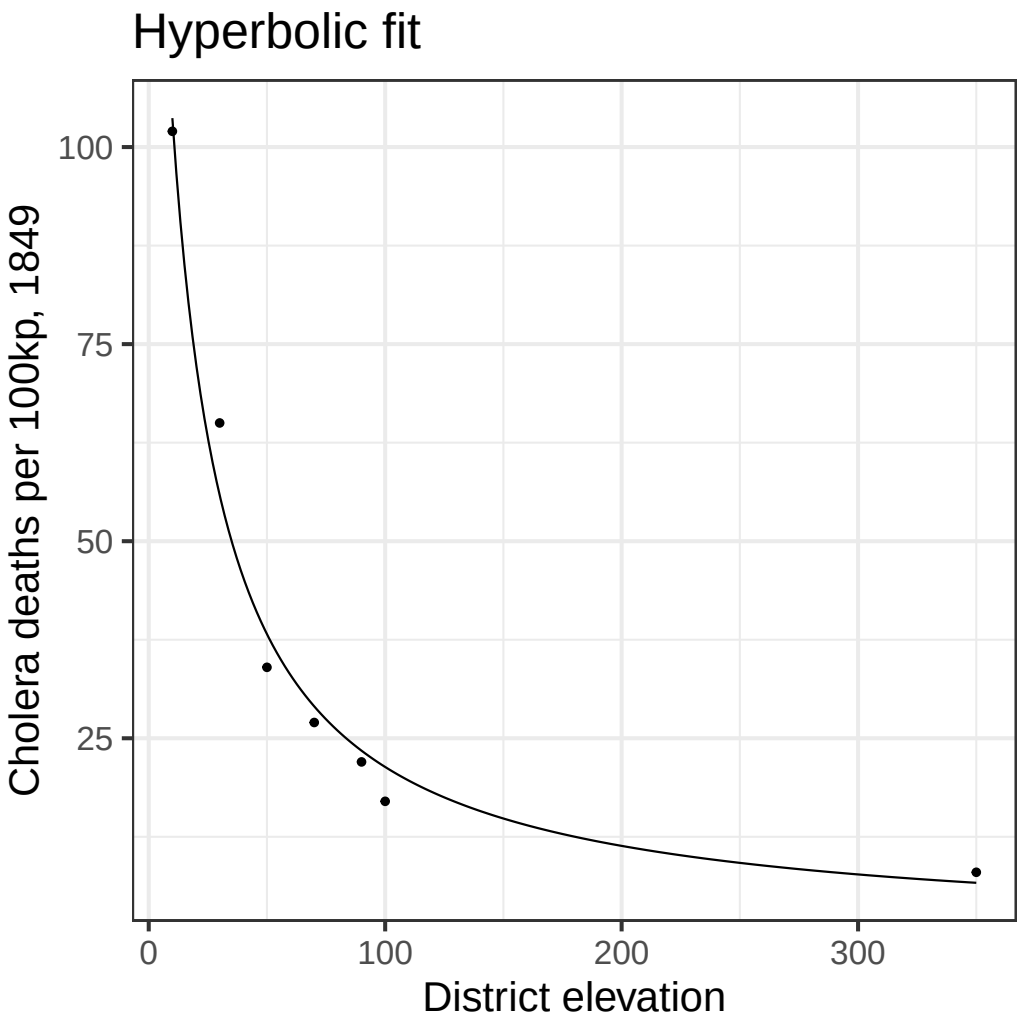
# Cholera

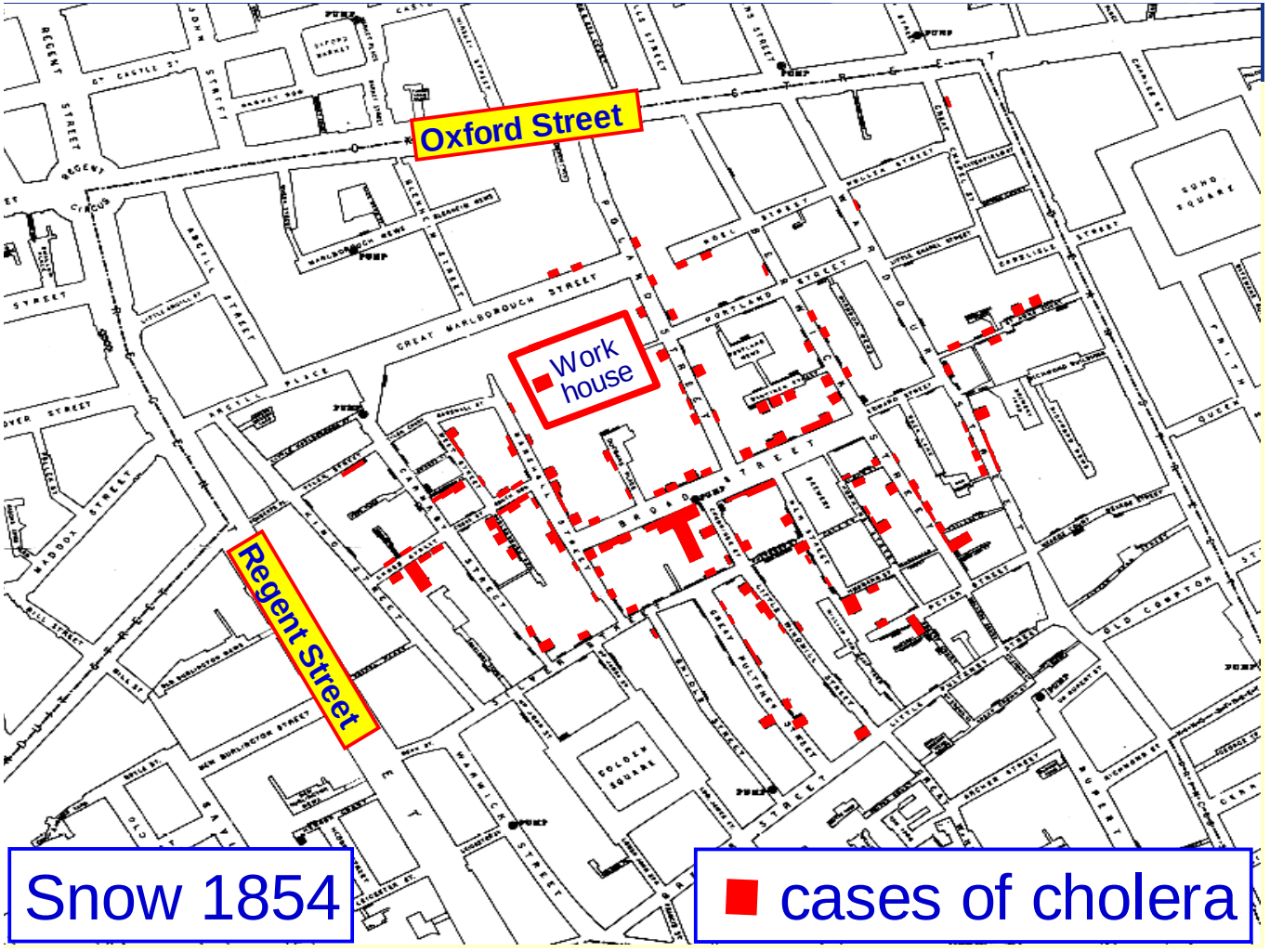
- Is it caused by bad air, or bad water?
- What's bad about it?

