Lecture 6
Inferring transmission dynamics from seroprevalence data

May 23rd, 2025

Seroanalytics Training
Blantyre, Malawi

Lecture Outline

- Understanding what age-specific patterns of seroprevalence can tell us about past transmission

- Non-endemic transmission
  - Timing and magnitude of past outbreaks

- Endemic transmission
  - Force of infection

Population transmission dynamics

How has the pathogen spread in a given population?

Population transmission dynamics

How has the pathogen spread in a given population

Endemic

- Sustained/persistent transmission
- Can have epidemic or seasonal cycles

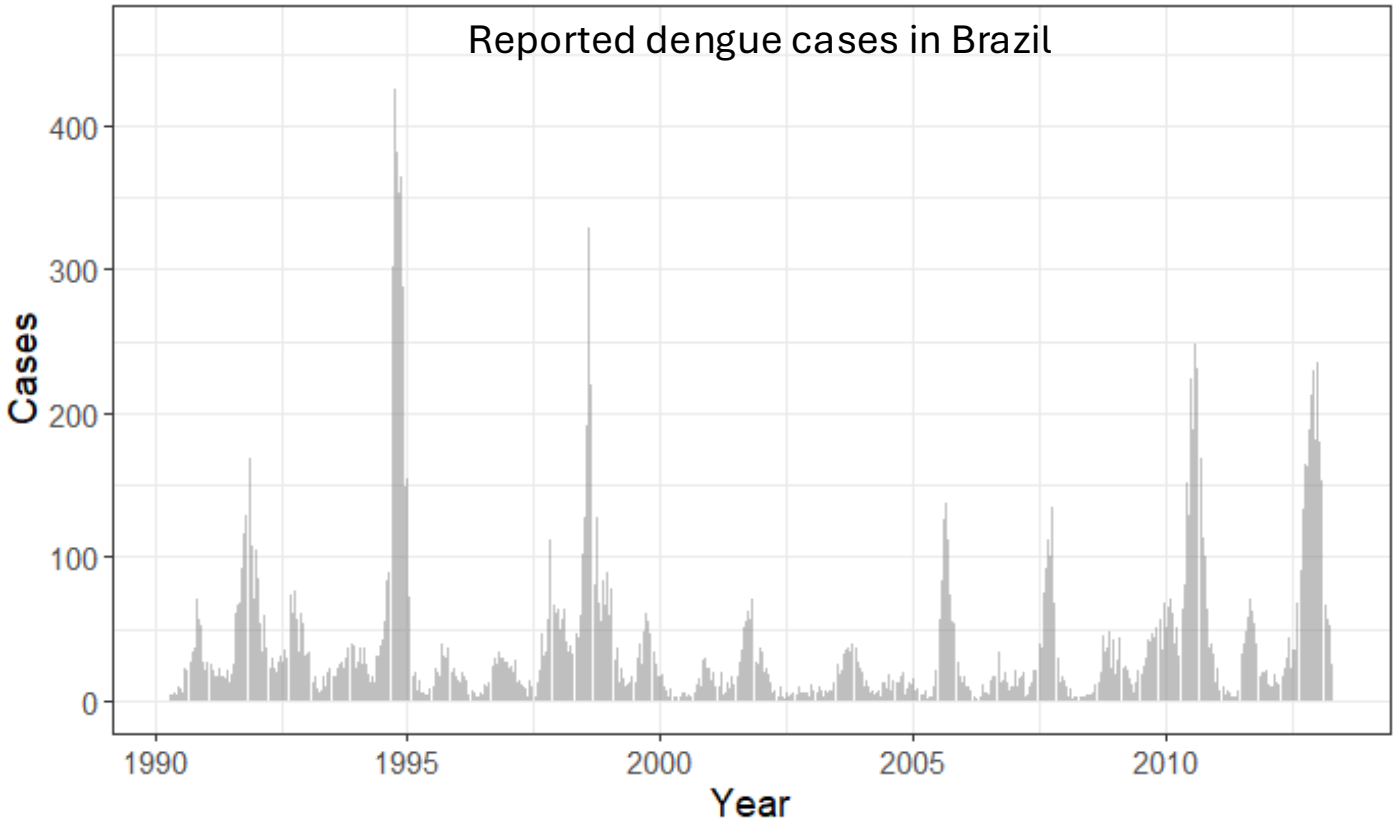
Non-Endemic

- Sporadic outbreaks
- Transmission is not continuous

Population transmission dynamics

Endemic

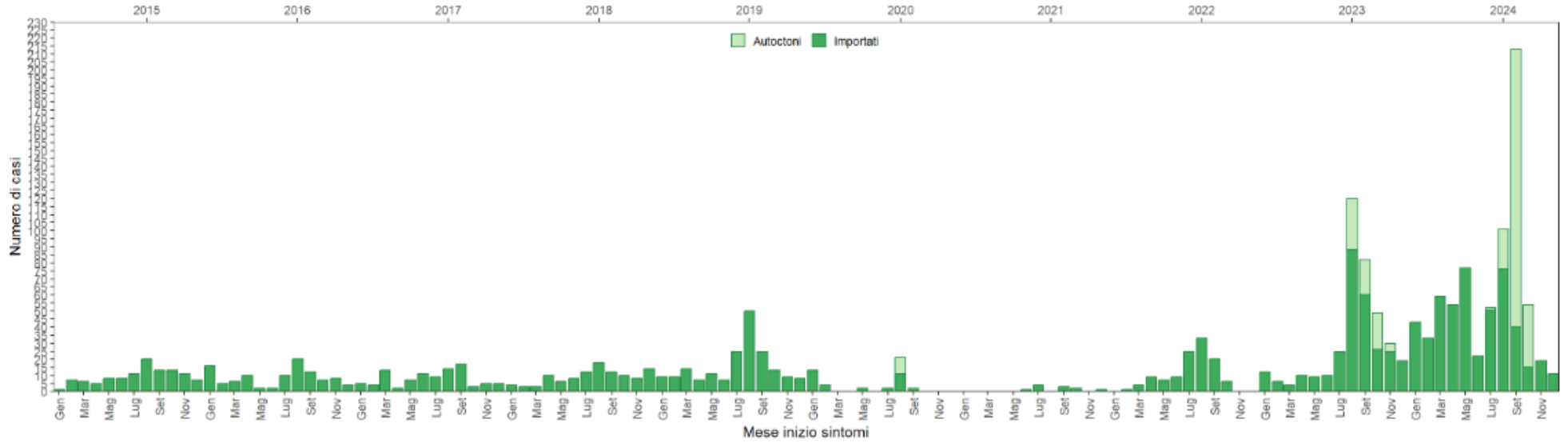
- Sustained/persistent transmission
- Can have epidemic or seasonal cycles



Population transmission dynamics

Non-Endemic

- Sporadic outbreaks
- Transmission is not continuous



Reported dengue cases in Italy

Why is this useful?

