E^2M^2: Ecological and Epidemiological Modeling in Madagascar

Study Design in Epidemiology and Ecology

Centre ValBio

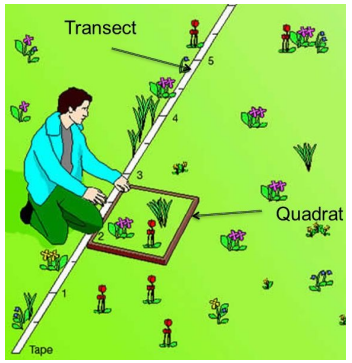
Ranomafana National Park, Madagascar

December 2022

Data: Sources of x and y

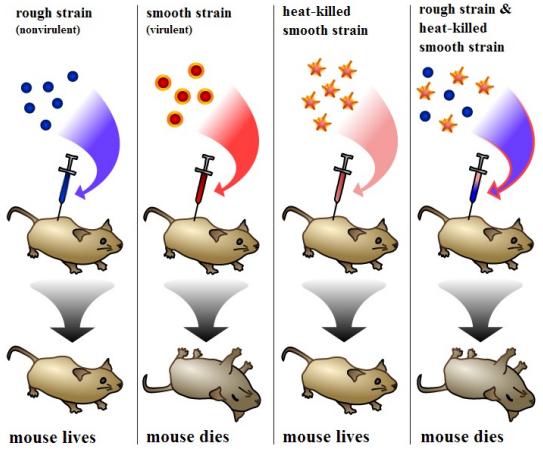
Observational

- Just measure x and y



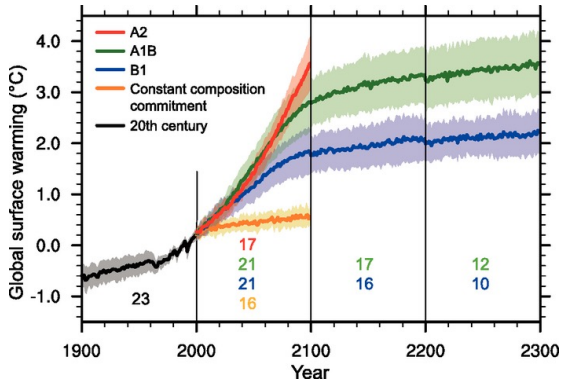
Experimental

- Interfere with x or the relationship between x and y



Simulated

- Create a relationship between x and y

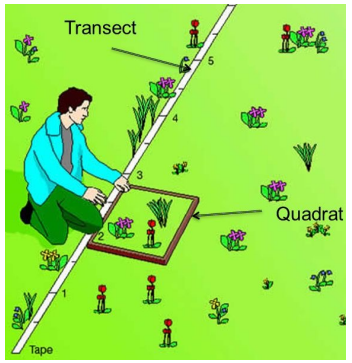


Empirical data

Data: Sources of x and y

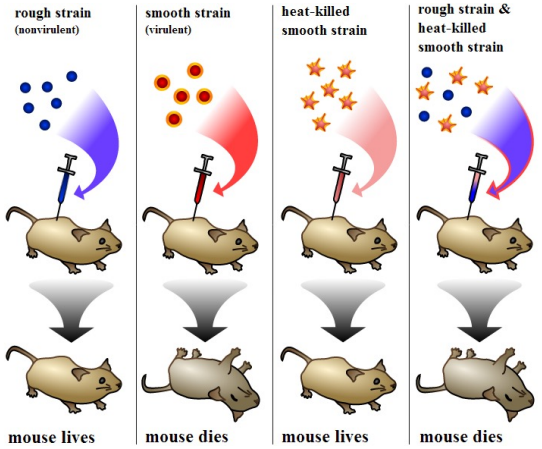
Observational

- Just measure x and y



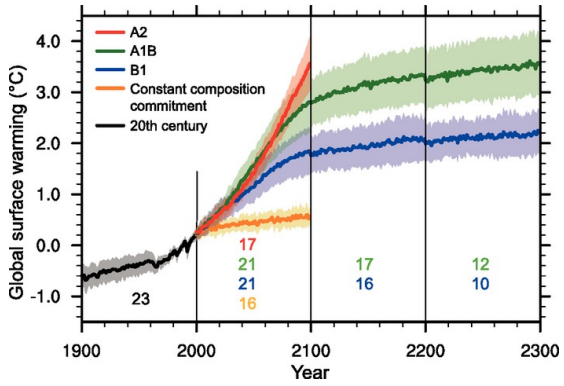
Experimental

- Interfere with x or the relationship between x and y



Simulated

- Create a relationship between x and y



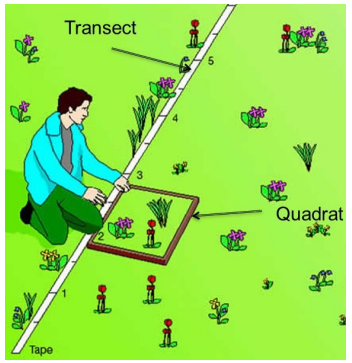
How do we choose the best method of data collection?

Empirical data

Data: Sources of x and y

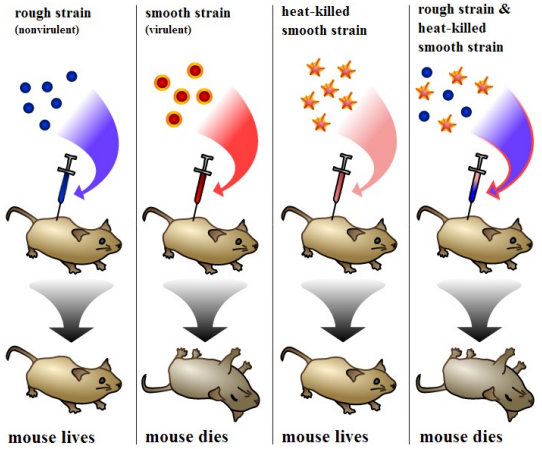
Observational

- Just measure x and y



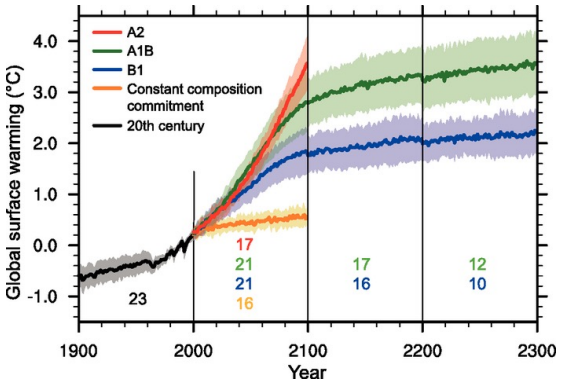
Experimental

- Interfere with x or the relationship between x and y



Simulated

- Create a relationship between x and y



Empirical data

How do we choose the best method of data collection?

The 'best method' will depend on the question!

Model-Guided Field Work

Model-Guided Field Work

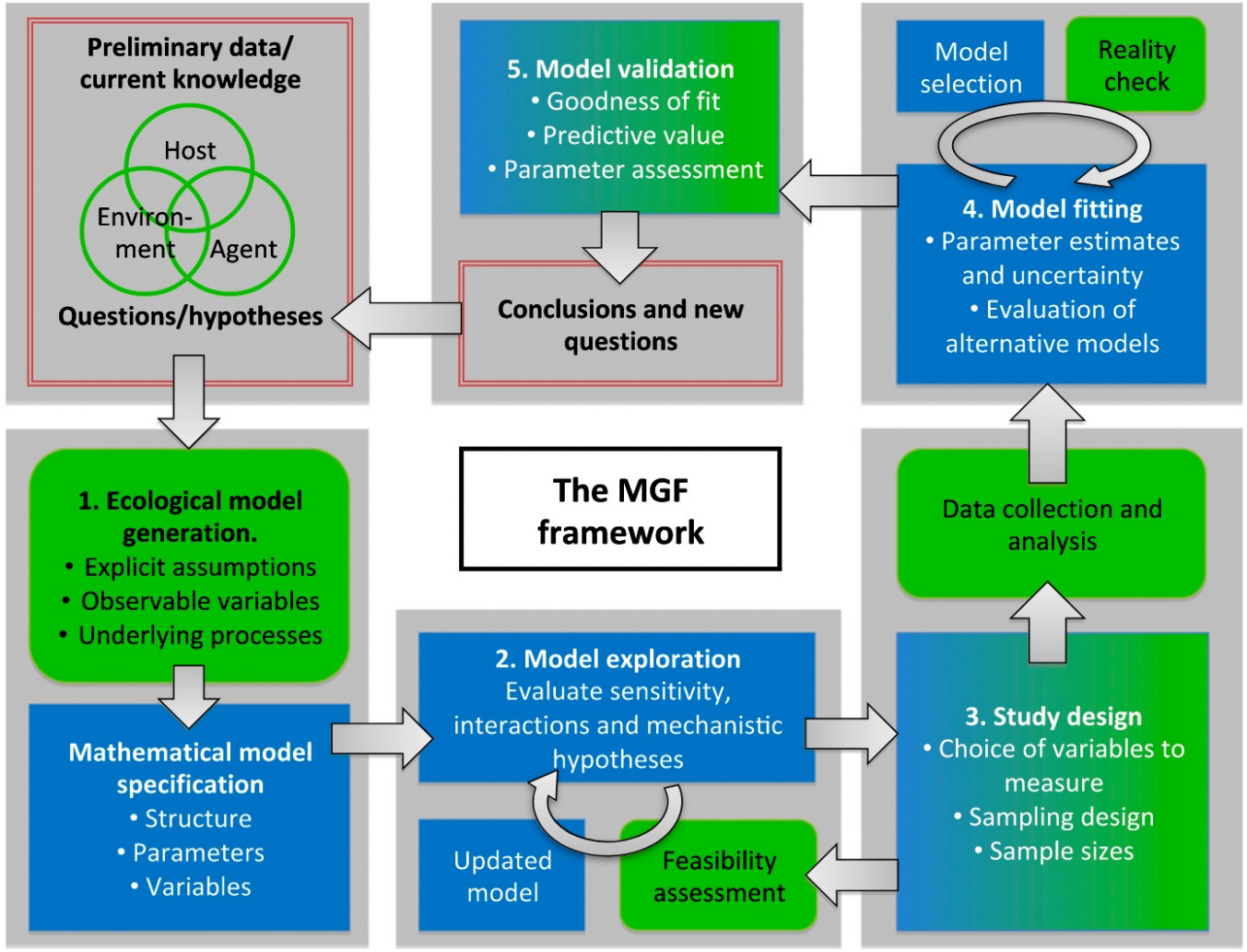
- a rational dialogue between researchers from multiple disciplines through a series of iterative steps, ultimately leading to improved causal inference and predictive power.

Model-Guided Field Work

- a rational dialogue between researchers from multiple disciplines through a series of iterative steps, ultimately leading to improved causal inference and predictive power.

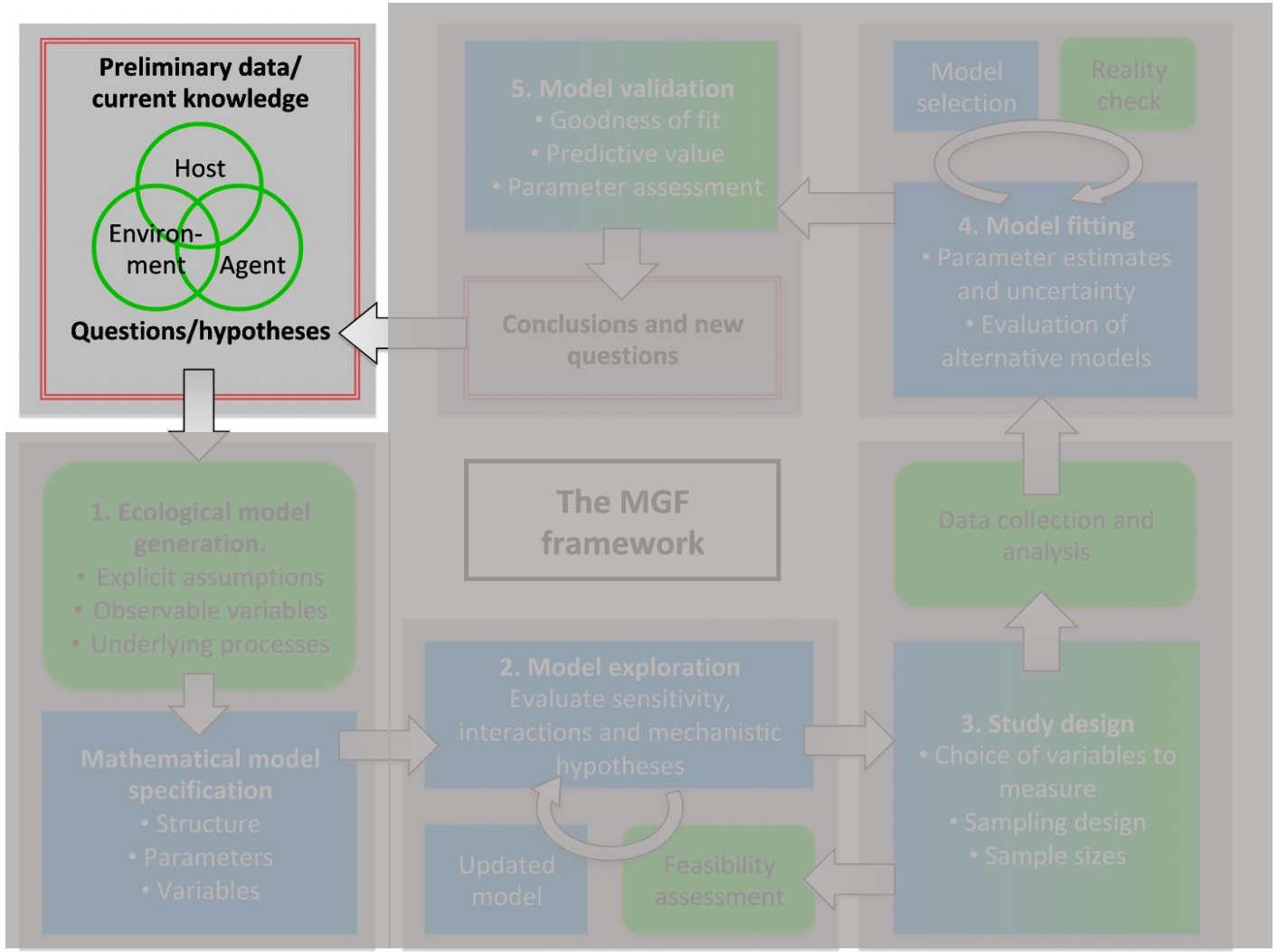
- biologists and modellers collaborate at all stages of the study, from initial model formulation and field study design, to data collection and analysis.

