Introduction to Modeling

Learning Objectives

By the end of this session you should learn:

- How does infectious disease modelling fit into the field of epidemiology
- What are the different goals of ID mechanistic models
- What are the different types of ID models
- What are common data relied upon for modelling
- How are ID data and ID models related

Contextualizing Infectious Disease Modelling

Disease vs Infectious Disease

Disease - A deviation from the normal physiological status of an organism that negatively affects its survival or reproduction

Infectious Disease - A disease in one organism (the host) that is caused by another organism (pathogen or parasite) which has entered the host's body

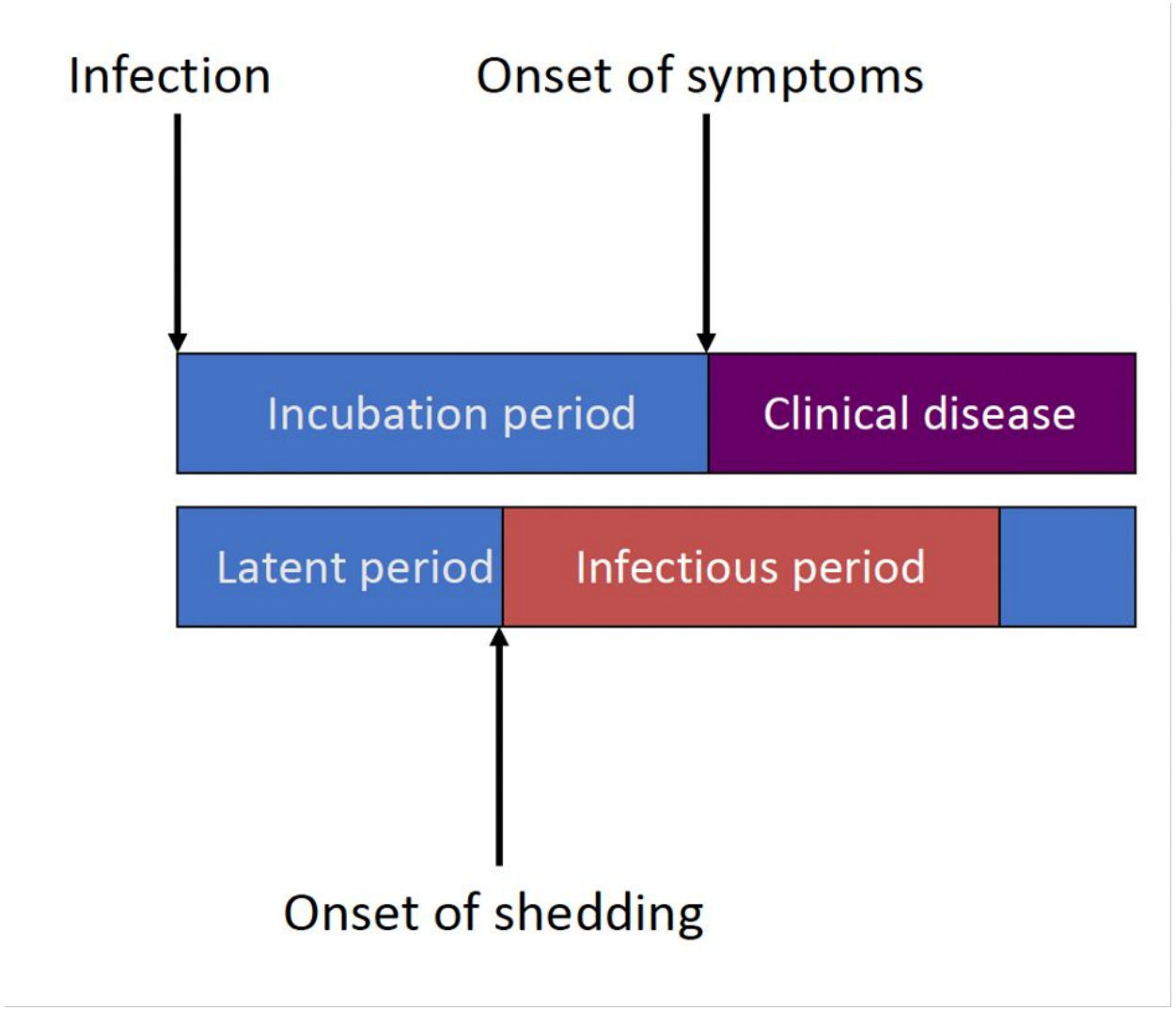
Agents of Infectious Diseases

Pathogens (agents) – organisms that are capable of producing diseases.

- Viruses (Examples: HIV -> AIDS, influenza -> flu, Measles morbillivirus -> measles)
- Bacteria (Examples: Vibrio cholerae -> Cholera, Yersinia pestis -> plague, Mycobacterium tuberculosis -> Tuberculosis)
- Fungi (Examples: Aspergillus -> Aspergillosis, tinea -> Athlete's foot)
- Protozoa (Examples: Plasmodium falciparum -> Malaria, Trypanosoma cruzi -> Chagas Disease)
- Helminths (Examples: Schistosoma mansoni -> Schistosomiasis, Hookworm -> hookworm infection)

parasites

From infection to disease



Clinical status

Infection status

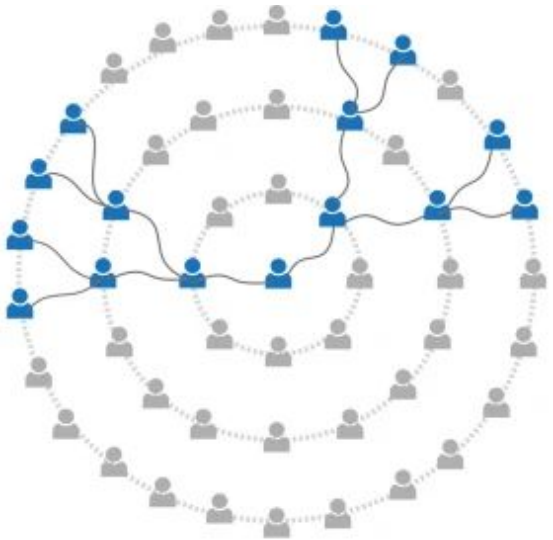
Tangent

We are modeling infections, although we actually observe clinical cases

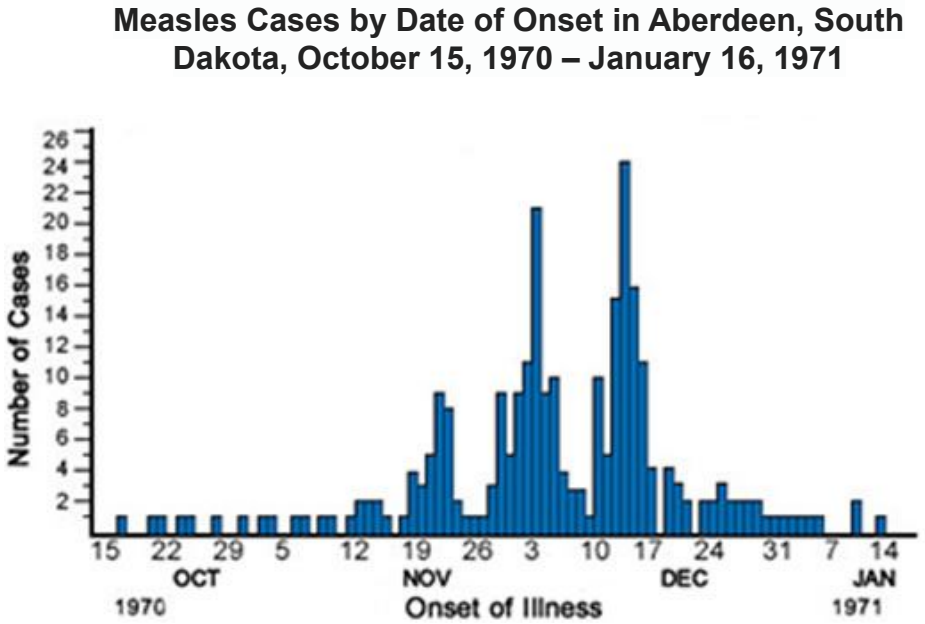
Epidemiology focus on population level

We can track the movement of pathogens throughout populations

Visualized easily via a transmission chain, which is the set of infection events that occur as a pathogen moves through a population

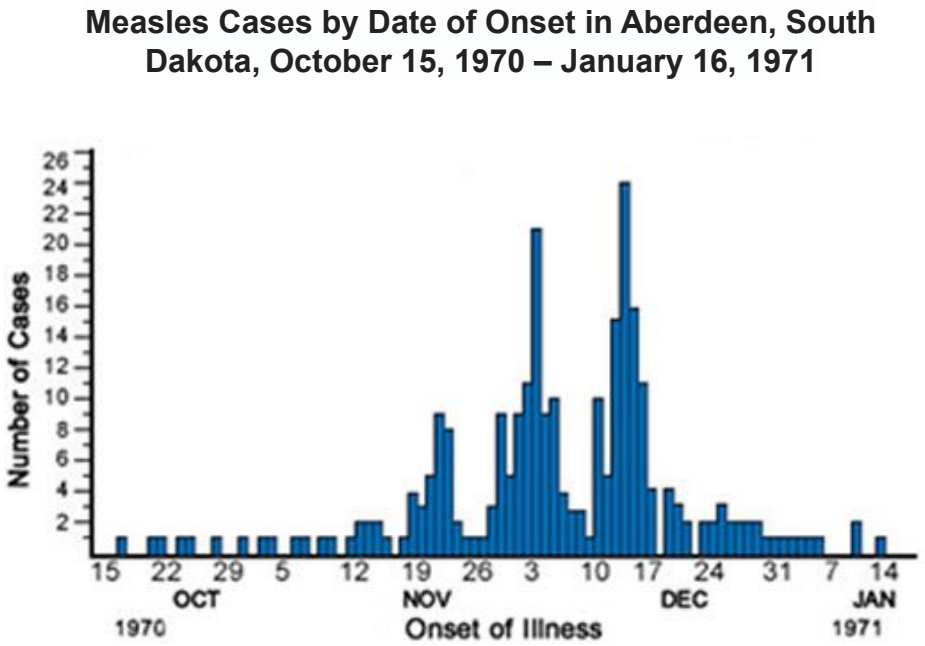


Wegehaupt et al. BMC Public Health 2023



https://www.cdc.gov/training/quicklearns/epimode/

Use cases from time series data (Epi Curve)



Epi Outbreak investigation:

- What pathogen is causing the illness?
- It is a novel pathogen?
- Who is infected?

Biological questions:

- What allows a pathogen to enter the population?
- What does the growth rate tell us?
- Why does the epidemic turn over?
- Why are there three peaks?

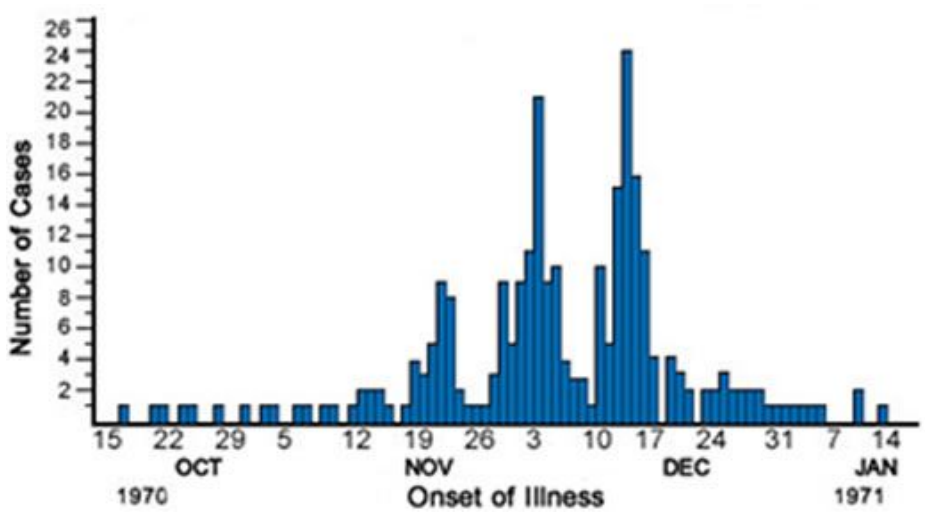
Intervention questions:

- How to prevent spread?
- When is best to implement control?
- Drugs, vaccines, or other control measures?