- Many infections can be missed by disease surveillance systems
  - Asymptomatic and mild infections
- To understand when in the past these infections occurred
- Quantifying annual risks of infection or average outbreak size
  - Valuable for planning control & prevention efforts

Age = time spent at risk of infection

As we get older, we are more likely to have encountered more pathogens

- Age = time spent at risk of infection
- Age is therefore a proxy for time

Age-specific seroprevalence can help tell us what has happened in the past

When is this applicable?

i.e. when can age tell us about past transmission

When seroprevalence ≈ ever infected

Pathogen:

- Must be a pathogen that induces long-term antibody response

Antibody isotype:

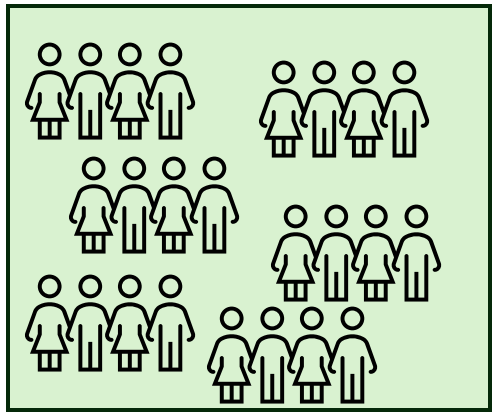
- IgG (longest-lived isotype)

Study type:

- Cross-sectional seroprevalence studies
  - One time-point, single sample per person
- Sometimes, data from cohort or other studies can be used
  - E.g. the first sample from each participant (if taken around the same timeframe)