Model-Guided Field Work



Restif et al. 2012 *Ecology Letters*

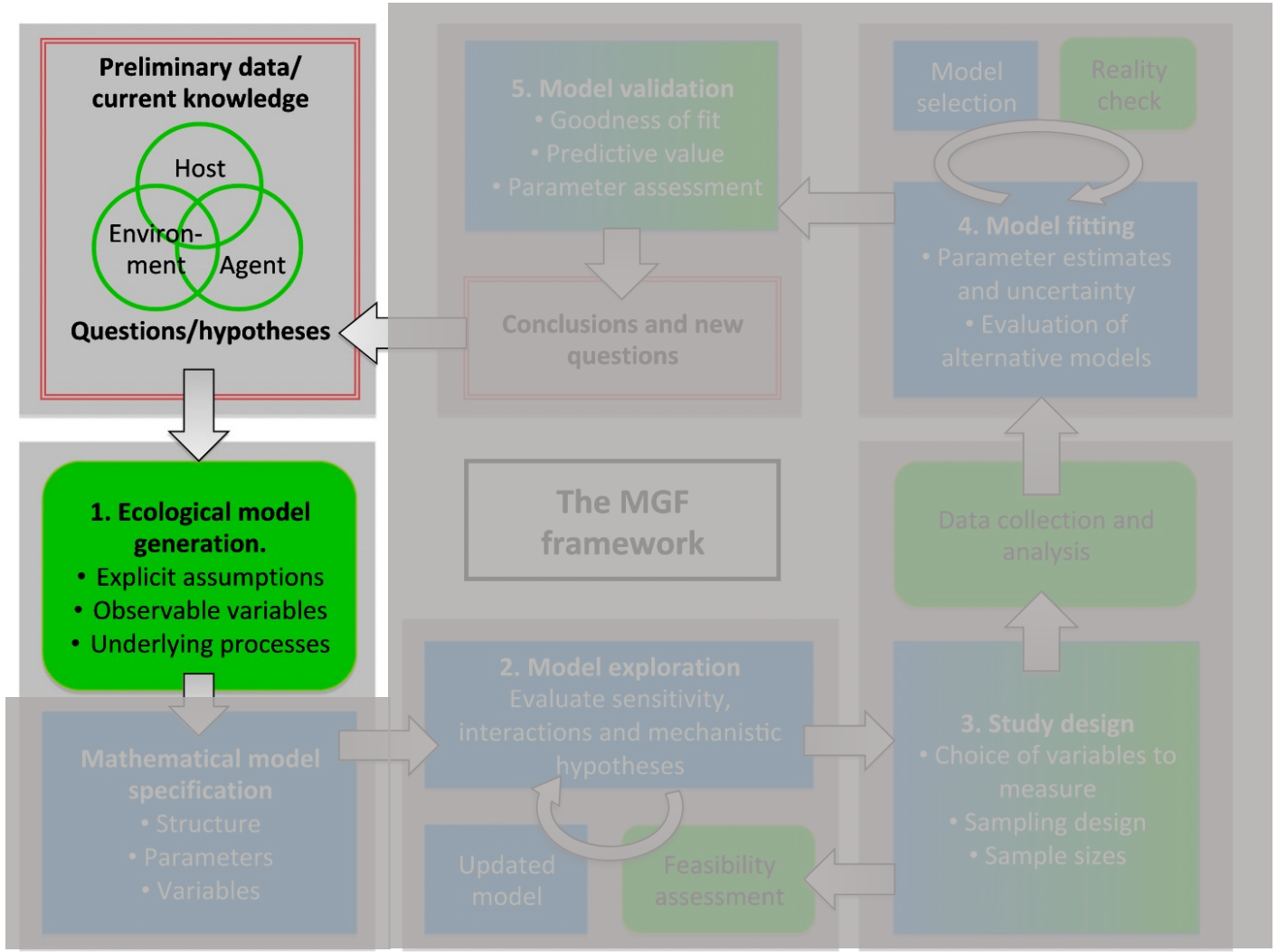
Model-Guided Field Work



Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work

Formulating research questions

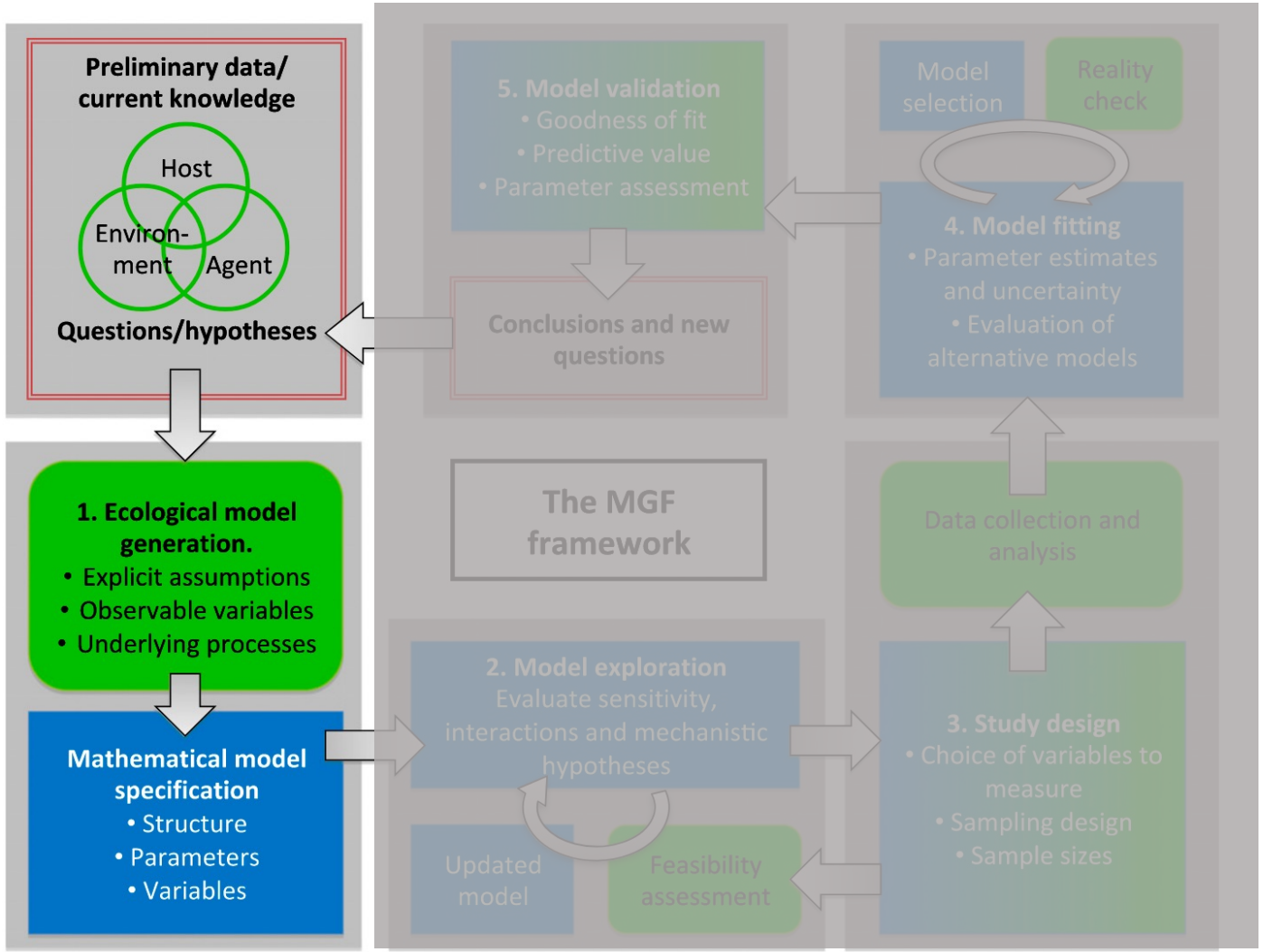


Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work

Formulating research questions

Model diagram

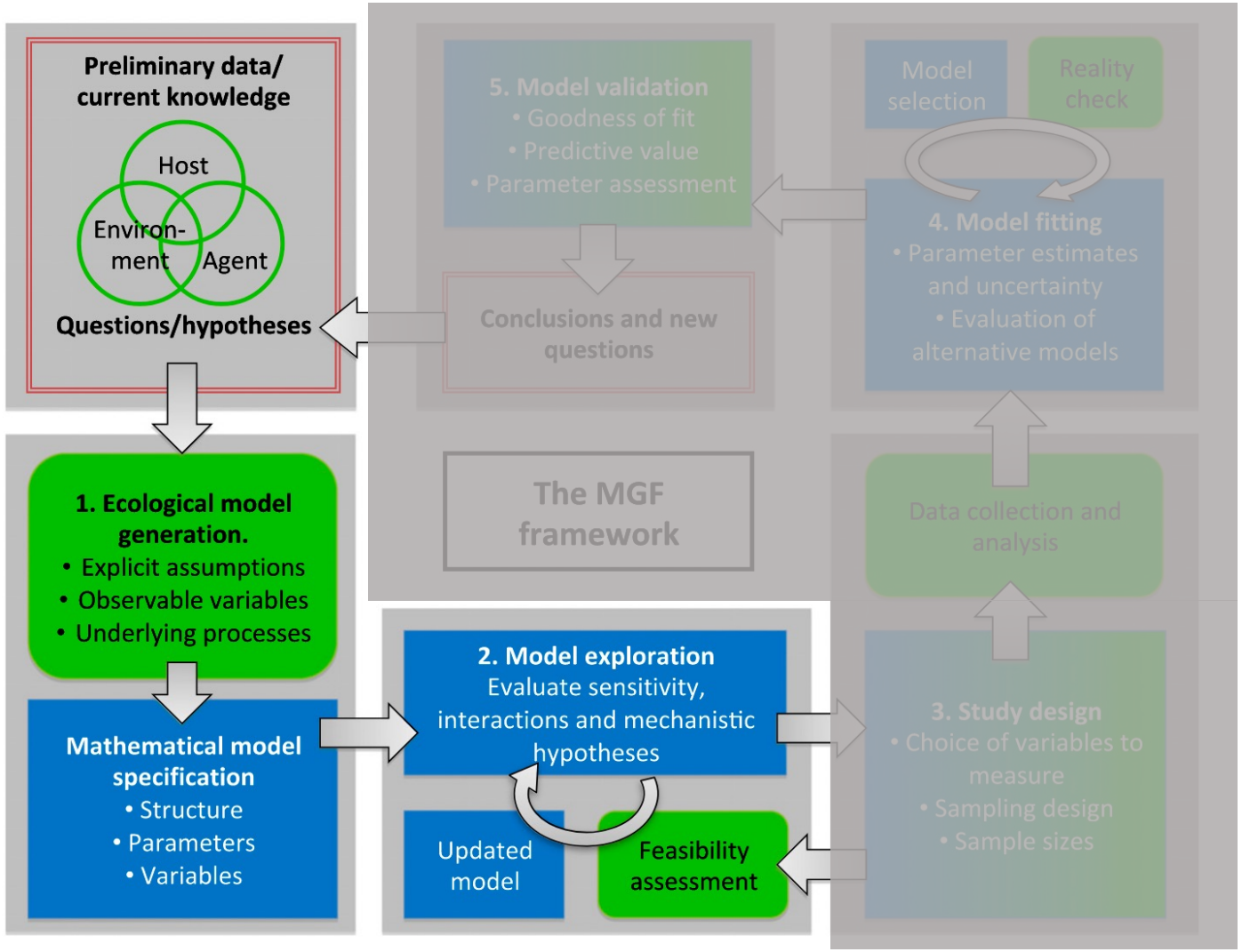


Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work

Formulating research questions

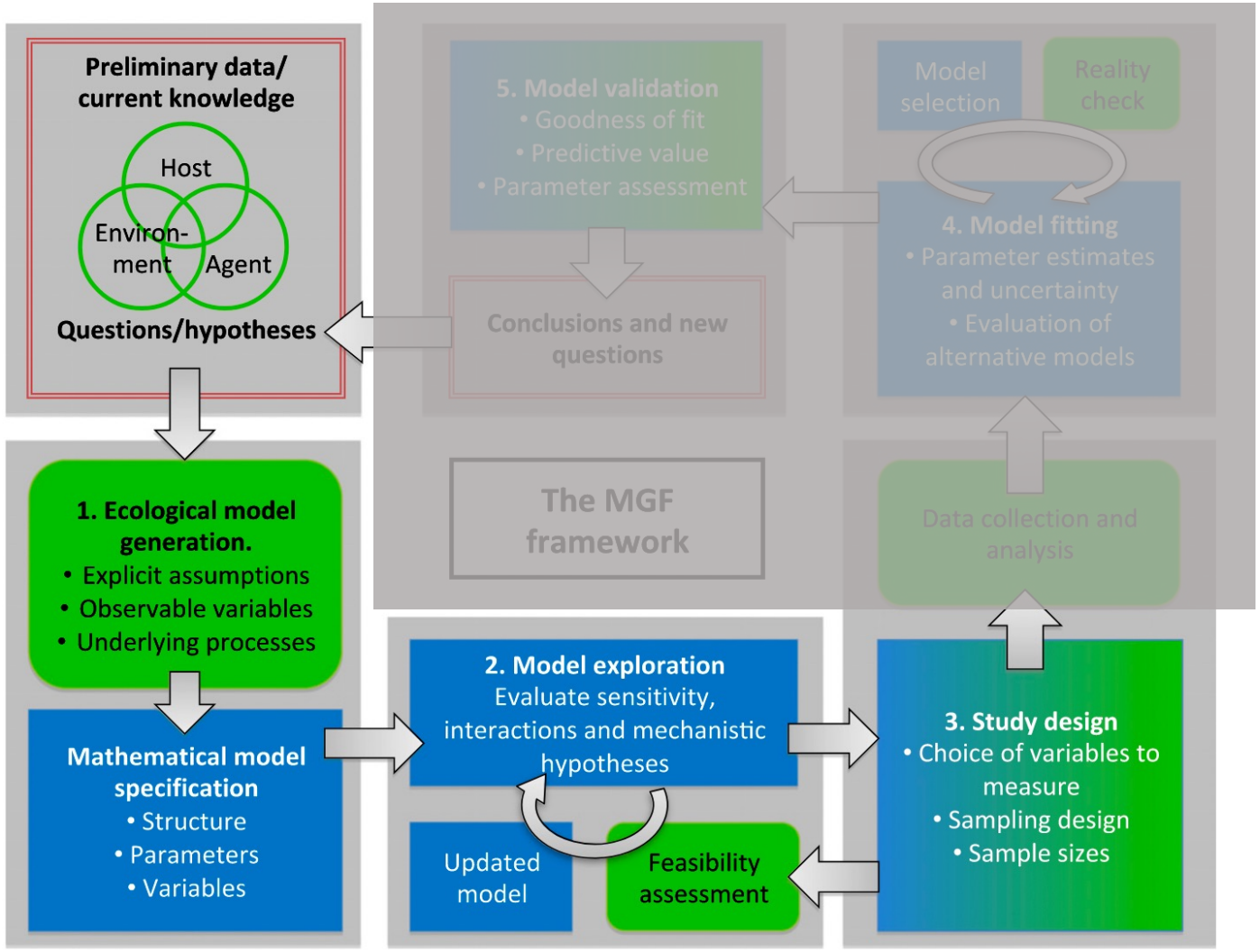
Model diagram



Model simulation

Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work



Formulating research questions

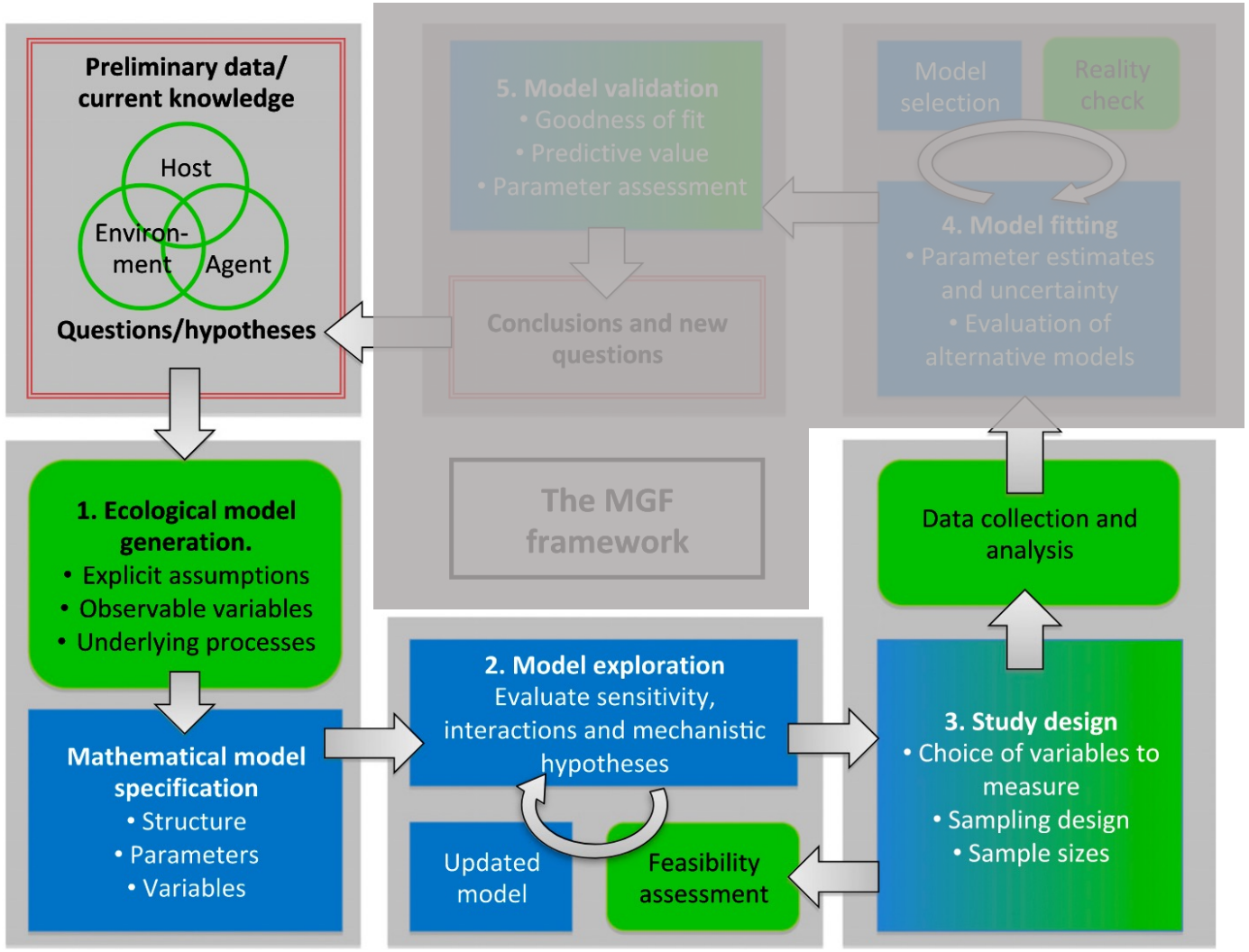
Model diagram

Study design

Model simulation

Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work



Formulating research questions

Model diagram

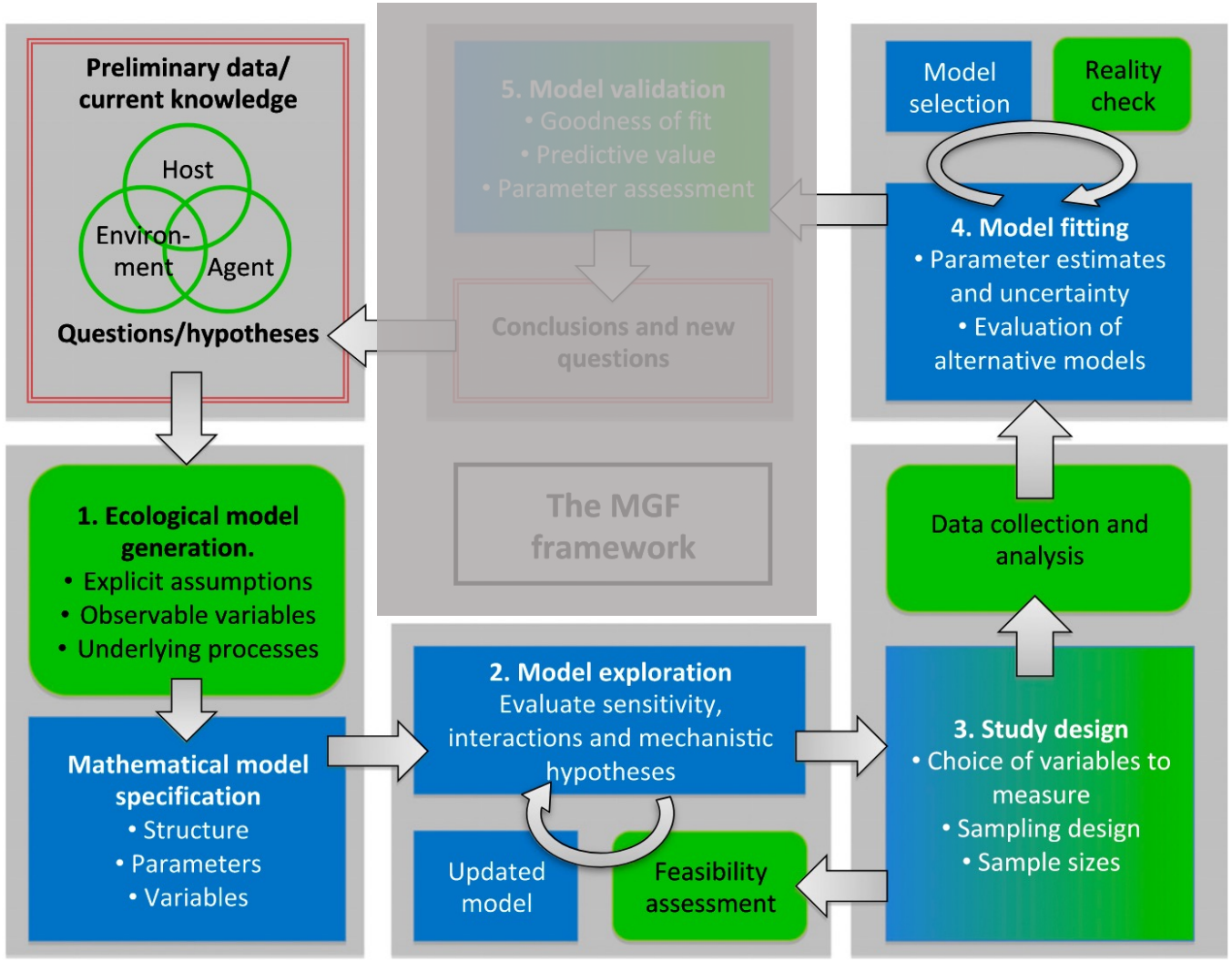
Model simulation

Data collection

Study design

Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work



Formulating research questions

Model diagram

Model fitting

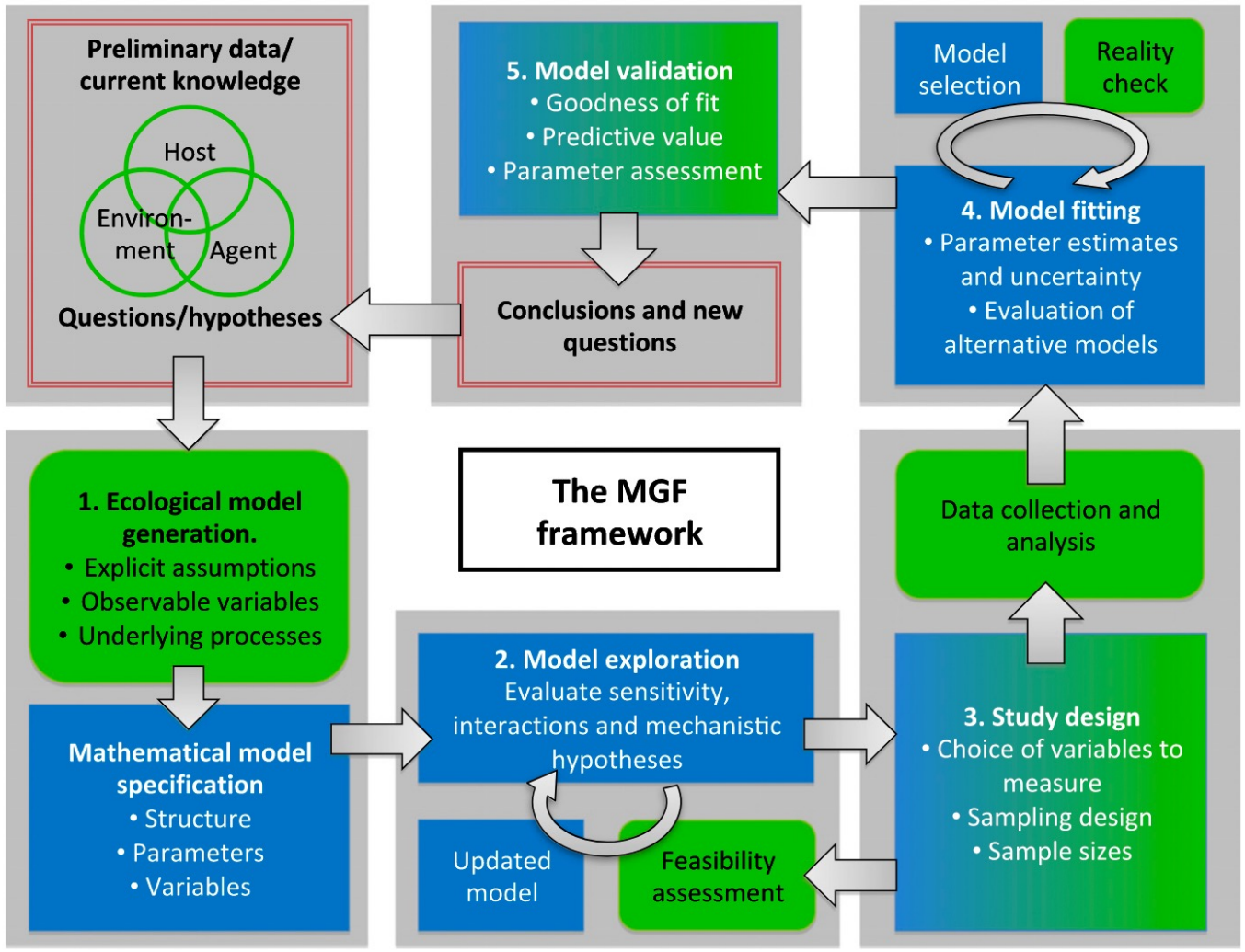
Data collection

Study design

Model simulation

Restif et al. 2012 *Ecology Letters*

Model-Guided Field Work



Restif et al. 2012 *Ecology Letters*

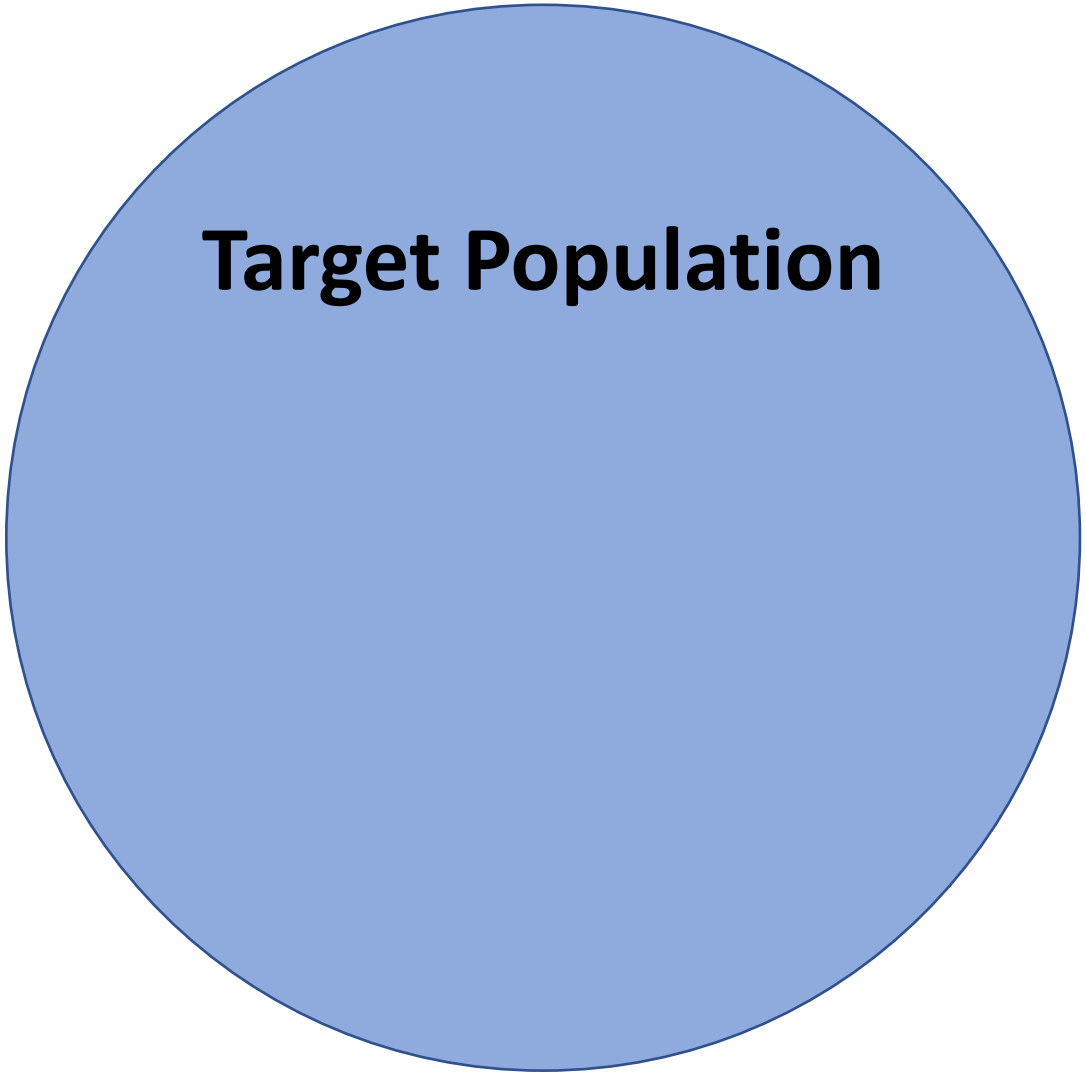
Model-Guided Field Work

- a rational dialogue between researchers from multiple disciplines through a series of iterative steps, ultimately leading to improved causal inference and predictive power.

- biologists and modellers collaborate at all stages of the study, from initial model formulation and field study design, to data collection and analysis.

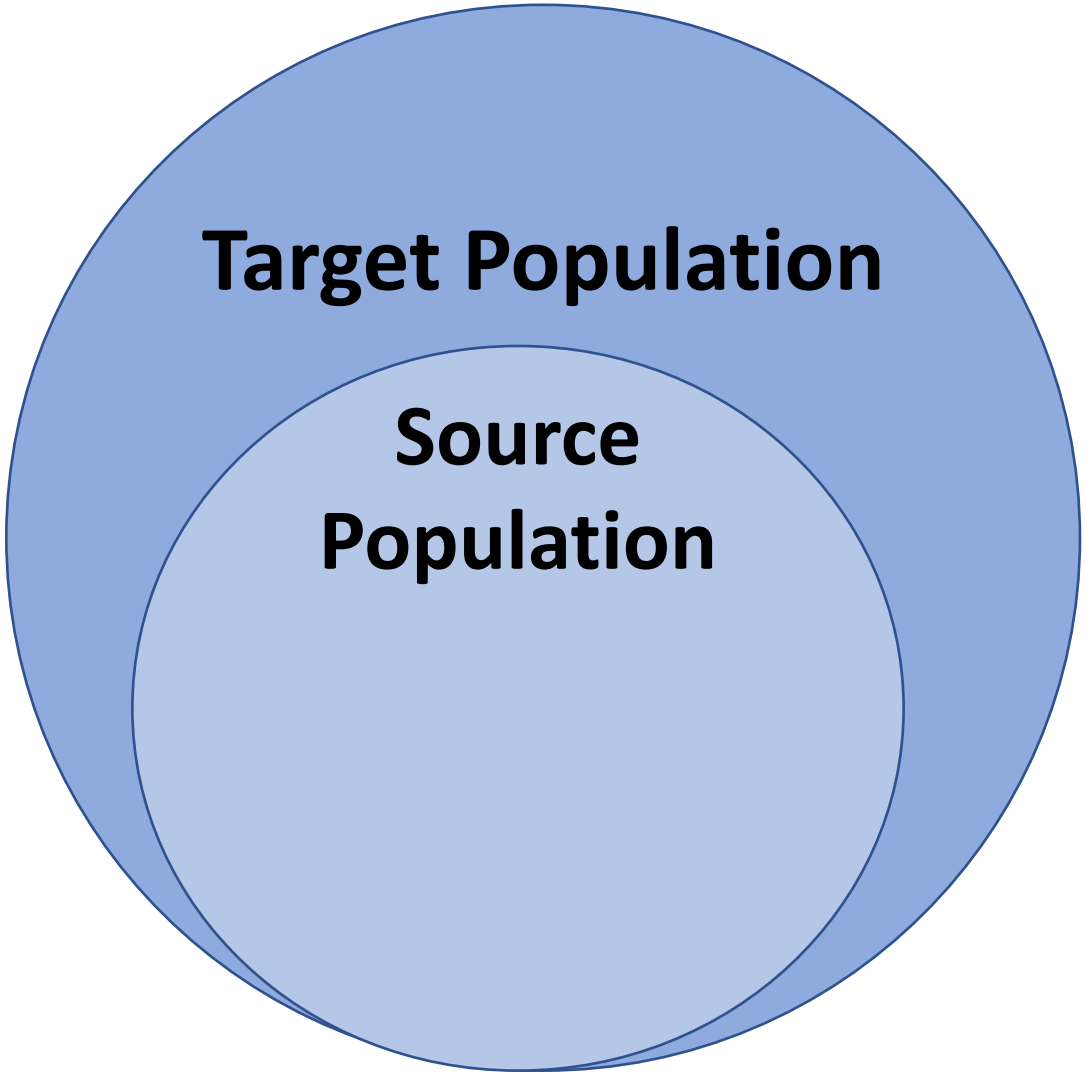
- applicable in both ecology AND epidemiology!