Single outbreak vs Persistence

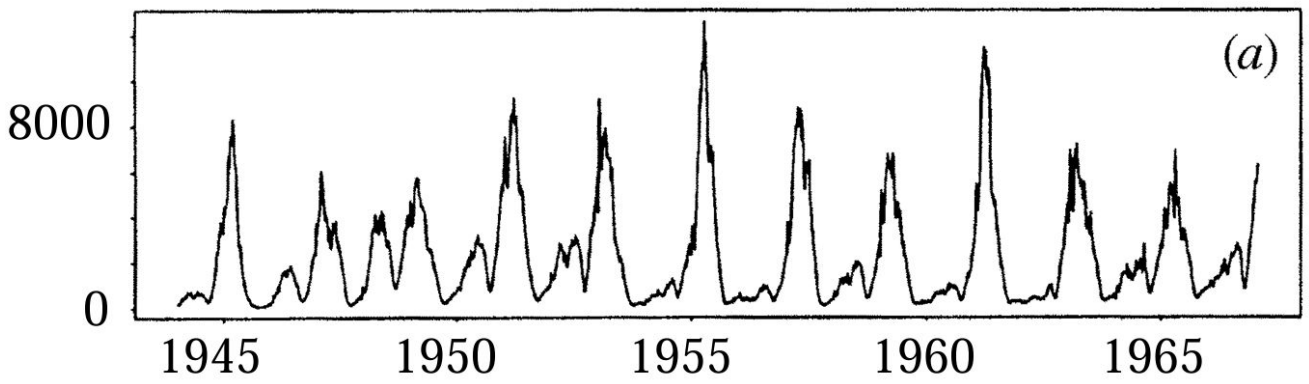
Closed population vs Open population

Measles Cases by Date of Onset in Aberdeen, South Dakota, October 15, 1970 – January 16, 1971



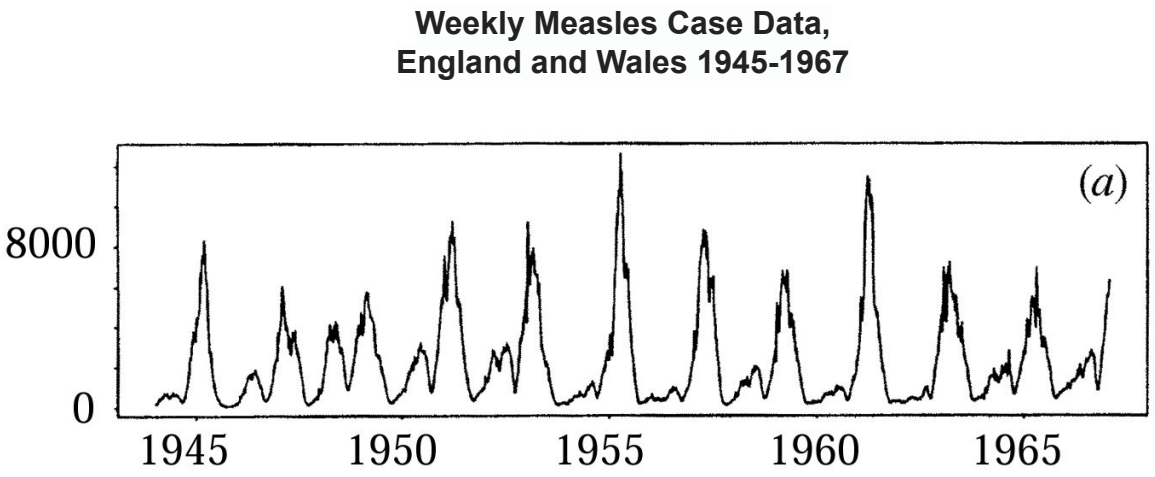
https://www.cdc.gov/training/quicklearns/epimode/

Weekly Measles Case Data, England and Wales 1945-1967



Finkenstädt B, Grenfell B. Proc Biol Sci. 1998

Use cases from long time series data



Epidemiological Dynamics

- What is the net reproduction number over time?
- How does seasonality shape transmission?
- Are there multi-annual cycles, and what might explain them?

Host-Pathogen Interactions

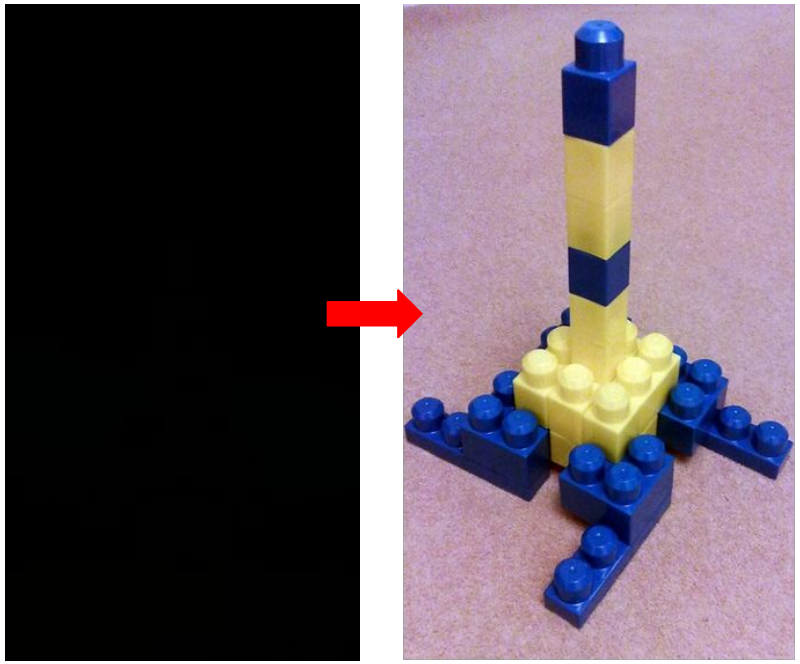
- What is the estimated duration of immunity following infection?
- Do re-infections contribute to sustained transmission?

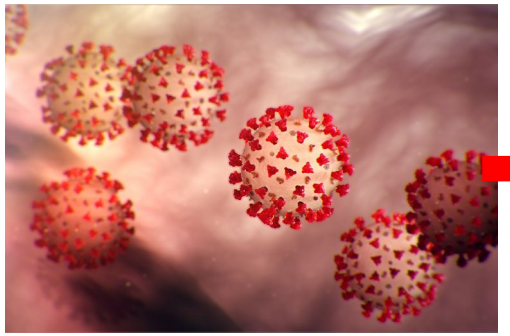
Forecasting and Interventions

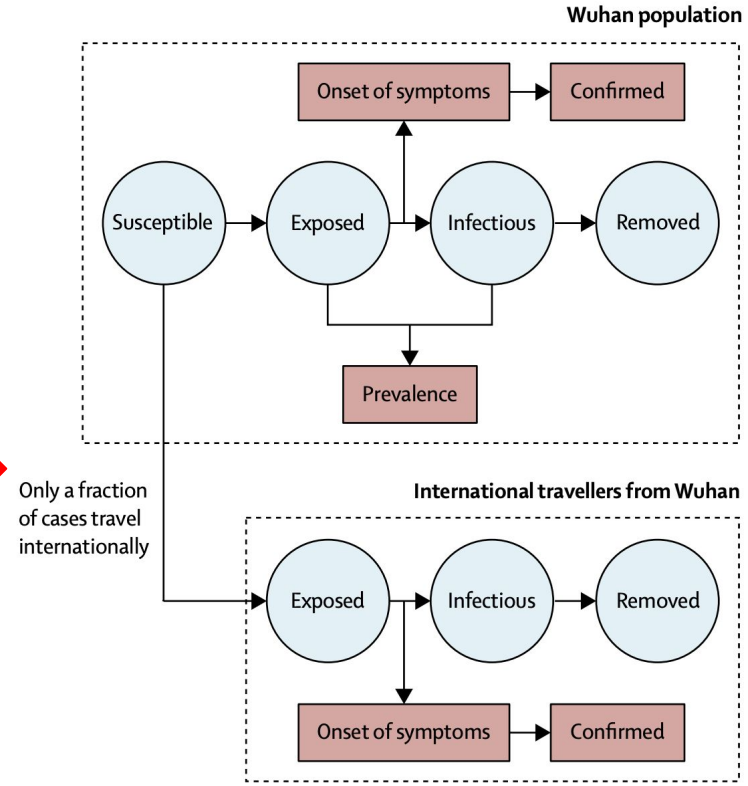
- Can we build a forecasting model to predict future peaks?
- When is best to implement control?

What is a model?

A model is a simplified representation of a more complex object/process, designed to address specific questions. It is an abstraction of reality.







Benh Lieu Song (CC 3.0 license), osde8info (CC-BY-SA 2.0 license)

https://www.nfid.org/infectious-disease/coronaviruses/

Kucharski et al. Lancet Inf Dis 2020

Infectious disease models are not new

Daniel Bernoulli's 1766 analysis of smallpox might be the first published model of an infectious disease, although not a transmission model. https://doi.org/10.1002/rmv.443

The first counterfactual

- Bernoulli worked out solutions for a system of differential equations to determine life expectancy at birth with inoculation and without inoculation to smallpox

Results

- Inoculating everyone at birth increased overall life expectancy by about 3 years
- Effective as long as the probability of dying from smallpox right after inoculation is less than ~ 11%, which Bernoulli estimated this risk to be ~1%

Where does modelling fit in epidemiology?

Classical Epidemiology	Mechanistic Epidemiology
Data-Centric	Process-Centric
Public Health	Disease Ecology
Risk Factors	Infectious Disease Dynamics
Biostatistics	Mathematical Modelling

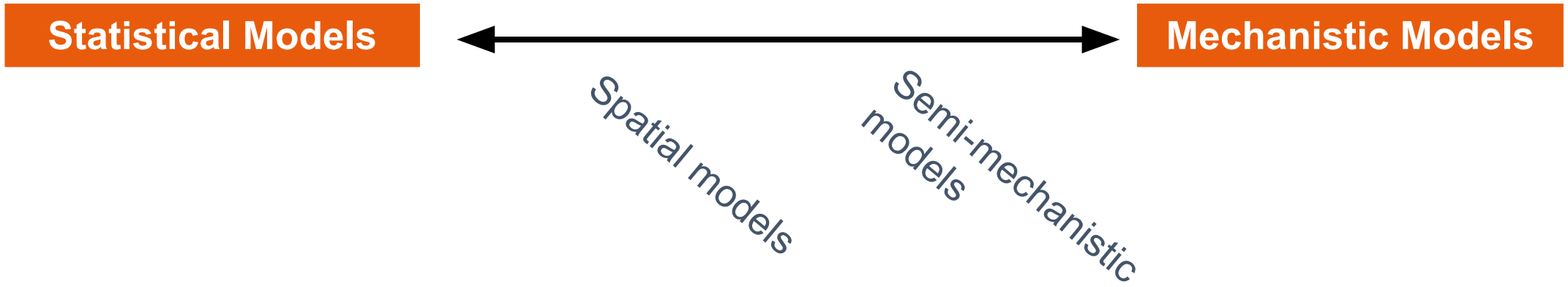
Where does modelling fit in epidemiology?