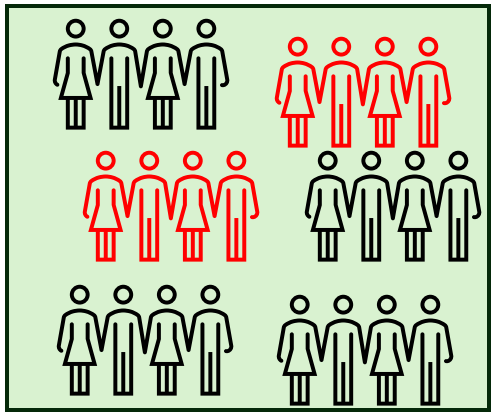
Example

<5 years



Seroprevalence=0%

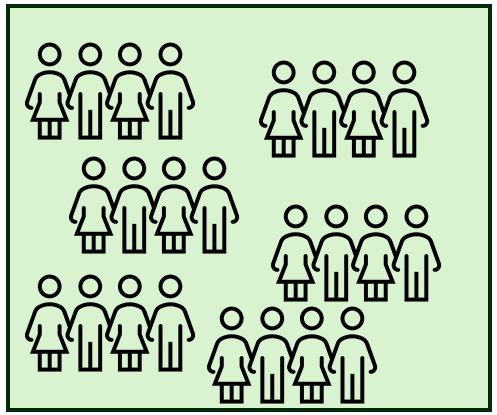
≥5 years



Seroprevalence=30%

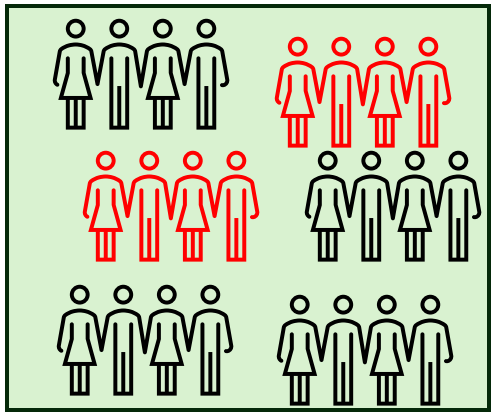
Example

<5 years old

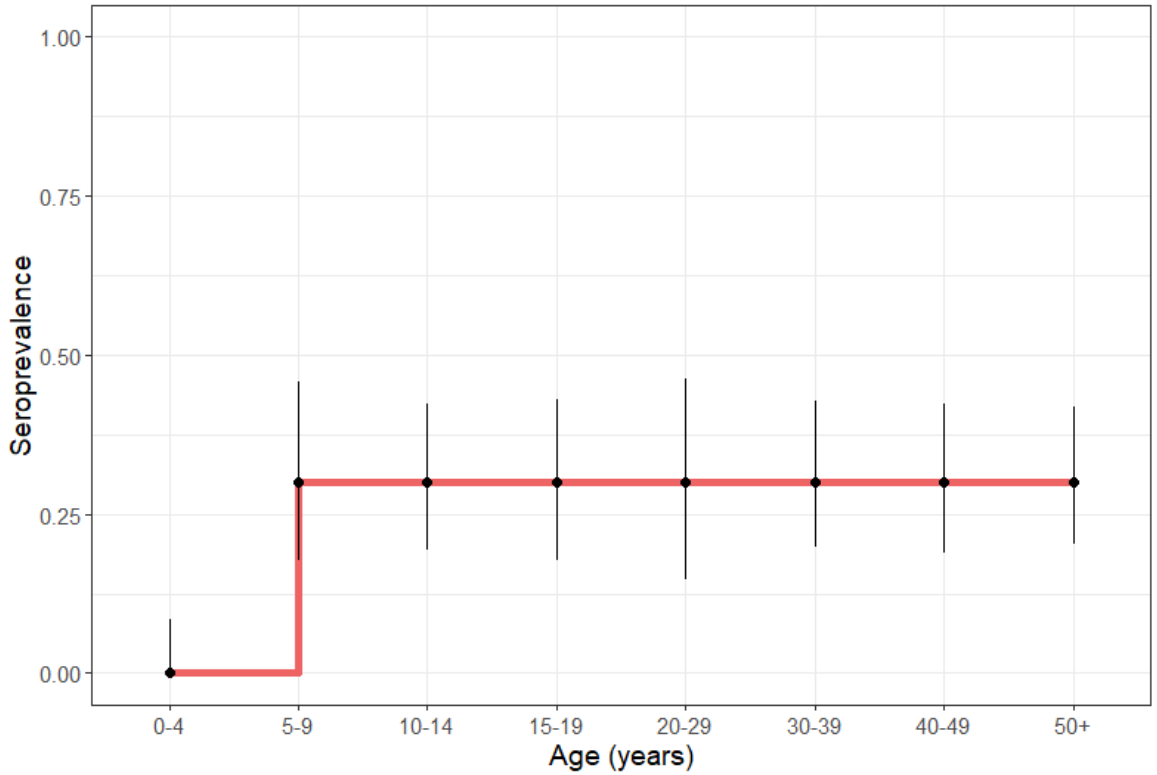


Seroprevalence=0%

≥ 5 years old



Seroprevalence=30%



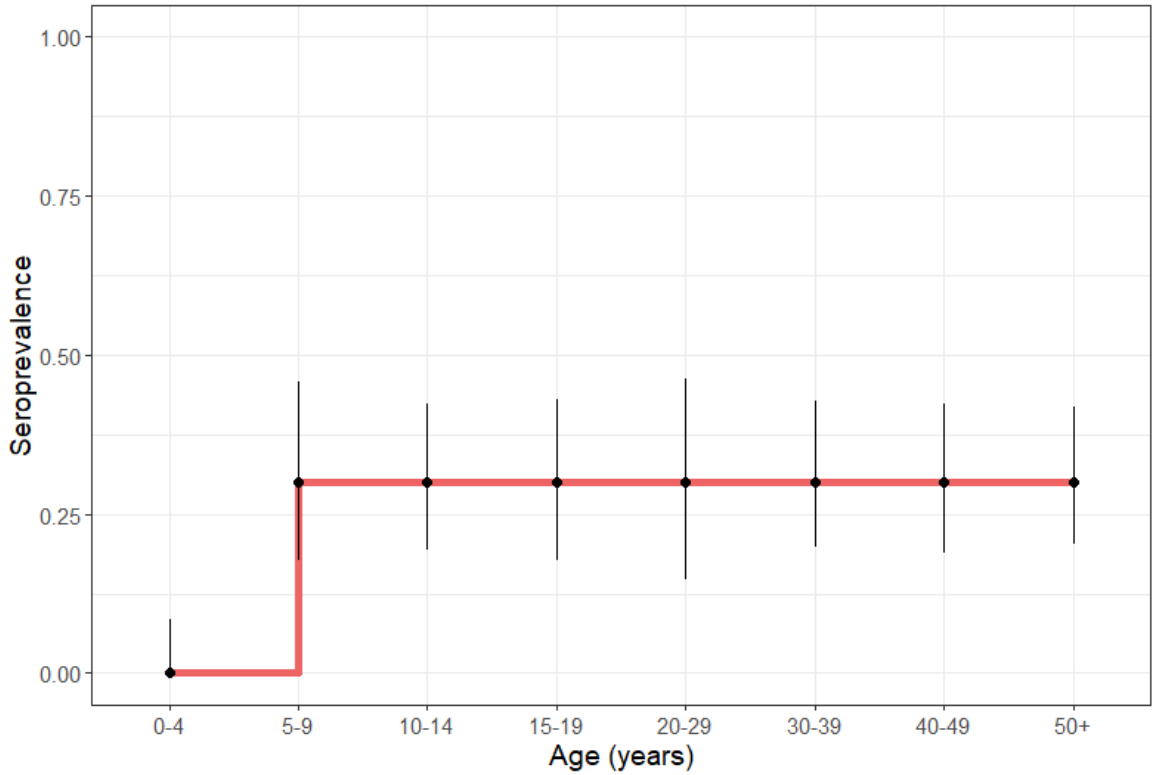
What kind of transmission has happened here?

Example

Non-endemic transmission

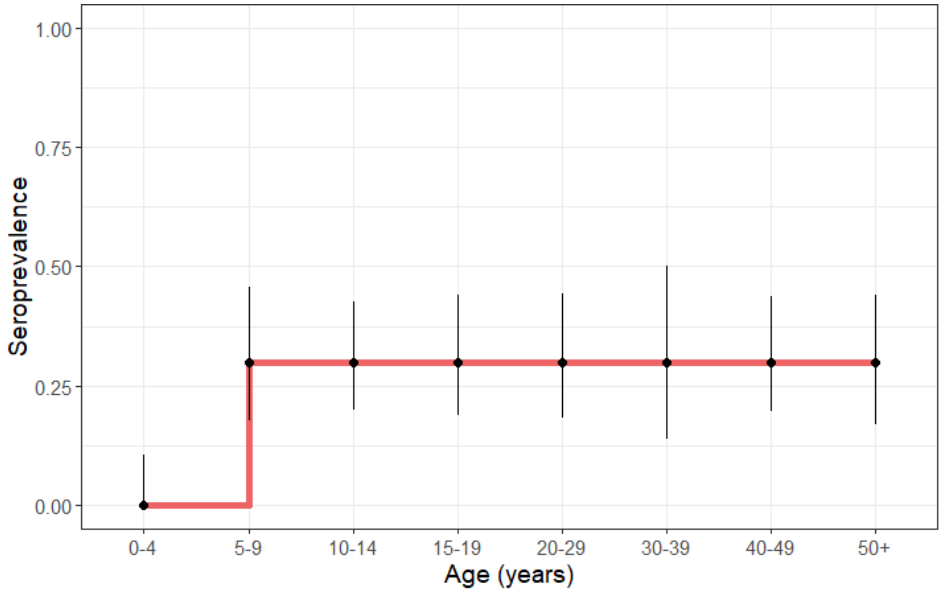
- 0% seroprevalence in children aged 0-4 means the pathogen has not transmitted in the past 4 years

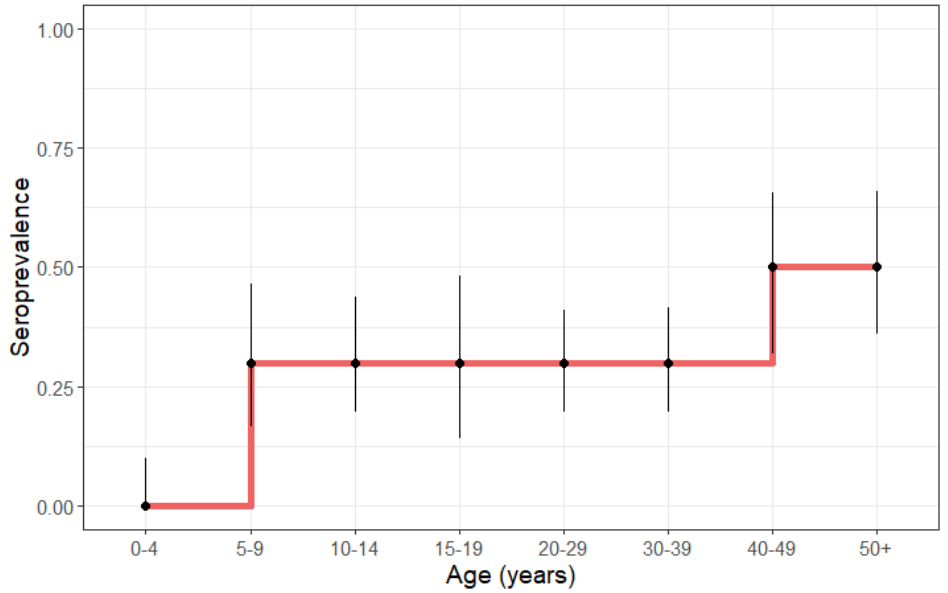
- An equal infection attack rate among those ≥ 5 years suggests:
  - An outbreak occurred ~ 5 years ago
  - 30% infection attack rate