Epidemiological Study Design



Target: Population to which it *might* be possible to extrapolate results of the study

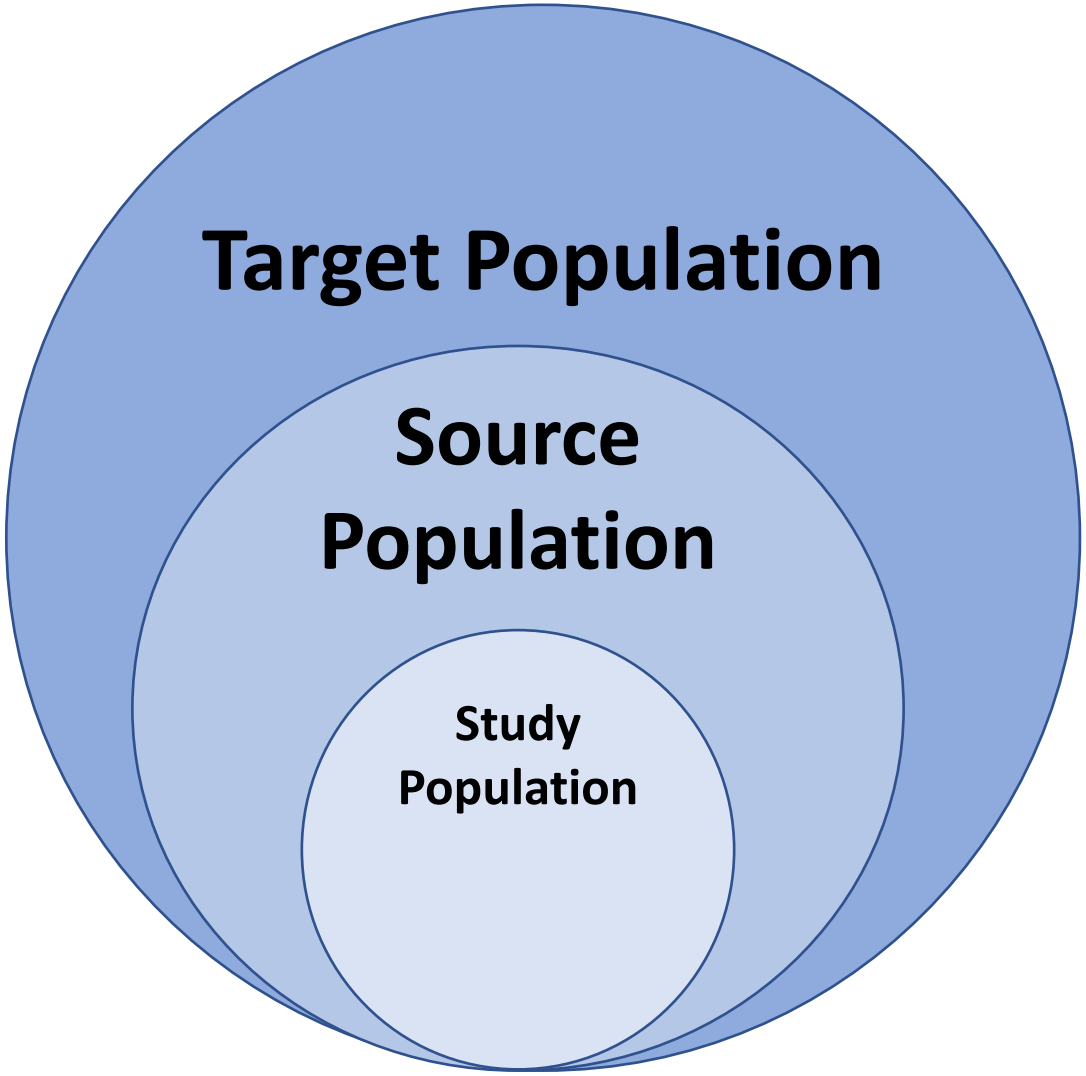
Epidemiological Study Design



Target: Population to which it *might* be possible to extrapolate results of the study

Source: Population from which study participants are drawn

Epidemiological Study Design

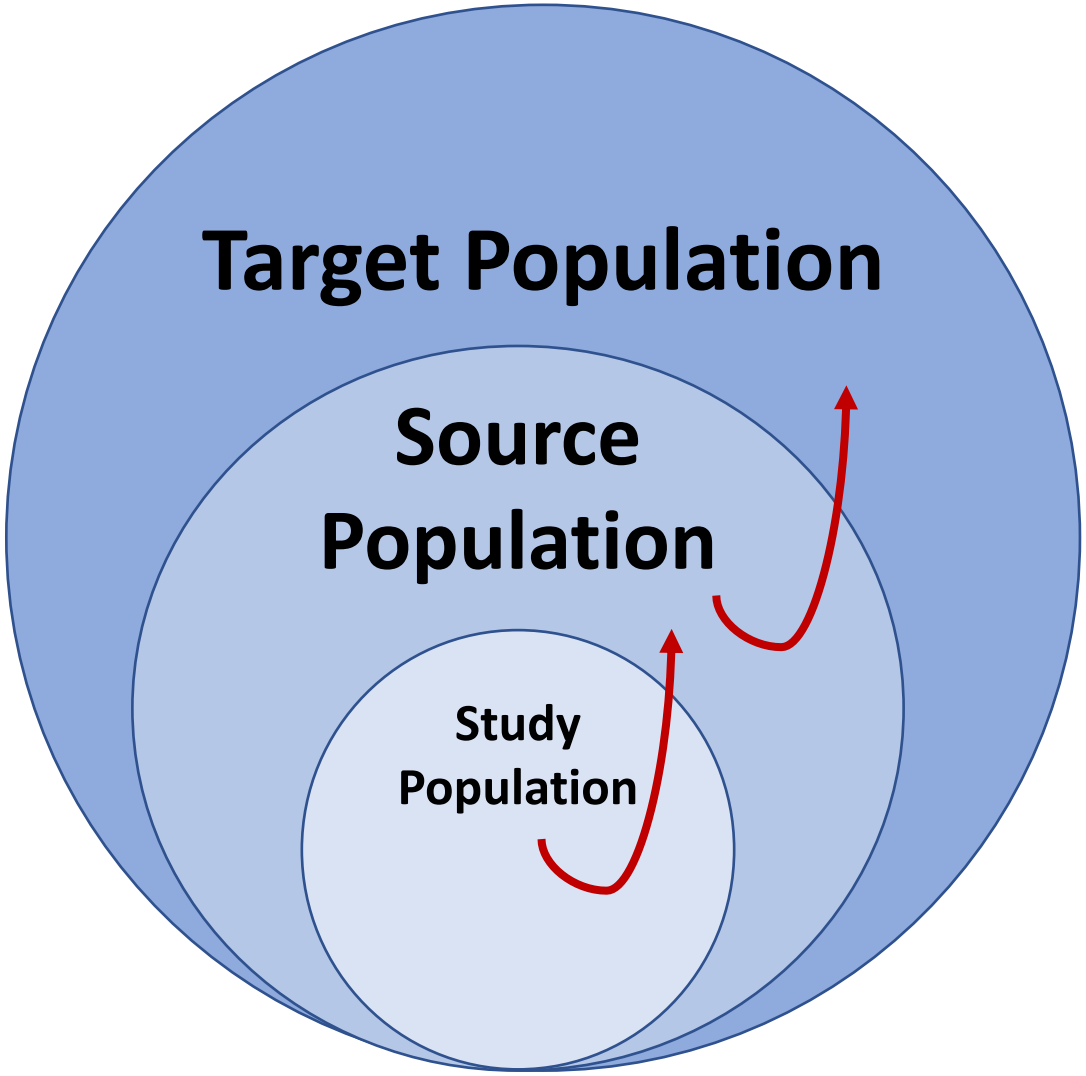


Target: Population to which it *might* be possible to extrapolate results of the study

Source: Population from which study participants are drawn

Study: Actual sampled population

Epidemiological Study Design



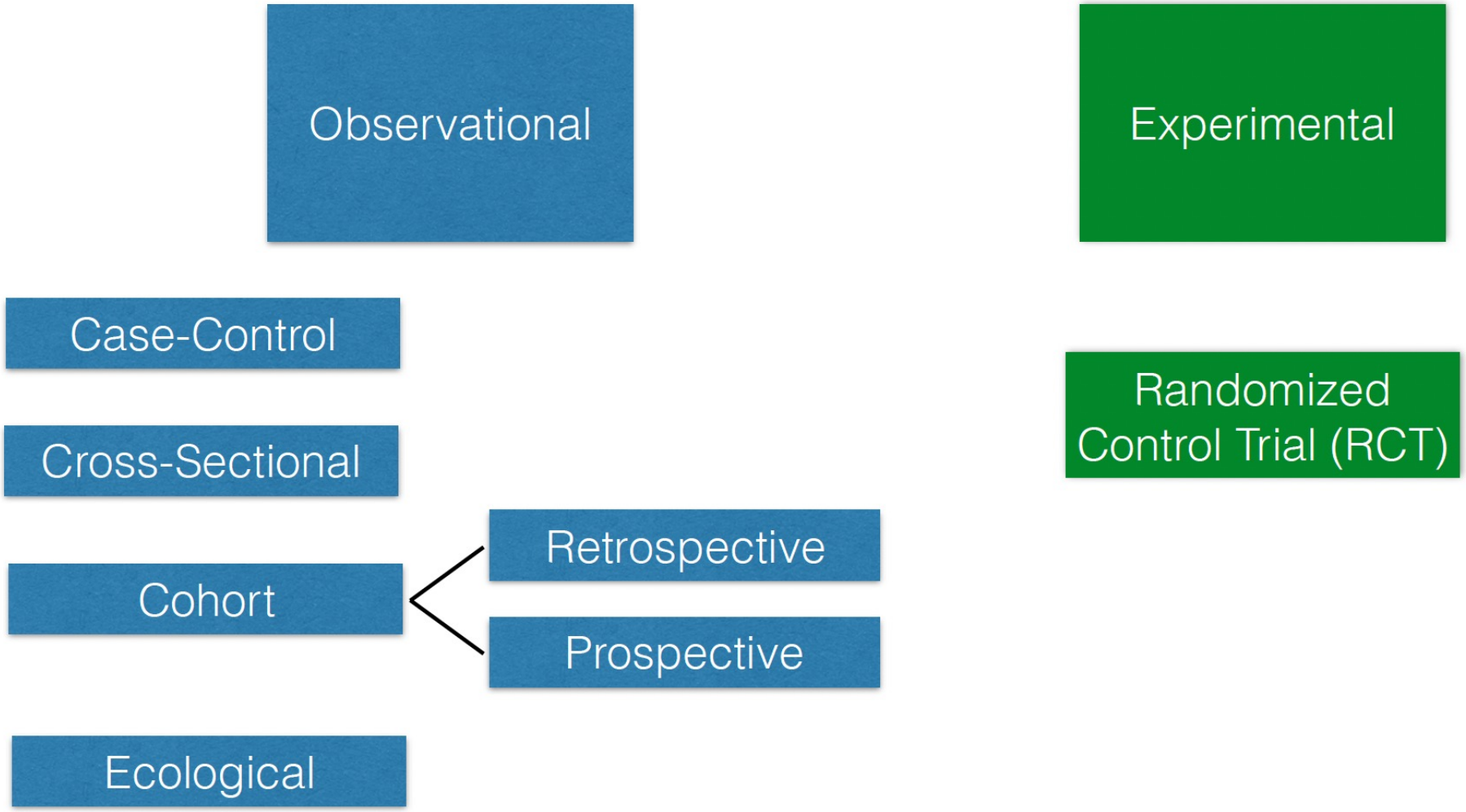
Target: Population to which it *might* be possible to extrapolate results of the study

Source: Population from which study participants are drawn

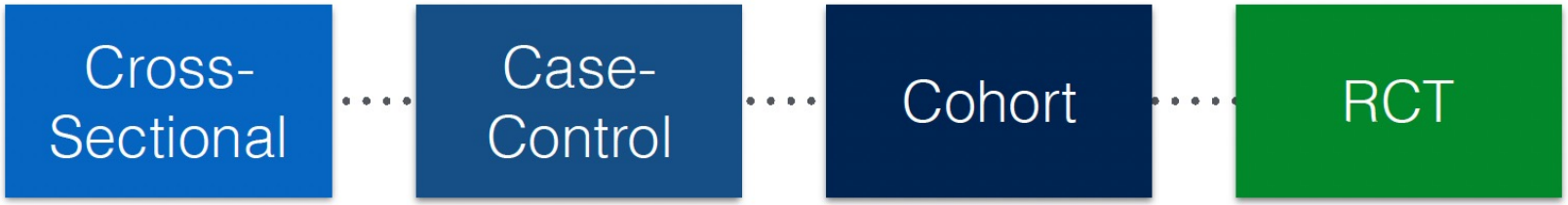
Study: Actual sampled population

Well-designed studies allow us to make inferences about the target population

Different Study Types in Epidemiology

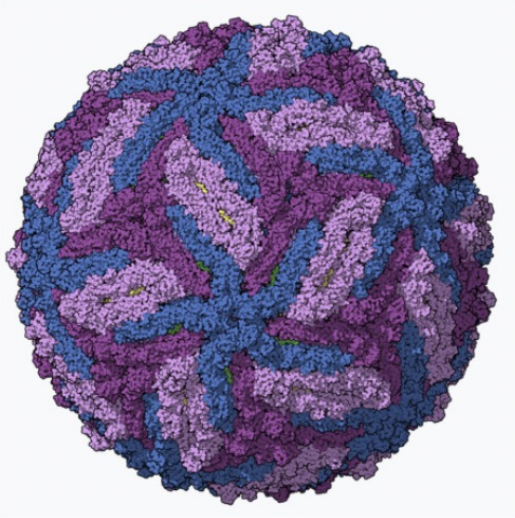


The Basics



Zika Virus: The Basics

Cross-Sectional ..... Case-Control ..... Cohort ..... RCT



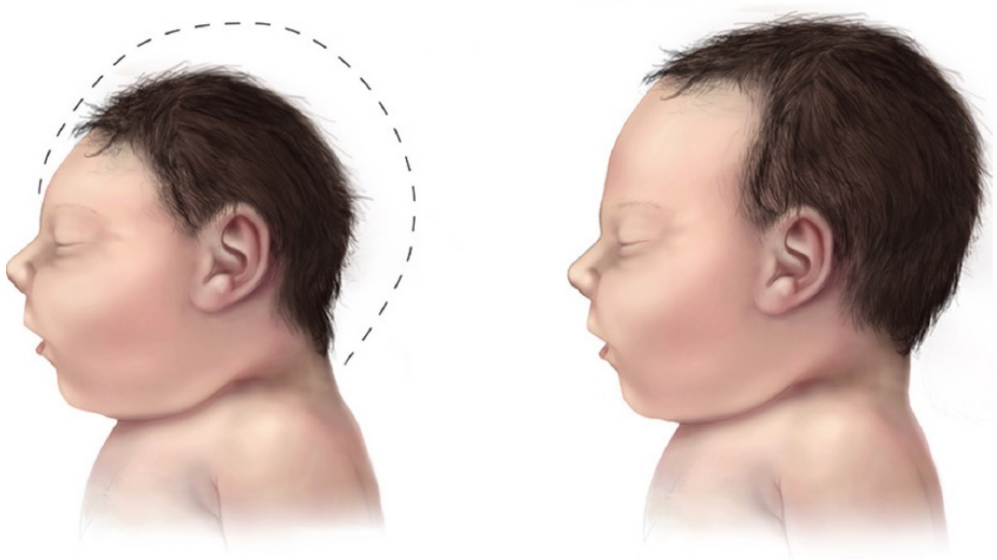
Virus in the family *Flaviviridae* (related to dengue, yellow fever)

Spread by Aedes mosquitoes

Multiple routes of transmission: vector, sexual, vertical

Zika Virus: The Basics

Cross-Sectional ..... Case-Control ..... Cohort ..... RCT

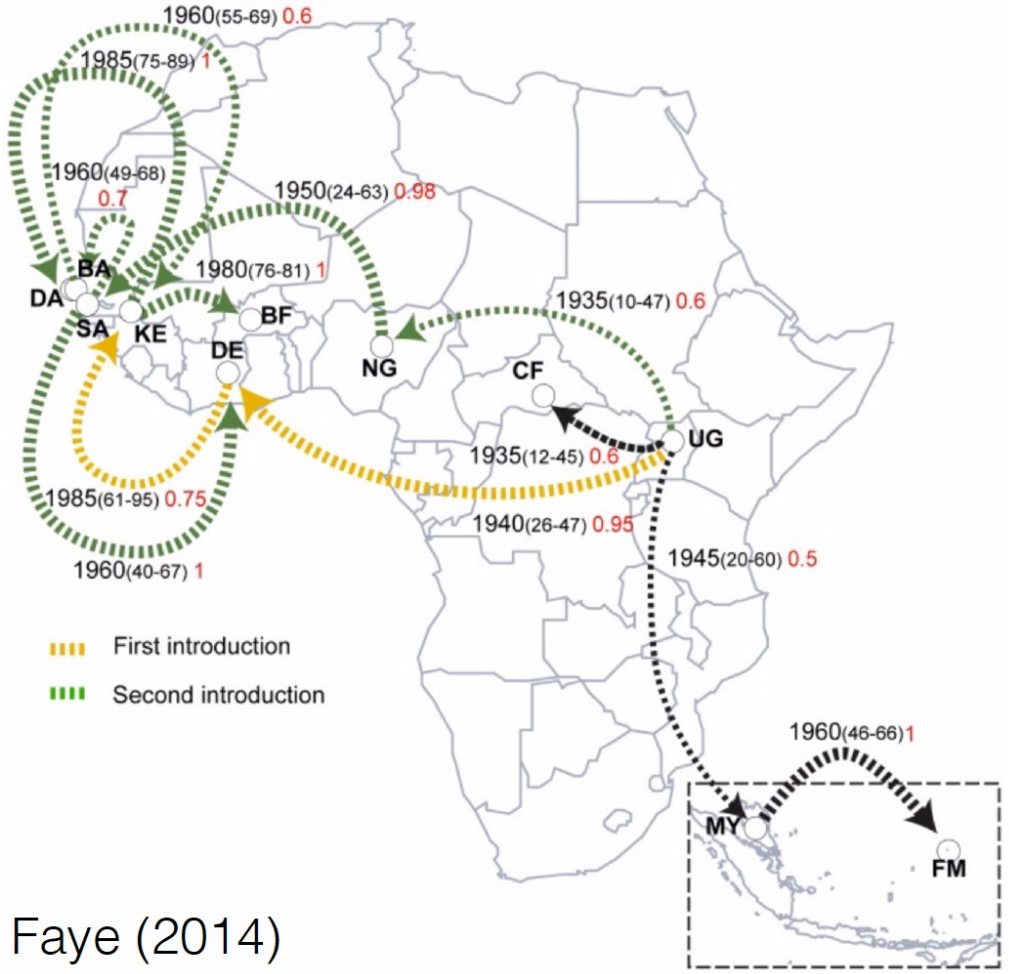


Often causes no or mild symptoms
But can spread from pregnant women to their fetuses and result in microcephaly, severe brain malformations, other birth defects

Currently cannot be prevented by medications or vaccines (current vaccine trials)

Zika Virus: The Basics

Cross-Sectional ···· Case-Control ···· Cohort ···· RCT

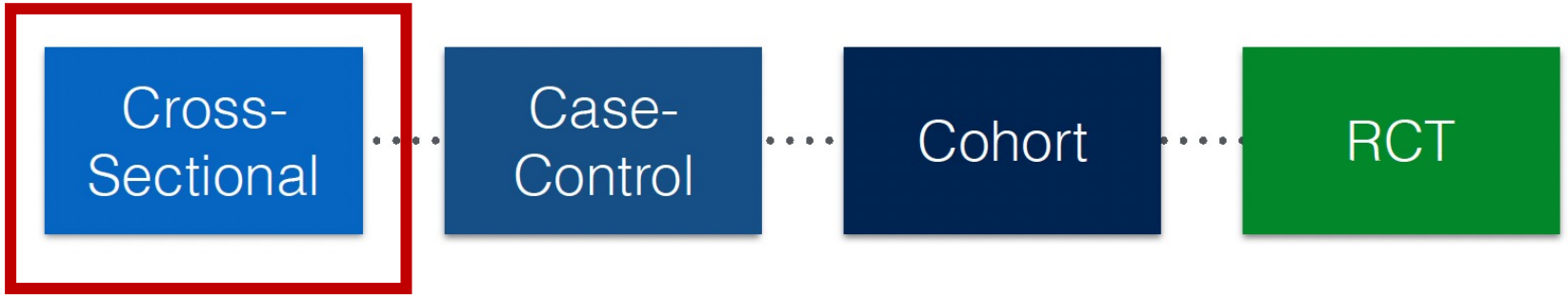


Faye (2014)

Originally isolated in 1947 (Ziika Forest, Uganda)

Sporadic outbreaks in Africa and Asia
Large outbreak in 2015-2016 (Americas, SE Asia, Pacific Islands, Brazil)

General: Study Design



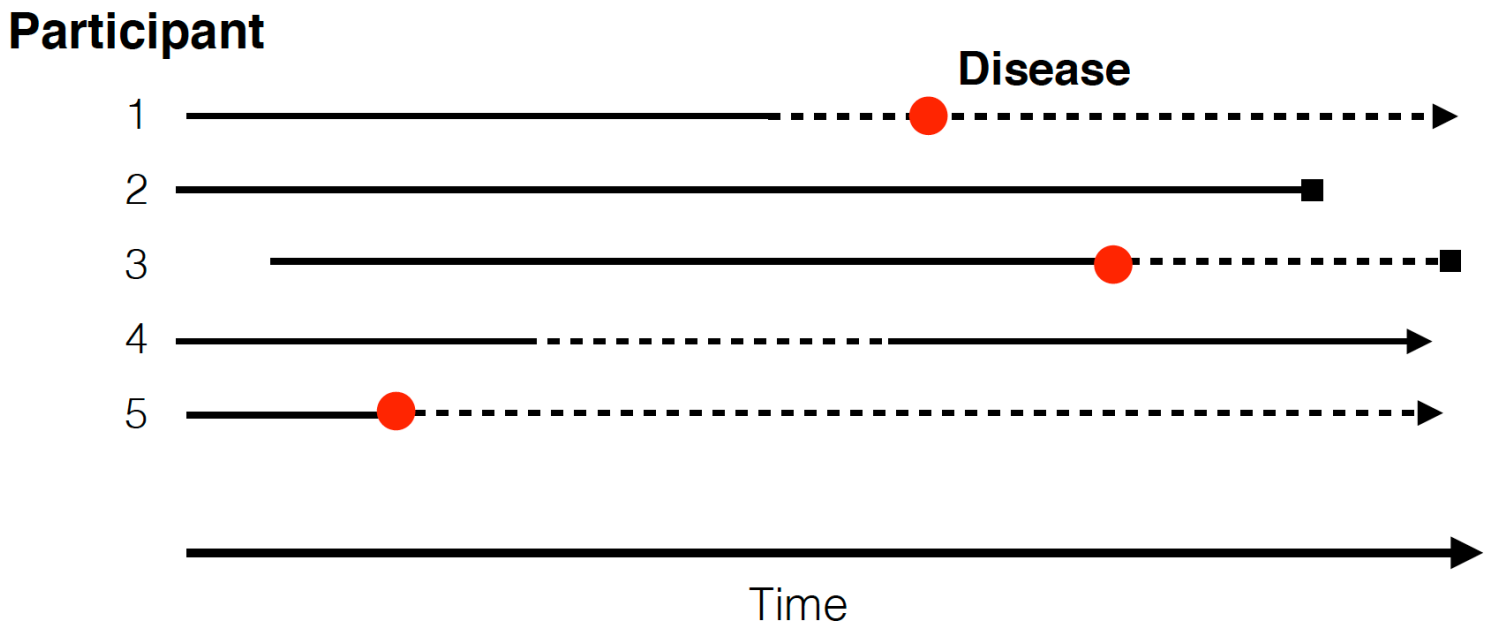
Cross-Sectional Study

- Examines relationship between diseases and other variables of interest (e.g. geographic distribution)
- Presence/absence of disease is determined for all members of a population
- Commonly used to estimate prevalence (rather than incidence)
- A single snapshot of the population at a moment in time
- Exposure and outcome are assessed simultaneously