(Bio)Statistical Epidemiology (data-centric)	Mechanistic Epidemiology (process-centric)
Account for bias and random error to find correlations that may imply causality	Systems Approach: Explicitly model multiple mechanisms to understand their interactions
Often the first step to assessing relationships	Links observed relationship at different scales
Assume independence of individuals (as some scale)	Explicitly focus on dependence of individuals

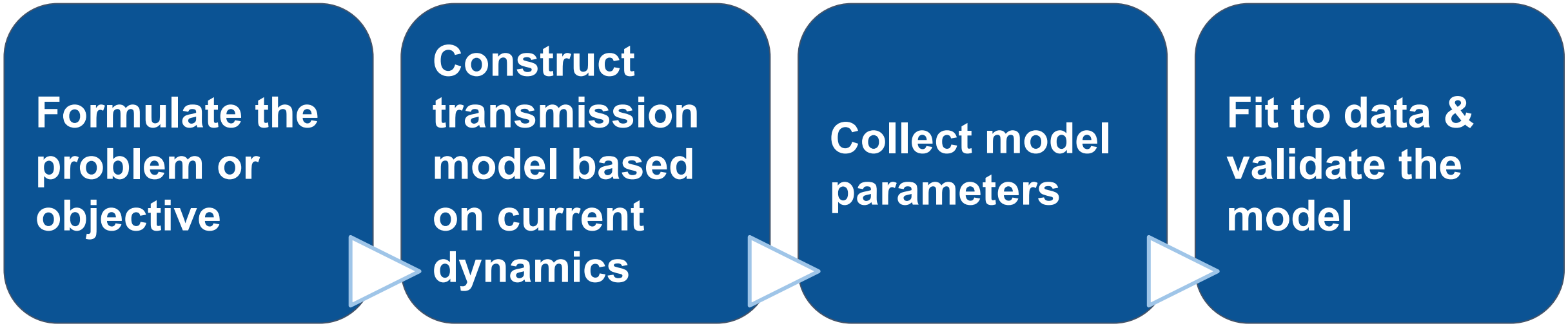
Where does modelling fit in epidemiology?

(Bio)Statistical Epidemiology (data-centric)	Mechanistic Epidemiology (process-centric)
Is HIV positively associated with the risk of TB infection?	Based on increased TB due to HIV, how much should we expect increase in TB to increase given HIV prevalence?
Are insecticide treated bednets (ITN) or indoor residual spraying (IRS) more effective for controlling malaria?	How do we expect the spatial distribution of malaria incidence to change after implementing ITB or IRS?
What are risk factors for dying from measles infection?	What is the impact of vaccination on the age profile of measles infection and deaths?

Where does modelling fit in epidemiology?

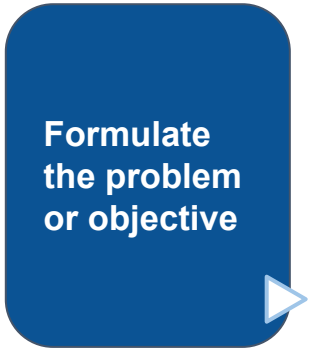


Steps of developing a model



Model Utility

Model utility



- We develop models based on a specific goal or objective
- When setting model goals with public health practitioners, it is critical to set and communicate realistic expectations about what the model can

Four goals for mechanistic ID models

Formulate the problem or objective

Theoretical Modelling

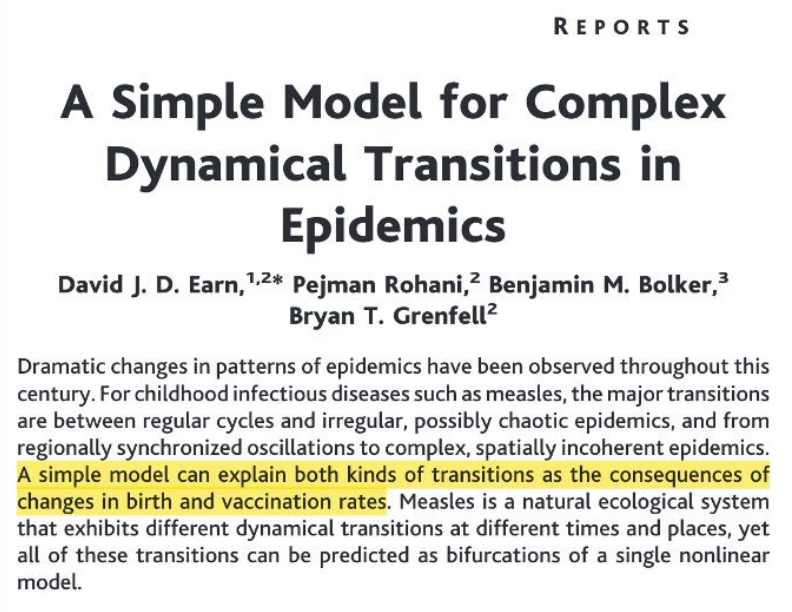
What type of emergence behavior is produced by disease systems with different properties?

- What-if type questions
- Highly abstract
- Explore consequences of hypothetical mechanisms

Manuscript in preparation Javier Perez-Saez & Justin Lessler

Theoretical model example

Earn et al. investigated the causes of transitions in measles epidemic patterns from regular (i.e., annual or biennial) to irregular outbreaks. The assumption under-investigation was the dynamical effect of changing birth rates and vaccination rates (i.e., changes in transmission) on incidence given term-time forcing. As a result, the authors displayed a bifurcation diagram looking across multiple transmission rates on incidence patterns and found more stochastic dynamics at lower transmission rates. [10.1126/science.287.5453.667]

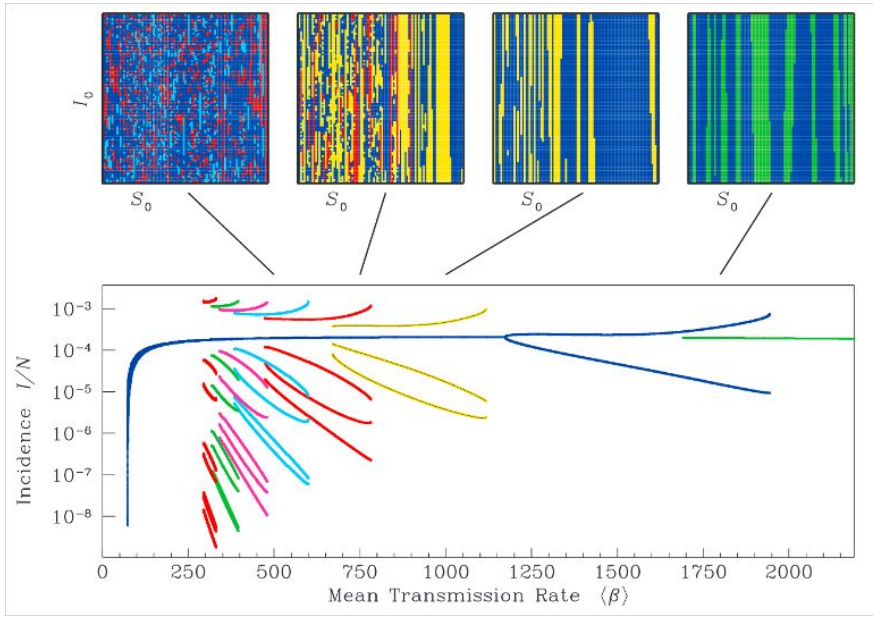


REPORTS

A Simple Model for Complex Dynamical Transitions in Epidemics

David J. D. Earn,[1,2*] Pejman Rohani,[2] Benjamin M. Bolker,[3] Bryan T. Grenfell[2]

Dramatic changes in patterns of epidemics have been observed throughout this century. For childhood infectious diseases such as measles, the major transitions are between regular cycles and irregular, possibly chaotic epidemics, and from regionally synchronized oscillations to complex, spatially incoherent epidemics. A simple model can explain both kinds of transitions as the consequences of changes in birth and vaccination rates. Measles is a natural ecological system that exhibits different dynamical transitions at different times and places, yet all of these transitions can be predicted as bifurcations of a single nonlinear model.



Four goals for mechanistic ID models

Formulate the problem or objective

Theoretical Modelling

What type of emergence behavior is produced by disease systems with different properties?

- What-if type questions
- Highly abstract
- Explore consequences of hypothetical mechanisms

Inference Modelling

What is the true nature of the disease processes that are producing the observed health metrics?

- Understand specific mechanisms of transmission
- Quantify value of specific parameters
- Account for epistemic and sampling process

Manuscript in preparation Javier Perez-Saez & Justin Lessler

Inference model example

Ecological Monographs, 72(2), 2002, pp. 185–202

DYNAMICS OF MEASLES EPIDEMICS: SCALING NOISE, DETERMINISM, AND PREDICTABILITY WITH THE TSIR MODEL

BRYAN T. GRENFELL,[1,4] OTTAR N. BJØRNSTAD,[2] AND BÄRBEL F. FINKENSTÄDT[1,3]

[1]*Department of Zoology, University of Cambridge, Cambridge CB2 3EJ UK*
[2]*Departments of Entomology and Biology, 501 ASI Bldg., Pennsylvania State University, University Park, Pennsylvania 16802 USA*
[3]*Department of Statistics, University of Warwick, Coventry CV4 7AL UK*

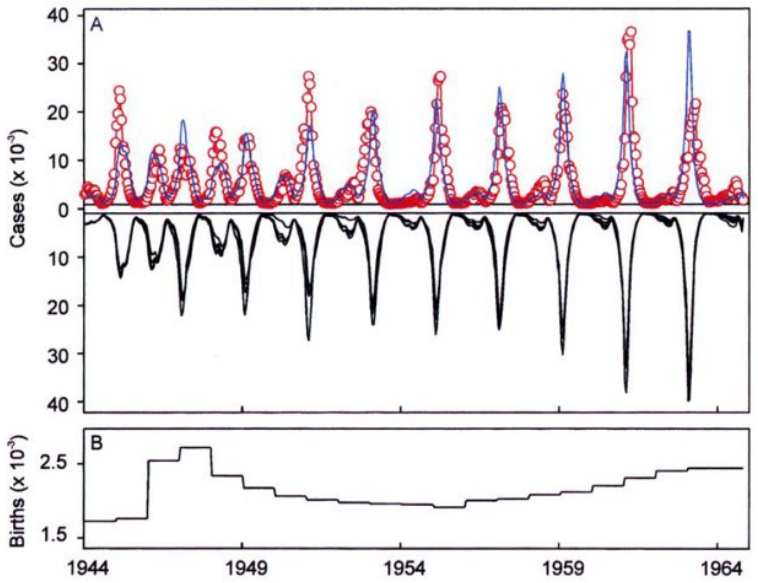


FIG. 1. (A) Measles incidence in 2-wk periods (in hundreds) in London from 1944 to 1965. The circles and the red line represent observed incidence (corrected for underreporting). The blue line represents the deterministic prediction from the TSIR model (using the susceptible and infected density in the first 2-wk period of 1944 as initial conditions). The black lines (and inverted scale) represent five stochastic realizations of the TSIR model. (B) The biweekly number of births (in hundreds) in London. The numbers are averaged within each year. The post-World War II baby boom in the late 1940s is associated with a period of annual cycles in measles incidence.

