Block 2.1: Model fitting approaches and challenges

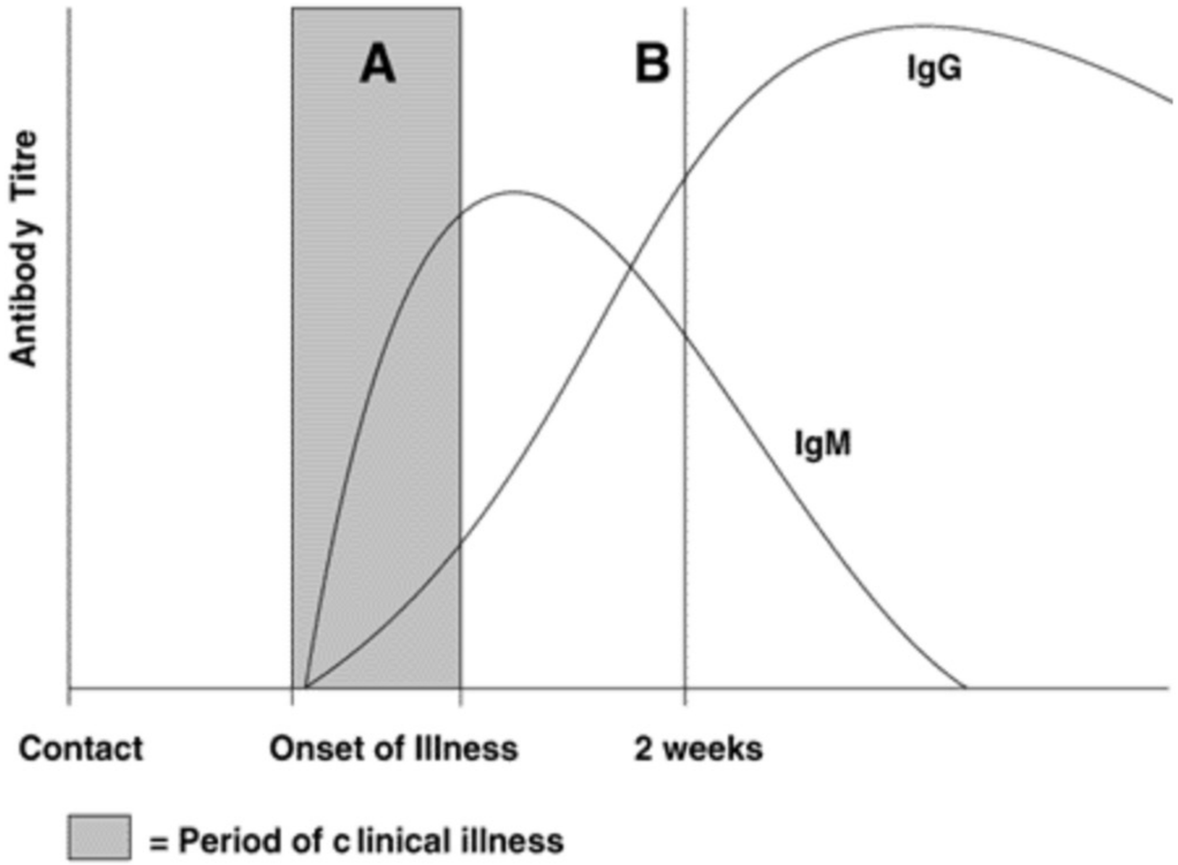
Fitting Catalytic Model to Serological Data using Least Squares

Objectives

By the end of this session you should:

- Know the equation for the catalytic model
- Be able to fit the catalytic model to data using least squares
- Understand the assumptions of the catalytic model
- Know one approach to estimate the force of infection and R_0
- Develop some intuition at interpreting cross-sectional serological data

Schematic of individual antibodies over time



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.

Cross-sectional binary serological data

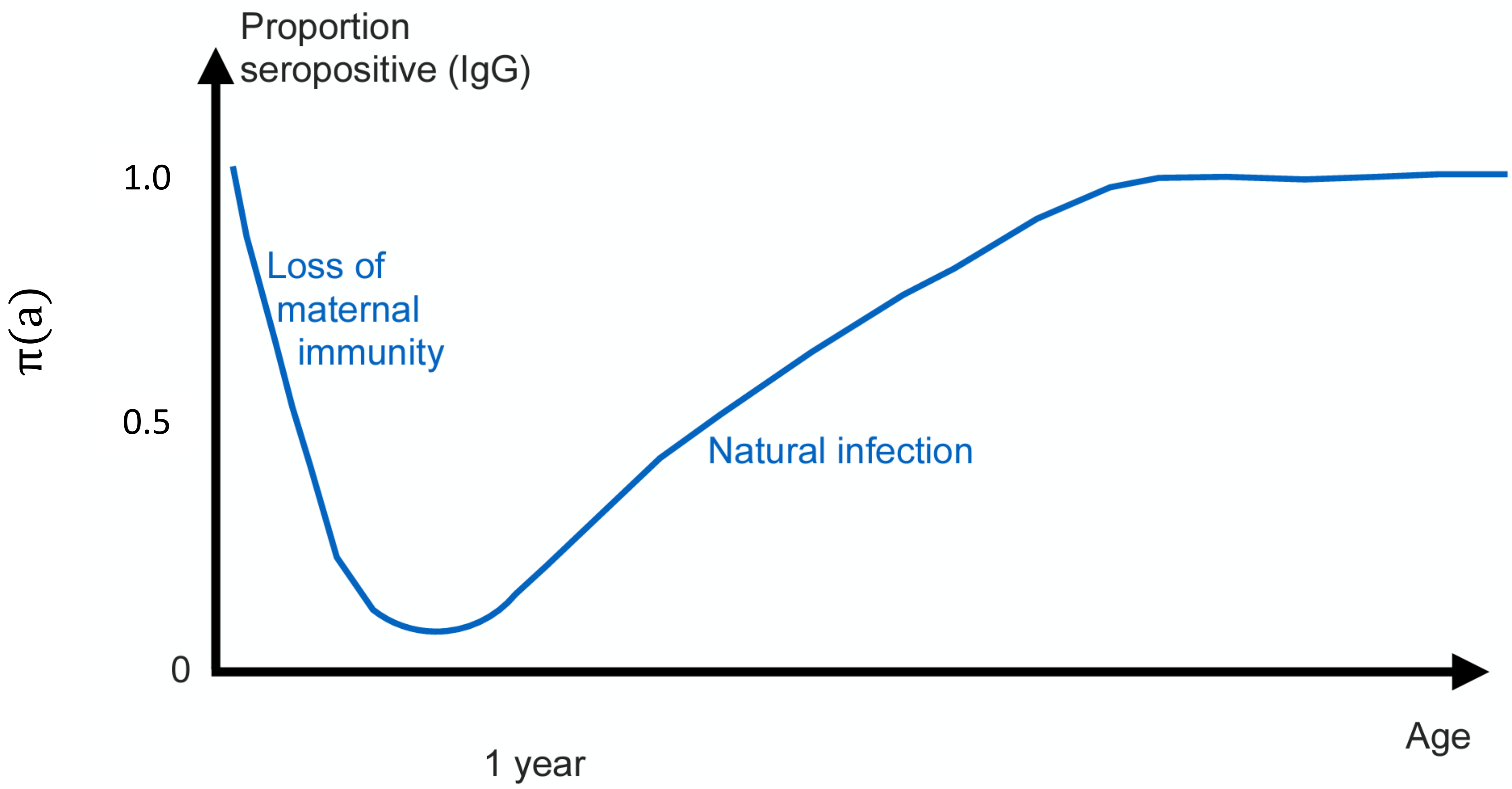
Assuming

1. We take a cross-sectional (one time point) sample from a population
   - Z_i, i = 1, . . . ,N, representing the IgG antibody titers (i.e., levels) for the i^{th} individual in the sample
2. And, take a binary classification of IgG antibody titers
   - Y_i, i = 1, . . . ,N, be an indicator variable representing the disease status for the i^{th} individual in the sample