General: Study Design

Cross-Sectional | Case-Control | Cohort | RCT

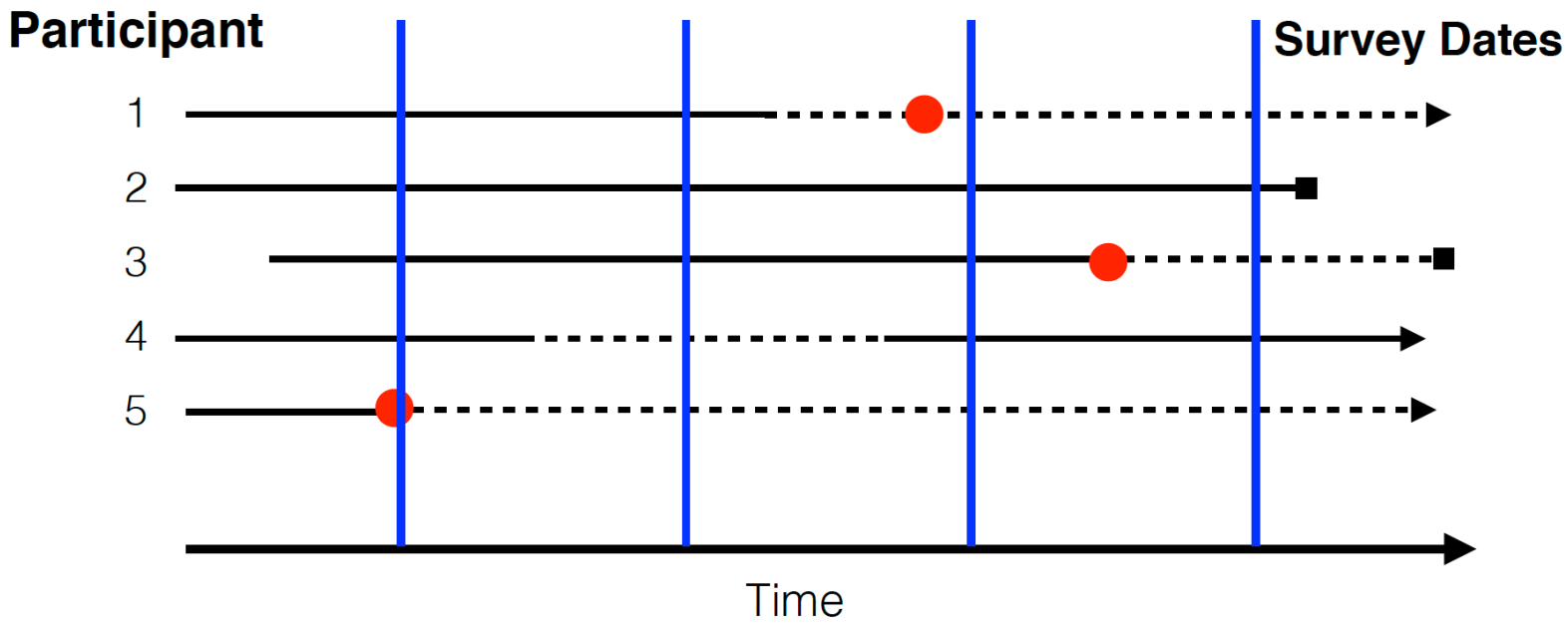
Cross-Sectional Study



General: Study Design

Cross-Sectional — Case-Control — Cohort — RCT

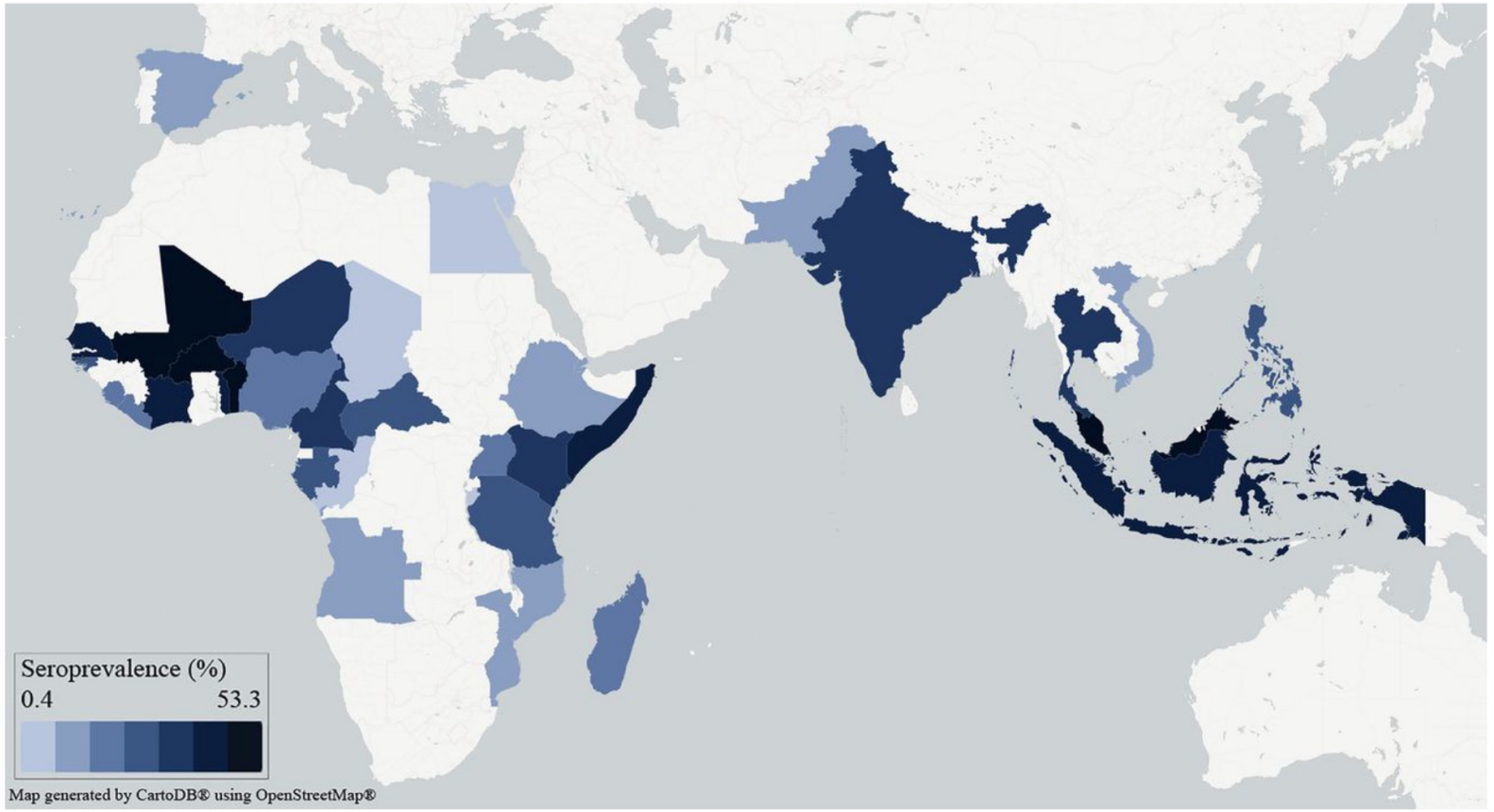
Cross-Sectional Study



Zika: Study Design

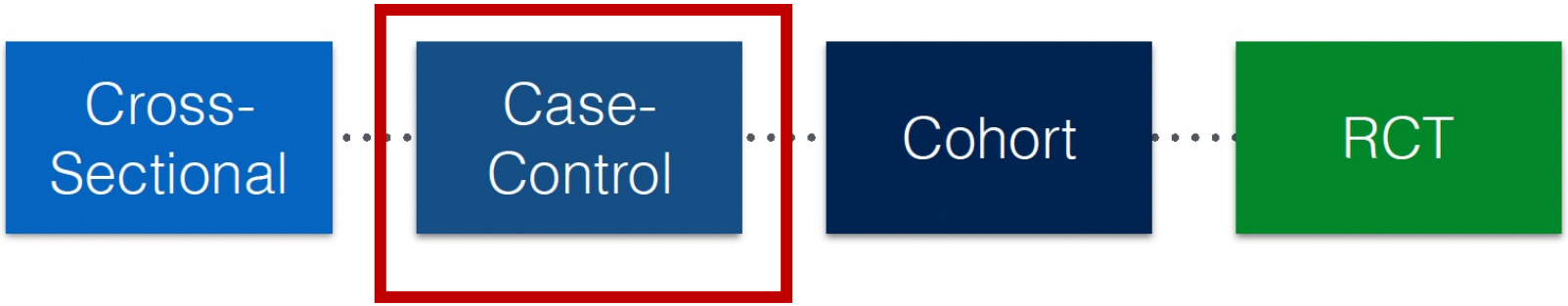
Cross-Sectional | Case-Control | Cohort | RCT

Cross-Sectional Study



Posen (2016)

General: Study Design



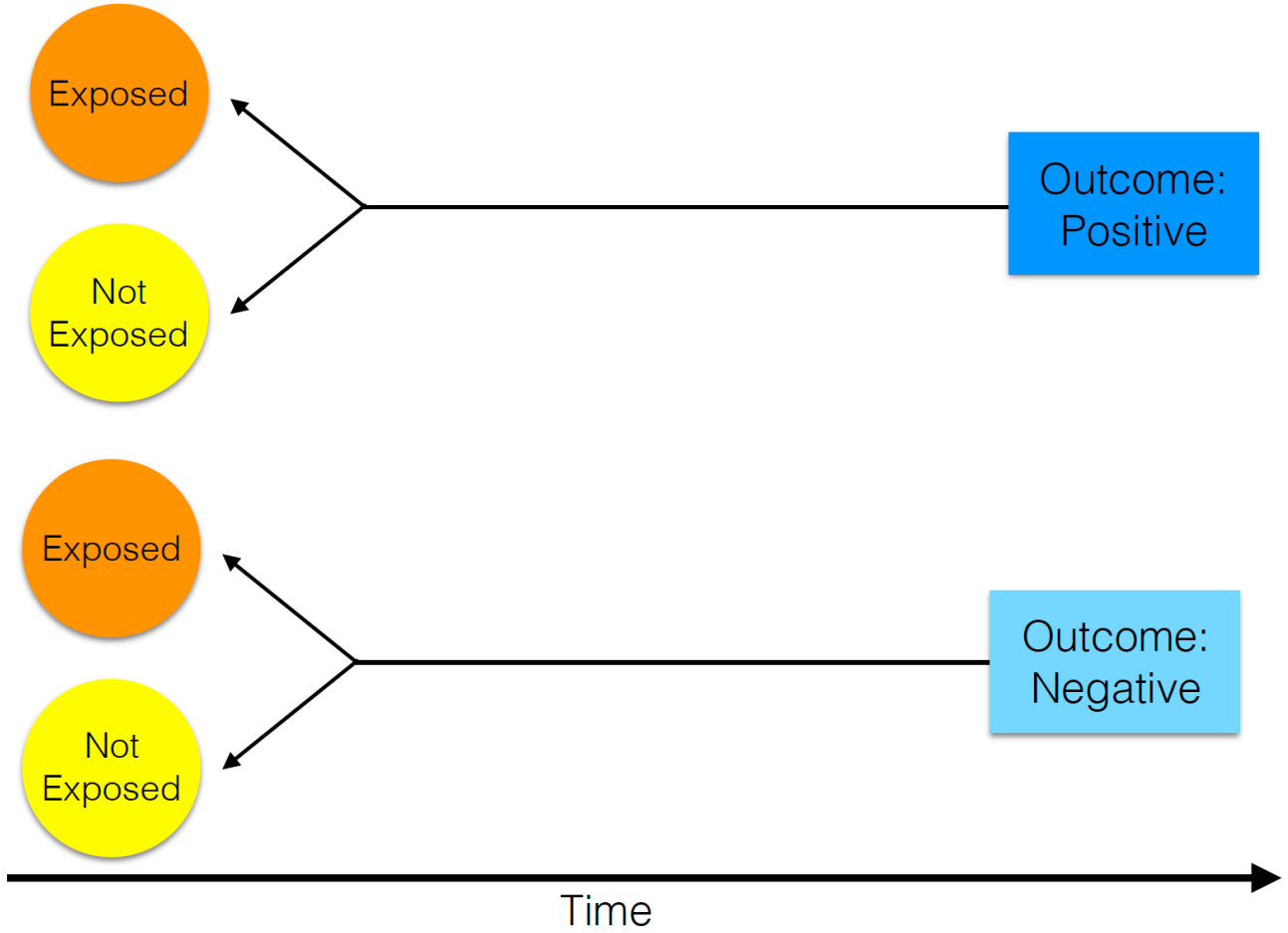
Case Control Study

- The observational epidemiological study of **persons with a disease of interest** and a suitable **control group of persons without the disease**
- Potential relationship of a **suspected risk factor** or an attribute to the disease is examined by **comparing the the disease and non-diseased subjects** with regard to how frequently the factor or attribute is present in each of these groups.

General: Study Design

Cross-Sectional ..... **Case-Control** ..... Cohort ..... RCT

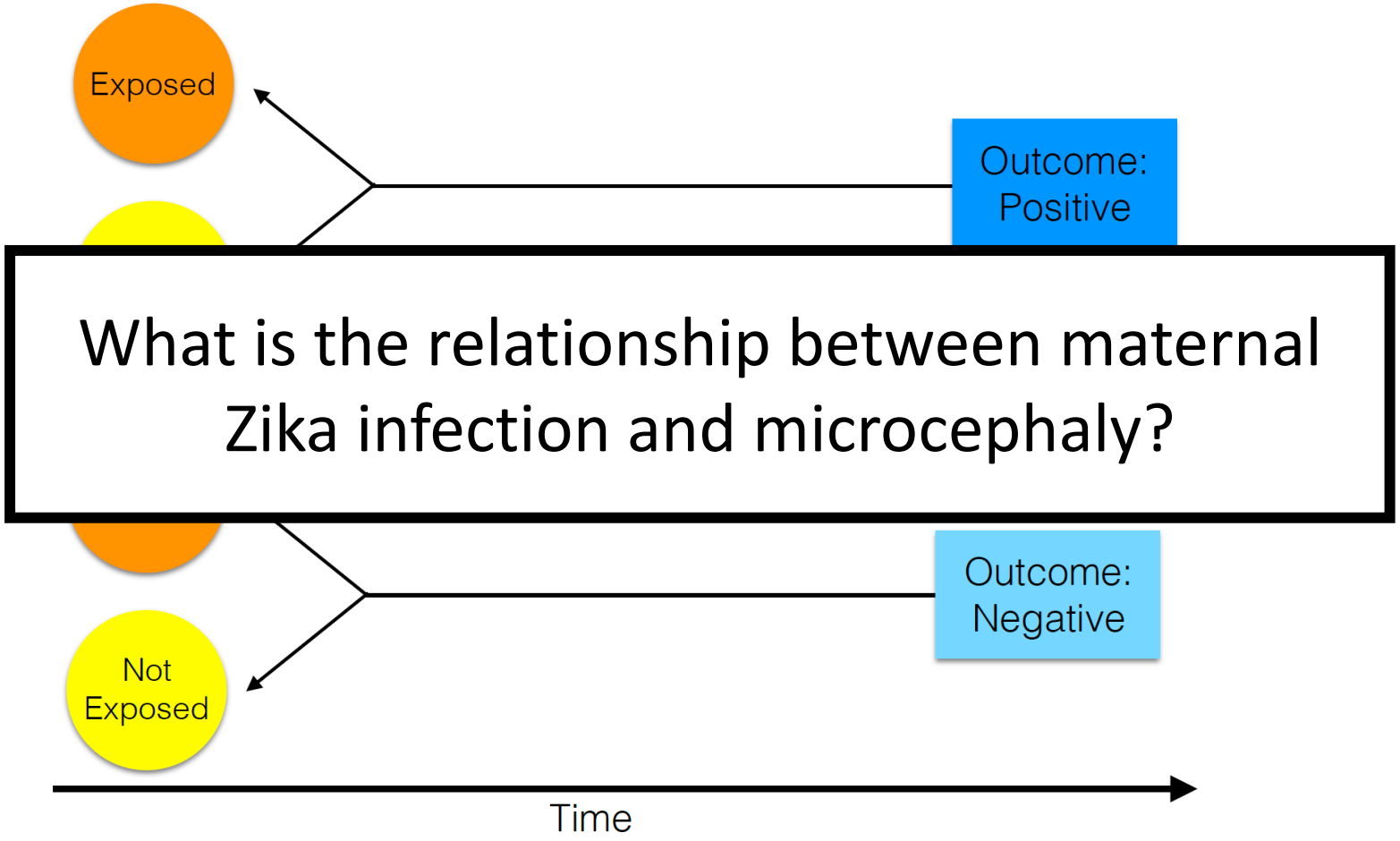
Case Control Study



General: Study Design

Cross-Sectional ..... **Case-Control** ..... Cohort ..... RCT

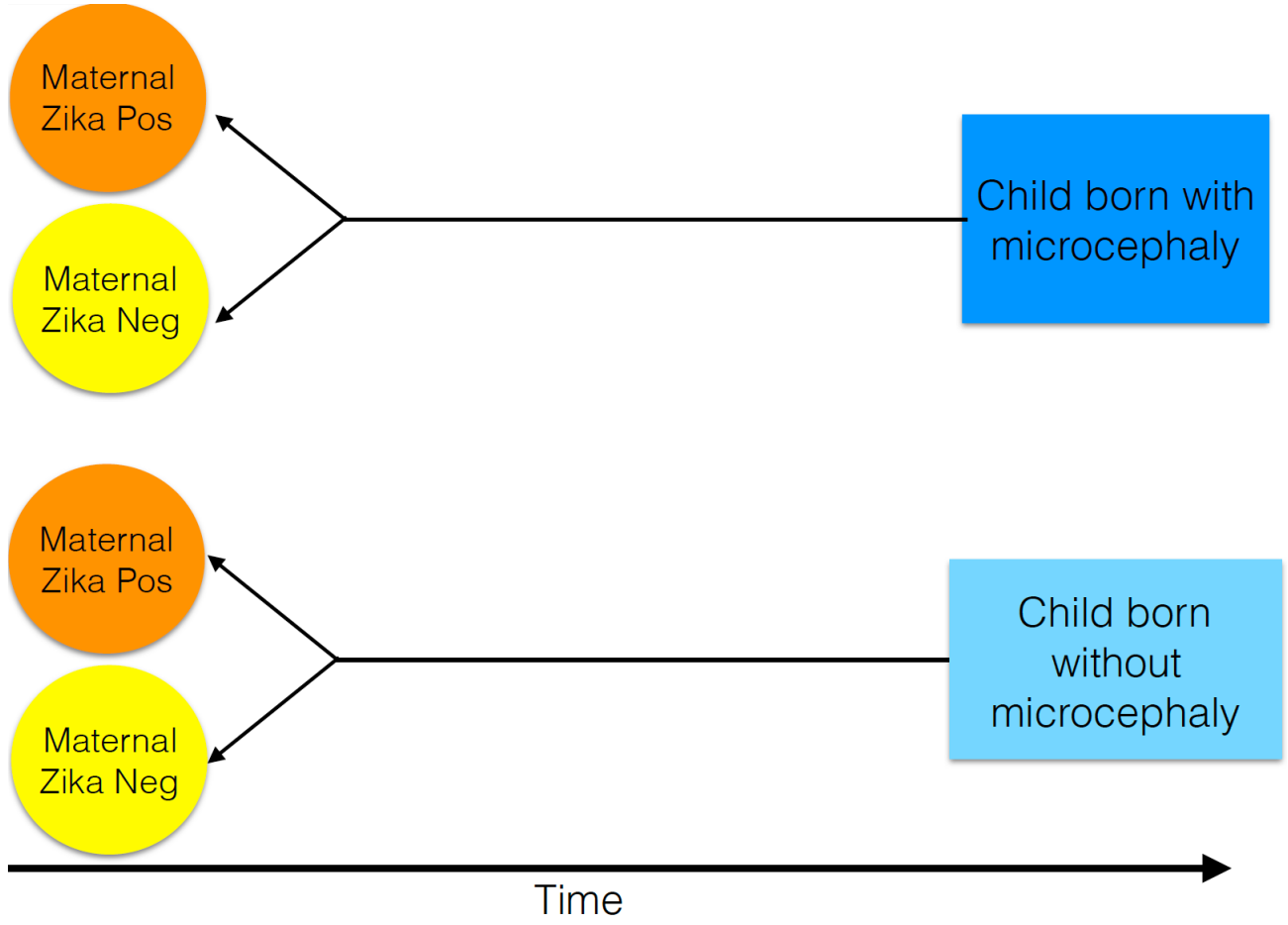
Case Control Study



General: Study Design

Cross-Sectional — Case-Control — Cohort — RCT

Case Control Study

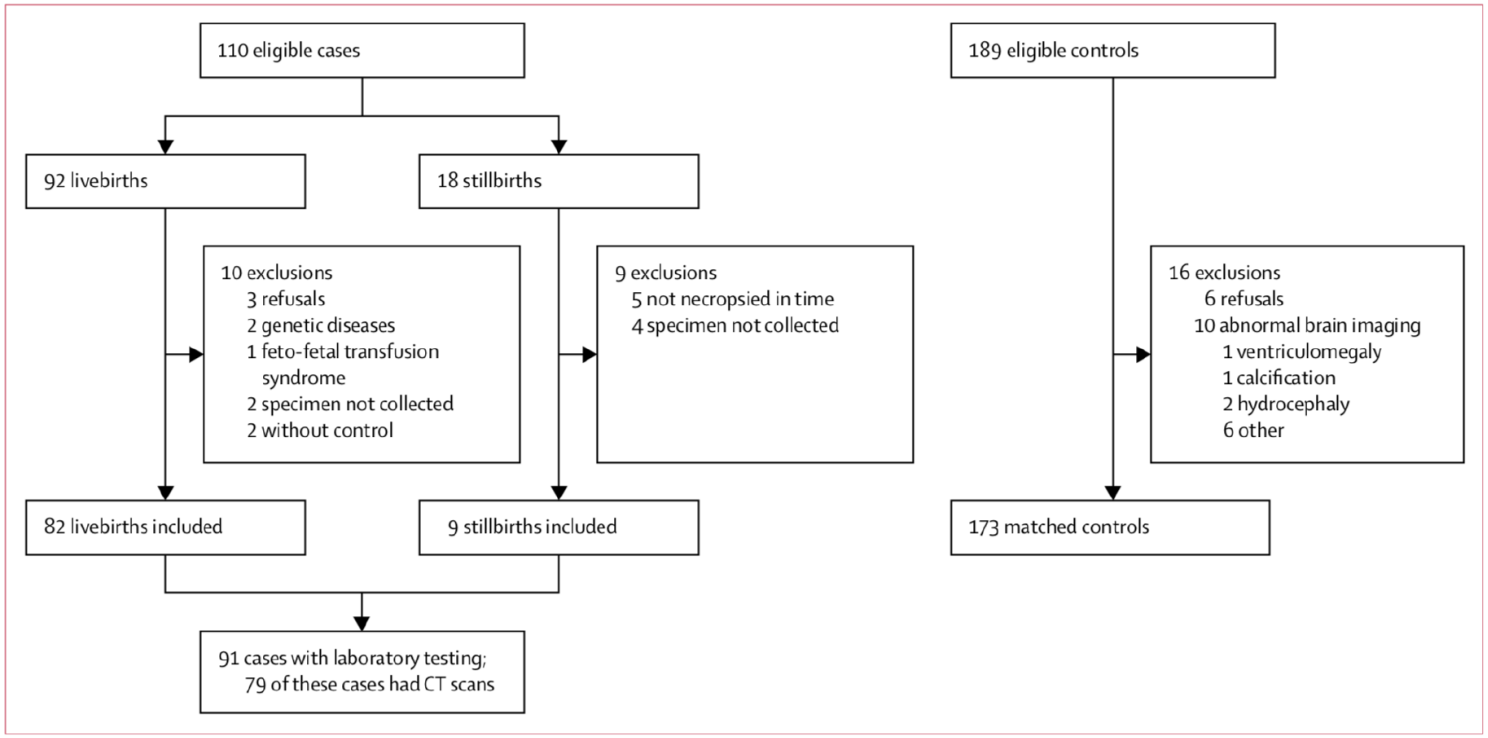


Zika: Study Design

Cross-Sectional · Case-Control · Cohort · RCT

Case Control Study

Case-Control study - 8 Brazilian hospitals



de Araujo (2018)

Zika: Study Design

Cross-Sectional — Case-Control — Cohort — RCT

Case-Control study - 8 Brazilian hospitals

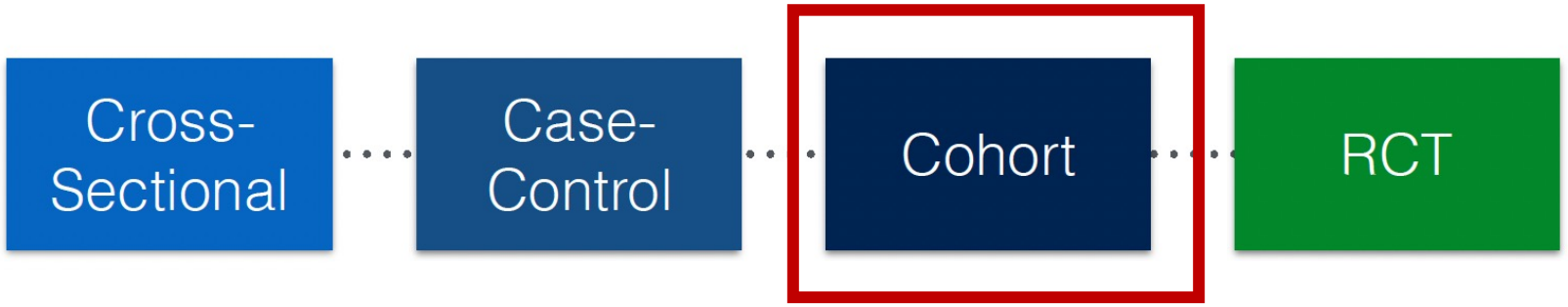
	Cases*	Controls*	Matched odds ratio (95% CI)
Serum, CSF samples, or macerated tissue			
Zika-positive, of total cases or controls	32/91 (35%)	0/173	87·0 (15·6–∞)
Zika-positive, of total cases or controls, adjusted†	..	..	73·1 (13·0–∞)
Cases, categorised by severity of microcephaly‡			
Severe	19/26 (73%)	0/51	52·4 (9·1–∞)
Not severe	13/65 (20%)	0/122	33·7 (5·6–∞)

*Data are the number of all cases or controls who were positive for Zika virus, assessed by qRT-PCR or Zika virus-specific IgM/total number of patients (%). †Odds ratio when adjusted by smoking during pregnancy, maternal vaccination against tetanus, diphtheria, and acellular pertussis during pregnancy, and skin colour. ‡Severe is defined as a head circumference of more than 3 SD smaller than the mean for their sex and gestational age.[10,14] Not severe was defined as a head circumference of 2–3 SD smaller than the mean for their sex and gestational age. Matched odds ratios in this subgroup are crude because of small numbers.