Four goals for mechanistic ID models

Formulate the problem or objective

Theoretical Modelling

What type of emergence behavior is produced by disease systems with different properties?

- What-if type questions
- Highly abstract
- Explore consequences of hypothetical mechanisms

Inference Modelling

What is the true nature of the disease processes that are producing the observed health metrics?

- Understand specific mechanisms of transmission
- Quantify value of specific parameters
- Account for epistemic and sampling process

Strategic Modelling

How will an epidemic unfold and different control strategies work under various conditions?

- Conditional predictions of what could happen under specific scenarios
- *Focus on contrast's between scenarios*

Manuscript in preparation Javier Perez-Saez & Justin Lessler

Strategic modeling example

vaccines

MDPI

Article

Rubella Vaccine Introduction in the South African Public Vaccination Schedule: Mathematical Modelling for Decision Making

Nkengafac Villyen Motaze [1,2,*], Ijeoma Edoka [3], Charles S. Wiysonge [2,4,5],

Table 1. Possible scenarios for rubella-containing vaccine (RCV) introduction in South Africa.

Scenario	Routine Vaccination in Expanded Program on Immunization (EPI)	Target Age Group for Routine Vaccination	Target Age Group for Initial Mass Campaign	Follow-Up Mass Campaigns	
				Target Age Group	Timing
1			No RCV in EPI		
2	RCV introduction	1 year	No initial campaign	No follow-up campaign	N/A
3	RCV introduction	1 year	1 to 14 years	No follow-up campaign	N/A
4	RCV introduction	1 year	1 to 14 years	1 to 4 years	One follow-up campaign 5 years after initial campaign
5	RCV introduction	1 year	1 to 14 years	1 to 4 years	Six follow-up campaigns every 5 years after initial campaign for 30 years
6	RCV introduction	1 year and 9 years	No initial campaign	No follow-up campaign	N/A

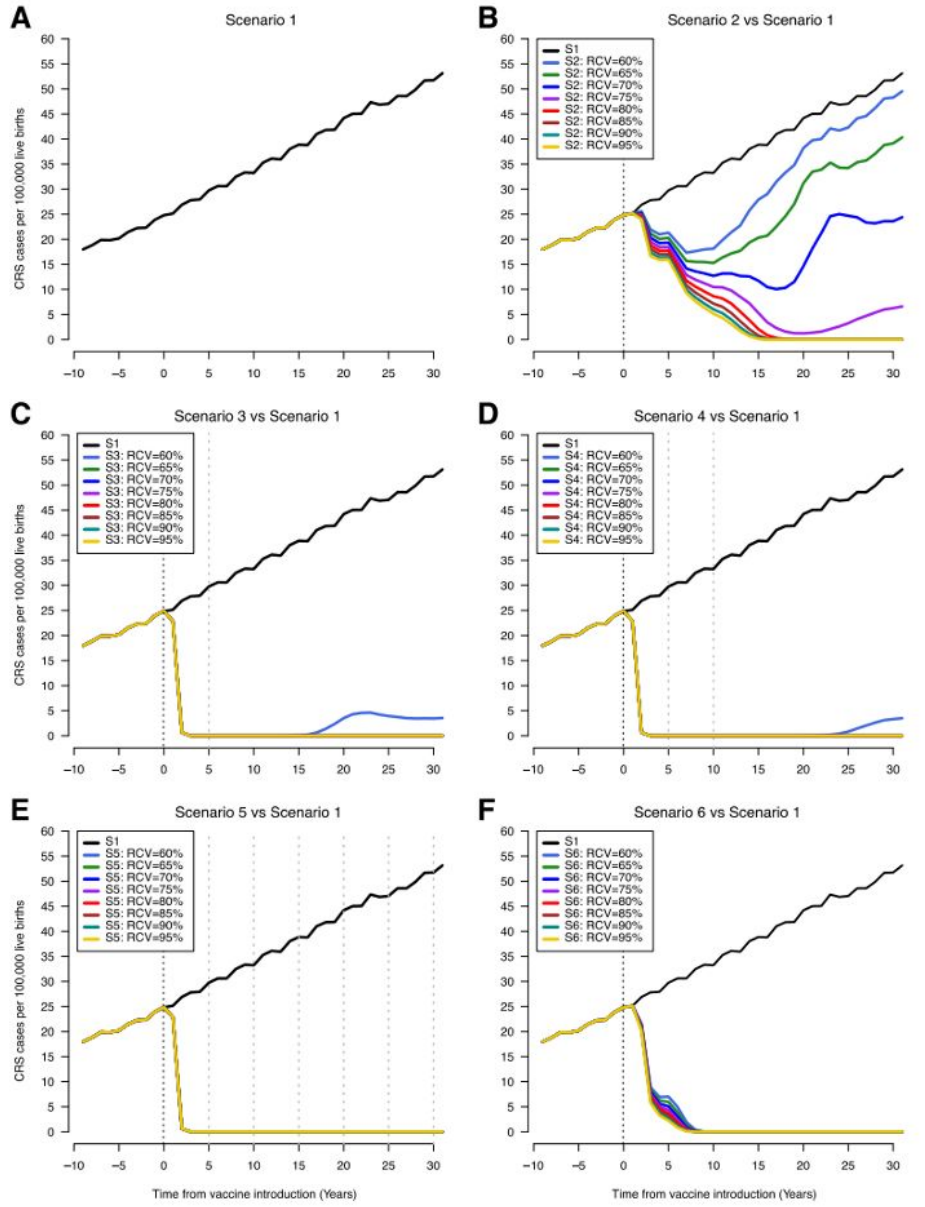


Figure 2. Time series of congenital rubella syndrome (CRS) incidence (CRS cases per 100,000 live births) showing scenario 1 (**A**) and comparing scenario 1 with scenarios 2–6 (**B**–**F**). The vertical black dotted

Four goals for mechanistic ID models

Formulate the problem or objective

Theoretical Modelling

What type of emergence behavior is produced by disease systems with different properties?

- What-if type questions
- Highly abstract
- Explore consequences of hypothetical mechanisms

Inference Modelling

What is the true nature of the disease processes that are producing the observed health metrics?

- Understand specific mechanisms of transmission
- Quantify value of specific parameters
- Account for epistemic and sampling process

Strategic Modelling

How will an epidemic unfold and different control strategies work under various conditions?

- Conditional predictions of what could happen under specific scenarios
- *Focus on contrast's between scenarios*

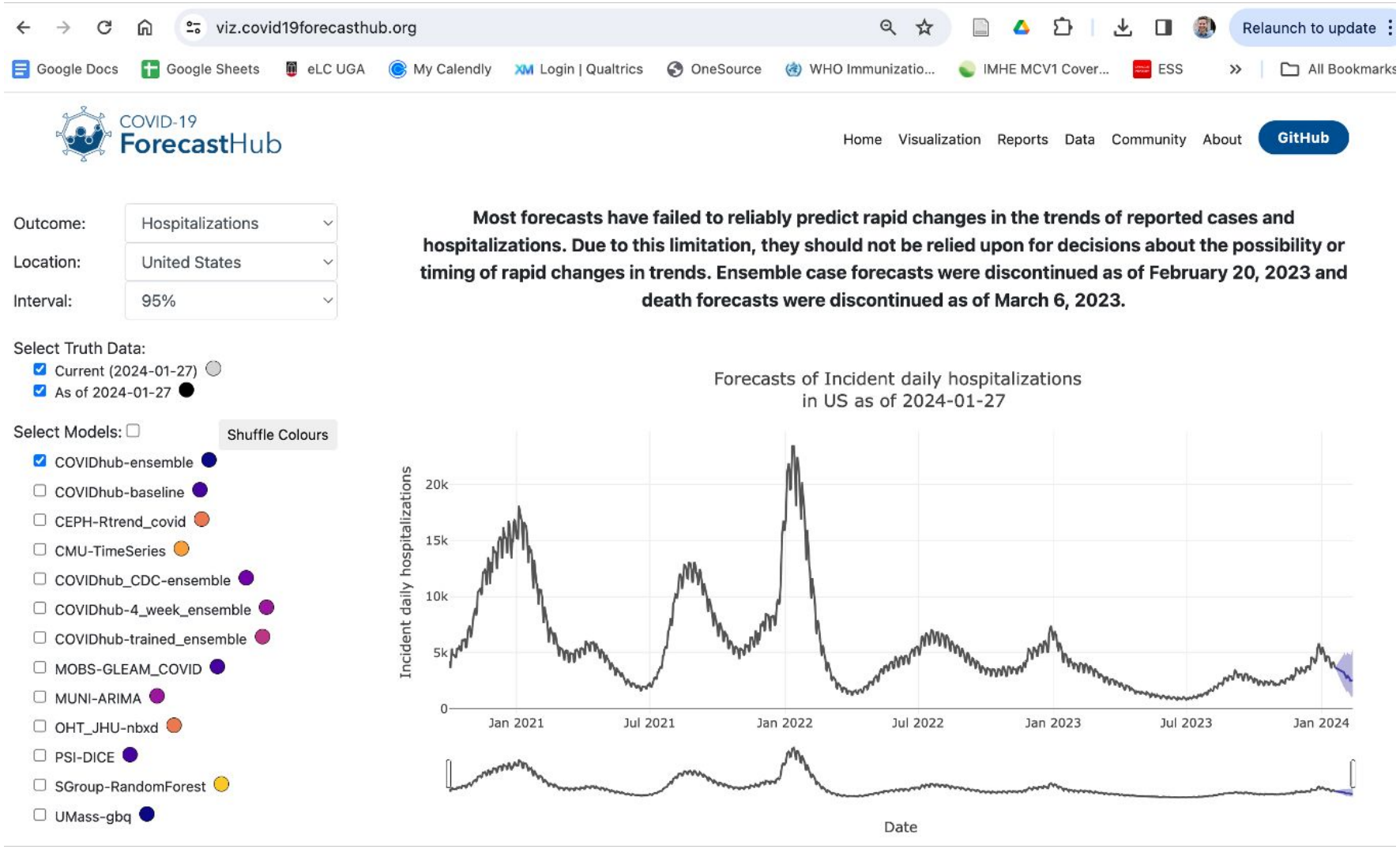
Forecast Modelling

How will an epidemic unfold in the coming weeks or months?