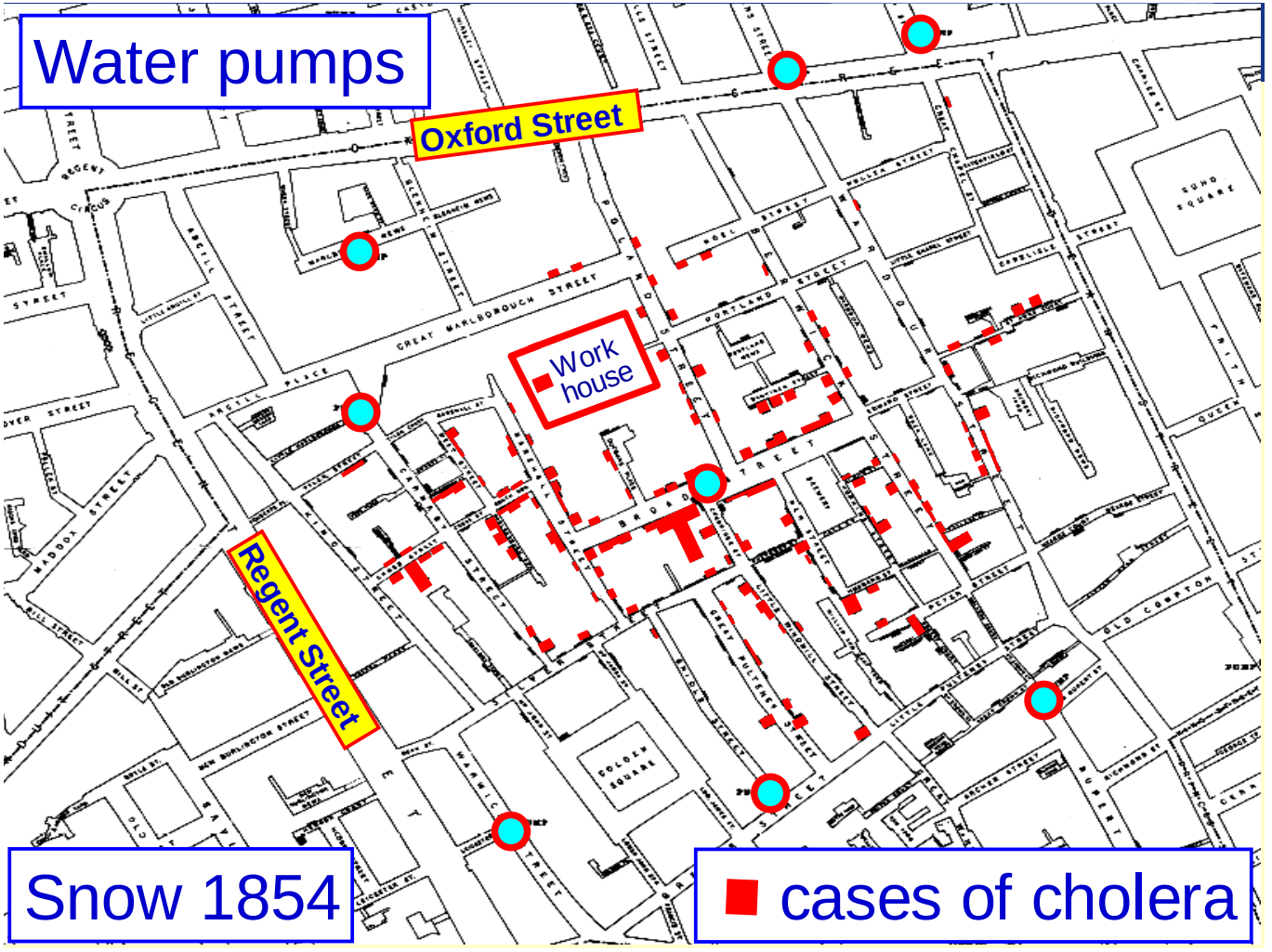
# *Cholera and air (present)*



# Cholera and air




Oxford Street
Work house
Regent Street
Snow 1854
cases of cholera


Water pumps
Oxford Street
Work house
Regent Street
Snow 1854
cases of cholera

# Outline

# Yellow fever and malaria

- Ross determined the cause of malaria primarily by experiments on mosquitoes
- Reed determined the cause of yellow fever primarily by experiments on human volunteers