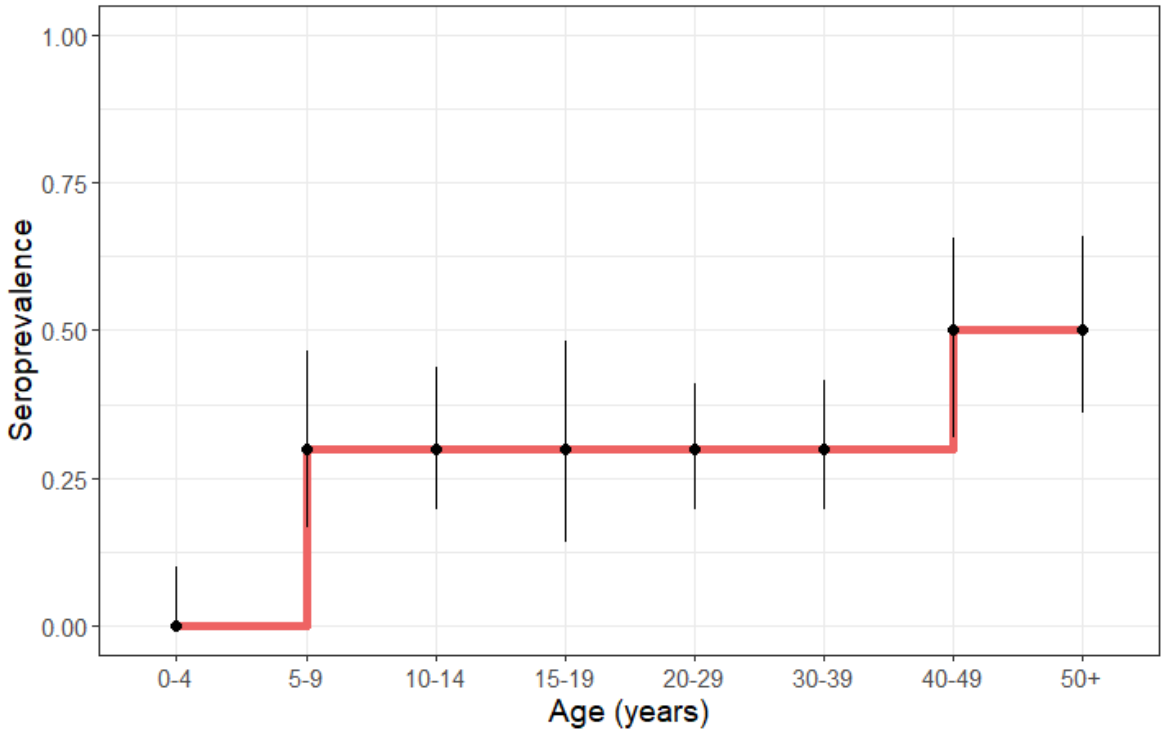
Non-endemic transmission

- Flat age-specific seroprevalence profiles
- Everyone alive during an epidemic has approximately the same infection risk
- The more epidemics an age-group has lived through, the higher the chance of infection





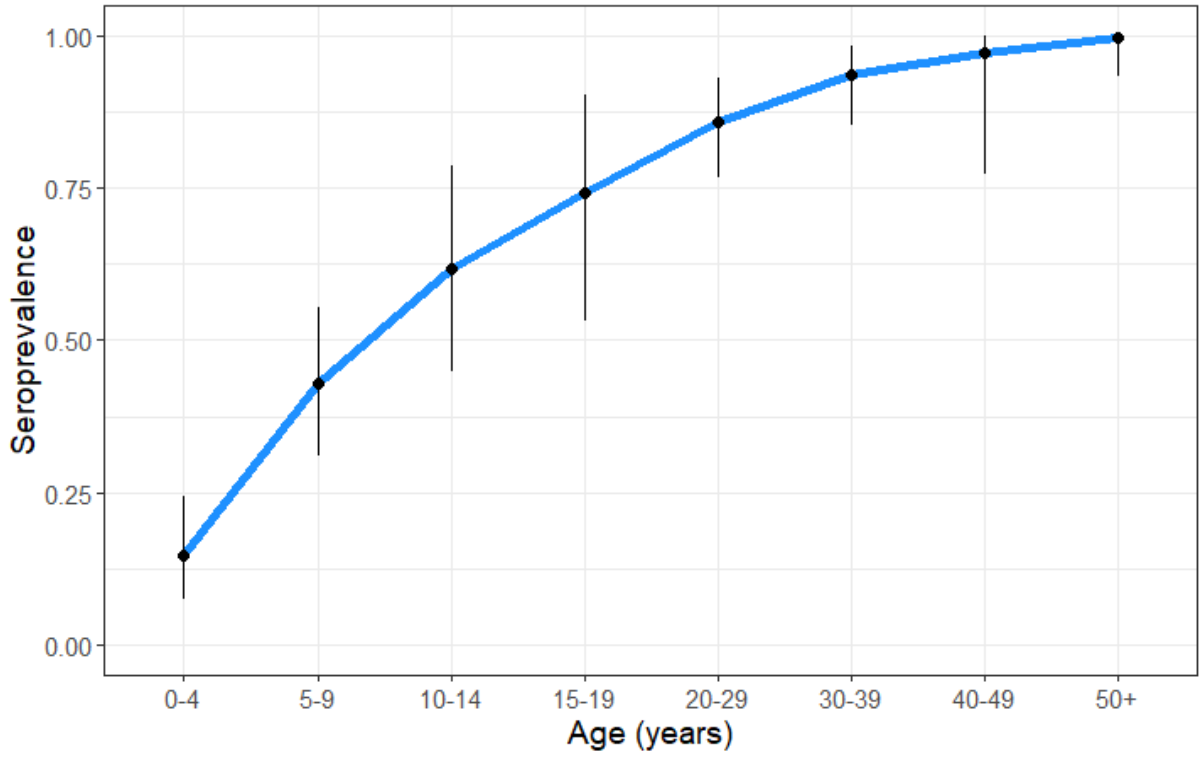
Non-endemic transmission



- Pathogen has been absent for past 4 years (0% seroprevalence in children <5)
- Participants aged 5-39 have equal attack rates
  - -> outbreak occurred ~ 5 years ago
  - -> 30% attack rate
- Participants aged ≥ 40 have equal attack rates
  - -> outbreak occurred ~40 years ago
  - -> 20% attack rate
  - These participants were alive for 2 outbreaks (20% + 30% attack rates = 50% seroprevalence)