Table 5: **Association between microcephaly and Zika virus infection**

de Araujo (2018)

General: Study Design



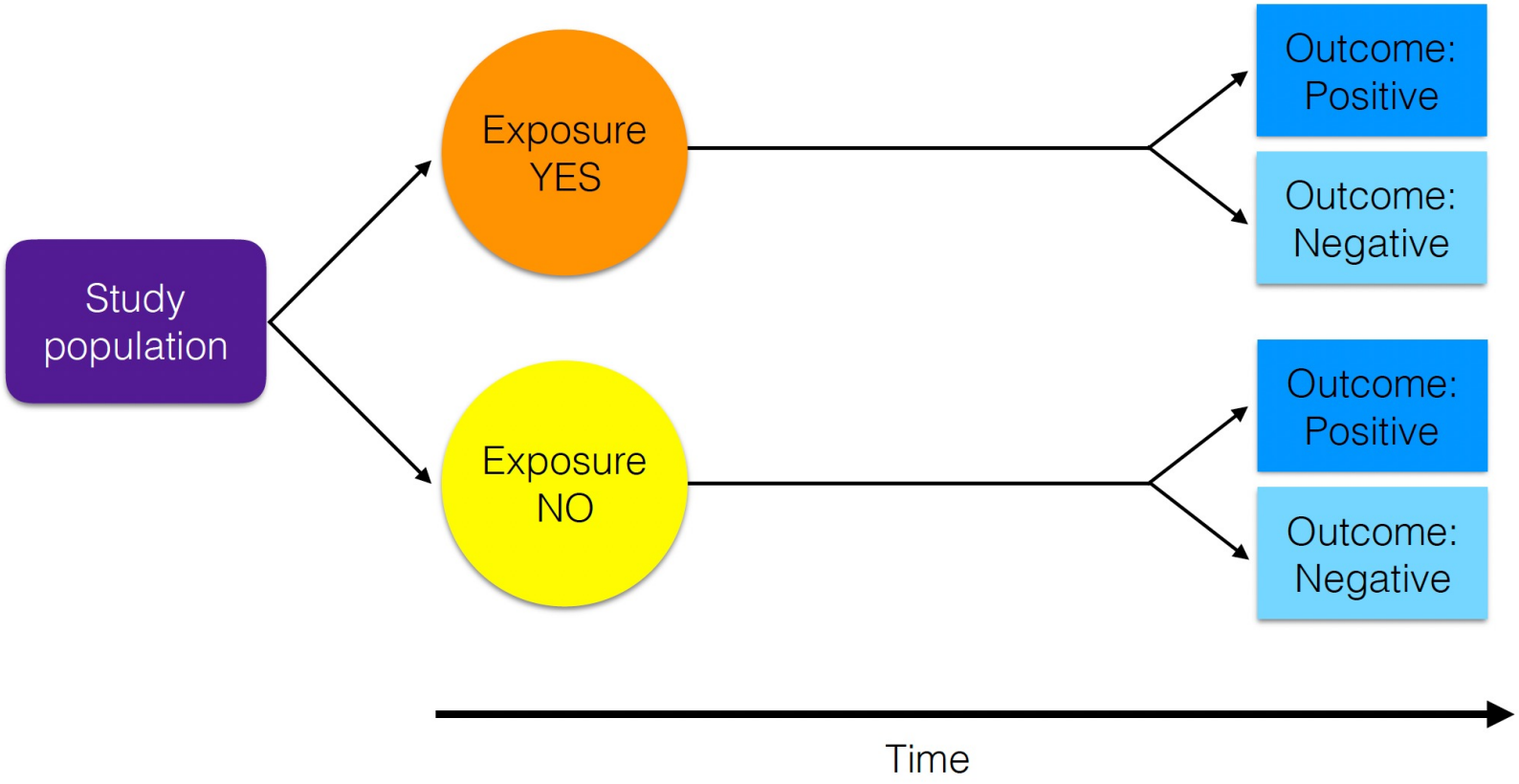
Cohort Study

- Enrolled group of people who have a common experience or grouping.
- Age cohort, risk cohort
- General population sample
- Clinic based
- Prospective or retrospective

General: Study Design

Cross-Sectional ..... Case-Control ..... **Cohort** ..... RCT

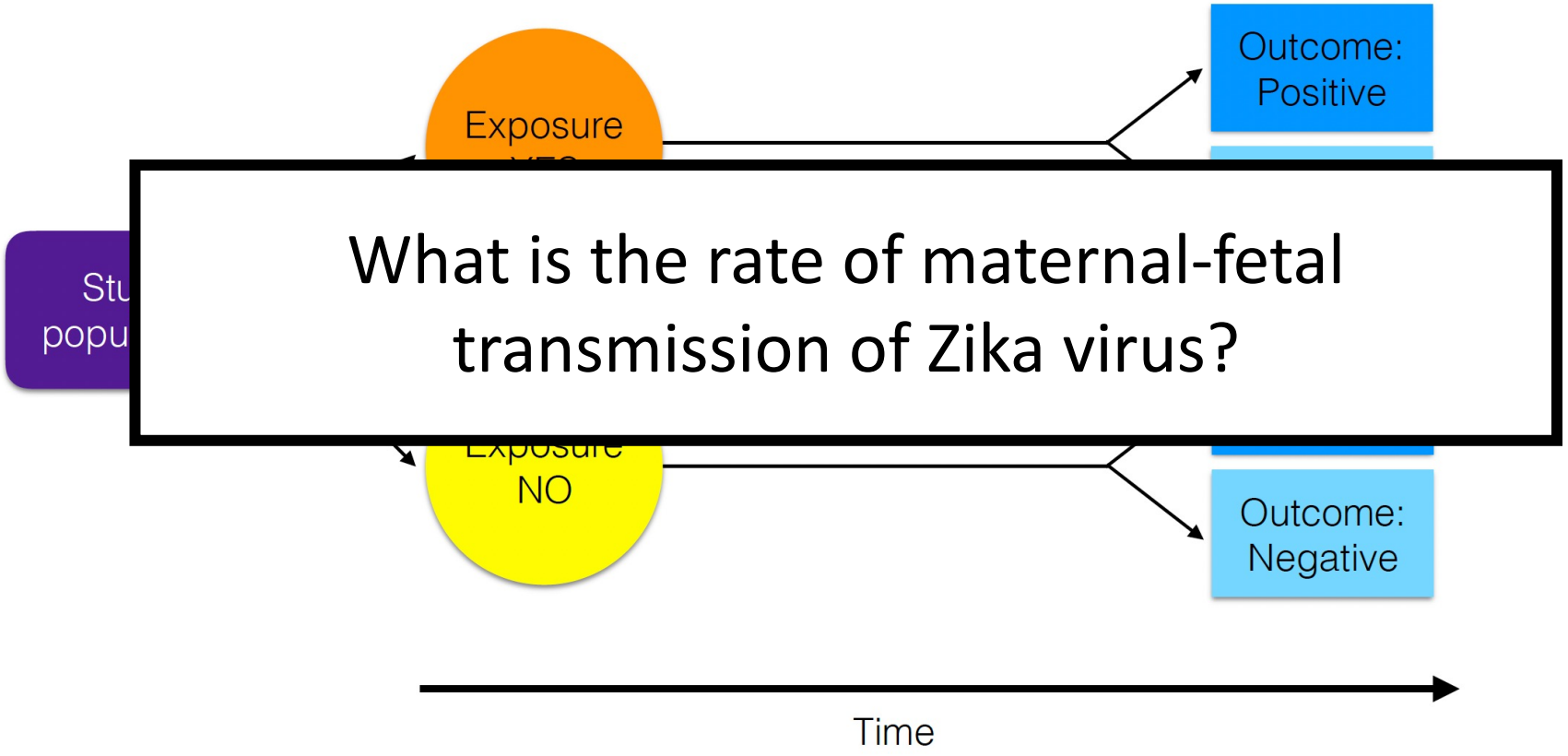
Cohort Study



Zika Virus: The Basics

Cross-Sectional ..... Case-Control ..... Cohort ..... RCT

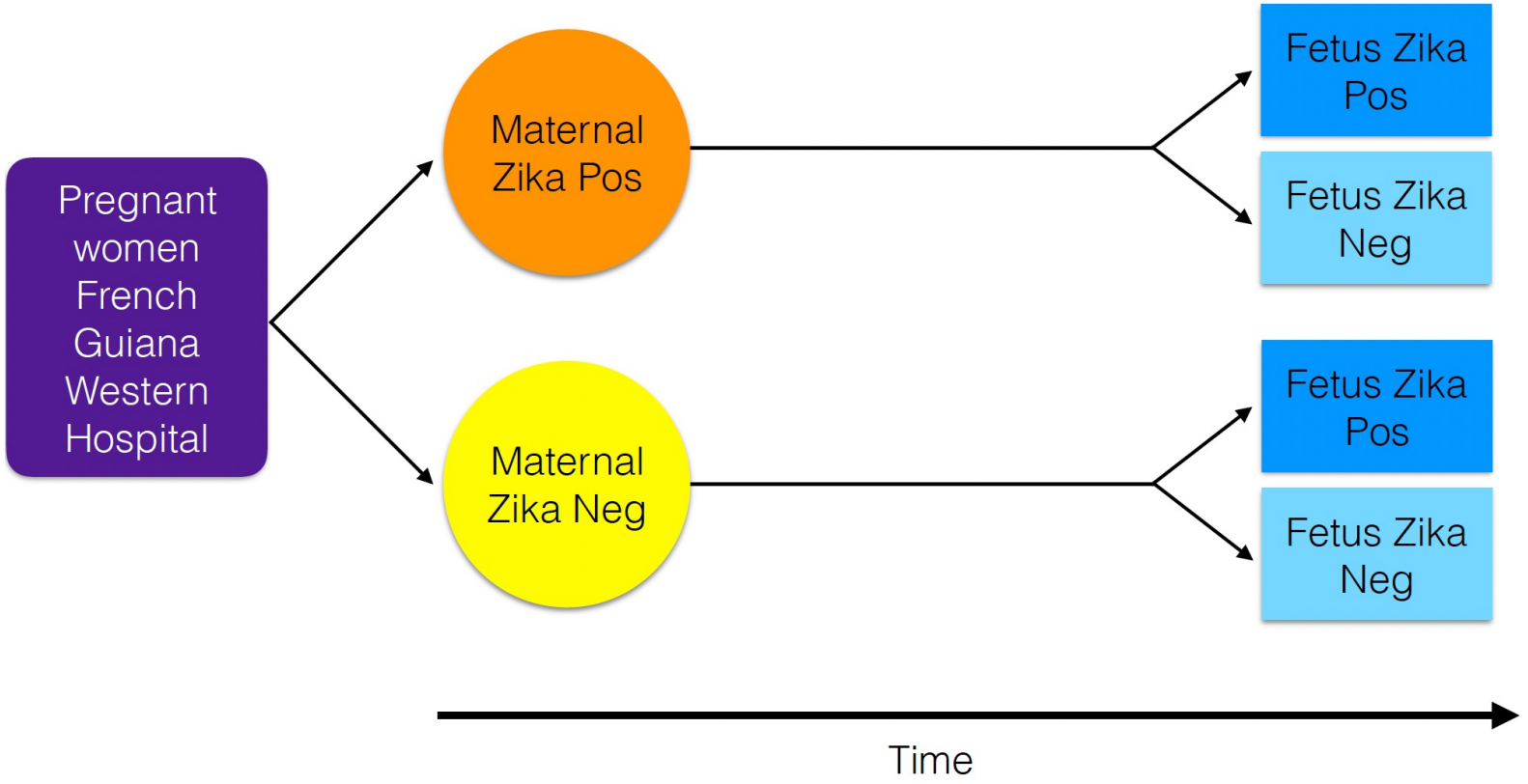
Cohort Study



Zika: Study Design

Cross-Sectional ..... Case-Control ..... **Cohort** ..... RCT

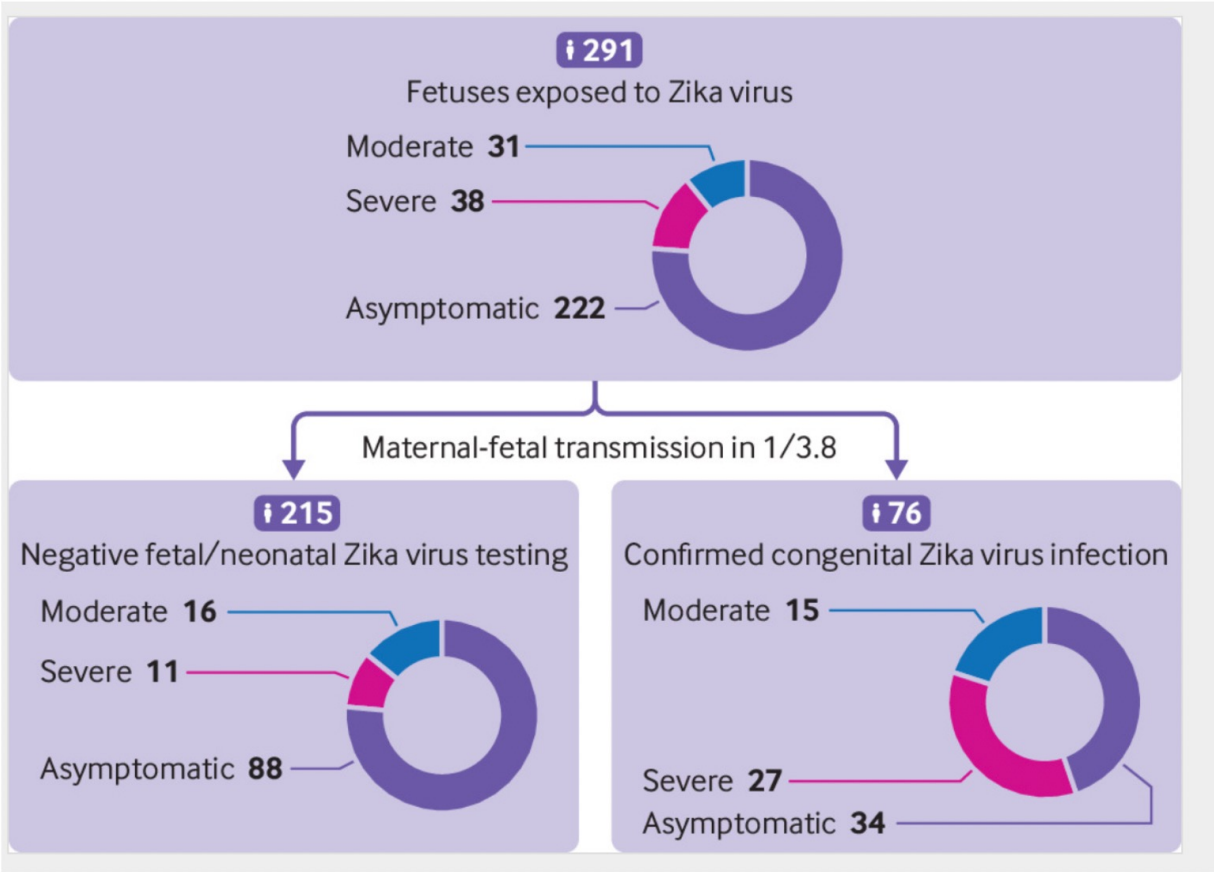
Cohort Study



Zika: Study Design

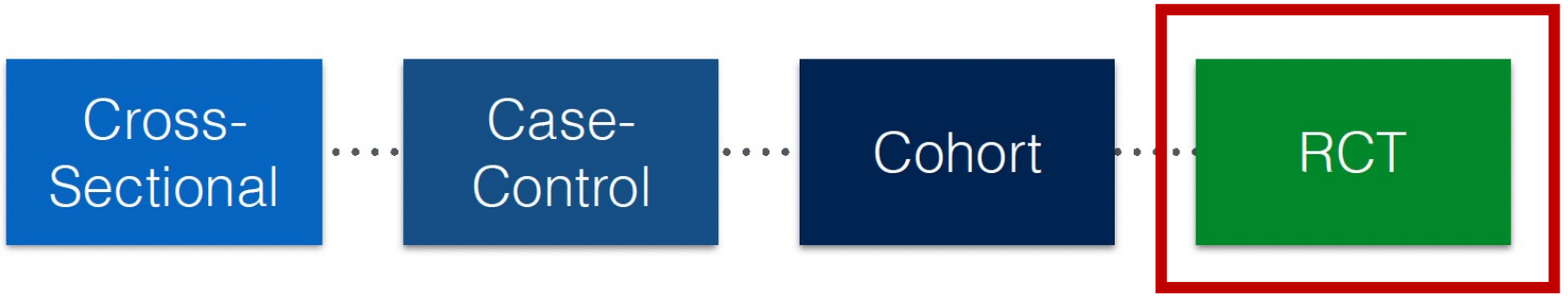
Cross-Sectional — Case-Control — **Cohort** — RCT

Cohort Study



Pomar (2018)

General: Study Design



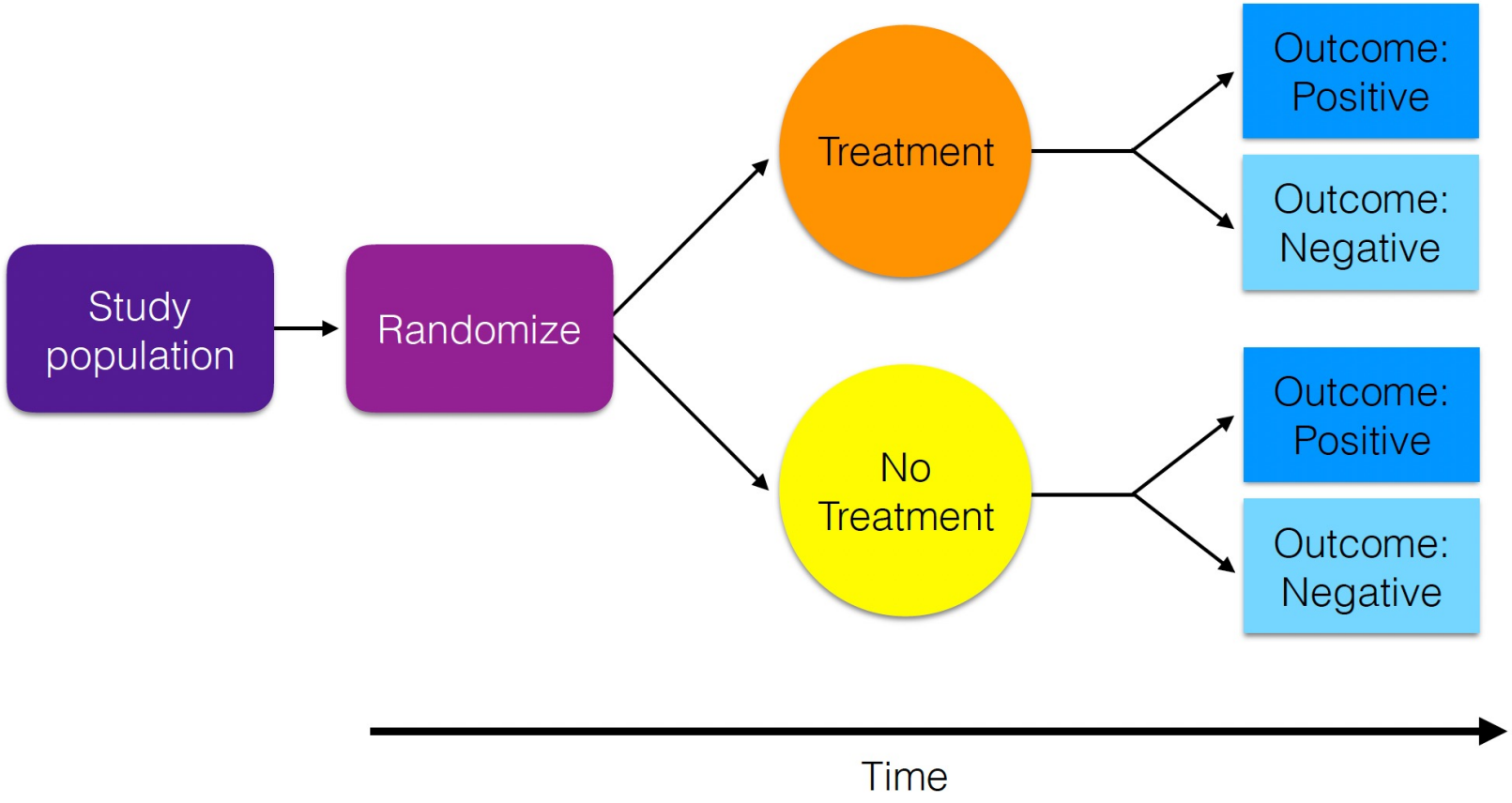
Randomized Controlled Trial

- Experiment in which subjects are **randomly allocated into groups** (test and control that are comparable) **to receive or not to receive** a preventative or a therapeutic procedure or intervention.
- Results are assessed by **comparison of rates** of disease, death, recovery, or other outcome **in the study groups**.
- Generally thought of as **the most rigorous method** of hypothesis testing.
- Randomization should be blinded!