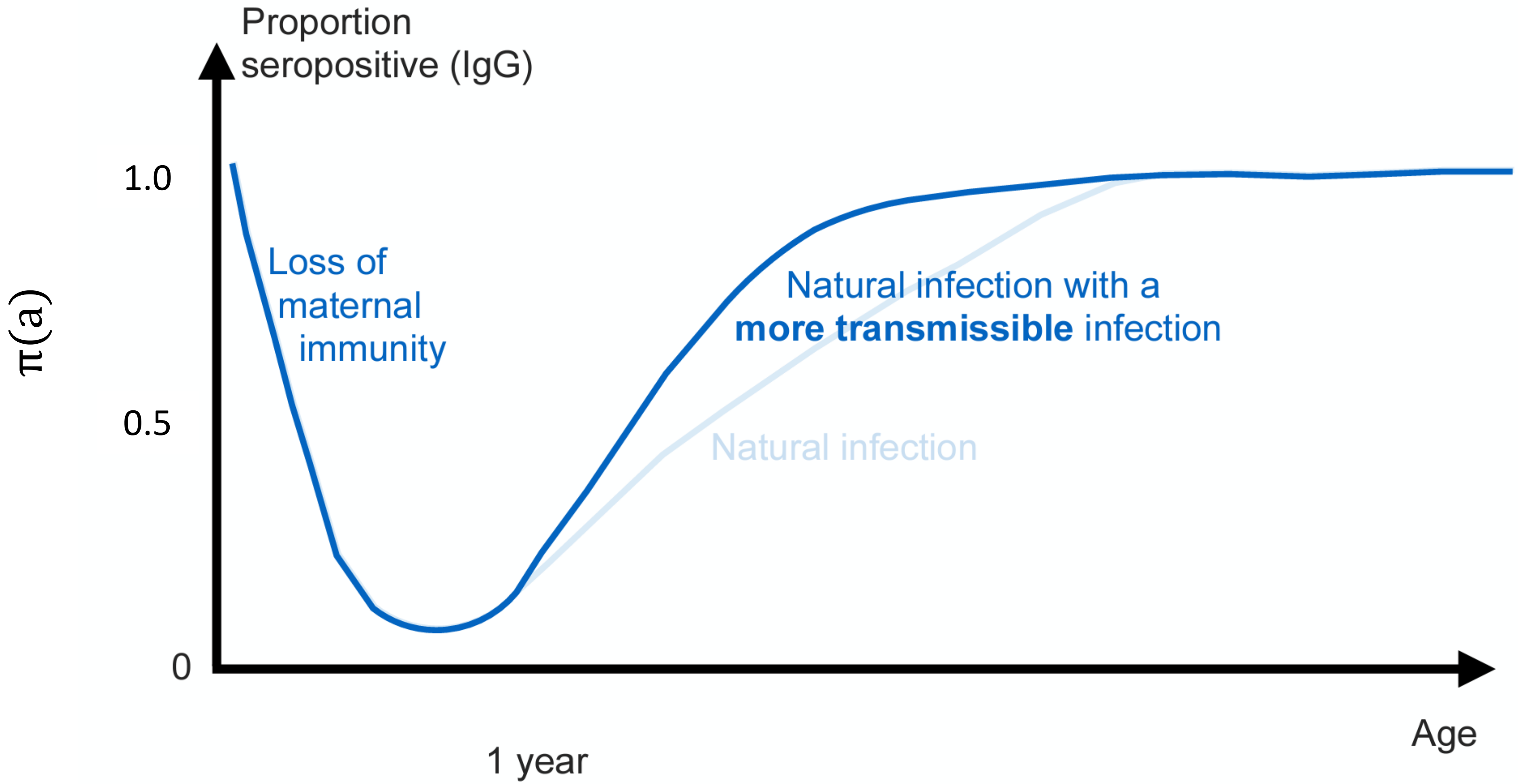
$$Y_i = \begin{cases} 1 & \text{if} \quad Z_i > \tau_u, \\ 0 & \text{if} \quad Z_i < \tau_\ell, \end{cases} \qquad Y_i = \begin{cases} 1 & \text{when seropositive (previously infected)}, \\ 0 & \text{when seronegative (susceptible to infection)}. \end{cases}$$

Lets calculate the proportion seropositive, as a function of corresponding one year age categories, i.e., $\pi(a)$