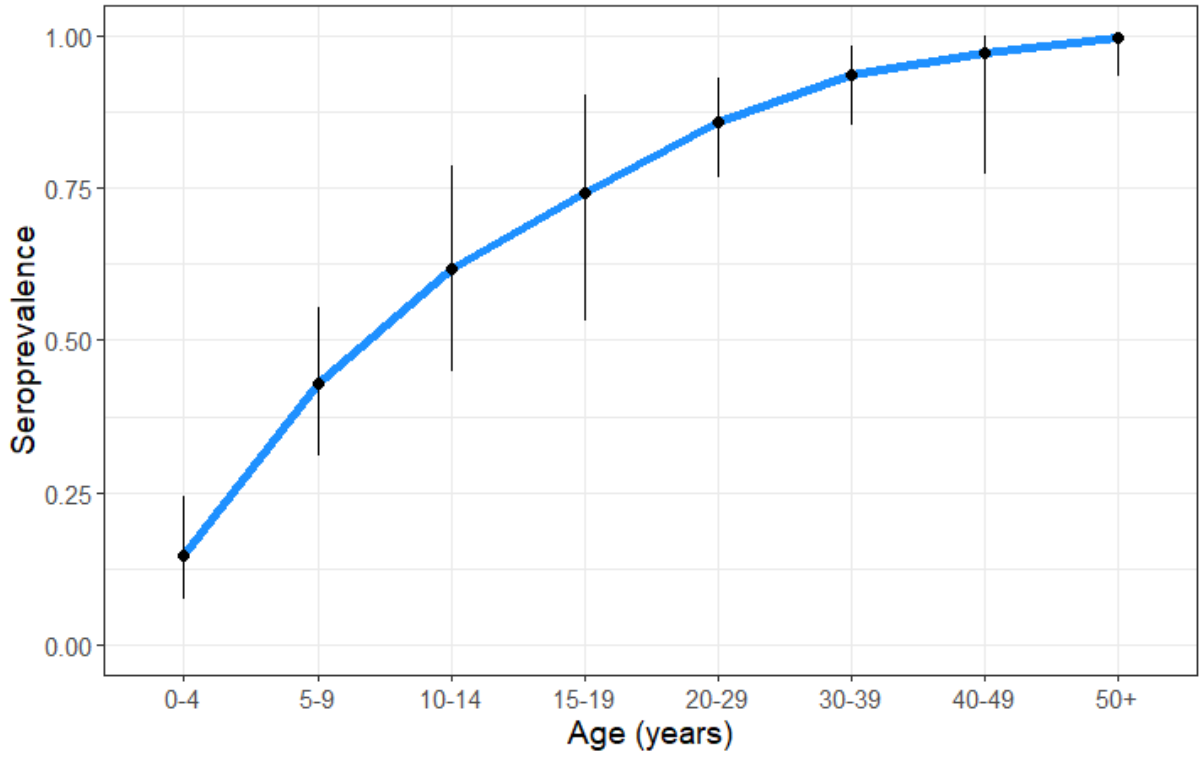
Endemic transmission

- Increasing seroprevalence with increasing age
- When a pathogen continuously transmits in a population, age is a direct measure of the time spent at risk of infection
- Valuable to quantify the force of infection (FOI)



Endemic transmission: FOI

- Force of infection (FOI) is also often called transmission intensity
- Defined as: “Rate at which susceptible individuals become infected”
- Can be estimated using a catalytic model
  - Fit to age-specific seroprevalence data



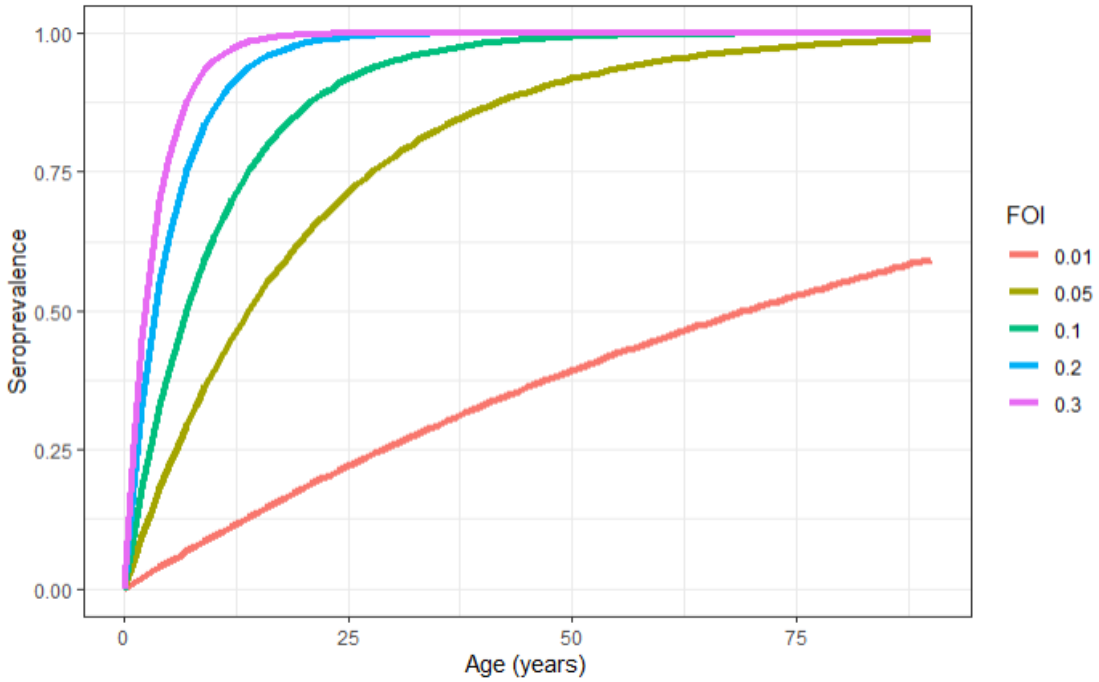
Catalytic model

$$\pi_a = 1-\exp(-\lambda * a)$$

- λ = FOI
- π_a = proportion positive at age a

- Assumptions:
  - Transmission is constant over time
  - Infection risk is independent of age