# Outline

# Data, models and science

- We're never finished, we compare models to data over and over again
- Data is what we use to develop and understand models
- Models are what we use to interpret data
  - and they can suggest what data we need to collect
- Complicated or hard-to-test theories may require *dynamical* models

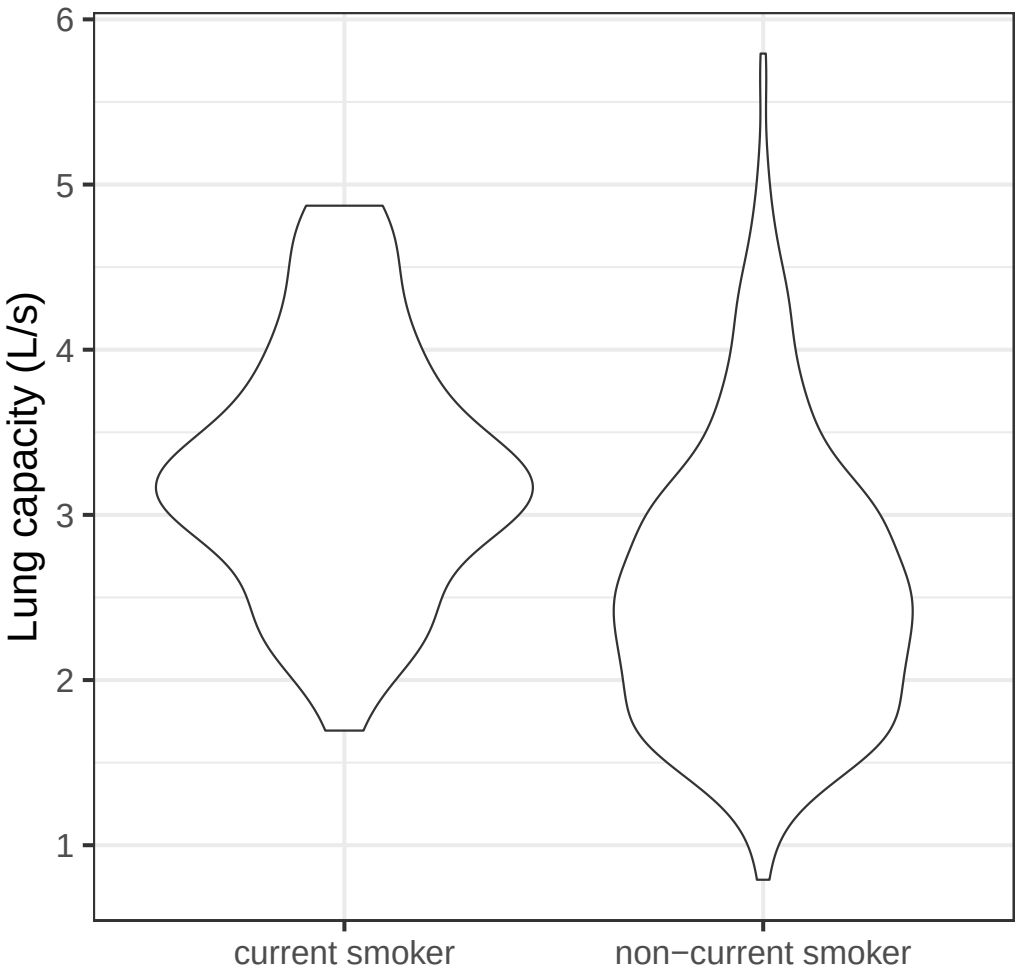
## Classical epidemiology

- Avoid mechanism
- Control for non-independence of "units"