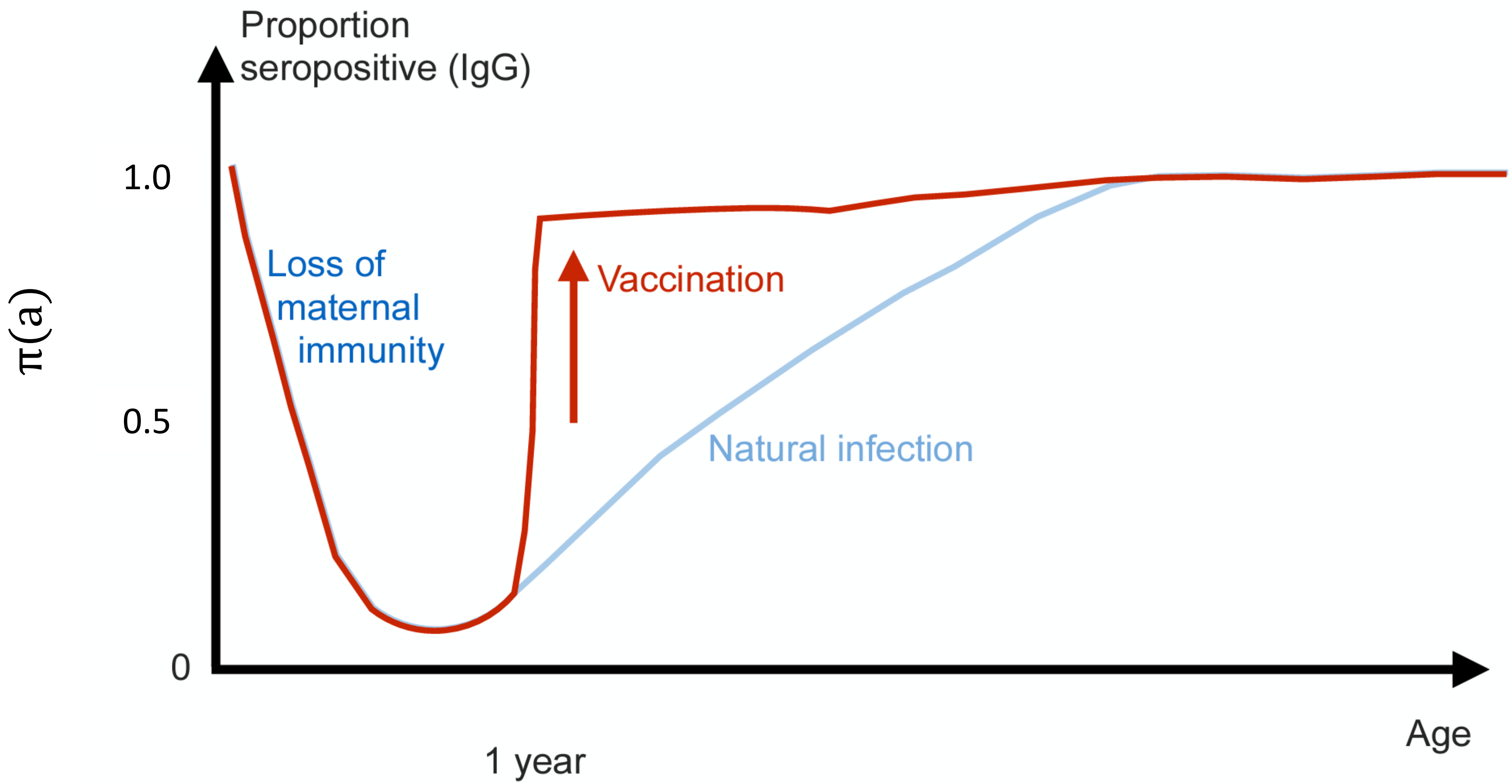
IgG serology and dynamical insights



IgG serology and dynamical insights



IgG serology and dynamical insights

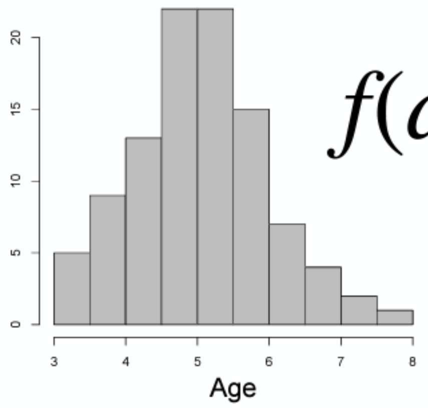


From time to age

$\lambda = \beta I$: force of infection, or rate at which susceptible individuals become infected

$A = 1/\lambda$:Average age of infection is the inverse of this rate

Age distribution of cases (IgM)



$f(a)$

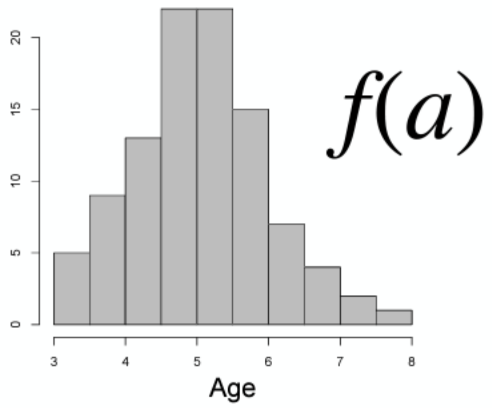
Fully immunizing or
SIR dynamical system
like measles and rubella

From time to age

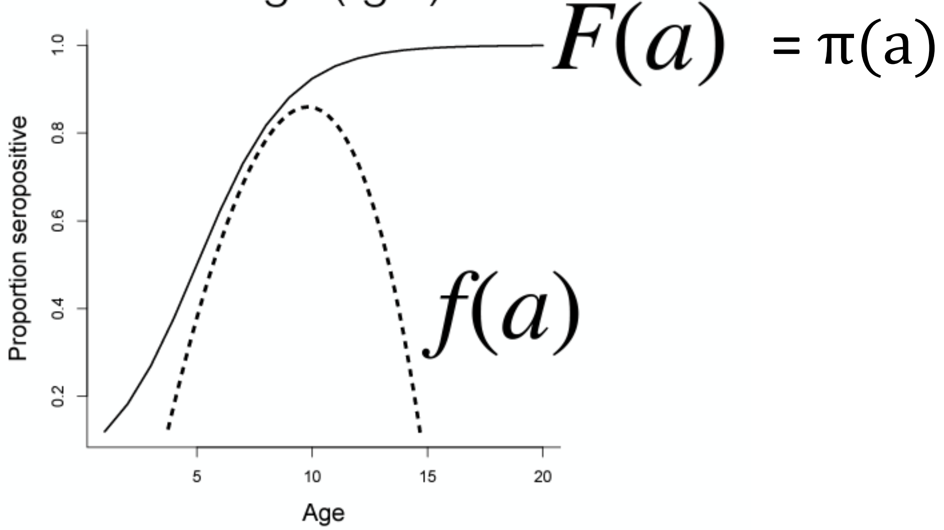
$\lambda = \beta I$: force of infection, or rate at which susceptible individuals become infected

$A = 1/\lambda$:Average age of infection is the inverse of this rate

Age distribution of cases (IgM)



Proportion seropositive across age (IgG)



Fully immunizing or SIR dynamical system like measles and rubella

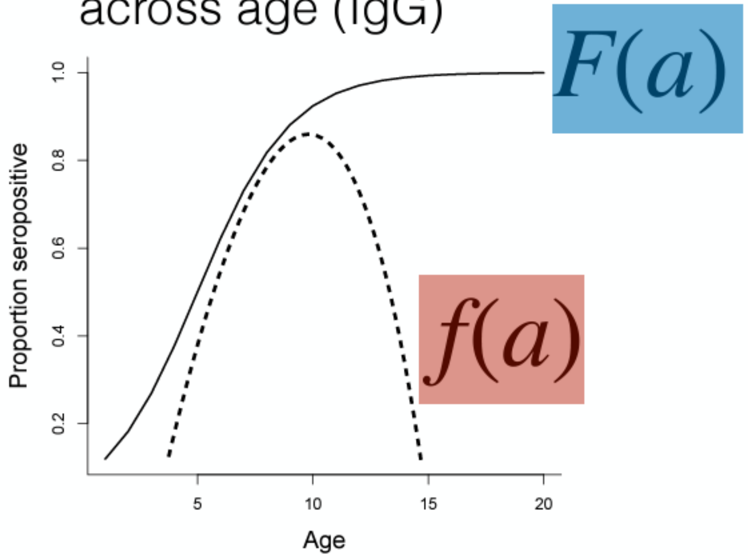
Force of infection from serology

To become seropositive at age a, you have to have avoided being seropositive up to that age, and then become infected

$$f(a) = (1 - F(a))\lambda(a)$$

$$I = S\lambda$$

Proportion seropositive across age (IgG)



Force of infection from serology

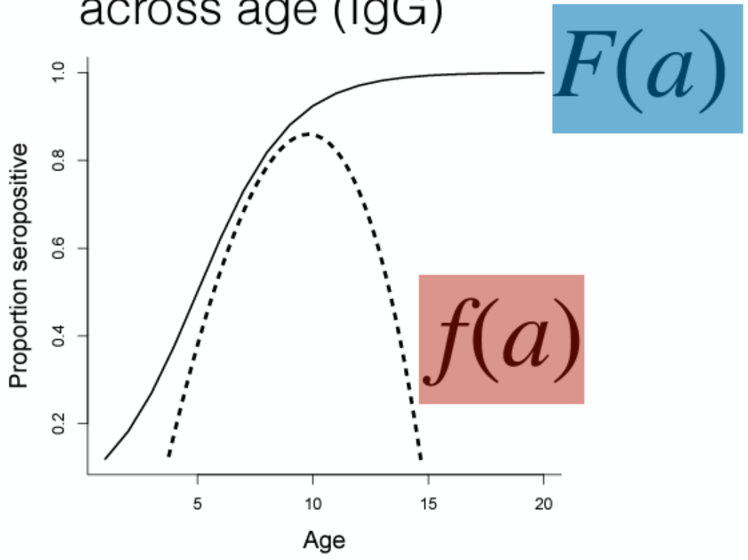
To become seropositive at age a, you have to have avoided being seropositive up to that age, and then become infected

$$f(a) = (1 - F(a))\lambda(a)$$

Dividing proportion seronegative by the derivative of the proportion seropositive yields the force of infection.