- The cumulative risk of ever having been infected accumulates with age



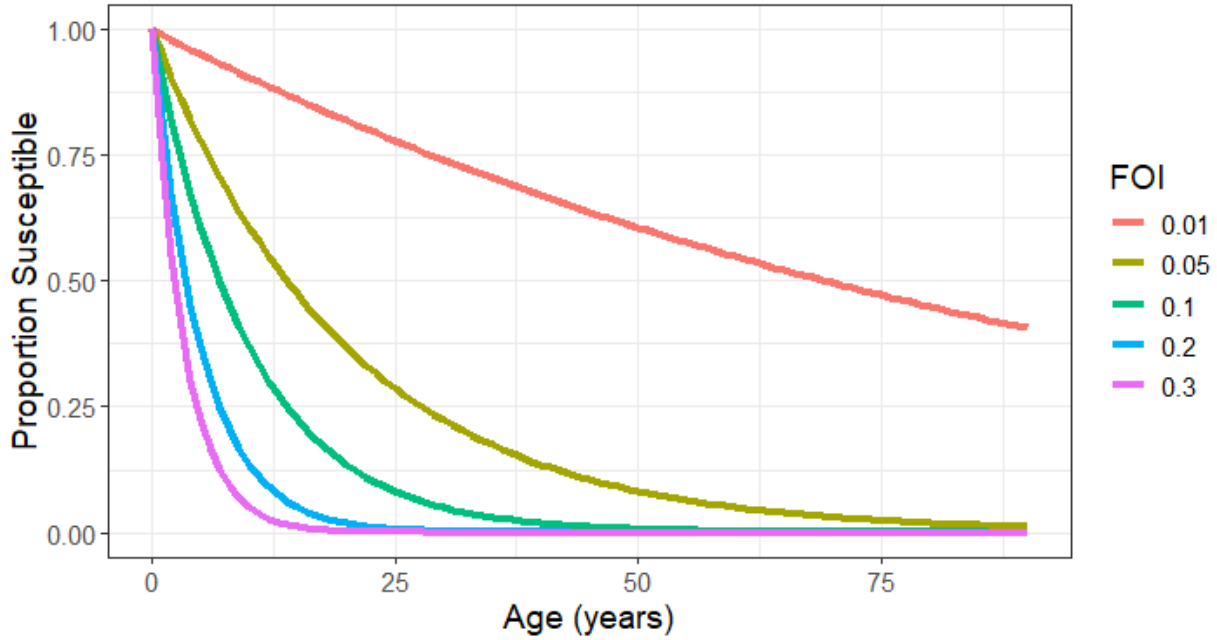
Catalytic model

$$\pi_a = 1-\exp(-\lambda * a)$$

- λ = FOI
- π_a = proportion positive at age a
- $Susceptible_a = 1 - \pi_a$

- Assumptions:
  - Transmission is constant over time
  - Infection risk is independent of age

- We will implement this using a simple regression model
  - glm function
  - complementary log-log link



Value of FOI

- Allows us to quantify the level of transmission
  - number of annual infections
- Prioritization of vaccines and other control strategies

SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

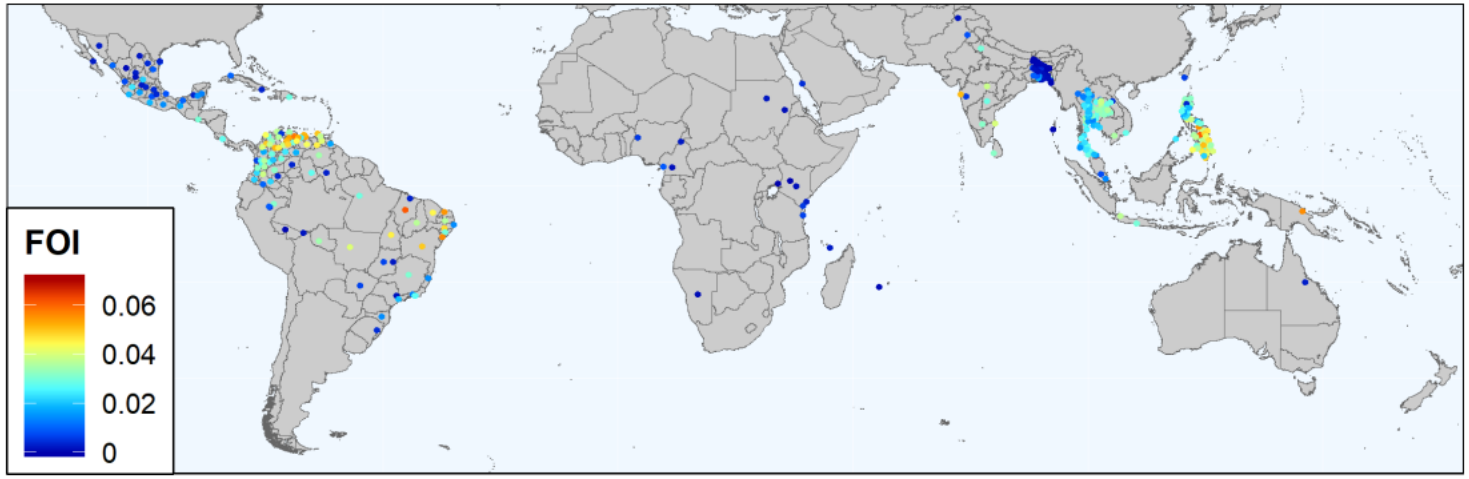
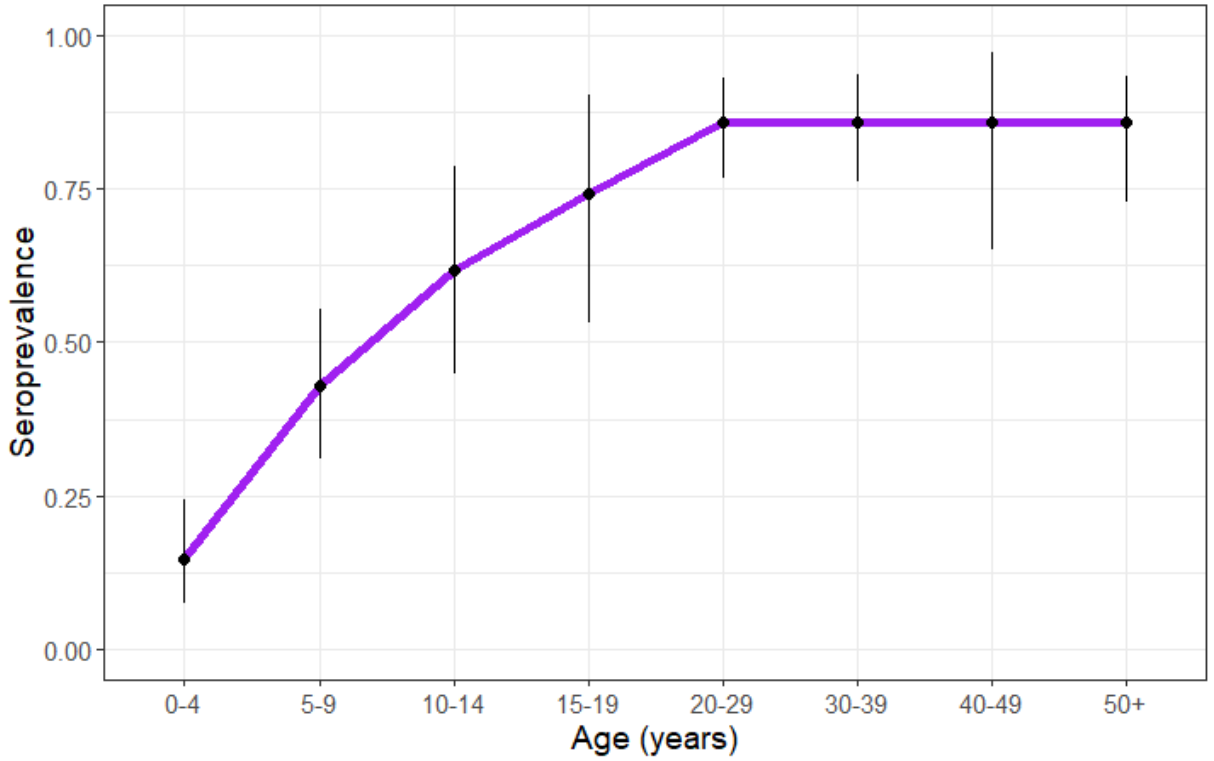