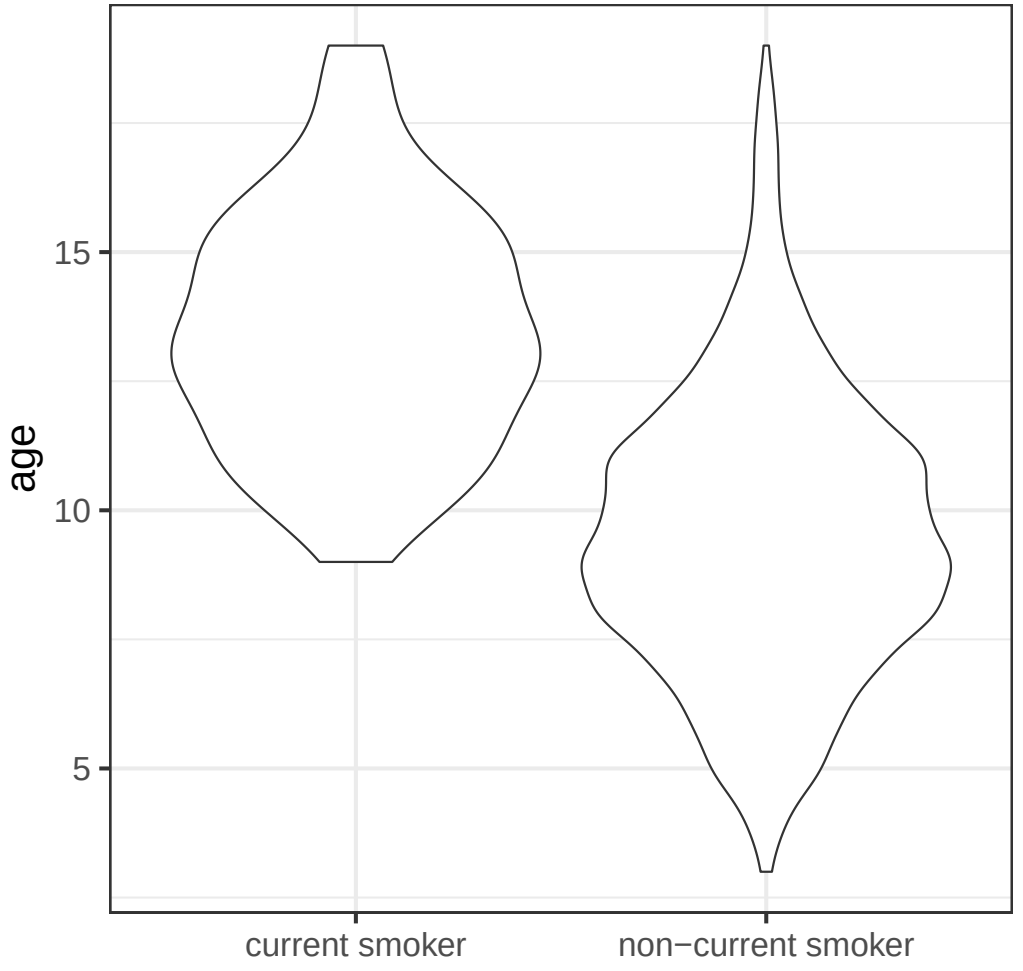
## Dynamical epidemiology

- Embrace mechanism
- Explicitly incorporate dependence between units
  - X is infected because Y infected them