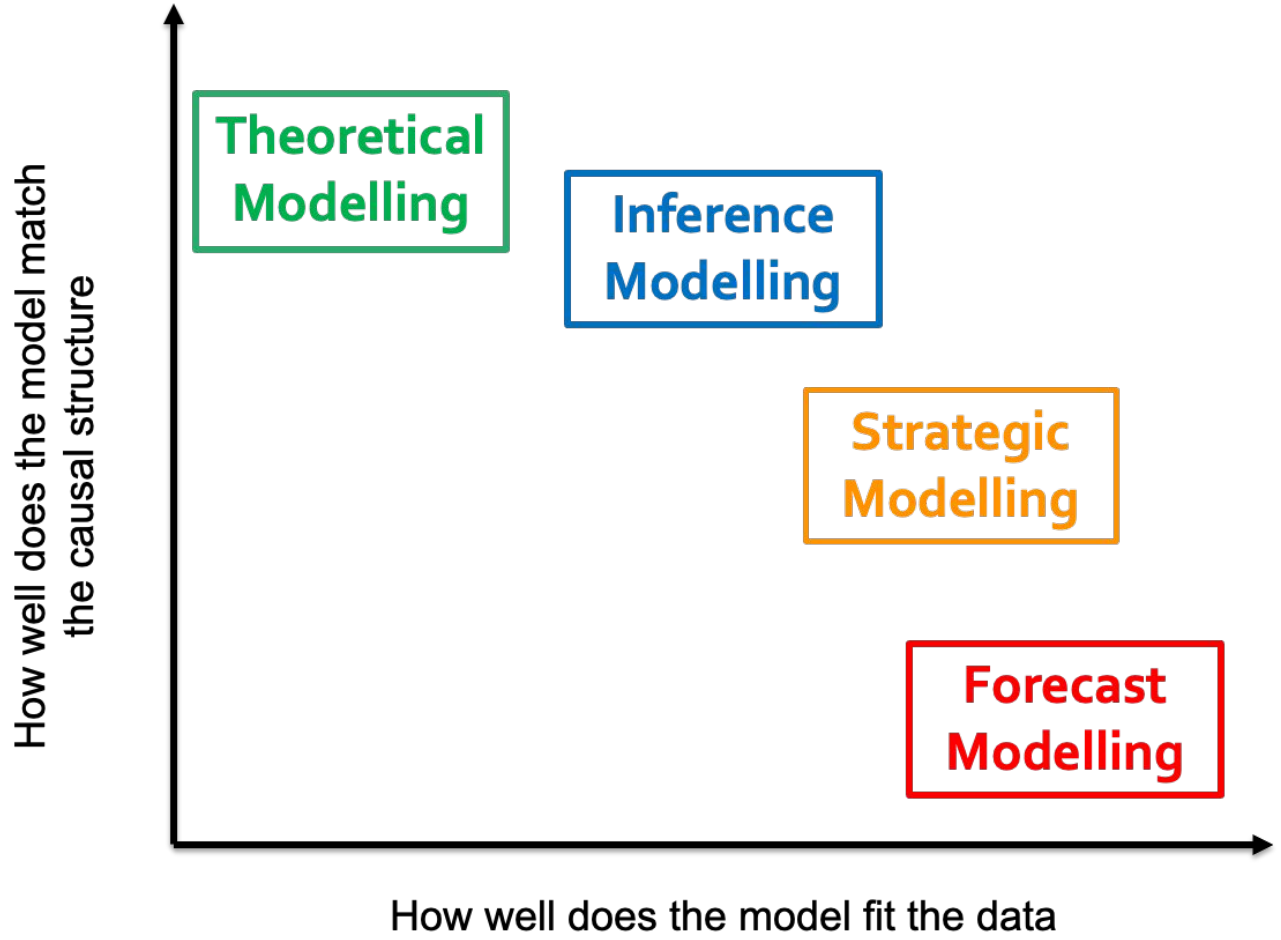
- Unconditional prediction of what will happen
- Choice of specific metrics

Manuscript in preparation Javier Perez-Saez & Justin Lessler

Forecast modeling example



Four goals for mechanistic ID models



Manuscript in preparation Javier Perez-Saez & Justin Lessler

Reporting standards based on model goals

Theoretical Modelling

- Clear description of mechanistic mapping b/w model and phenomena
- Clear description of assumptions and model structure
- Evaluation of how non-target model components influence results

Inference Modelling

- Clear description of how models are linked to data
- Description of sources of uncertainty
- Description of how threats to inference addressed (e.g., confounding)

Strategic Modelling

- Clear description of how interventions are mechanistically modeled
- Definition and justification of scenarios
- Clear definition of what outcomes will be used to make contrasts b/w scenarios

Forecast Modelling (Pollet et al 2021)

- Describe forecasting targets
- Define time horizon
- Describe how forecasts were validated

Manuscript in preparation Javier Perez-Saez & Justin Lessler

Model Type

What type of model is most useful?

Construct transmission model based on current dynamic

"**Selecting the correct level of detail is one of the most difficult decisions a modeler faces**. Models that are too simple may lose face validity because they do not incorporate aspects that content experts feel are required, but models that are too complex may be difficult to build, debug, analyze, understand, and communicate."

Mark Roberts *et al*, Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. *Medical Decision Making* (2012). Available at: https://doi.org/10.1177/0272989X12454941

Model structure

Construct transmission model based on current dynamic

Model structure is usually determined by considering the **relationship between:**

- **Inputs** relevant to the natural history of disease, clinical pathways, intervention effectiveness etc.
- **Outputs** most useful to decision makers eg. cases of disease, deaths, hospital admissions, life years gained, QALYs, DALYs.

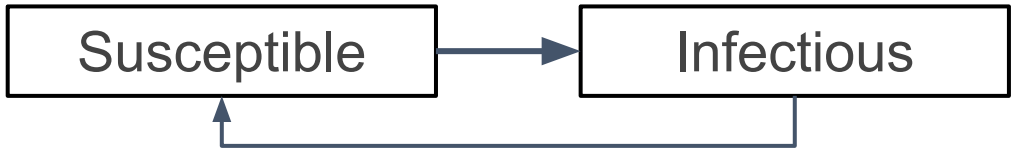
Model structure

Construct transmission model based on current dynamic

SI

Chronic infection (e.g., HIV, Hep B, Hep C)

SIS



Reinfection possible (e.g., Gonorrhea, Chlamydia, Malaria)

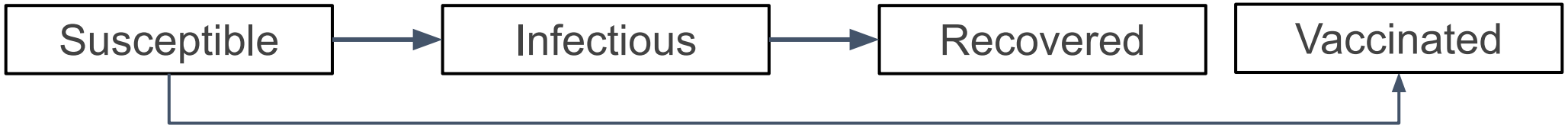
SIR

Lifelong immunity (e.g., Measles, Rubella, Mumps, Varicella zoster, Ebola)

SIRS

Temporary immunity (e.g., Pertussis, Flu, RSV, Coronaviruses)

SIRV



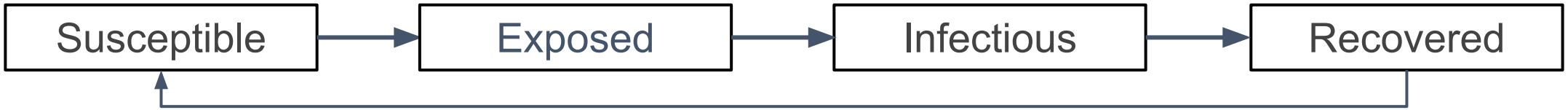
VPDs

Model structure - Common mistake

Construct transmission model based on current dynamic

SEIR

SEIRS

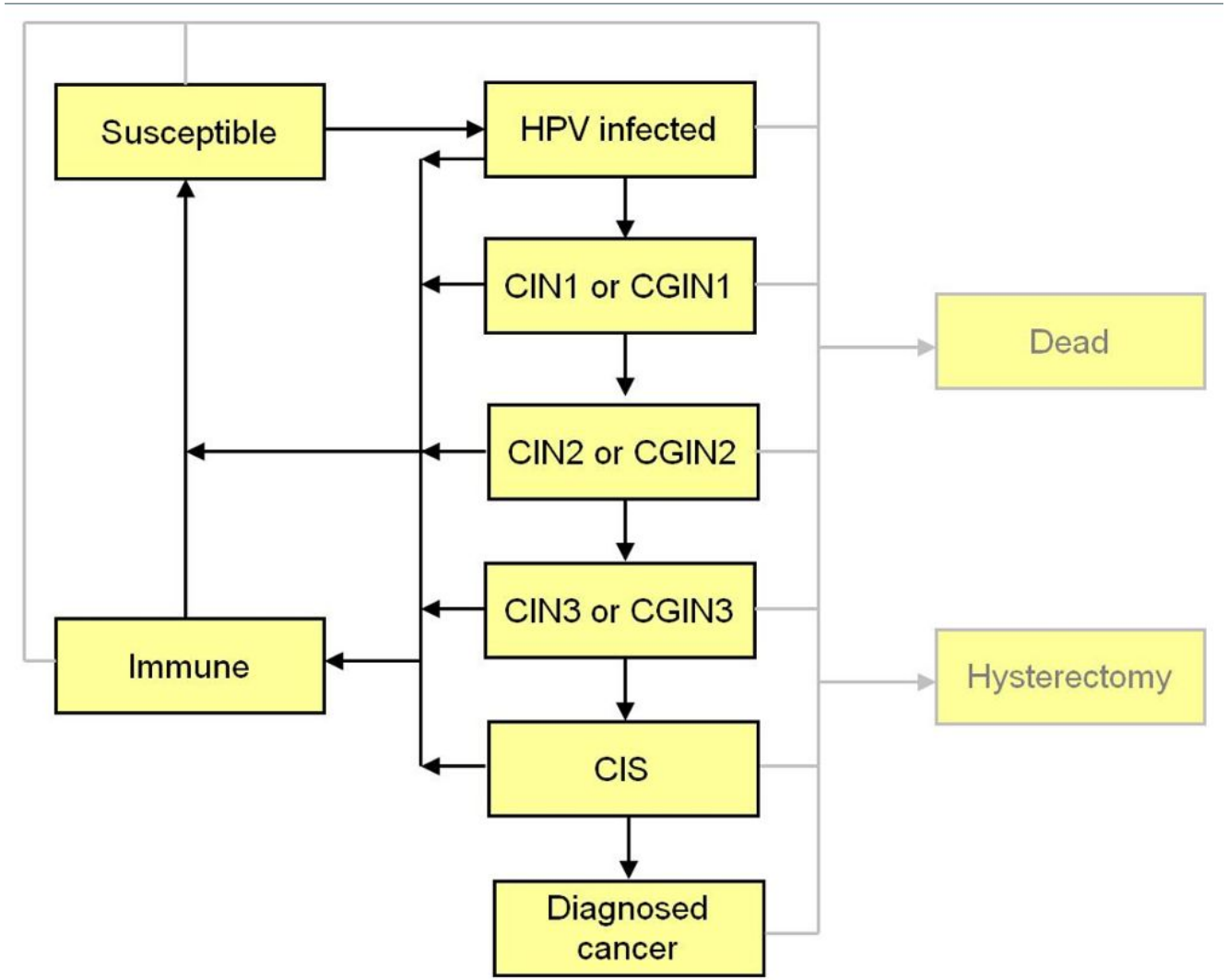


Tangent

"Exposed" is misleading. Really it means infected but not yet infectious. Think "pre-infectious"

Model structure - Can get complicated

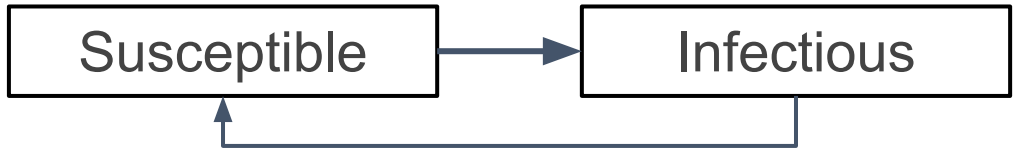
Construct transmission model based on current dynamic