Fig. 1. Geographic location of FOI data. A total of 382 data points are shown comprising the national administrative division of all countries and the second national administrative division (admin 2) of Mexico, Colombia, Venezuela, Brazil, India, and Australia. Color scale shows FOI value (average per-serotype).

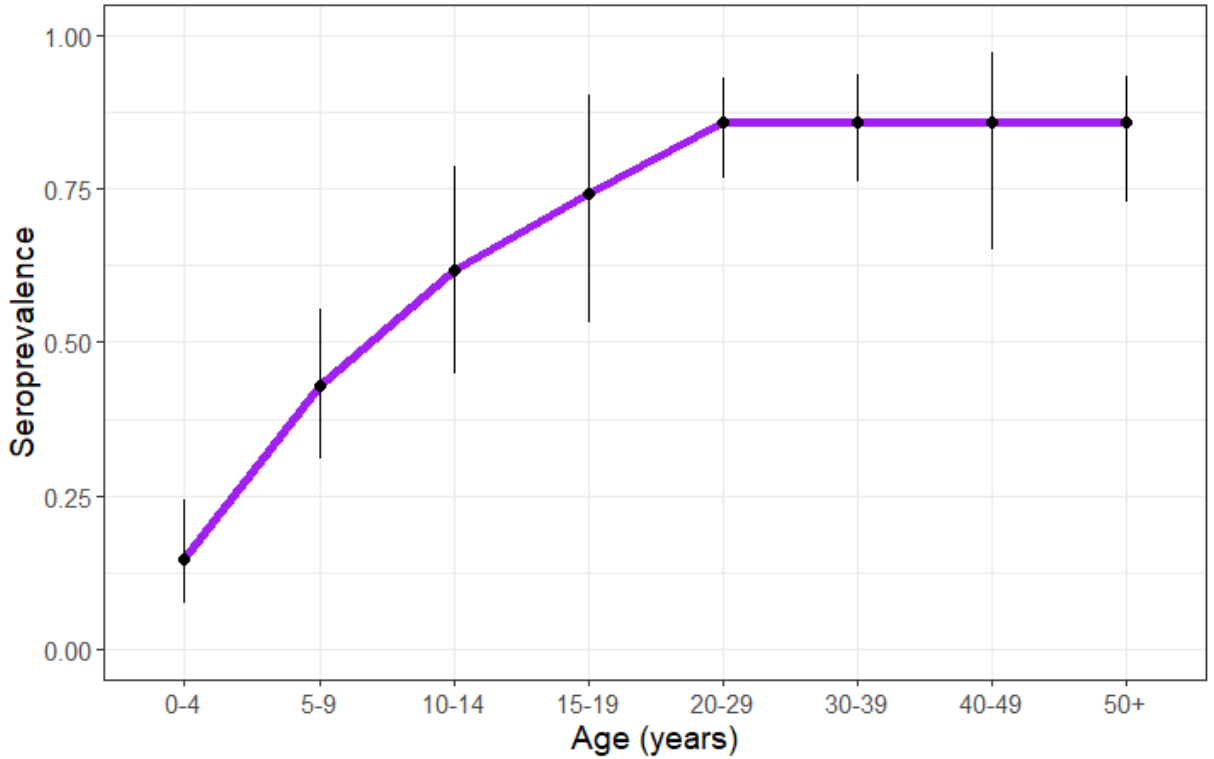
Catterino et al., Science Translational Medicine, 2020

?



What kind of transmission has happened here?