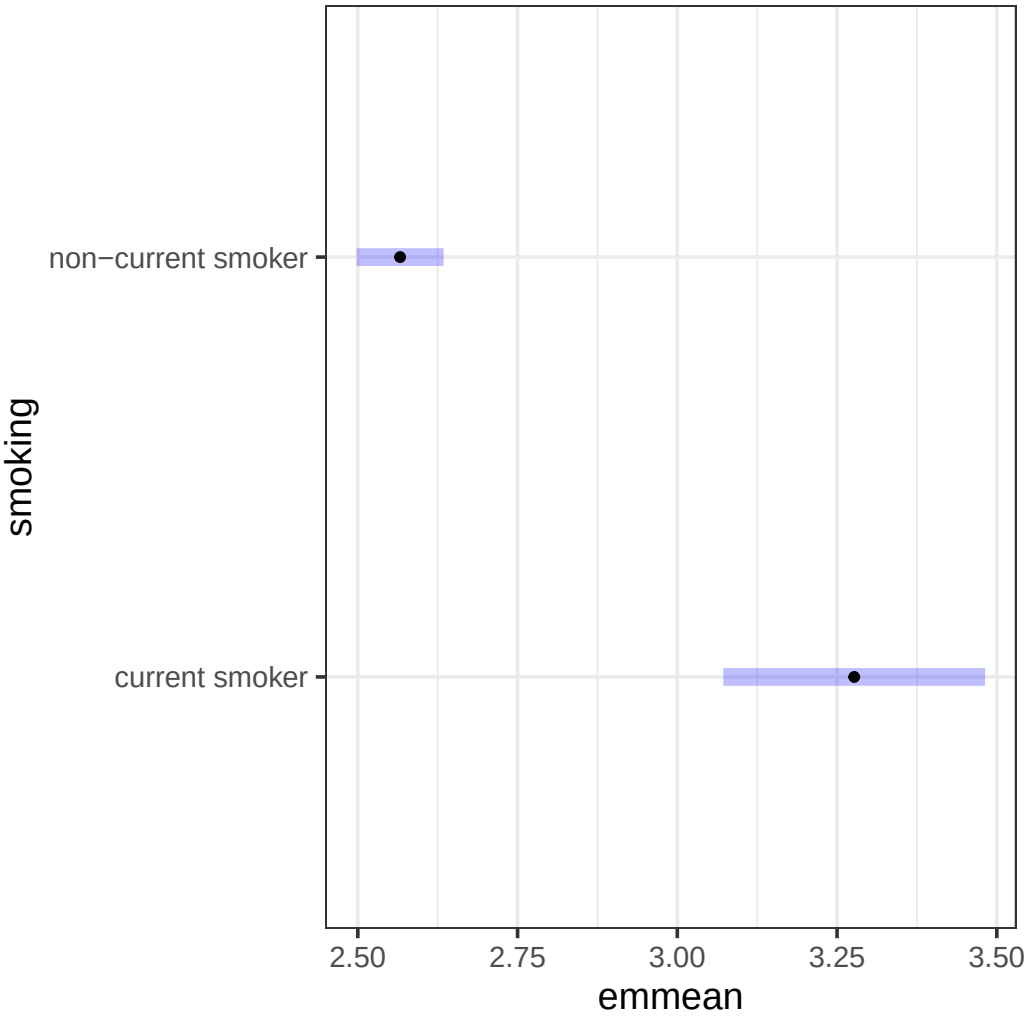
# Classical example



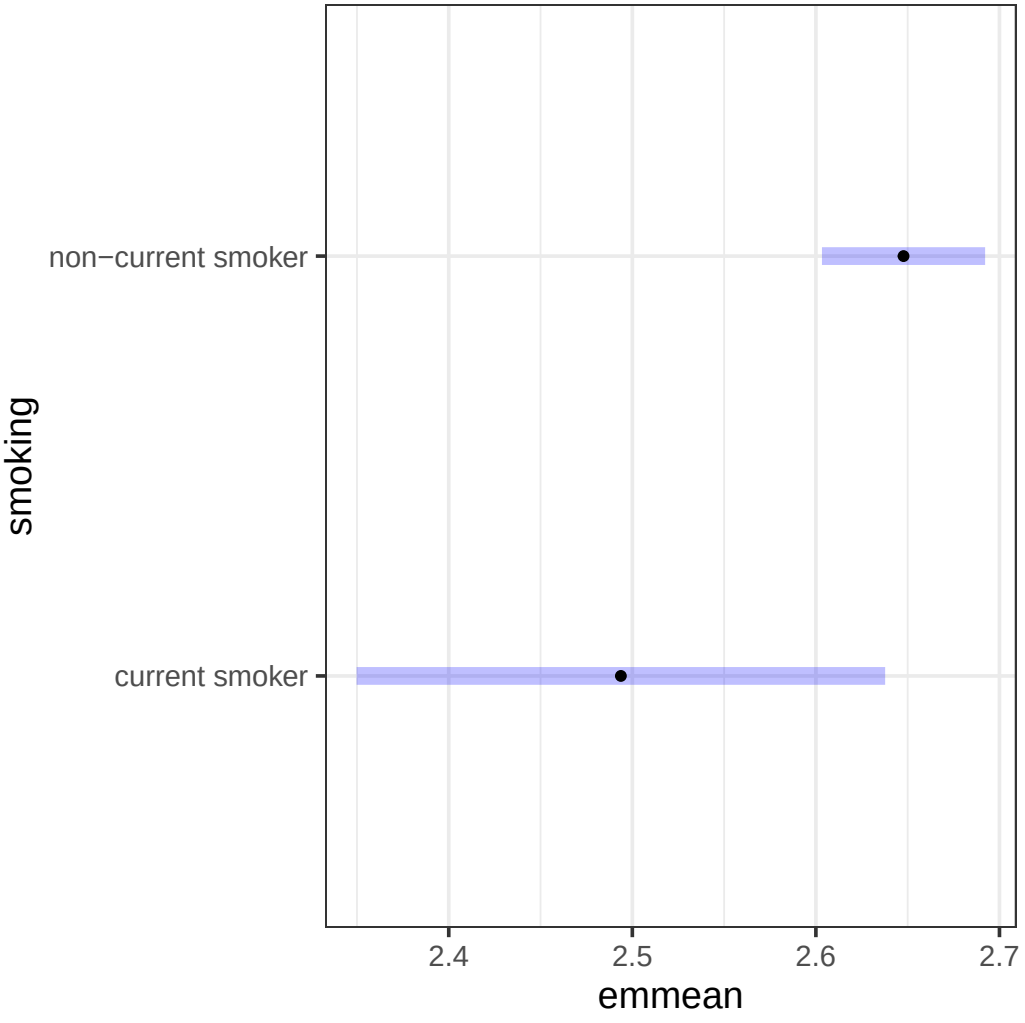
# Classical example



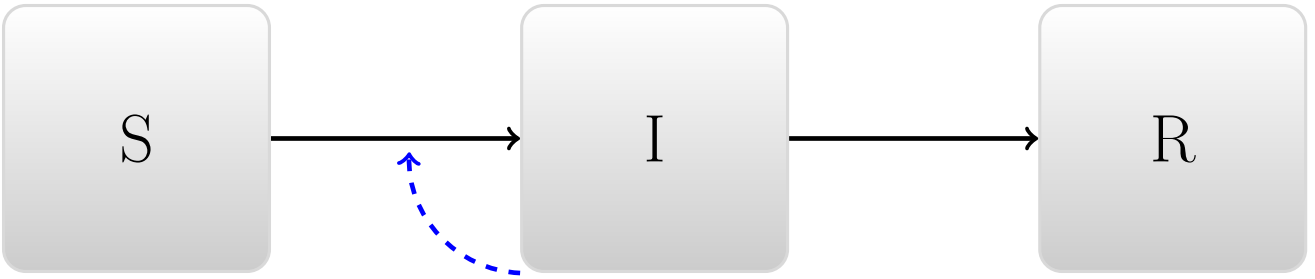
# Univariate means



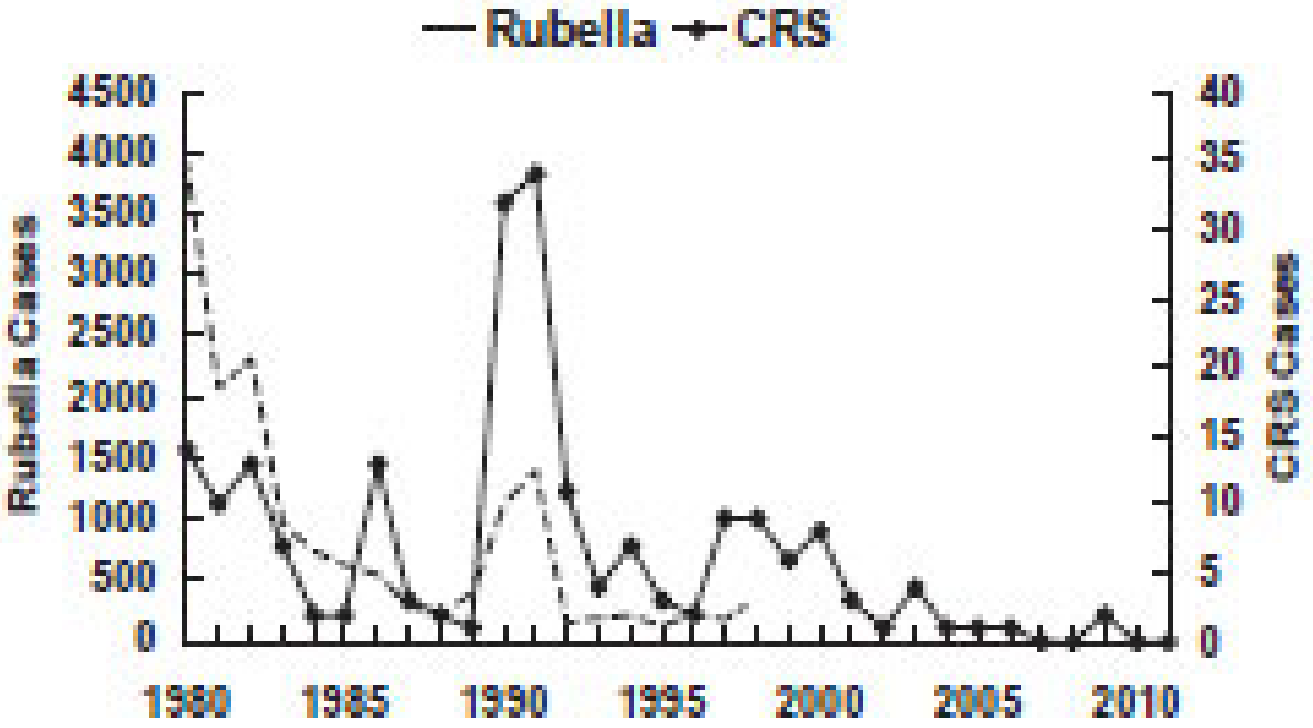
# Multivariate means



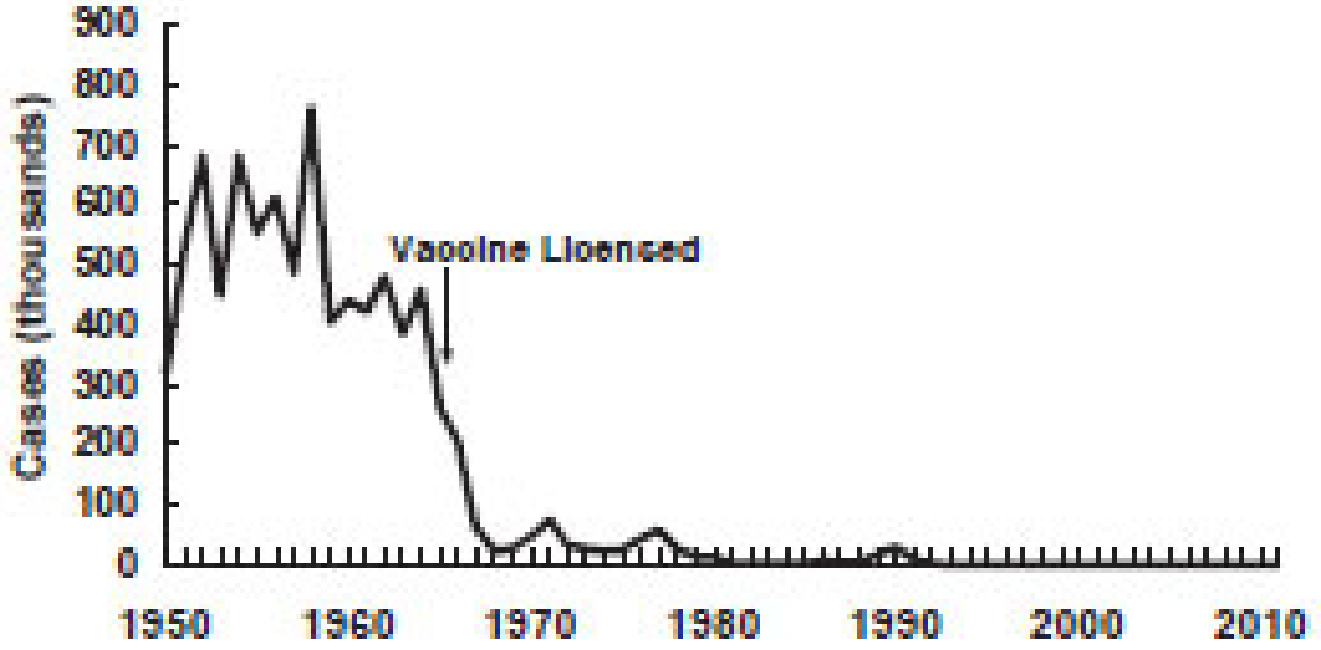