Types of models

Construct transmission model based on current dynamic

Compartmental model



Non-homogenous mixing

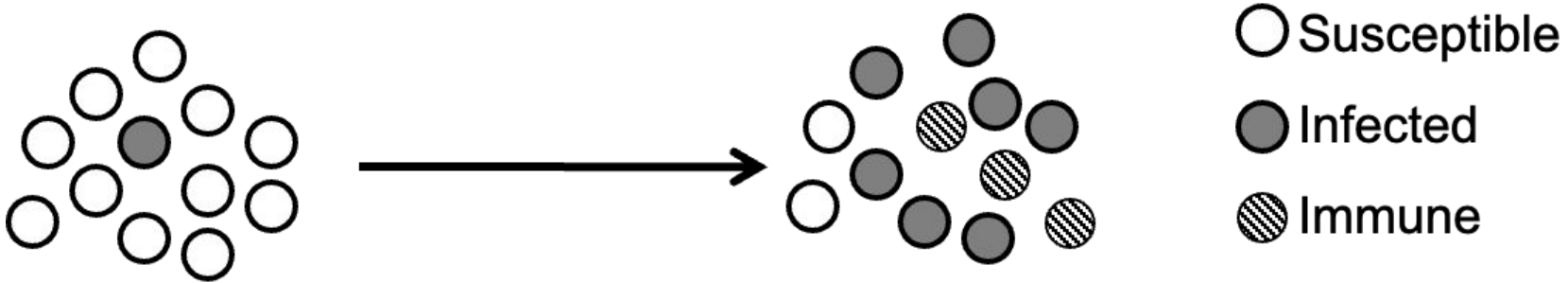
Structured compartmental model (e.g. age, sex)

Meta-population model (i.e., spatially structured)

Network model

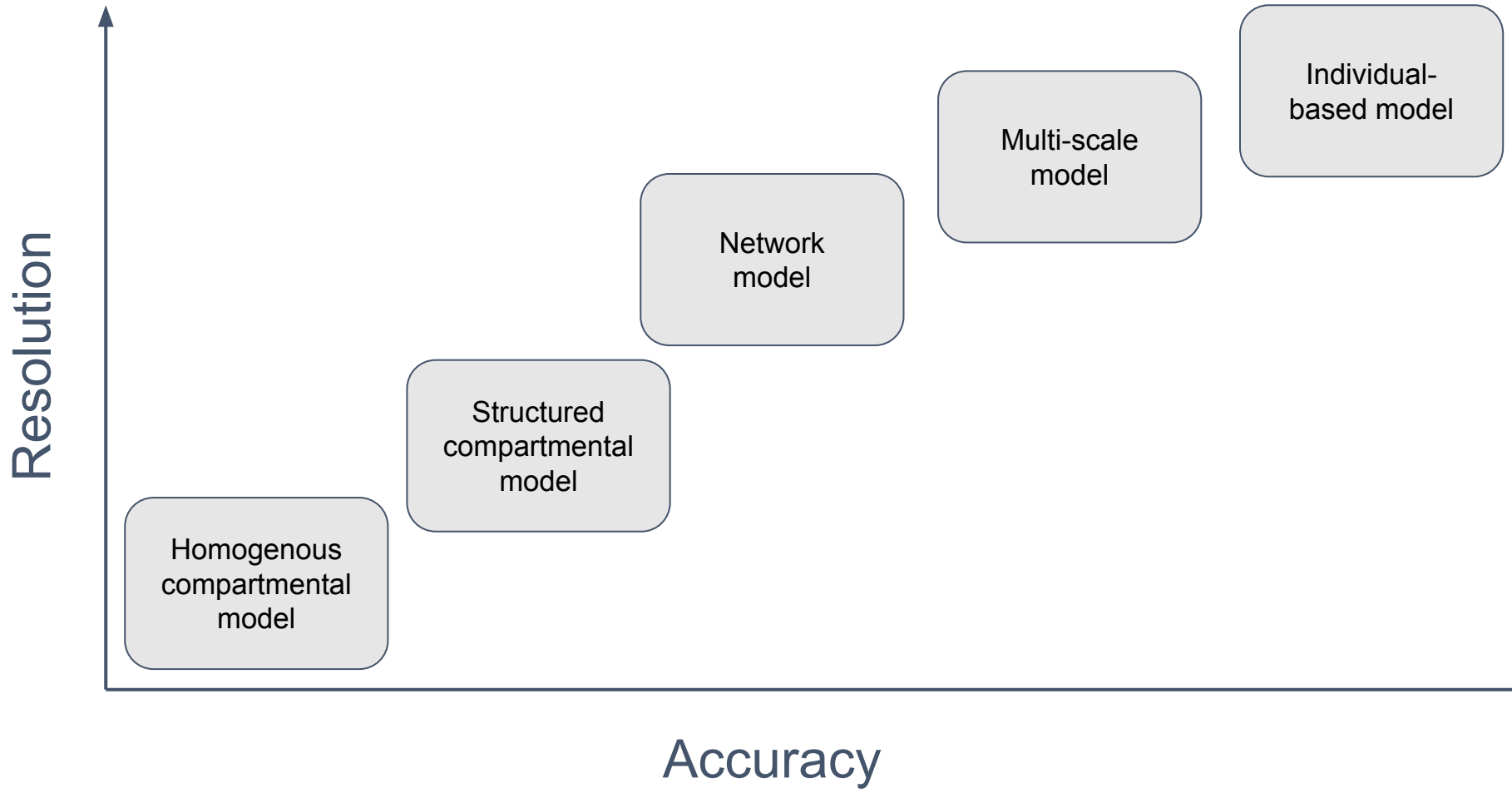
Multi-scale model

Individual- (agent-) based model



Model type depends on modelling objective

Construct transmission model based on current dynamic



What makes a good model?
Three model features…

Construct transmission model based on current dynamic

1. **Accuracy:** ability to capture observed patterns (qualitative or quantitative) and make predictions
2. **Transparency:** ability to understand model components. Decreases with model complexity
3. **Flexibility:** How easily the model can be adapted to new scenarios. Decreases model complexity

2 Types of Modeling Methods

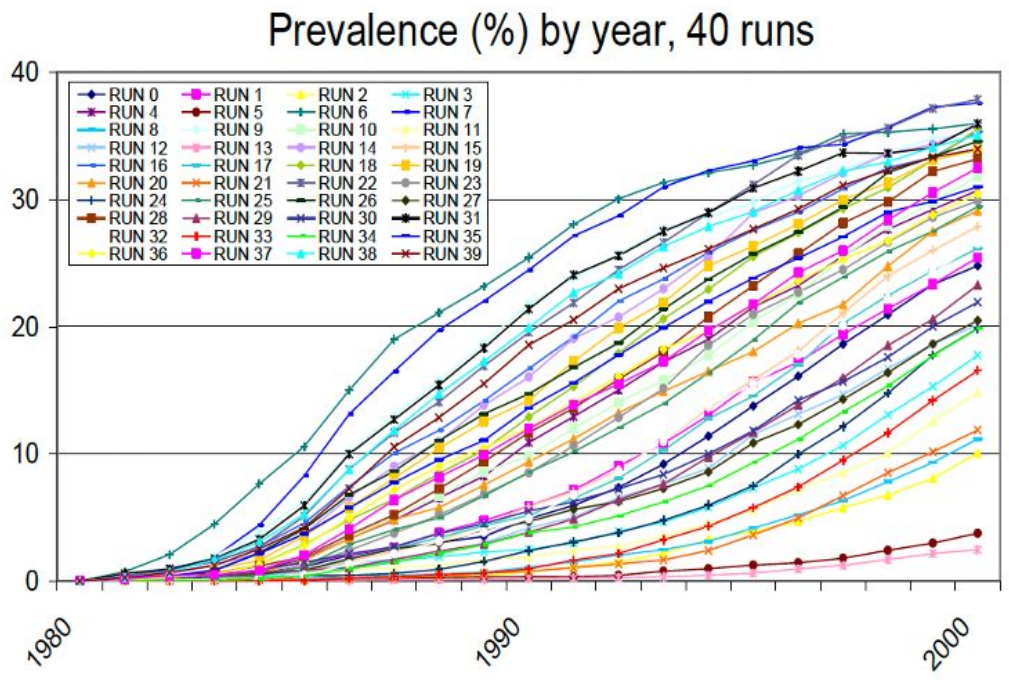
Construct transmission model based on current dynamic

1. **Deterministic:**
   - **No randomness**
   - Models give you the **same output** every time you run/solve (assuming same starting parameters/conditions)
   - Relies on average rates to describe what will happen on **average** in a population
2. **Stochastic:**
   - Incorporate chance variation or **randomness**
   - Running the model multiple times can give **different output**
   - Provide the probability of a given outcomes of range in which the outcome is likely to occur (e.g., probability that transmission ceases)

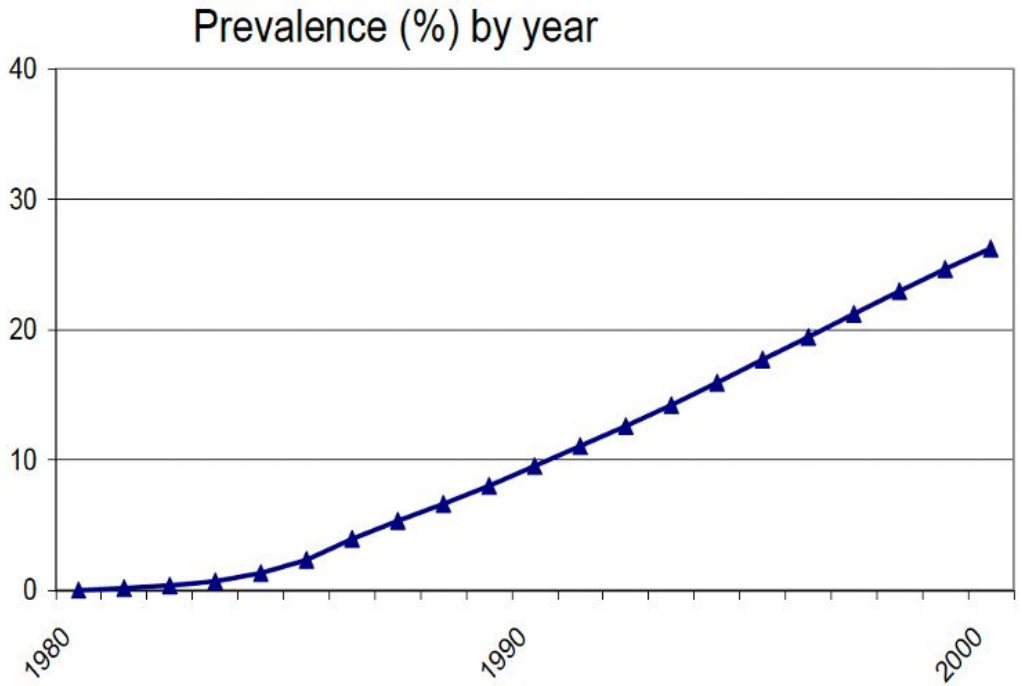
2 Types of modeling methods

Construct transmission model based on current dynamic

Stochastic model



Deterministic model



Deterministic Models:
Difference vs Differential Equations

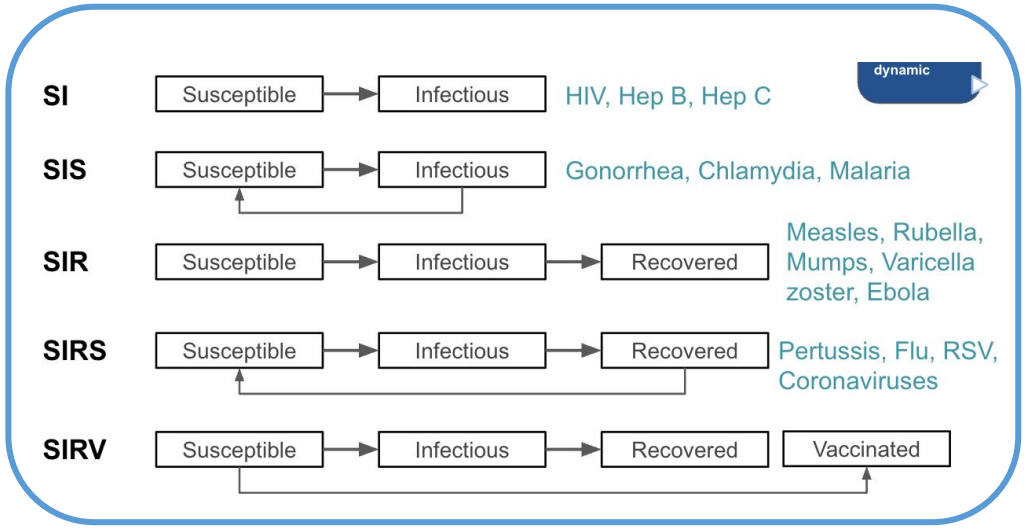
Construct transmission model based on current dynamic