Recently established endemic transmission

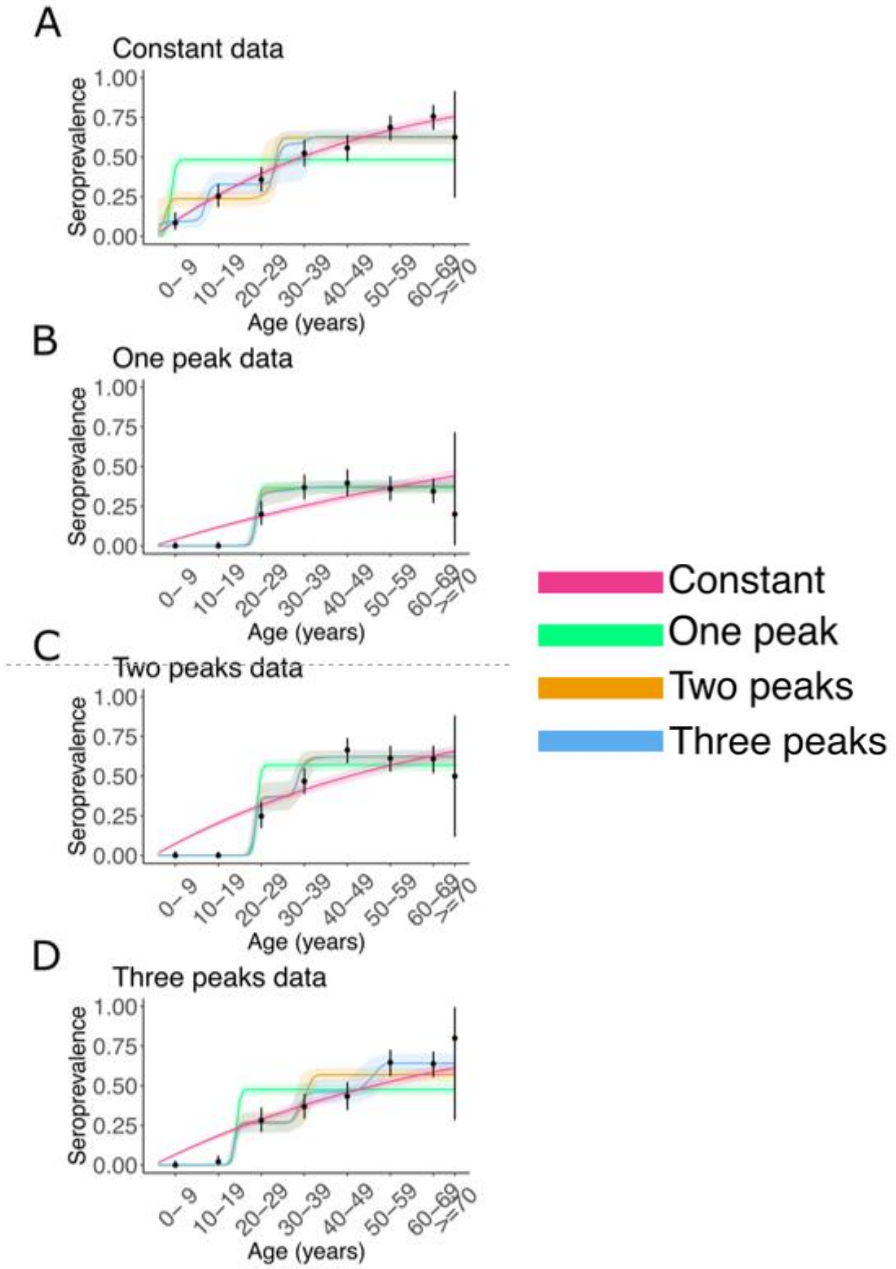


- Pathogen began circulating ~20 years ago
  - Everyone >20 years has same attack rate
- Increasing seroprevalence with increasing age in younger age groups
  - -> recent endemic transmission

RSero: R package

RSero: A user-friendly R package to reconstruct pathogen circulation history from seroprevalence studies

Nathanaël Hozé, Margarita Pons-Salort, C. Jessica E. Metcalf, Michael White, Henrik Salje, Simon Cauchemez



Additional resources

- RSero package for reconstructing past transmission
  - https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1012777

- Seroincidence estimation (pathogens that induce only short-lived antibody responses)
  - Cholera: https://pmc.ncbi.nlm.nih.gov/articles/PMC7738617/
  - Typhoid: https://pubmed.ncbi.nlm.nih.gov/35750069/

- Reversible catalytic model (when antibody responses wane significantly over time)
  - Malaria: https://www.pnas.org/doi/10.1073/pnas.0408725102
  - Enteropathogens: https://elifesciences.org/articles/45594#s4

- Joint fitting of mixture and catalytic models
  - https://journals.sagepub.com/doi/10.1177/1471082X12457495
  - https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0010592

Conclusions

- Age-specific seroprevalence data can help us reconstruct past transmission dynamics
    - Seroprevalence that continuously increases with age indicates endemicity
    - Staggered, step-wise increases in seroprevalence indicate sporadic epidemics

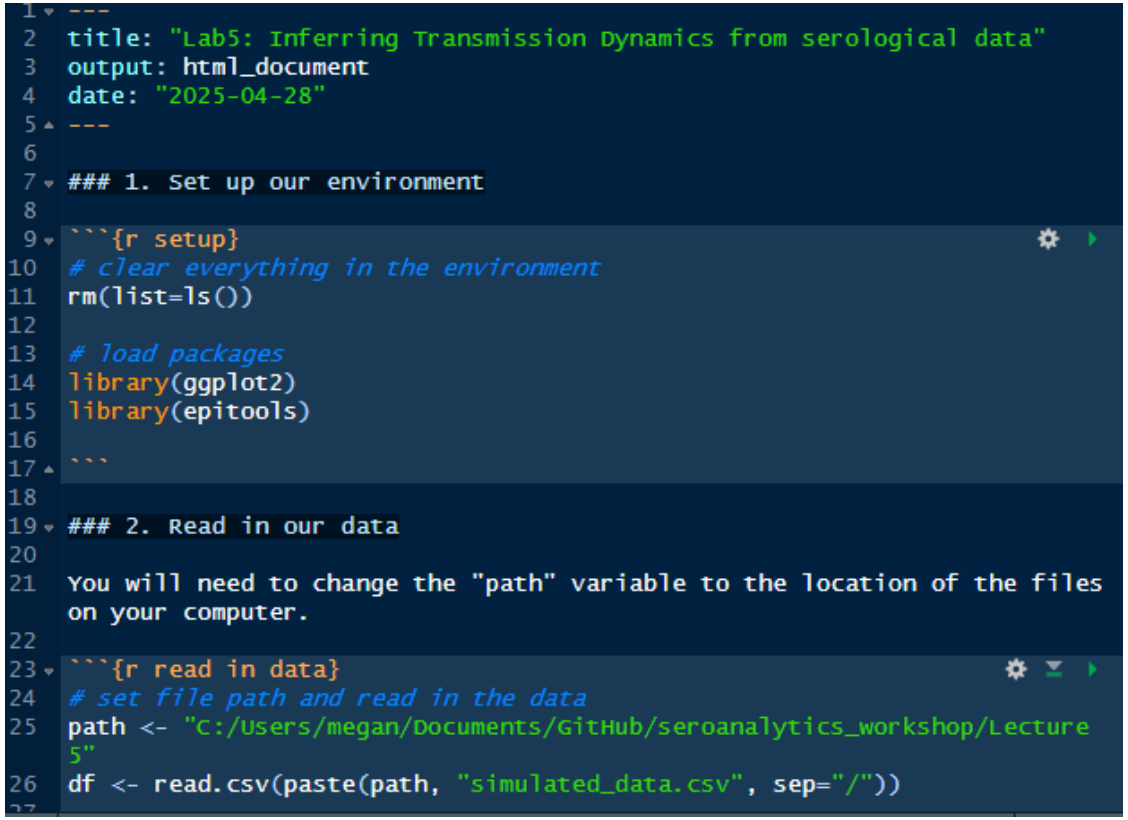
- Understanding if a pathogen is endemic or not has important public health implications

- Past infection rates can be quantified
    - Non-endemic settings: frequency & magnitude of outbreaks
    - Endemic settings: force of infection

email: ***megan.odriscoll@hug.ch***

Tutorial/Lab

- Calculate age-specific seroprevalence
- Investigate the transmission dynamics of two pathogens
- Fit a catalytic model to estimate FOI



```
---
title: "Lab5: Inferring Transmission Dynamics from serological data"
output: html_document
date: "2025-04-28"
---

### 1. Set up our environment

```{r setup}
# clear everything in the environment
rm(list=ls())

# load packages
library(ggplot2)
library(epitools)

```

2. Read in our data

You will need to change the "path" variable to the location of the files on your computer.

```{r read in data}
# set file path and read in the data
path <- "C:/Users/megan/Documents/GitHub/seroanalytics_workshop/Lecture 5"
df <- read.csv(paste(path, "simulated_data.csv", sep="/"))
```
```