# Dynamical examples



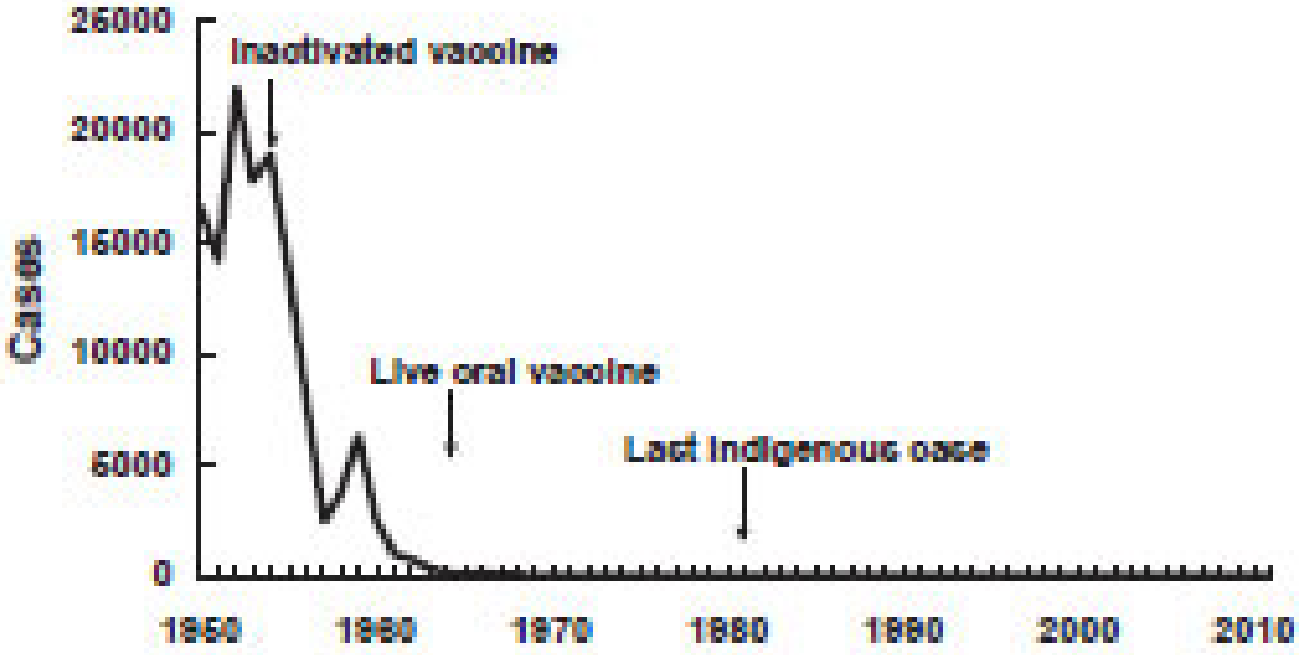
# *Rubella (present)*



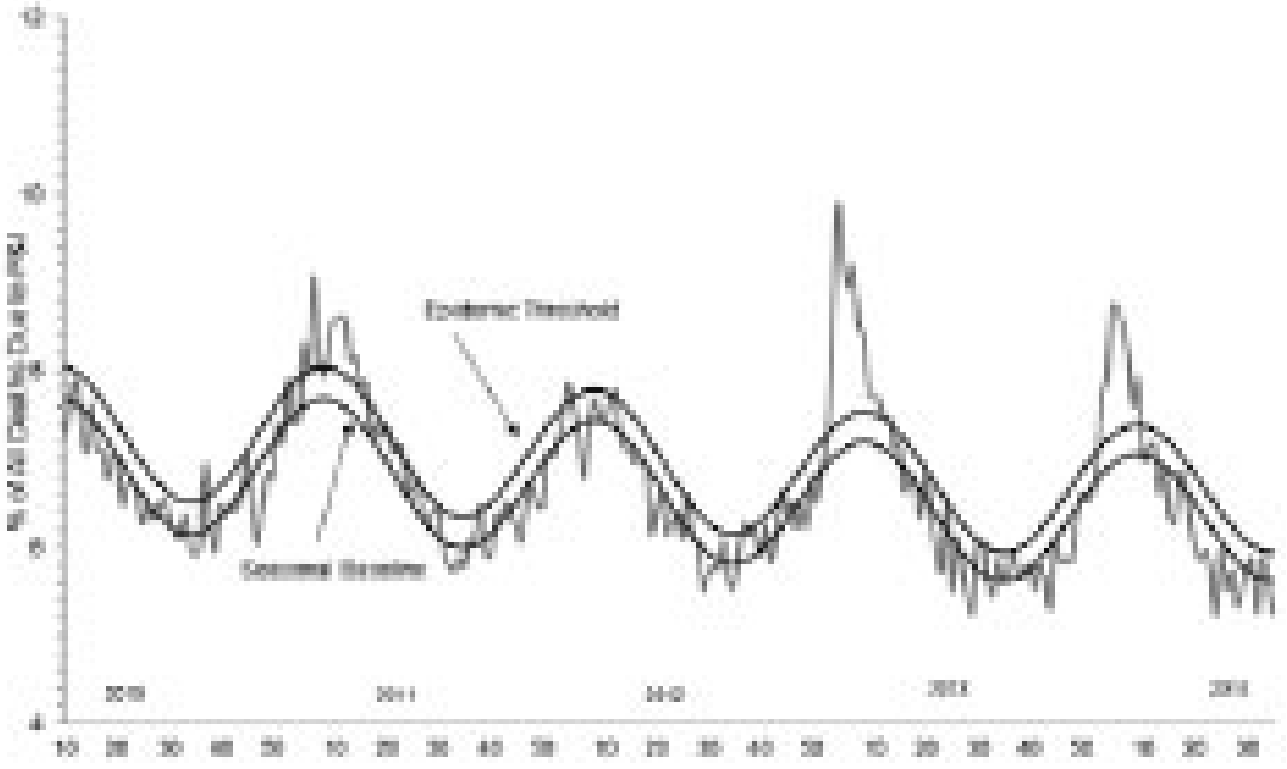
# Measles (present)



# *Polio (present)*



# *Influenza (present)*



# Bridging

- Classical epidemiology relies on statistics, avoids mechanism
- Mathematical epidemiology (the traditional approach to dynamical epidemiology) explores mechanism, avoids statistics
- Much modern dynamical epidemiology seeks ways to put dynamical mechanisms into a statistical framework
  - This is hard

# Outline

# Summary

- Science is an ongoing process
- Models are the way that we bridge between theory and reality
- Dynamical models have a key role
  - When we can't do experiments
  - When mechanisms are complex
- We should work to combine dynamics with statistical approaches




Title: Data, models and science
Attribution: Jonathan Dushoff, McMaster University, DAIDD 2020