We can leverage this relationship with a **range of parametric forms** fitted to serology data.

Proportion seropositive across age (IgG)



Catalytic model

Assuming endemic equilibrium in which the **force of infection is constant over time** (and age) we can rely on the catalytic model, Muench (1959).

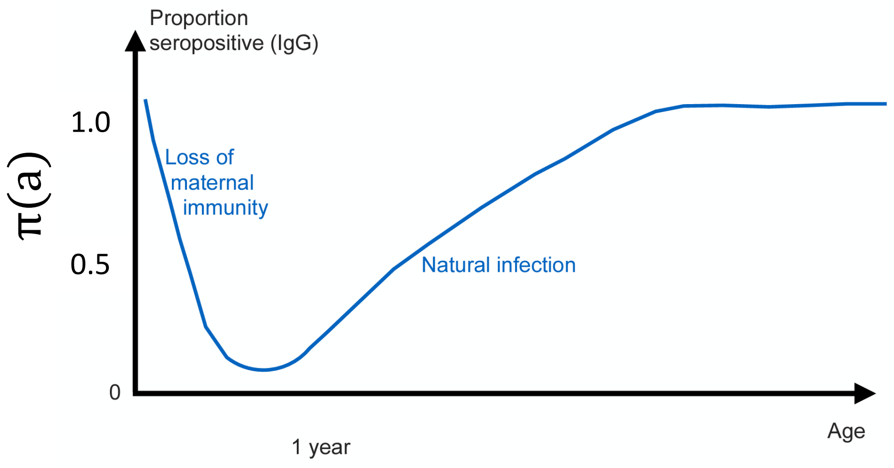
The "**catalytic model**" was first developed in the field of chemistry to understand how molecules convert over time.

In the field of sero-epidemiology it was developed to model the infection process for a Susceptible-Infected-Recovered model where susceptible movement to the infected class is exponential with rate λ:

$$\pi(a) = 1 - \exp(-\lambda a)$$

the proportion of seropositive individuals by age a corresponds directly to the cumulative proportion of infected individuals

Lets fit a catalytic model

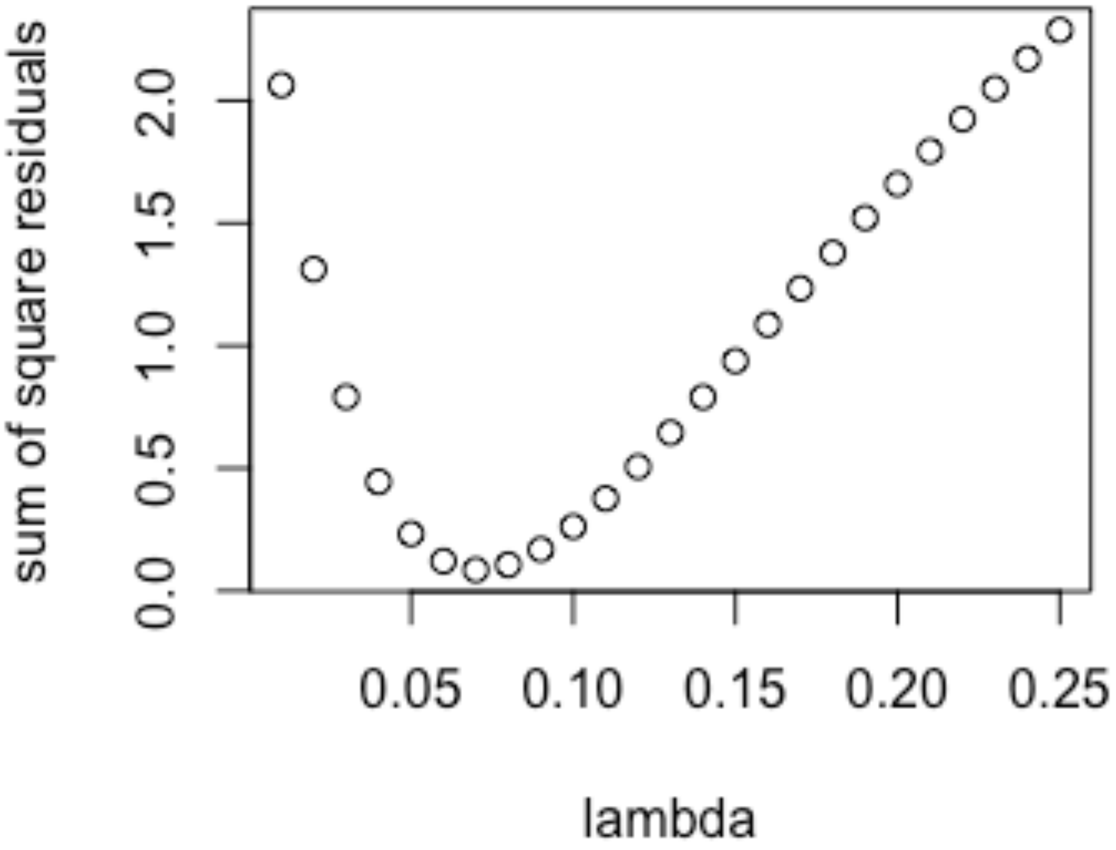


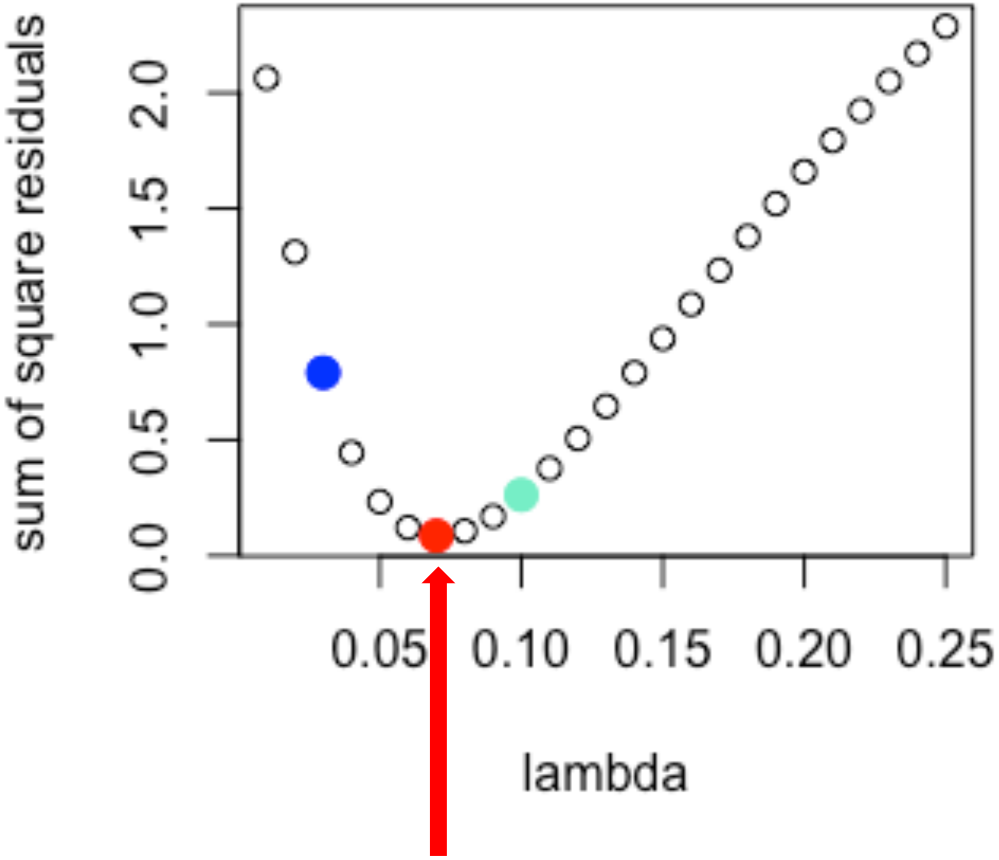
The Data:

- Rubella serological data
- Assam 2018 among 9 months – 15 years old
  - Why did I drop individuals <9 months old?
- Cross-sectional sample
- Tested using Euroimmun EIA test for rubella-specific IgG antibodies

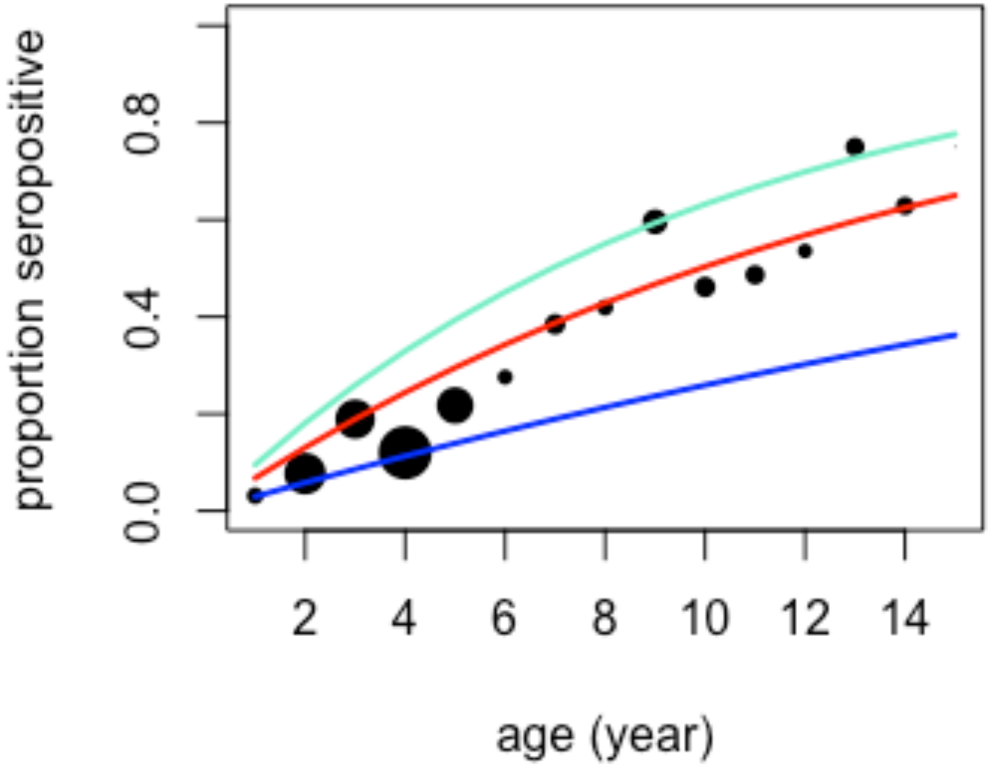
$$\pi(a) = 1 - \exp(-\lambda a)$$

Residual based on range of lambda





$\lambda = 0.07$



Points represent data
Lines represent model fit

Estimating R_0 and Average Age of Infection

In addition to the assumptions of the catalytic model, if you also assume a population has type 1 mortality (i.e., everyone survives to age e_0 and then dies), one can estimate:

1. The basic reproduction number, $R_0 = \lambda e_0$, where e_0 is life expectancy at birth. In the more general case of a growing population, e_0 can be replaced with the reciprocal of the per capita birth rate, $R_0 \approx \lambda$G, in which G = 1/CBR = N/B, where N is the population size and B are the total number of births
2. The average age of infection, A = 1/ λ