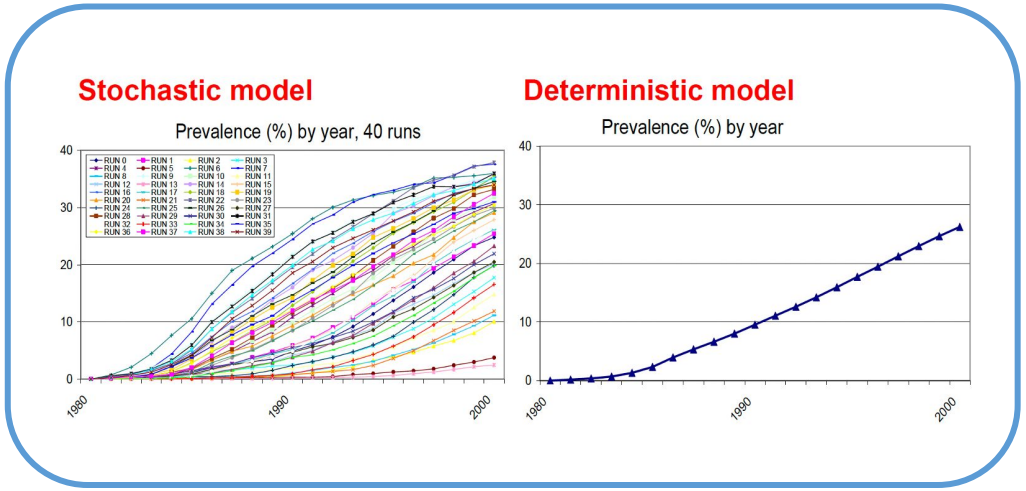
1. **Difference Equations:**
   - Calculate the number in each infected category using discrete time steps

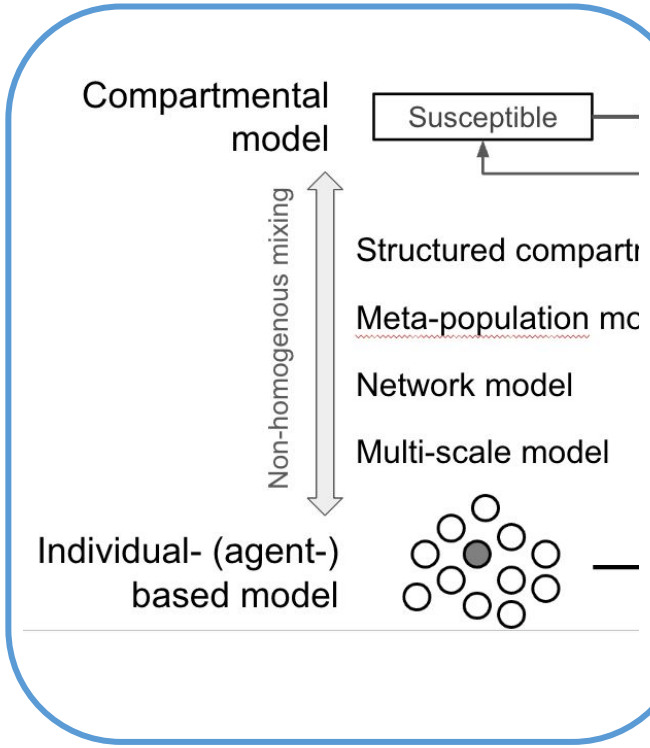
2. **Differential Equations:**
   - Calculate the number in each infection category using time steps with are "infinitesimally" small (i.e., continuous time)

Model Type Overview

Construct transmission model based on current dynamic







How to implement (run/solve) a model

Analytical Models (i.e. **pure math**)

- Often working with differential equations
- Aim to find exact solutions or expressions: e.g., what is final outbreak size
- Good for **fundamental principles**
- Can’t easily handle complexity like heterogeneity or time-specific changes

Computational Models (i.e., **computer simulations**)

- You can incorporate **real-world data,** changes over time, randomness, networks, spatial structure, etc

Data

Data Types

- Case Surveillance Data (passive vs active; suspected vs epi linked vs clinical vs confirmed)
- Genetic Surveillance Data
- Serological Surveillance Data
- Mortality Surveillance Data
- Vaccination Data (admin, survey)
- Socio-Demographic Data
- Epidemiological studies

Data Scales

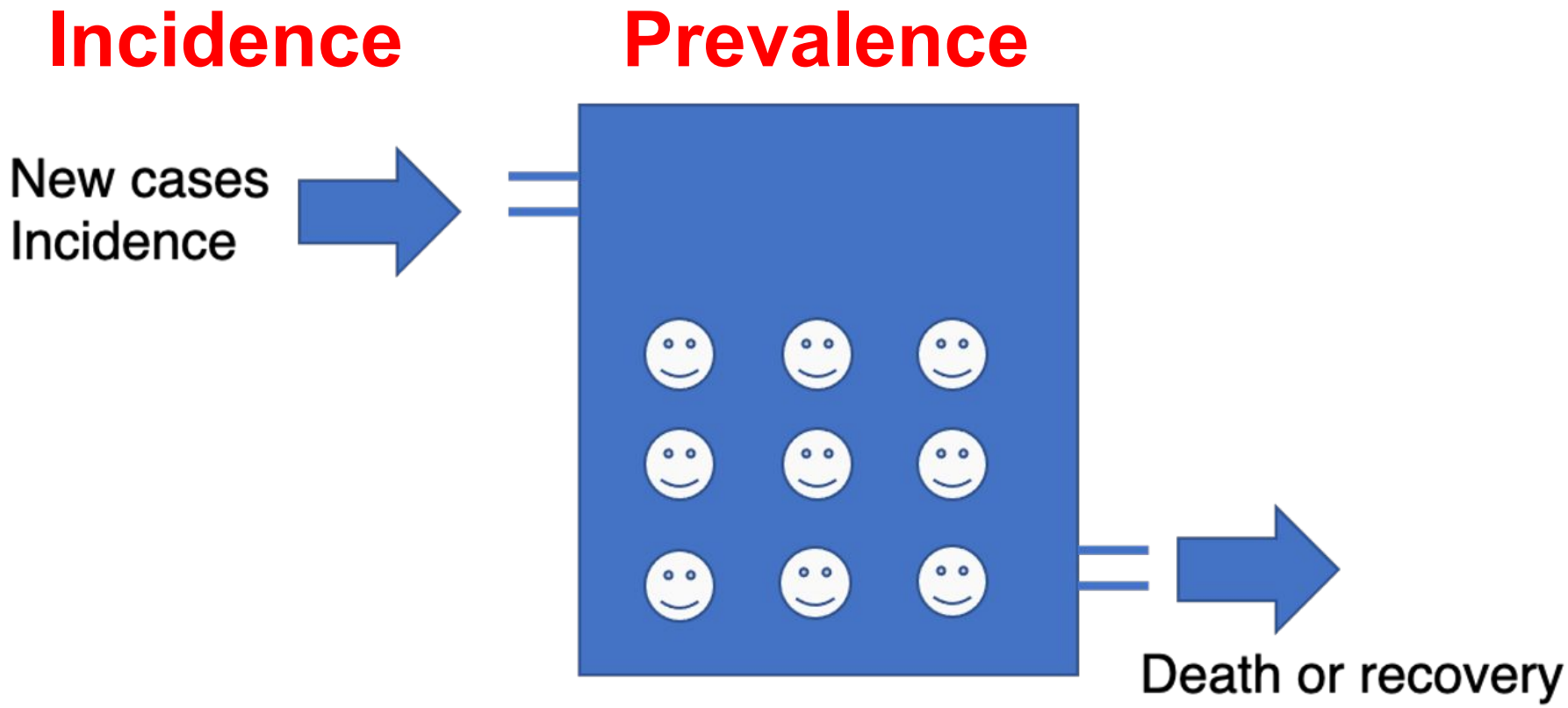
- Individual data (e.g., line list, survey observations)
- Aggregate data
- Spatial data

How are data related to ID models?

Discovery comes from testing ideas (models) against observations (data)

Models drive empirical development and vice versa

Measures of disease frequency

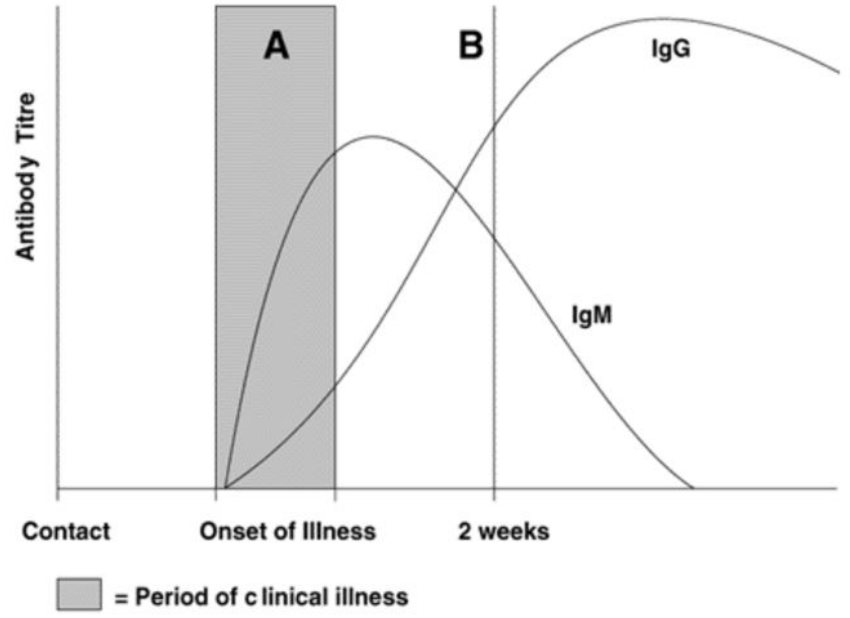


Disease **prevalence** is influenced by:

- **Incidence** of disease
- Duration of disease (time to recovery or death)

Measures of disease frequency

- Incidence of infection
- Prevalence of infection
- Seroprevalence (prevalence of antibody)

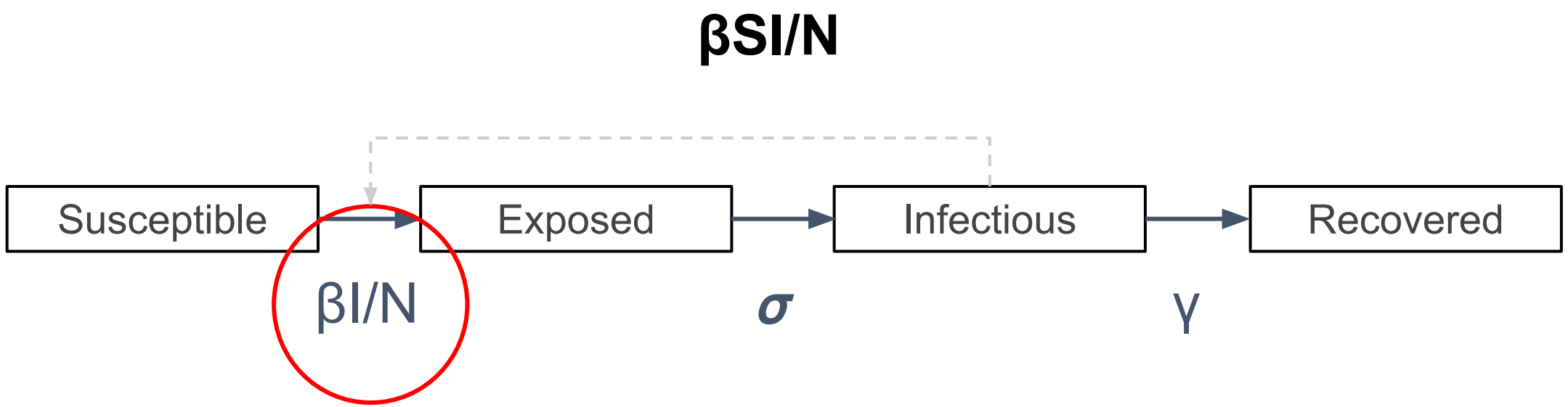


IgG antibodies persist from years to decades and are a correlate of immunity

IgM antibodies persist for a few weeks, and are often used as to confirm infection.