General: Study Design

Cross-Sectional ..... Case-Control ..... Cohort ..... RCT

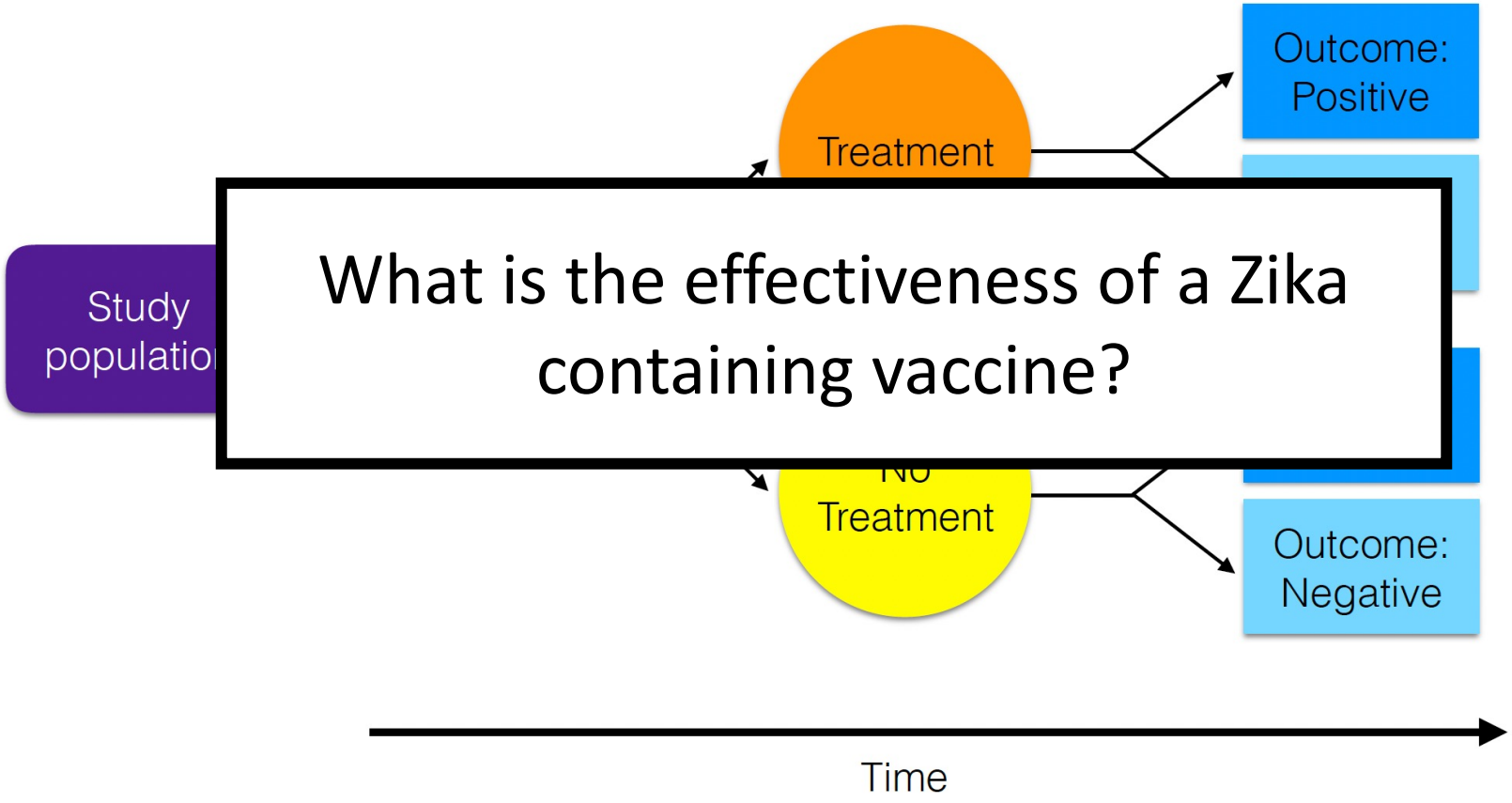
Randomized Controlled Trial



Zika: Study Design

Cross-Sectional ···· Case-Control ···· Cohort ···· RCT

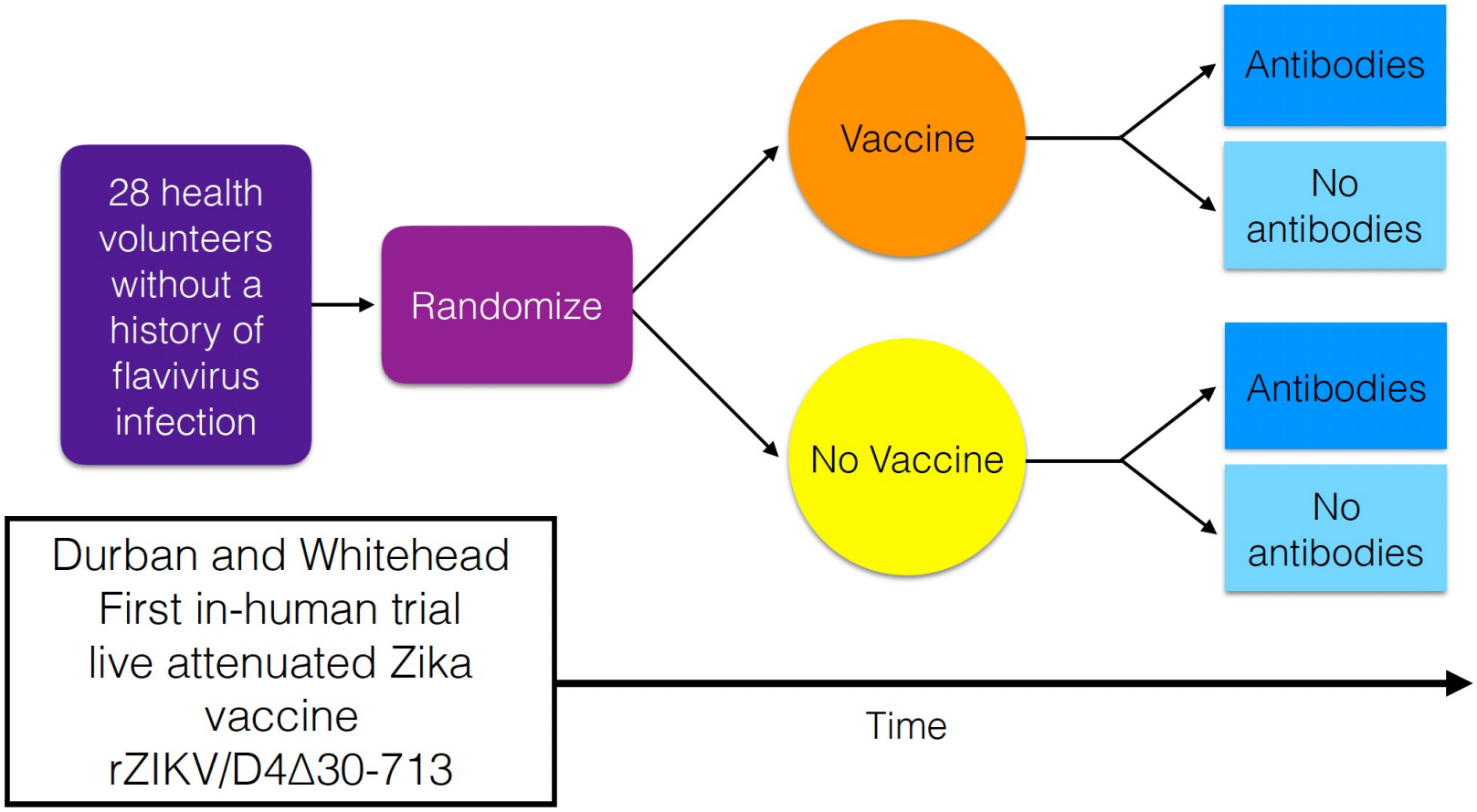
Randomized Controlled Trial



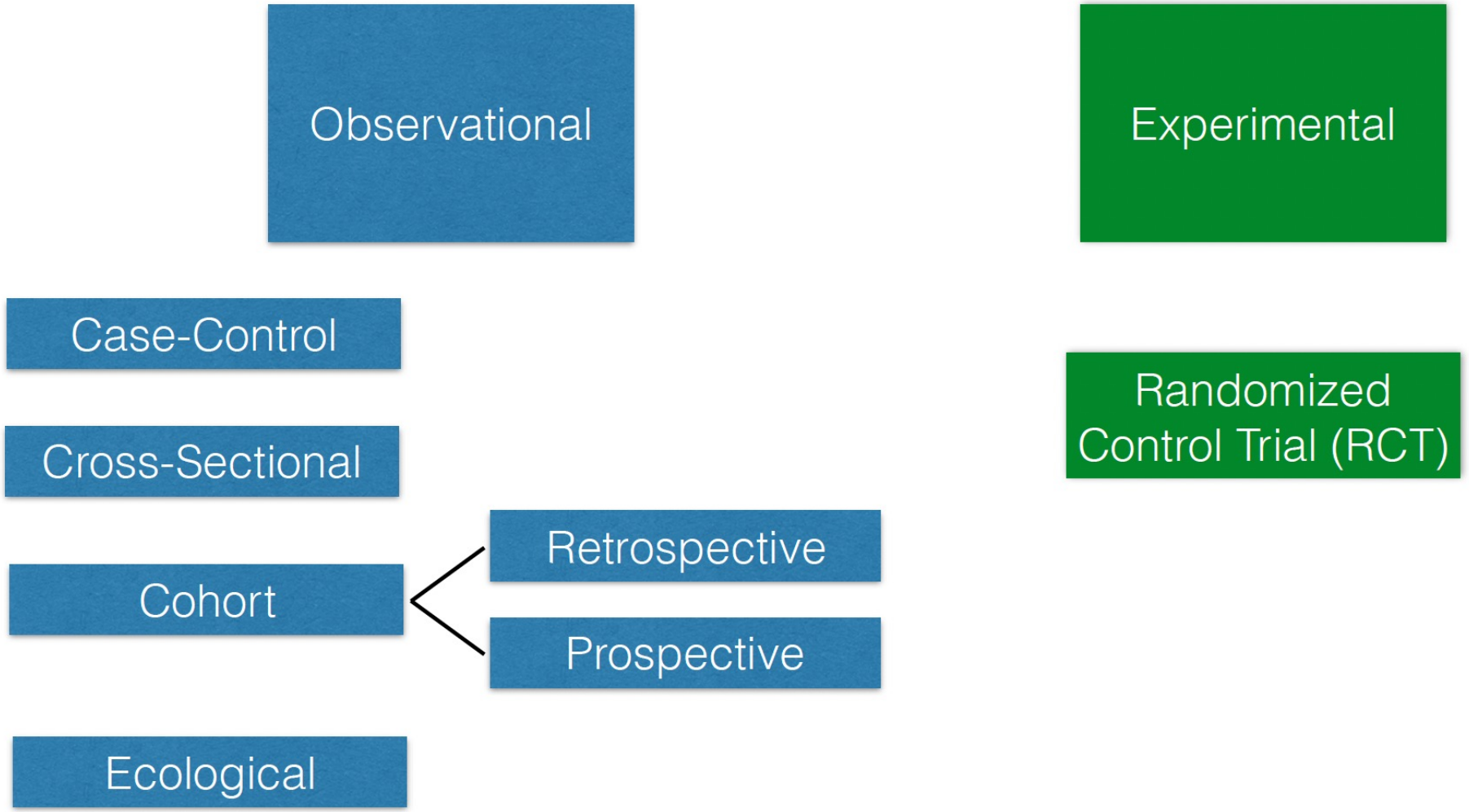
Zika: Study Design

Cross-Sectional ..... Case-Control ..... Cohort ..... RCT

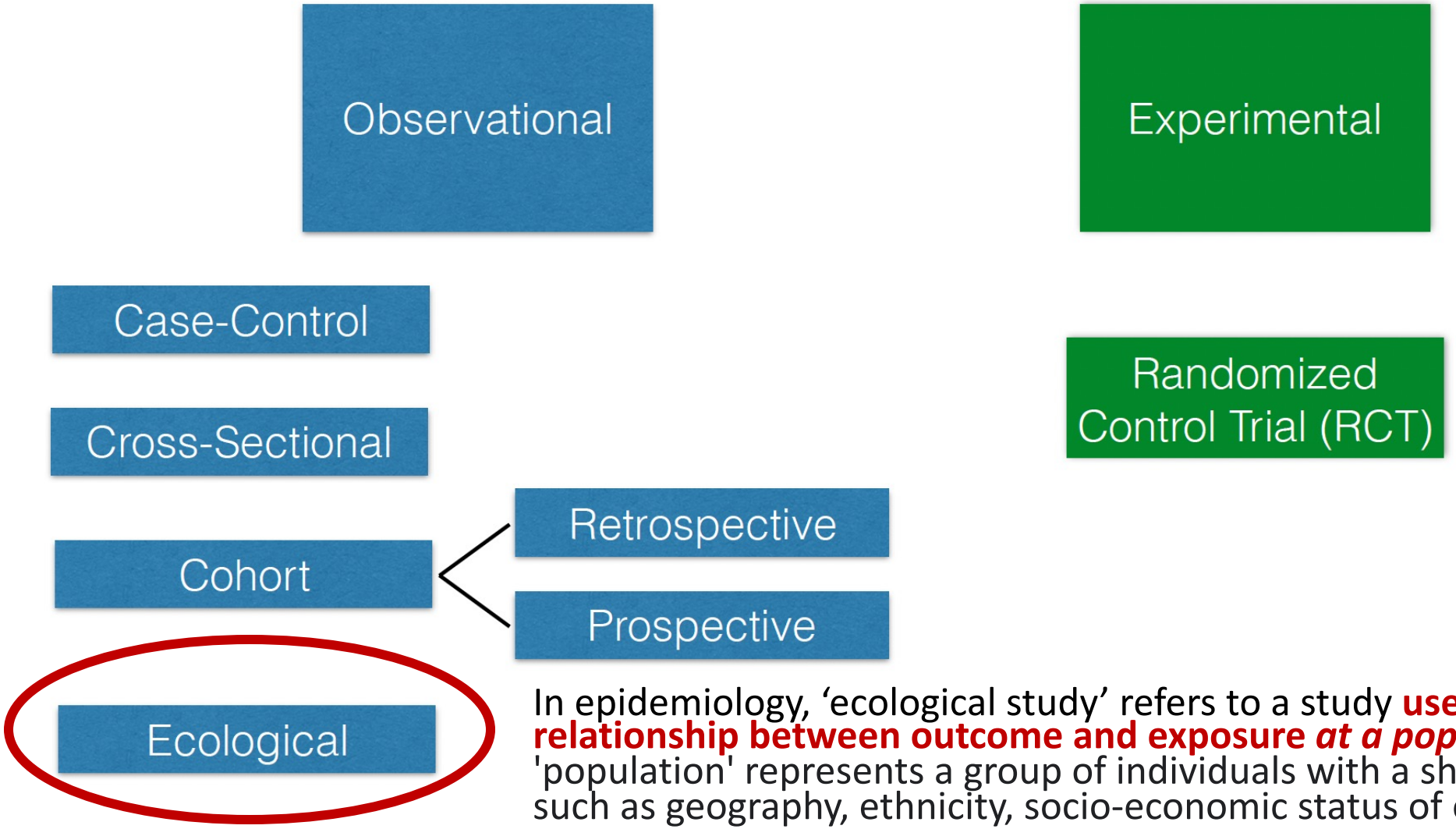
Randomized Controlled Trial



Different Study Types in Epidemiology



Different Study Types in Epidemiology



In epidemiology, 'ecological study' refers to a study **used to understand the relationship between outcome and exposure *at a population level***, where 'population' represents a group of individuals with a shared characteristic such as geography, ethnicity, socio-economic status of employment.

Ex: What is the correlation between prevalence of TB and country-level GDP?

Ecological Study (Sampling) Design

Observational

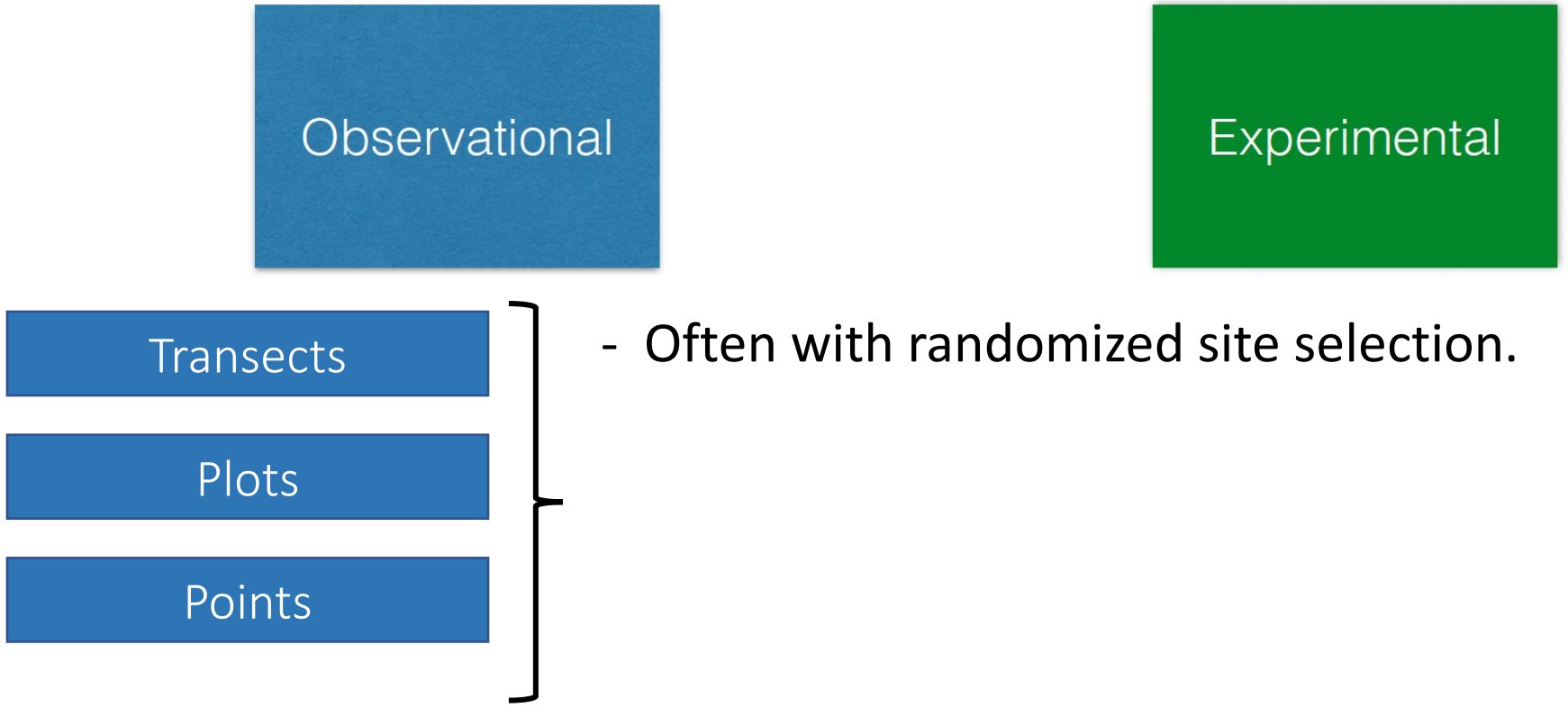
Experimental

Transects

Plots

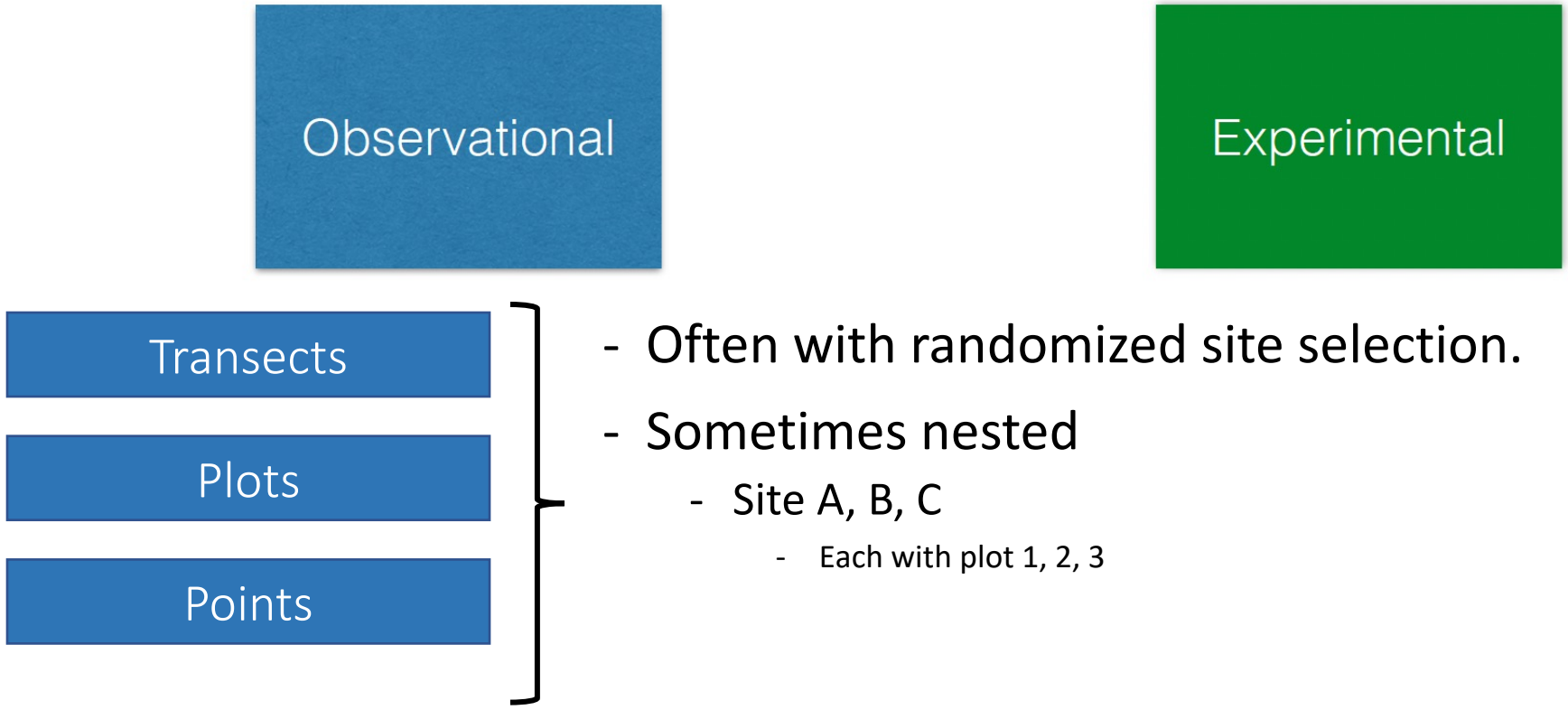
Points

Ecological Study (Sampling) Design



- Often with randomized site selection.

Ecological Study (Sampling) Design



- Often with randomized site selection.
- Sometimes nested
  - Site A, B, C
    - Each with plot 1, 2, 3

Ecological Study (Sampling) Design

Observational

Experimental

Transects

Plots

Points

- Often with randomized site selection.
- Sometimes nested
  - Site A, B, C
    - Each with plot 1, 2, 3
- Or, sometimes in a paired design.
  - 3 x pairs of (a) degraded vs. (b) pristine sites

Ecological Study (Sampling) Design

Observational

Experimental

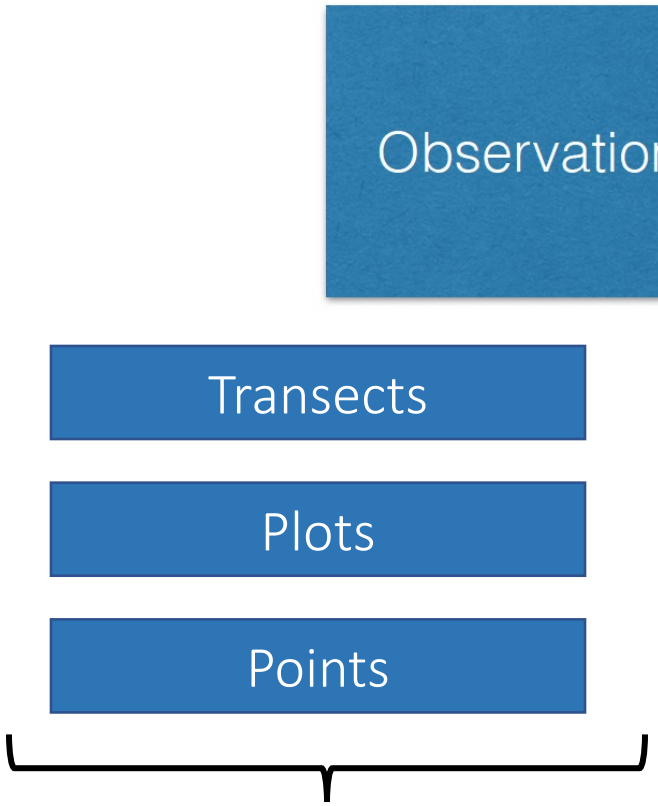
Transects

Plots

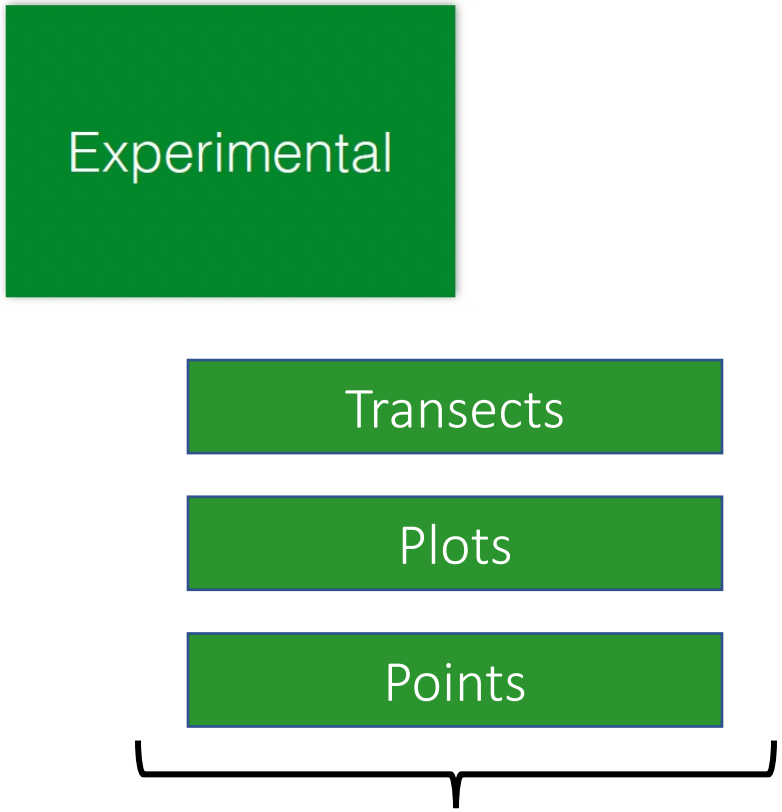
Points

- Often with randomized site selection.
- Sometimes nested
  - Site A, B, C
    - Each with plot 1, 2, 3
- Or, sometimes in a paired design.
  - 3 x pairs of (a) degraded vs. (b) pristine sites
- Replication is key!