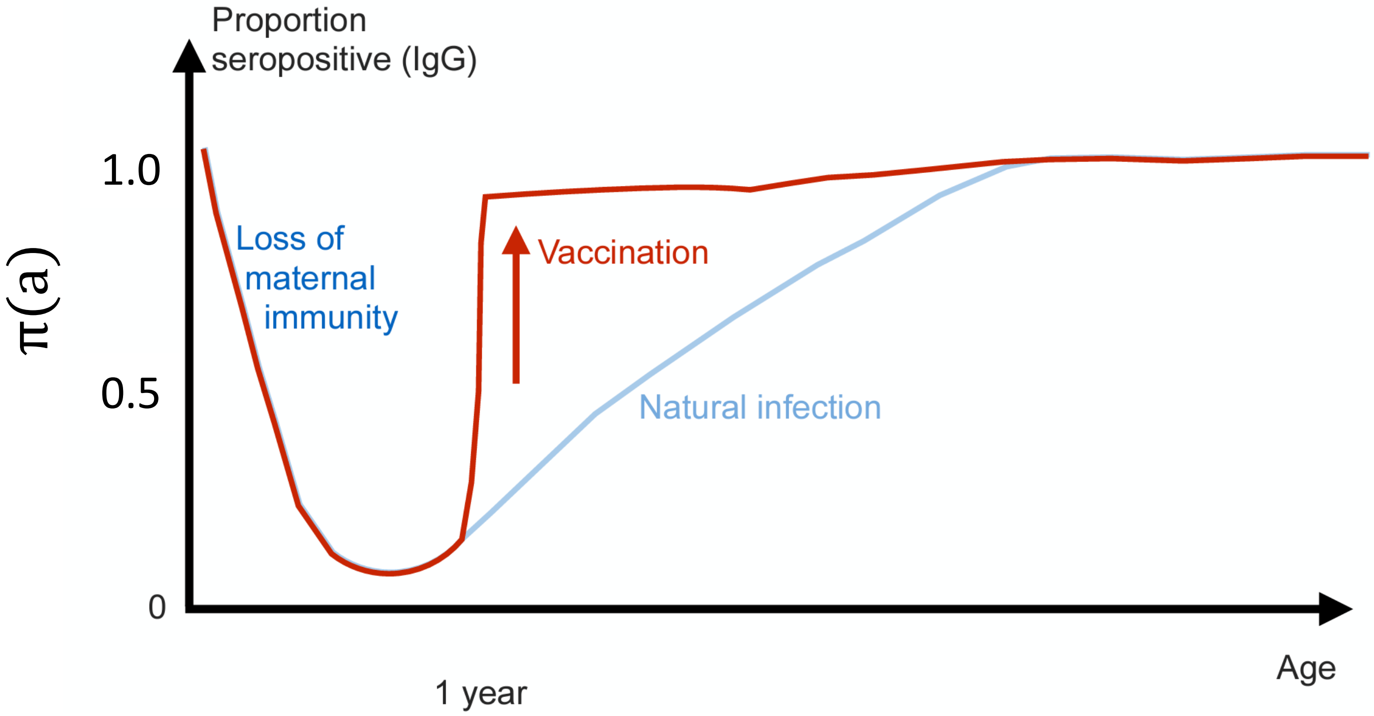
You try

Estimate R_0 and Average age of infection

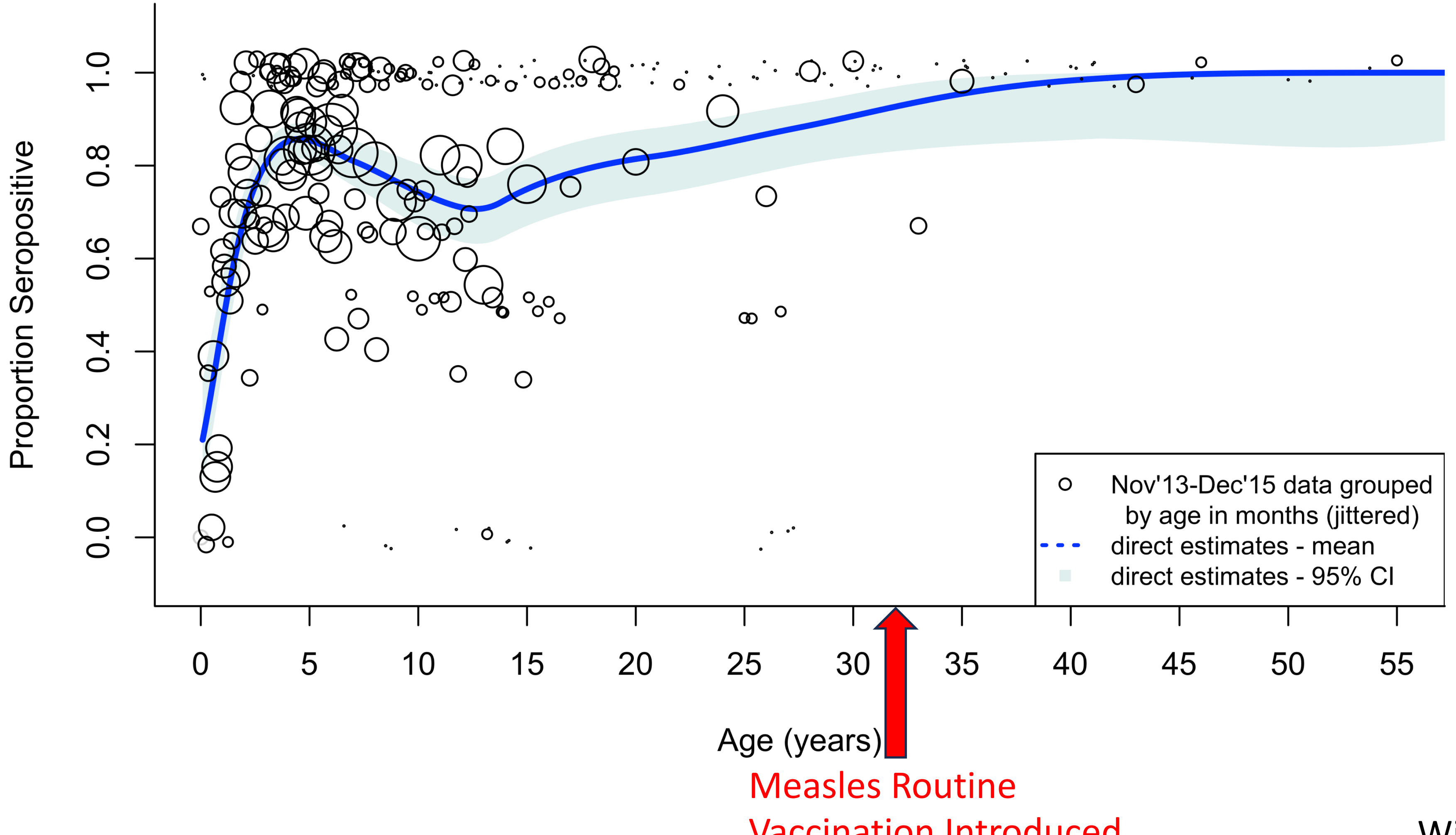
Given the crude birth rate for Assam in 2016 was 21.7 births per 1000 population

Catalytic Model Intuition

- What about after we introduce vaccine:
  - Do you expect the catalytic model to be a good fit?
  - What does the λ mean now?