Measures of disease frequency

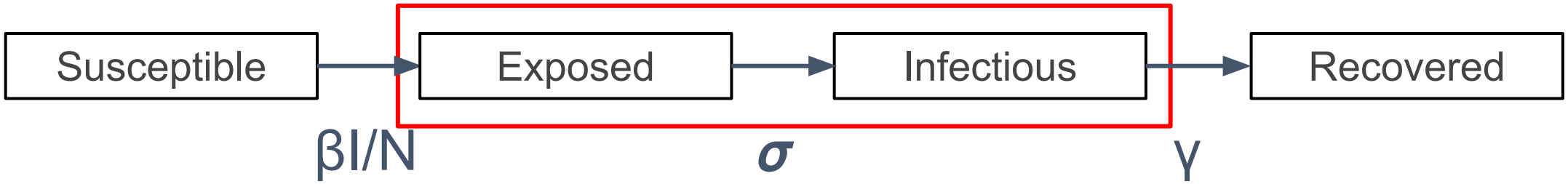
- Incidence of infection
- Prevalence of infection
- Seroprevalence (prevalence of antibody)



Measures of disease frequency

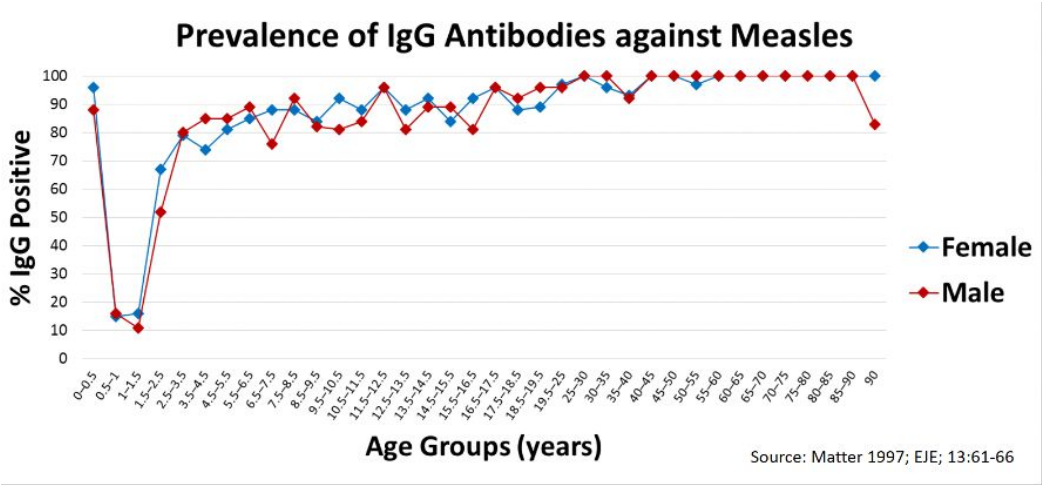
- Incidence of infection
- Prevalence of infection
- Seroprevalence (prevalence of antibody)

(E+I)/N

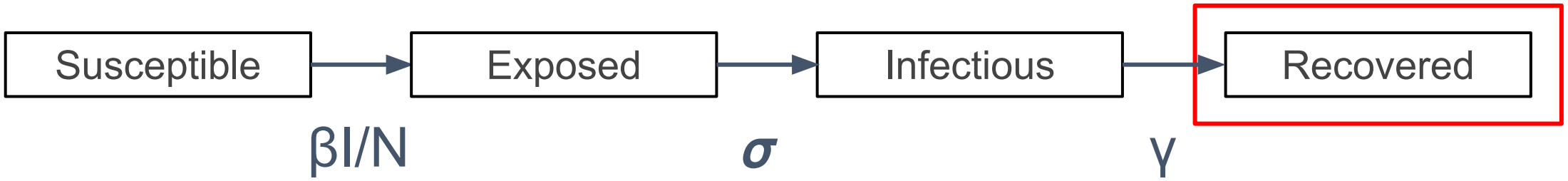


Measures of disease frequency

- Incidence of infection
- Prevalence of infection
- Seroprevalence (prevalence of antibody)



R/N



Risk vs Rate

$$\text{Incidence risk} = \frac{\text{Number of new cases in a time period}}{\text{Population at risk at the start}}$$

Cumulative incidence

Attack rate

$$\text{Incidence rate} = \frac{\text{Number of new cases}}{\text{Total person-time at risk}}$$

Incidence density rate

Risk vs Rate

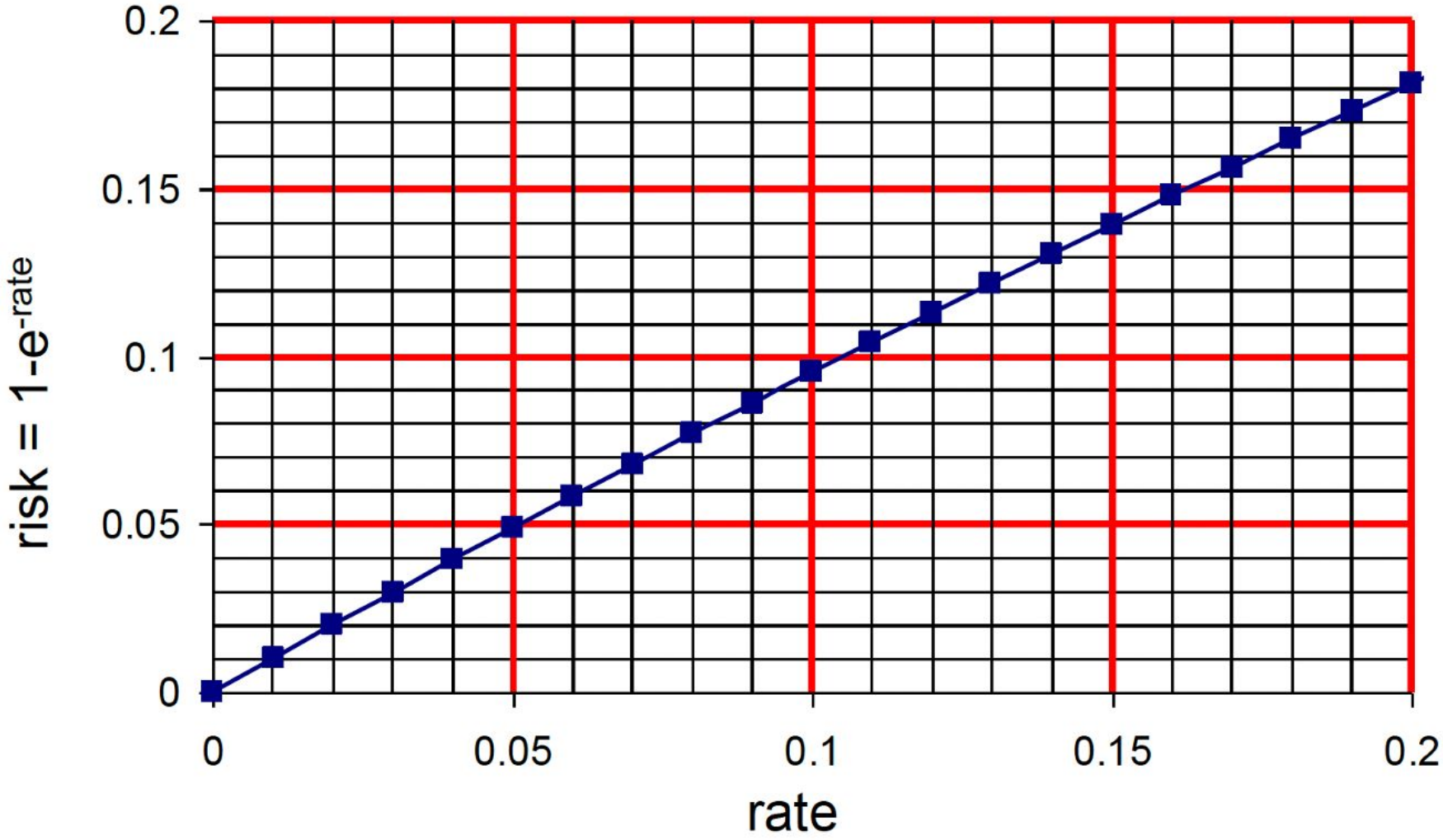
The parameters that go into **difference equations** should be **risks**

Risk and rate are related via the following expression:

$$risk = 1 - e^{-rate}$$

However, if rate is small then $e^{-rate} = 1-rate$ therefore, **risk ≈ rate**

Risk vs Rate



Rates and Average Time

The **rate** at which something occurs

= 1/{**average time** to the event}

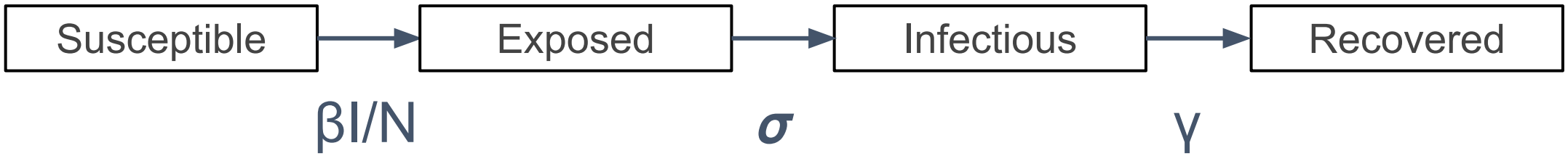
The **average time** to event

= 1/{**rate** at which event occurs}

Rates and Average Time Examples

Examples:

- The rate at which individuals recover from being infectious (γ)
  = 1/{average duration of infectiousness}
- The average duration of infected but not yet infections period
  = 1/σ



Acknowledgments

Emilia Vynnycky

Mark Jit

Learning Objectives

You should have learned:

- How does infectious disease modelling fit into the field of epidemiology
- What are the different goals of ID mechanistic models
- What are the different types of ID models
- What are common data relied upon for modelling
- How are ID data and ID models related

Steps of developing a model