Ecological Study (Sampling) Design

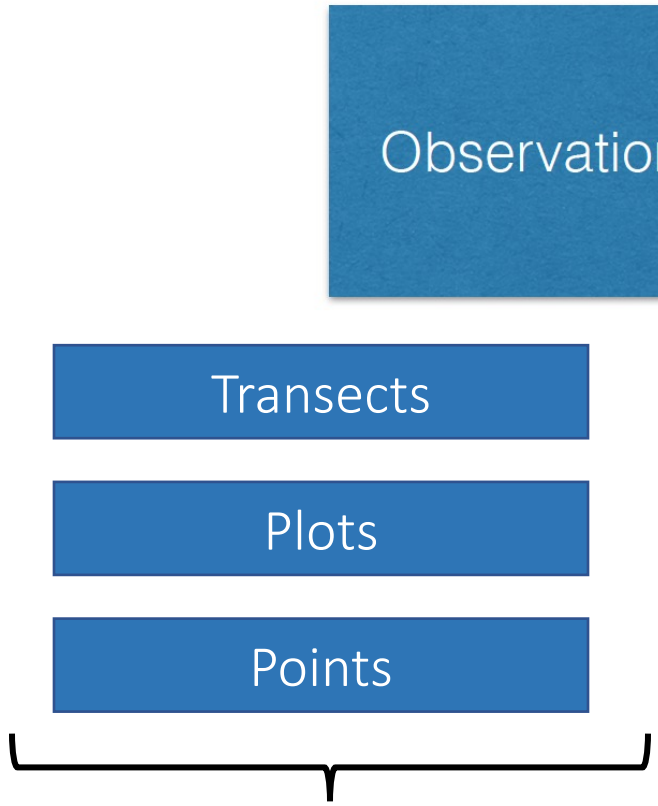


- Often with randomized site selection.
- Sometimes nested.
- Sometimes paired.
- Replication is key!

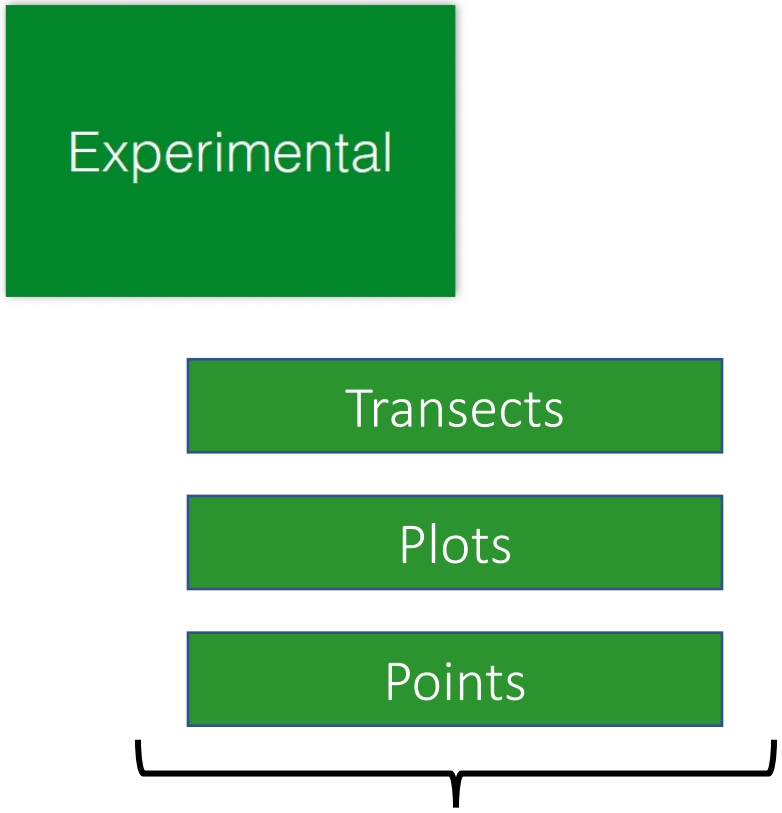


- Ecological manipulation

Ecological Study (Sampling) Design



- Often with randomized site selection.
- Sometimes nested.
- Sometimes paired.
- Replication is key!

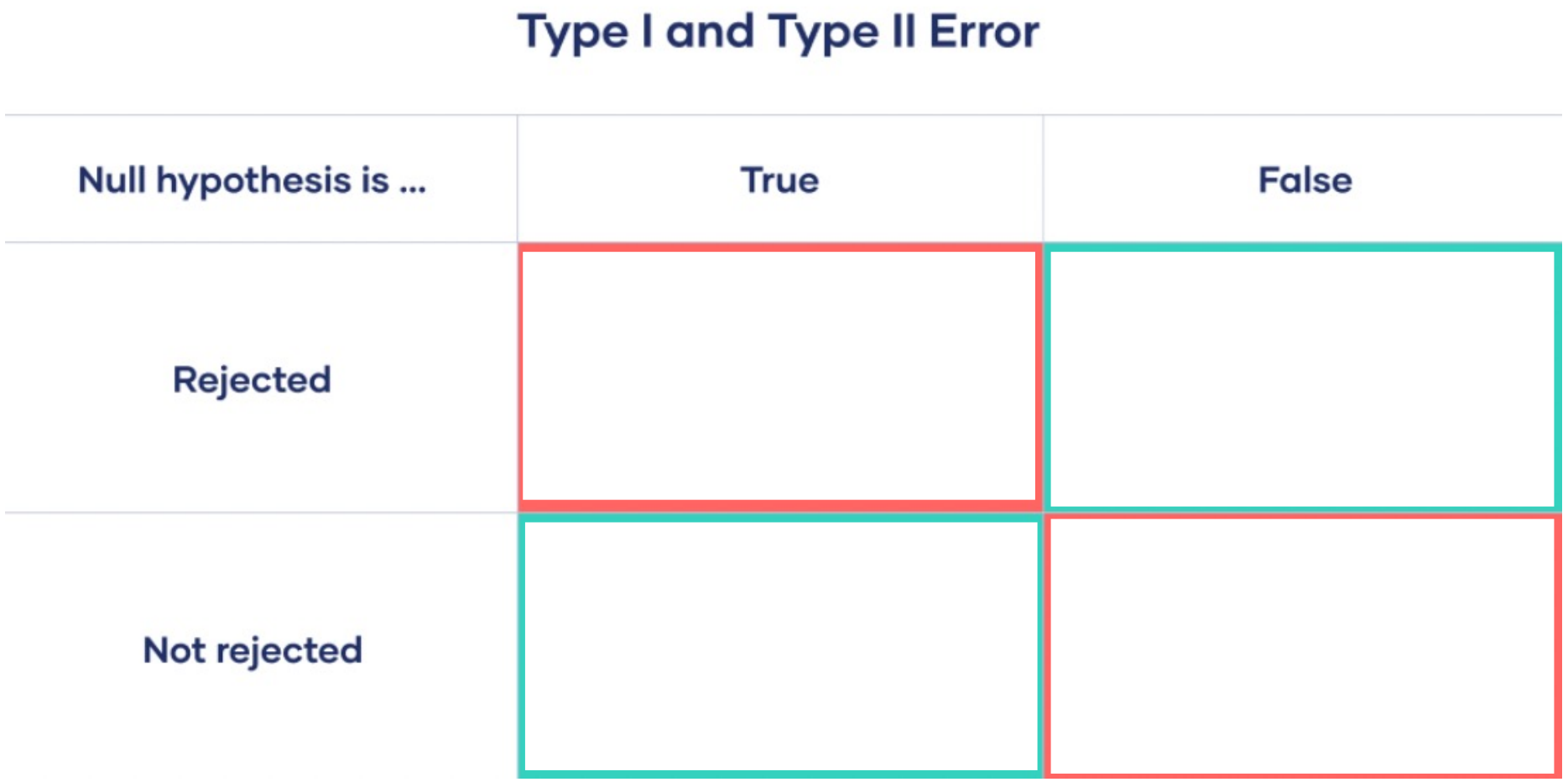


- Ecological manipulation

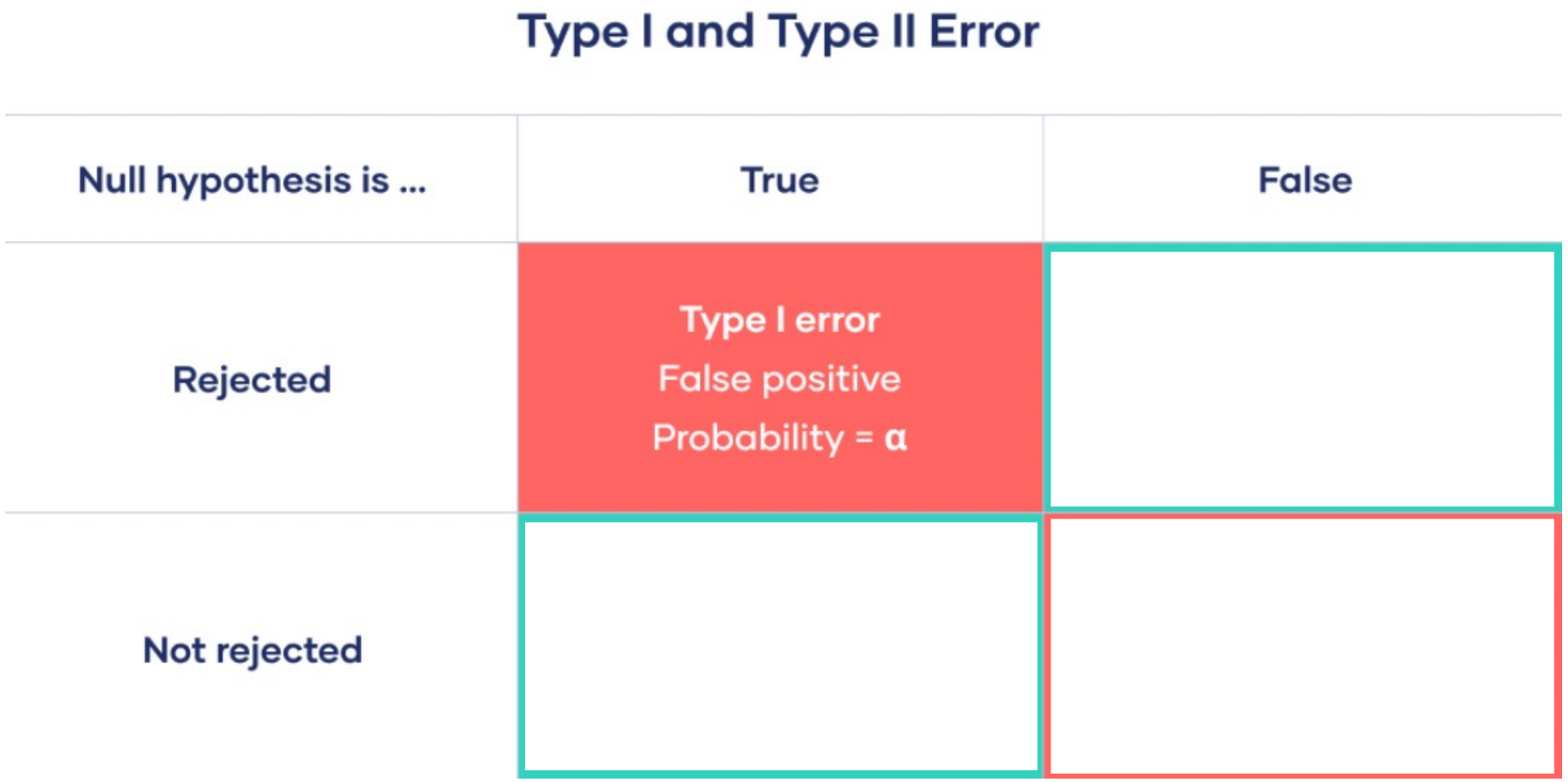
The most appropriate sampling design will depend on your study system and research question. *You should simulate data before carrying out your study to test your analytical approach.*

Power Analysis for Statistical Inference

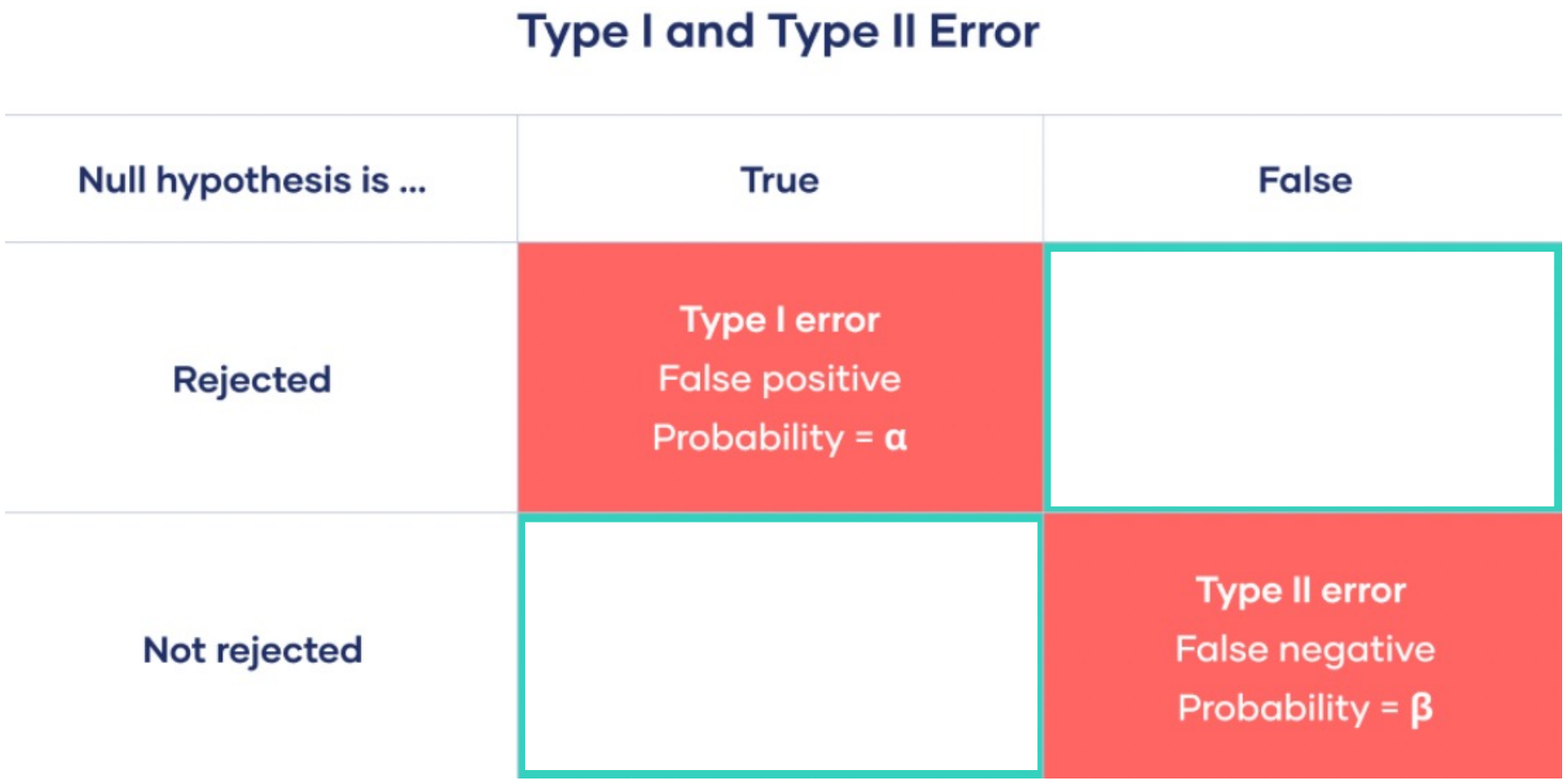
Power Analysis for Statistical Inference



Power Analysis for Statistical Inference



Power Analysis for Statistical Inference



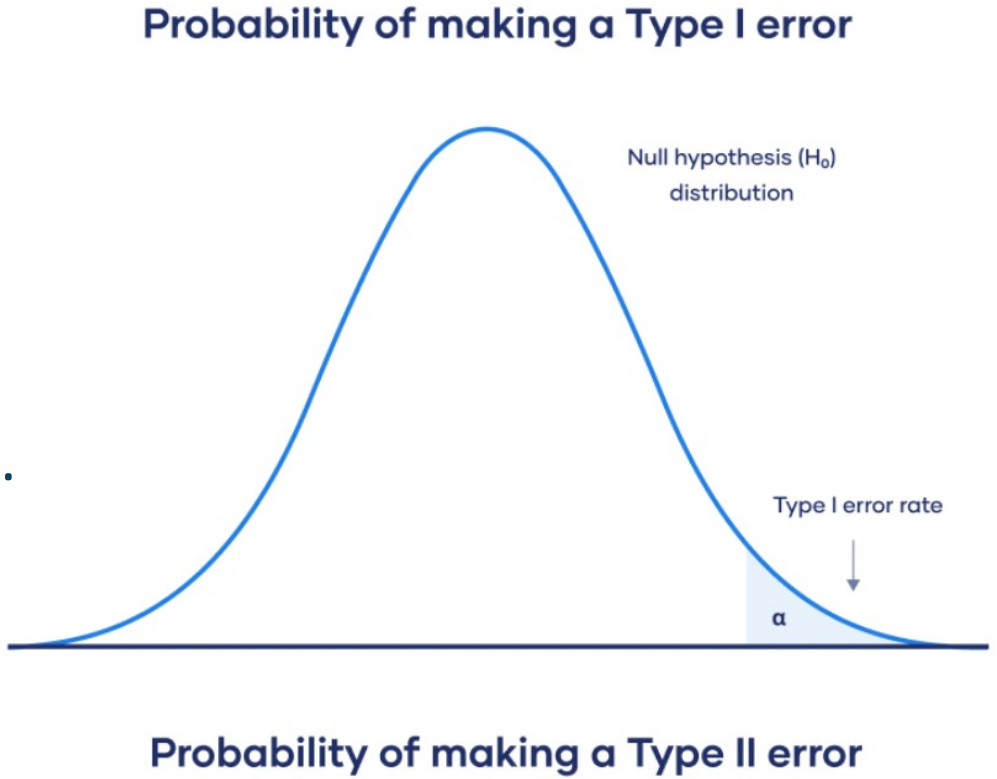
Power Analysis for Statistical Inference

Type I and Type II Error

Null hypothesis is ...	True	False
Rejected	Type I error False positive Probability = α	Correct decision True positive Probability = $1 - \beta$
Not rejected	Correct decision True negative Probability = $1 - \alpha$	Type II error False negative Probability = β

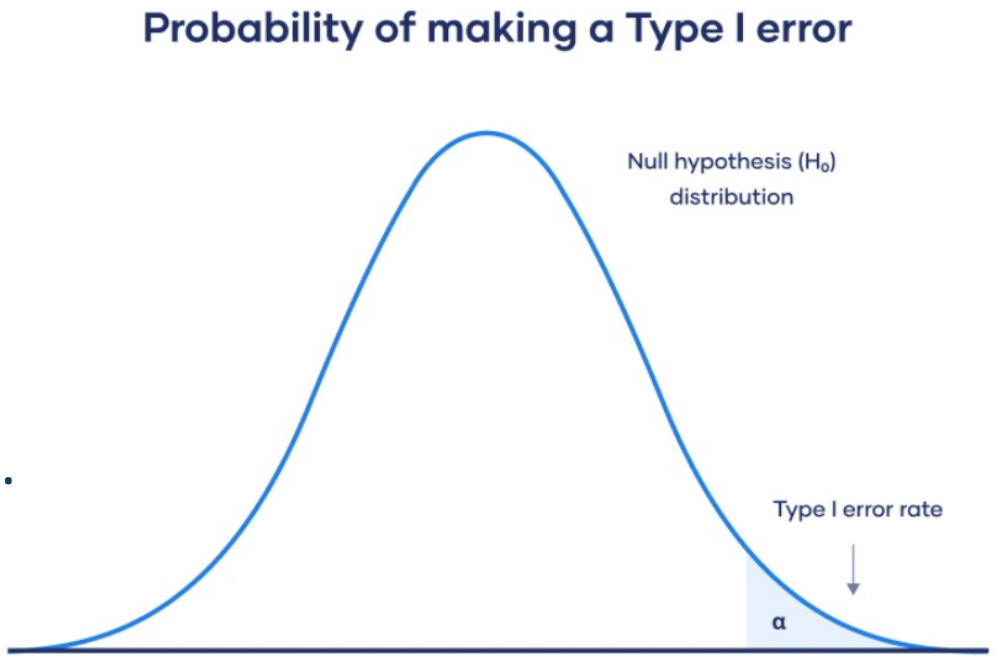
Power Analysis for Statistical Inference

- Type I error =false positive
  - rejecting the null hypothesis when it's actually true.
  - *the test result says you have coronavirus, but you actually don't.*

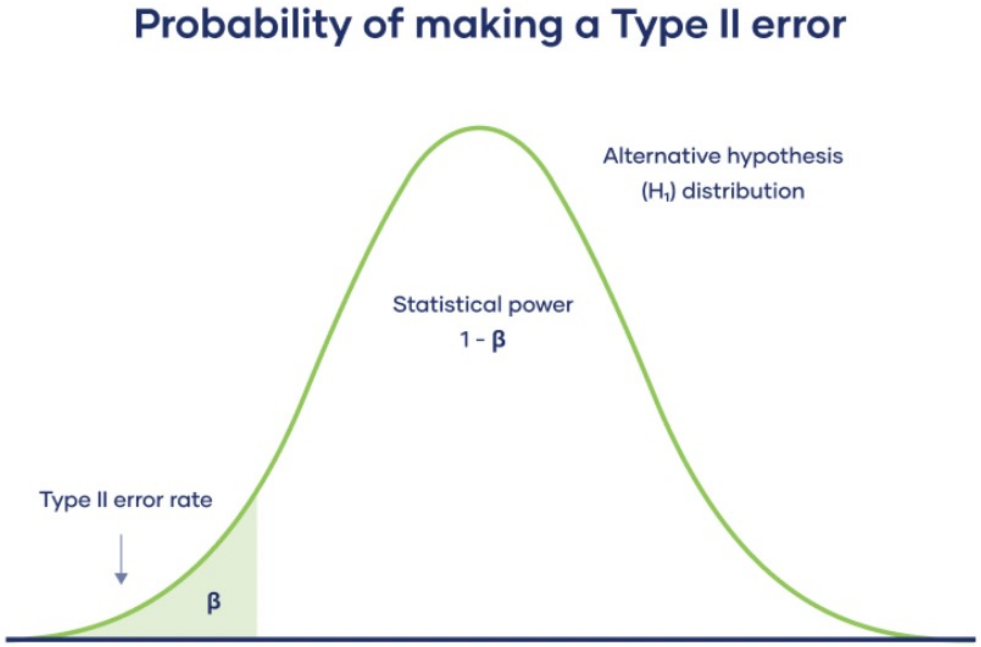


Power Analysis for Statistical Inference

- Type I error =false positive
  - rejecting the null hypothesis when it's actually true.
  - *the test result says you have coronavirus, but you actually don't.*



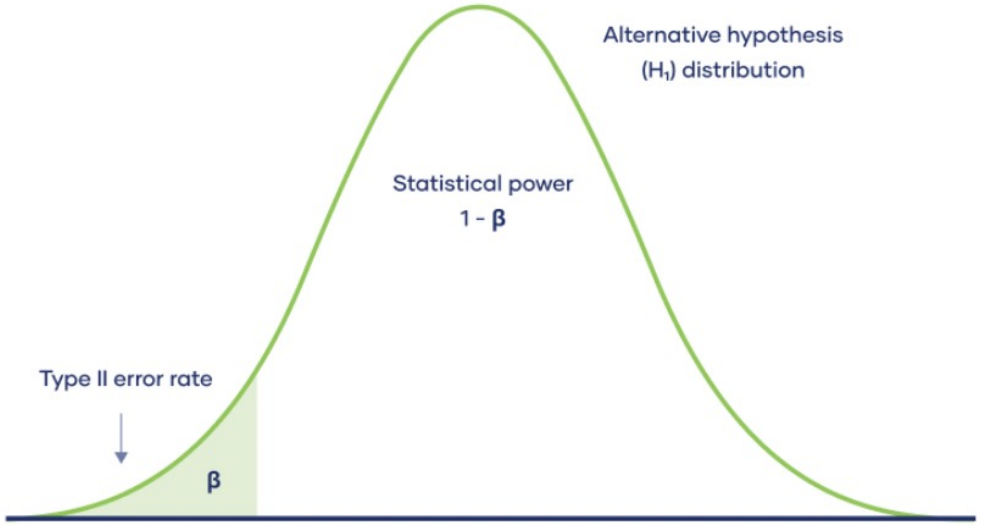
- Type II error = false negative
  - failing to conclude there was an effect when there actually was
  - *the test result says you don't have coronavirus, but you actually do.*



Power Analysis for Statistical Inference

Power analysis investigates the probability of detecting an effect if it is actually there.