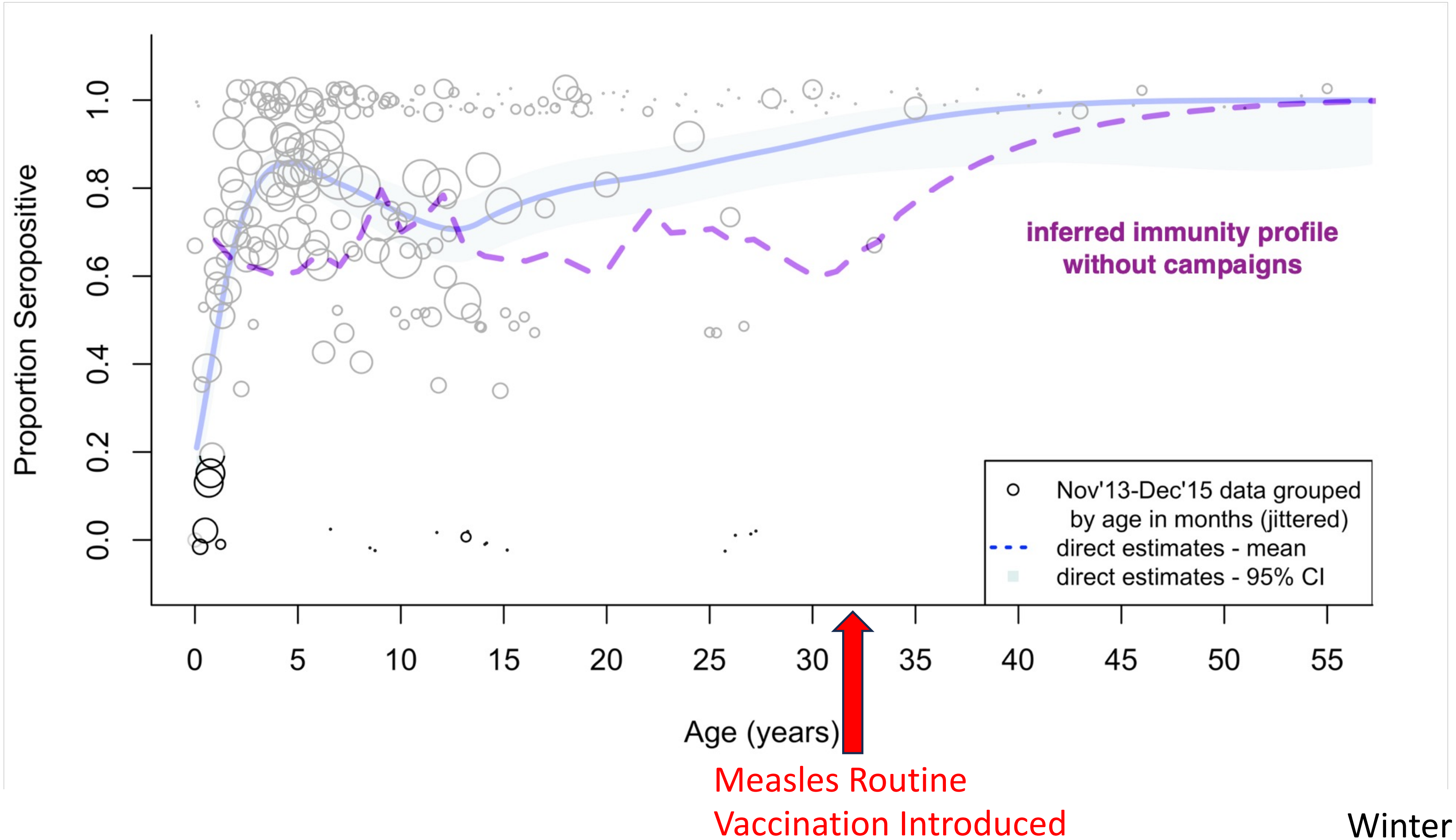


Realistic Post-Vaccination (Measles) Serology



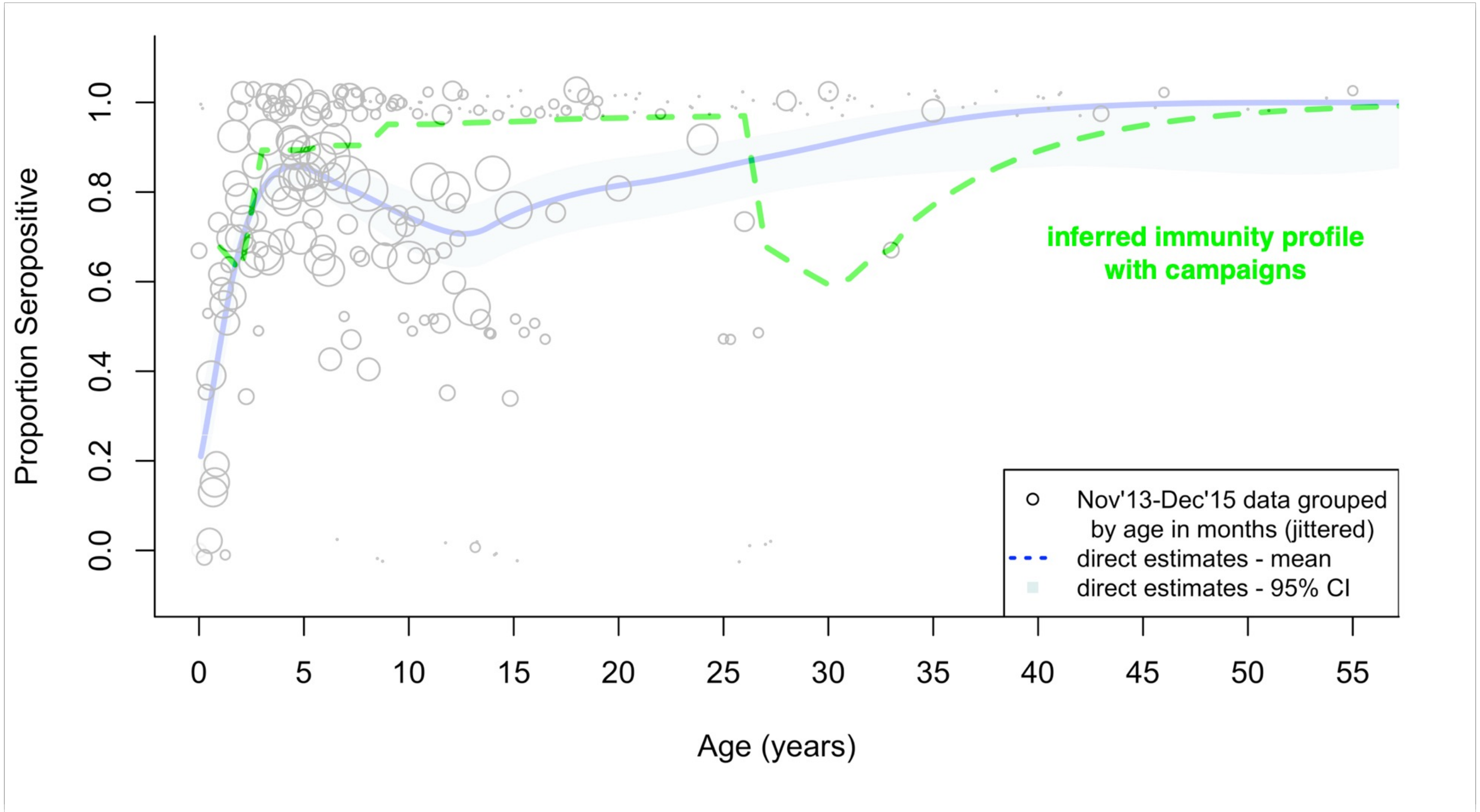
Winter et al, AJE, 2018

Realistic Post-Vaccination (Measles) Serology



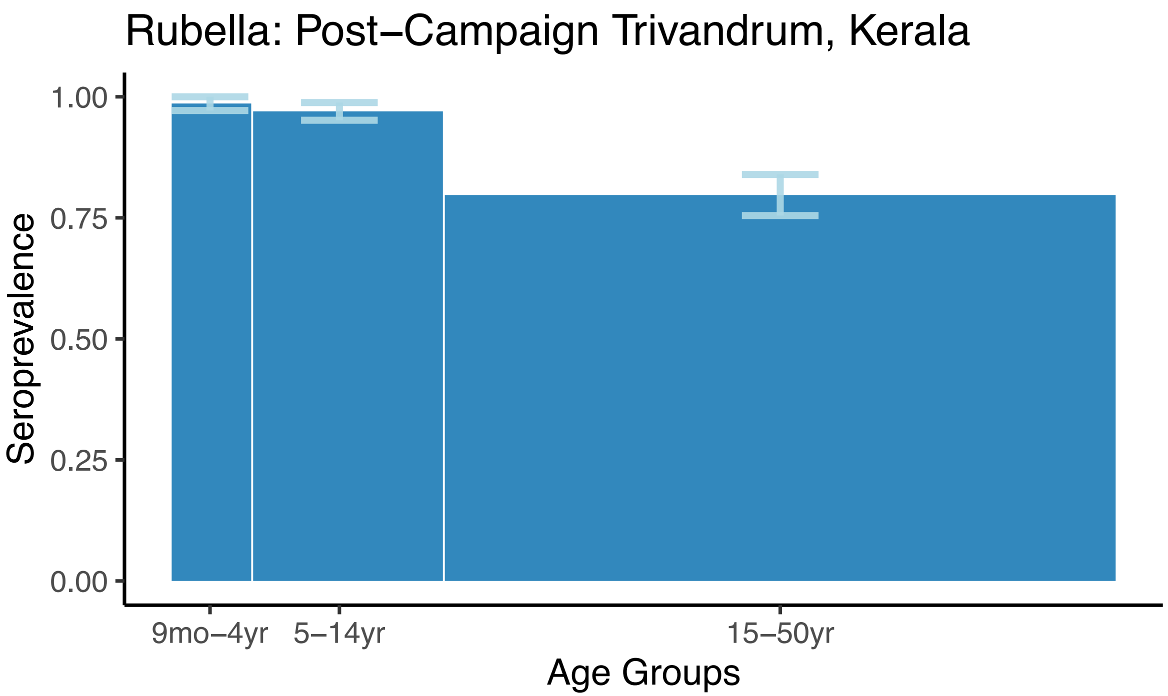
Winter et al, AJE, 2018

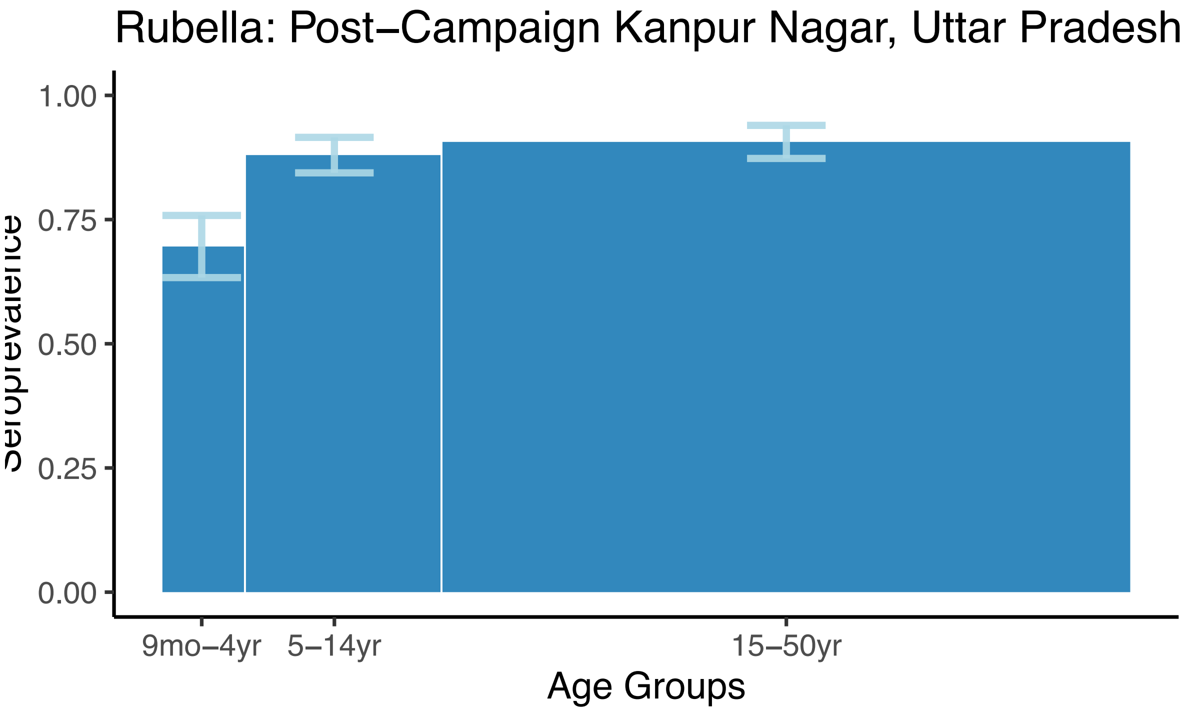
Realistic Post-Vaccination (Measles) Serology

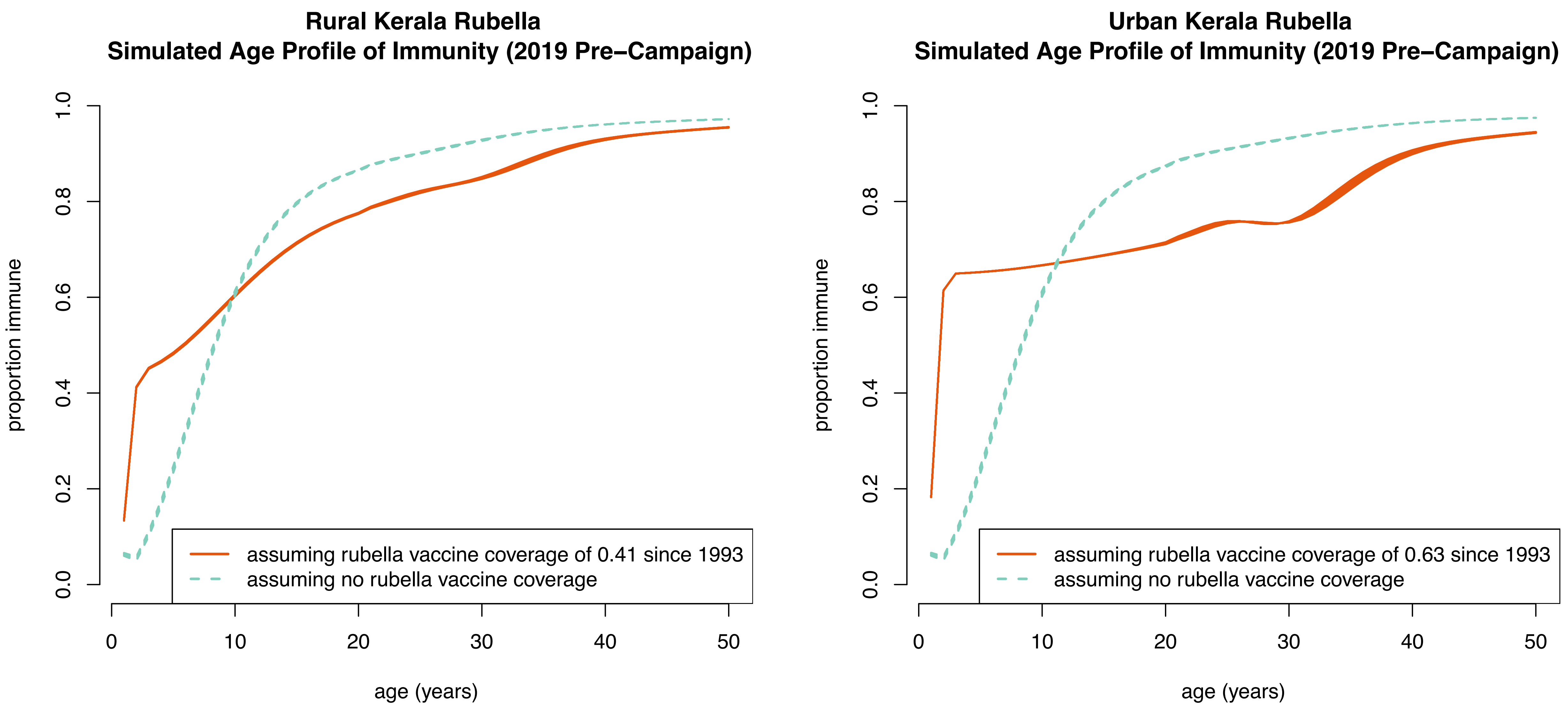


Winter et al, AJE, 2018

Realistic Post-Vaccination Rubella Serology






Rural Kerala Rubella
Simulated Age Profile of Immunity (2019 Pre–Campaign)
proportion immune
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
age (years)
assuming rubella vaccine coverage of 0.41 since 1993
assuming no rubella vaccine coverage
Urban Kerala Rubella
Simulated Age Profile of Immunity (2019 Pre–Campaign)
proportion immune
assuming rubella vaccine coverage of 0.63 since 1993
assuming no rubella vaccine coverage
age (years)

Objectives

Hopefully from this session you now:

- Know the equation for the catalytic model
- Are able to fit the catalytic model to data using least squares
- Understand the assumptions of the catalytic model
- Know one approach to estimate the force of infection and R_0
- Have developed some intuition at interpreting cross-sectional serological data