- Type II error = false negative
  - failing to conclude there was an effect when there actually was
  - *the test result says you don't have coronavirus, but you actually do.*

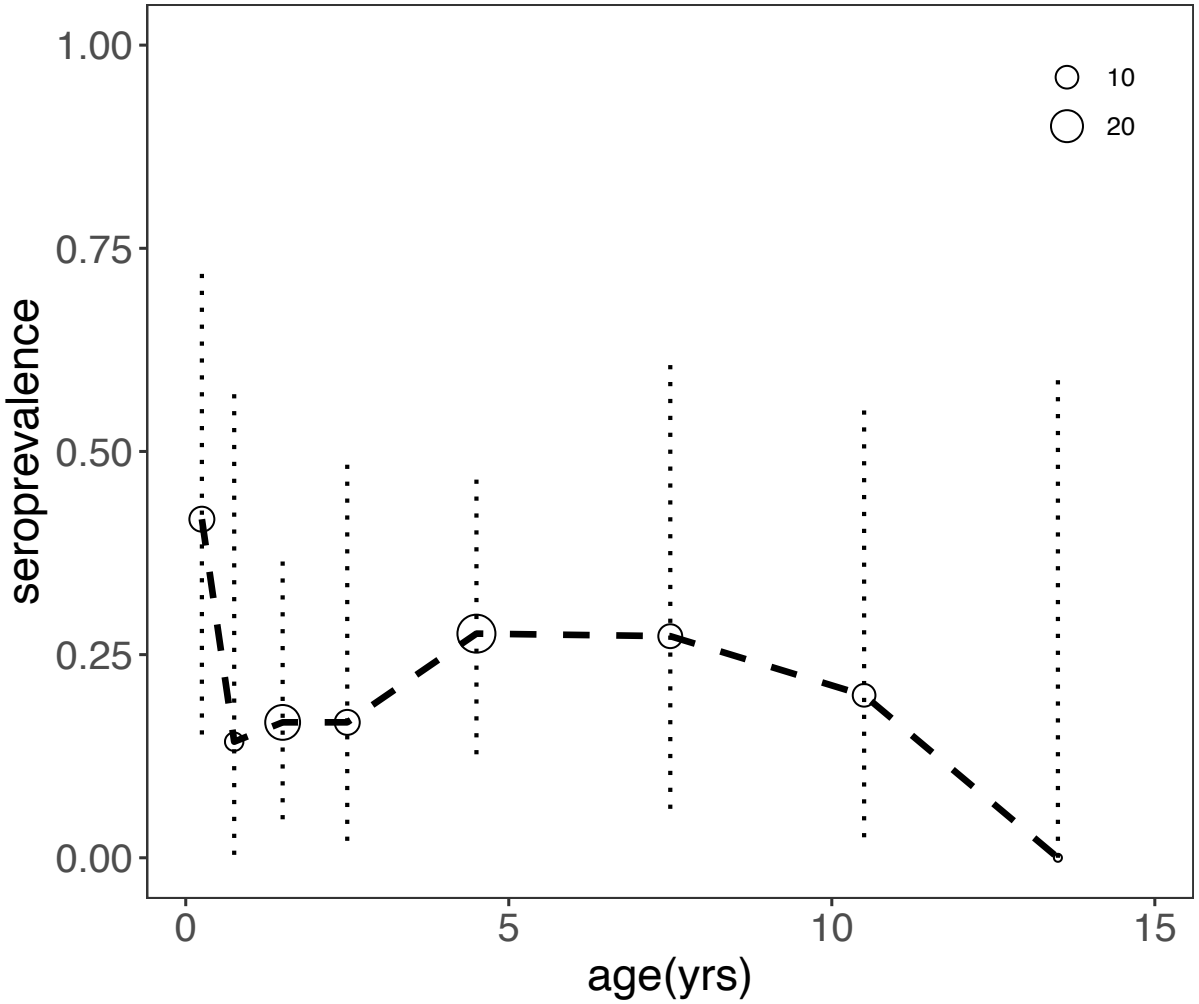


Power Analysis in R

- Program ‘pwr’ for statistical models
- Simulation and evaluation for mechanistic models

Simulation and evaluation for mechanistic models

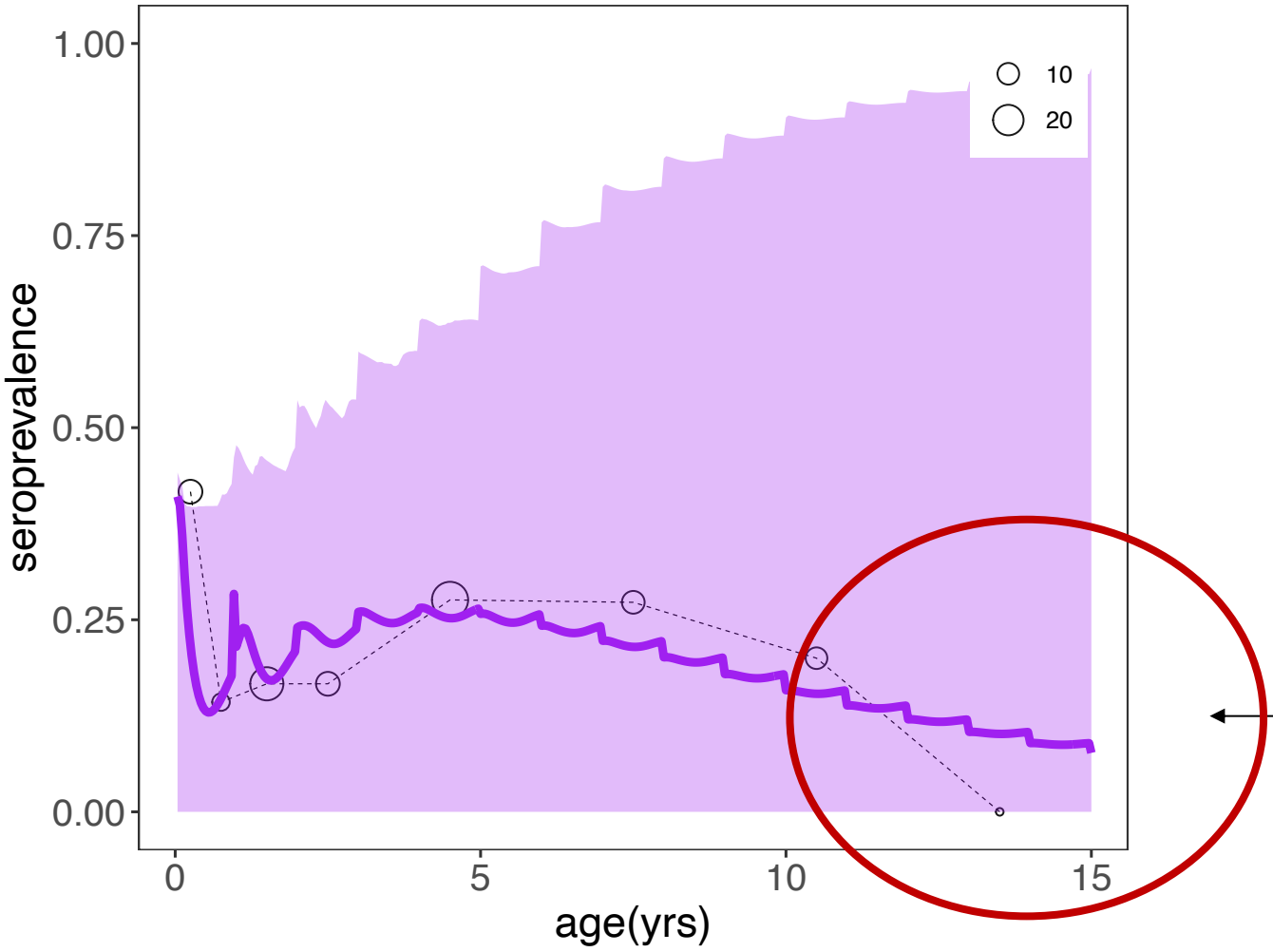
age-seroprevalence data for E. dupreanum Nipah virus



Brook et al. 2019 *J of Animal Ecology*

Simulation and evaluation for mechanistic models

fitted model to age-seroprevalence data for E. dupreanum Nipah virus

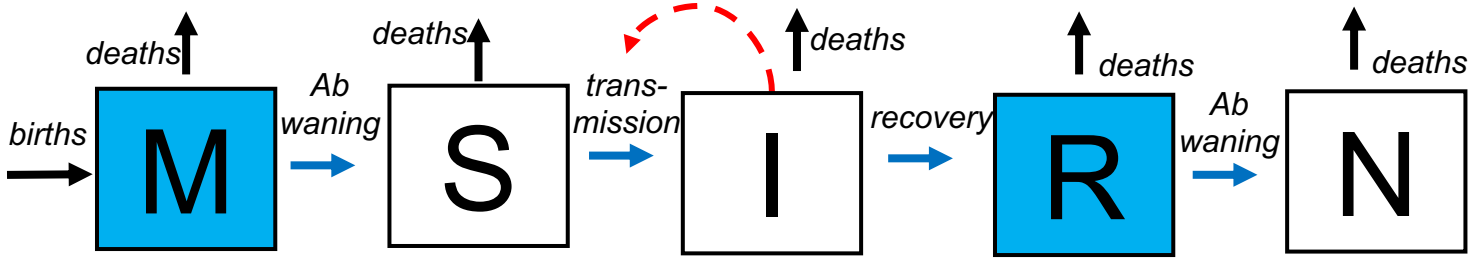


Fit could be better in **late age classes**….

Brook et al. 2019 *J of Animal Ecology*

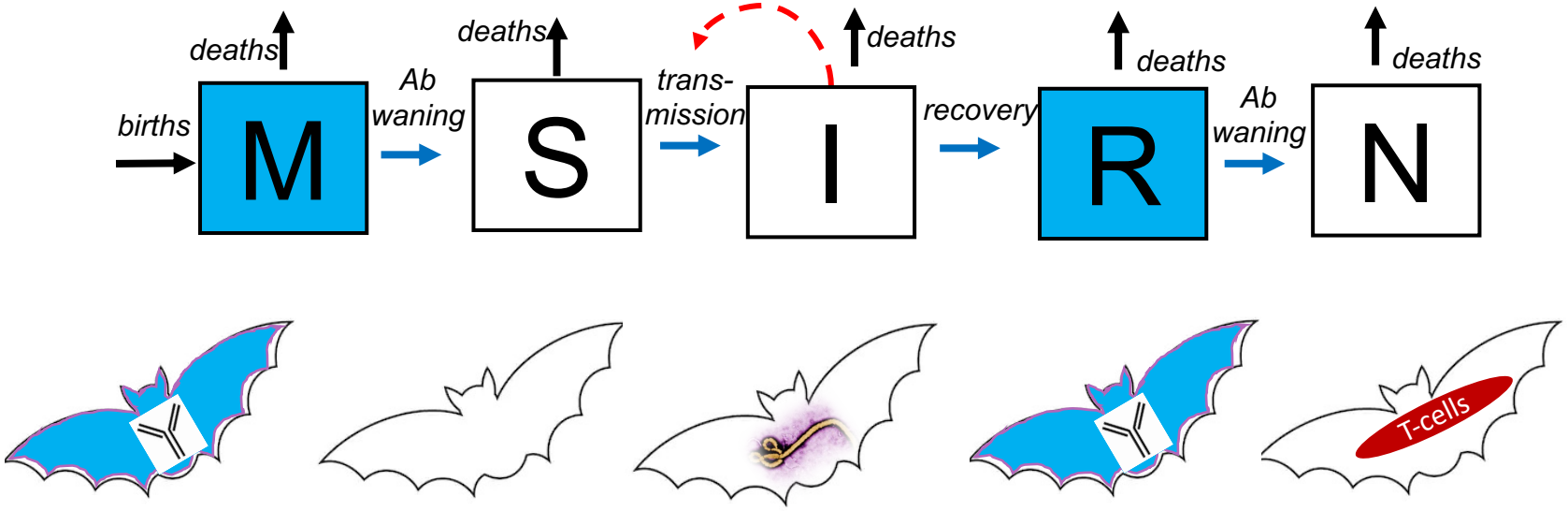
Simulation and evaluation for mechanistic models

Hyp 1: Lifelong Immunity

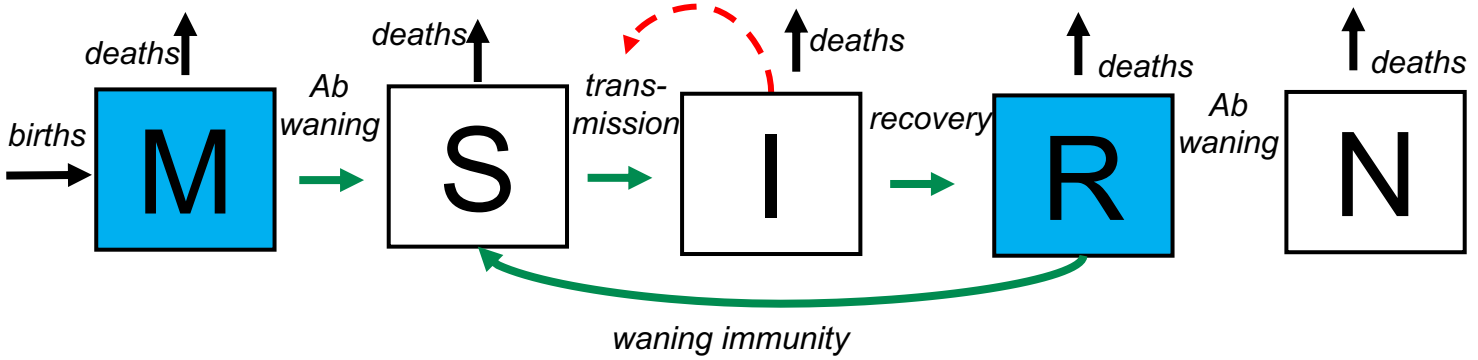


Simulation and evaluation for mechanistic models

Hyp 1: Lifelong Immunity



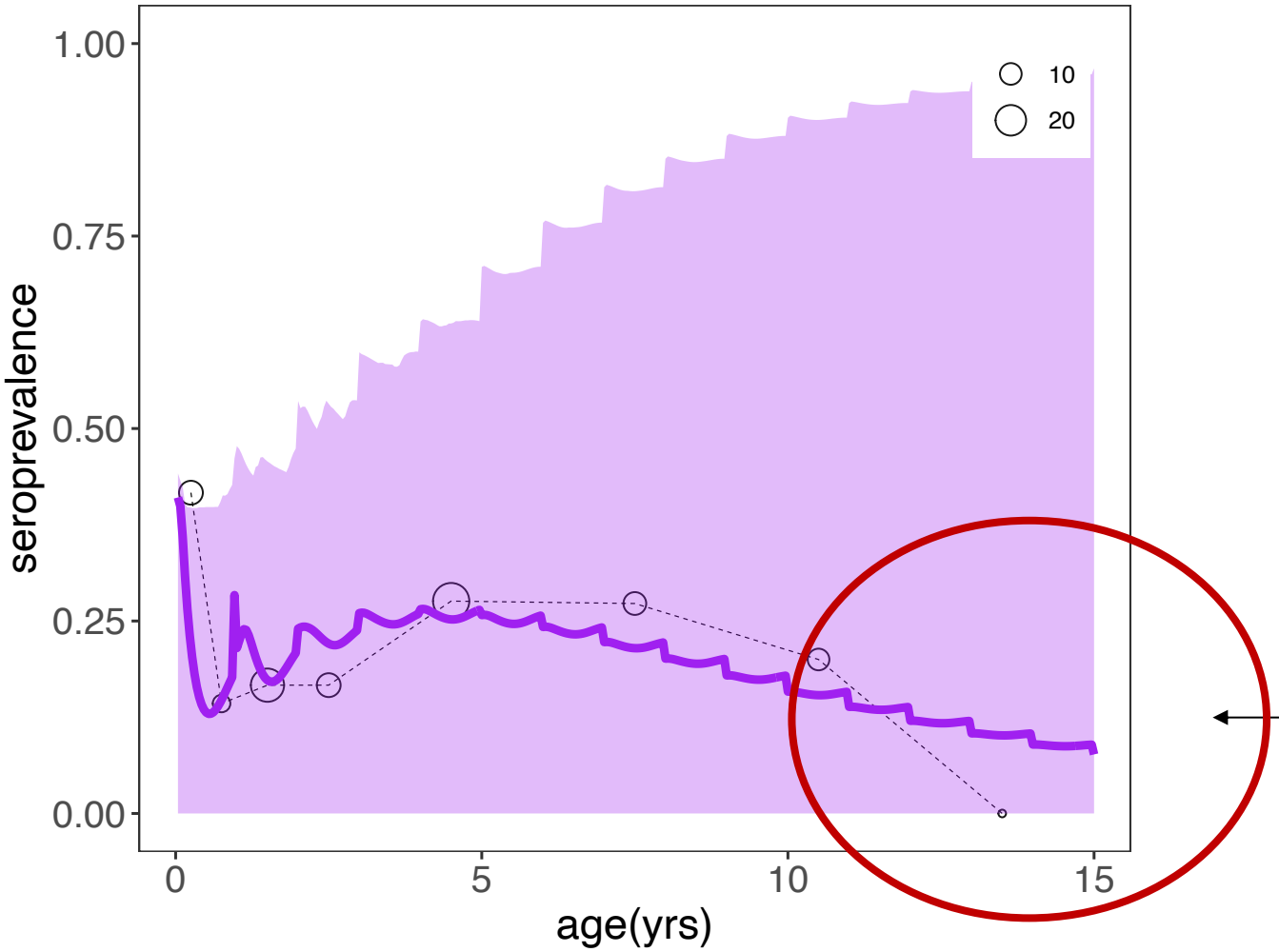
Hyp 2: Waning Immunity



Simulation and evaluation for mechanistic models

fitted model to age-seroprevalence data for E. dupreanum Nipah virus

We need more field data to evaluate fit!



Fit could be better in **late age classes**….

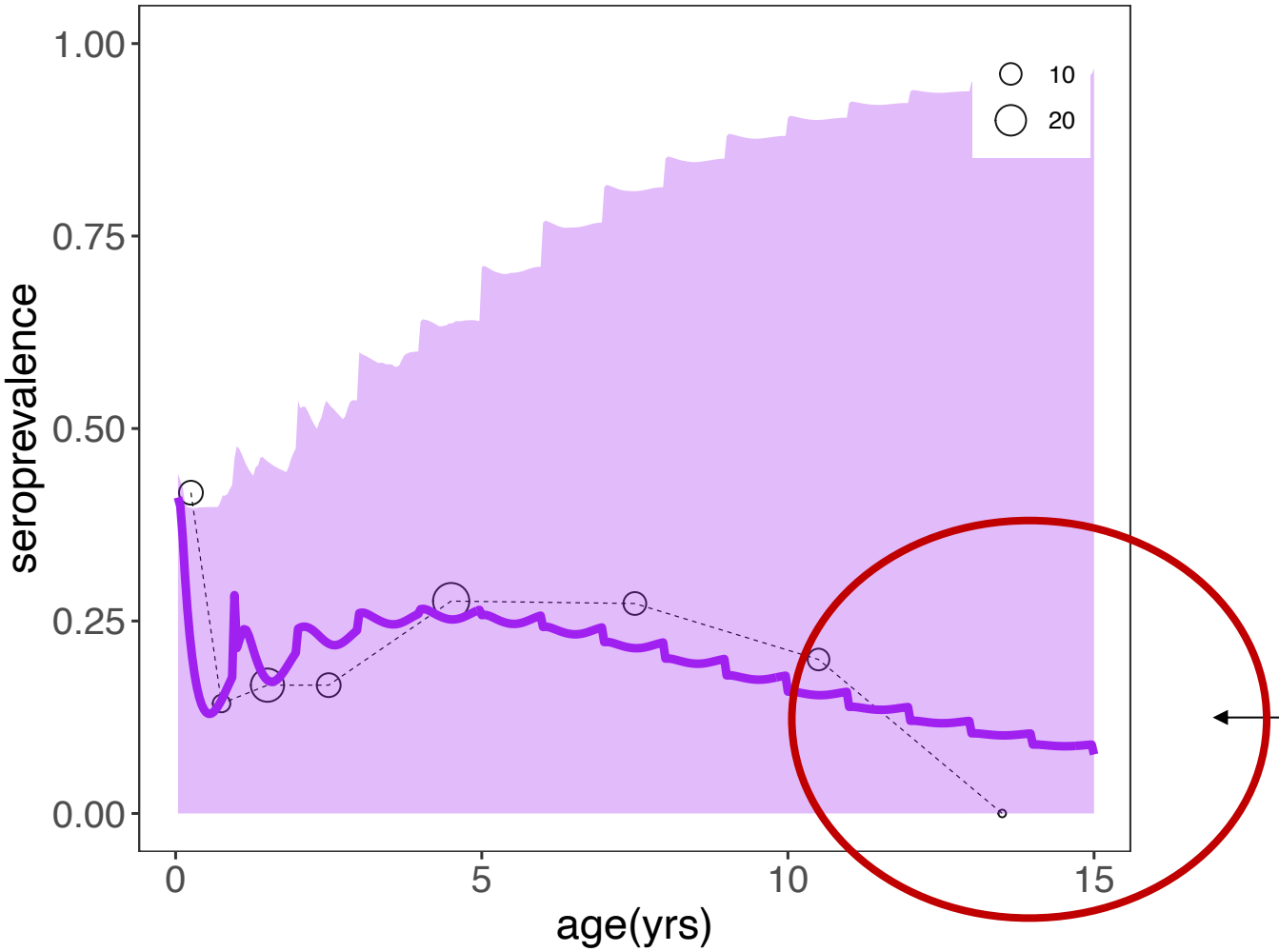
Brook et al. 2019 *J of Animal Ecology*

Simulation and evaluation for mechanistic models

fitted model to age-seroprevalence data for E. dupreanum Nipah virus

We need more field data to evaluate fit!

Simulation can help us evaluate how much more.

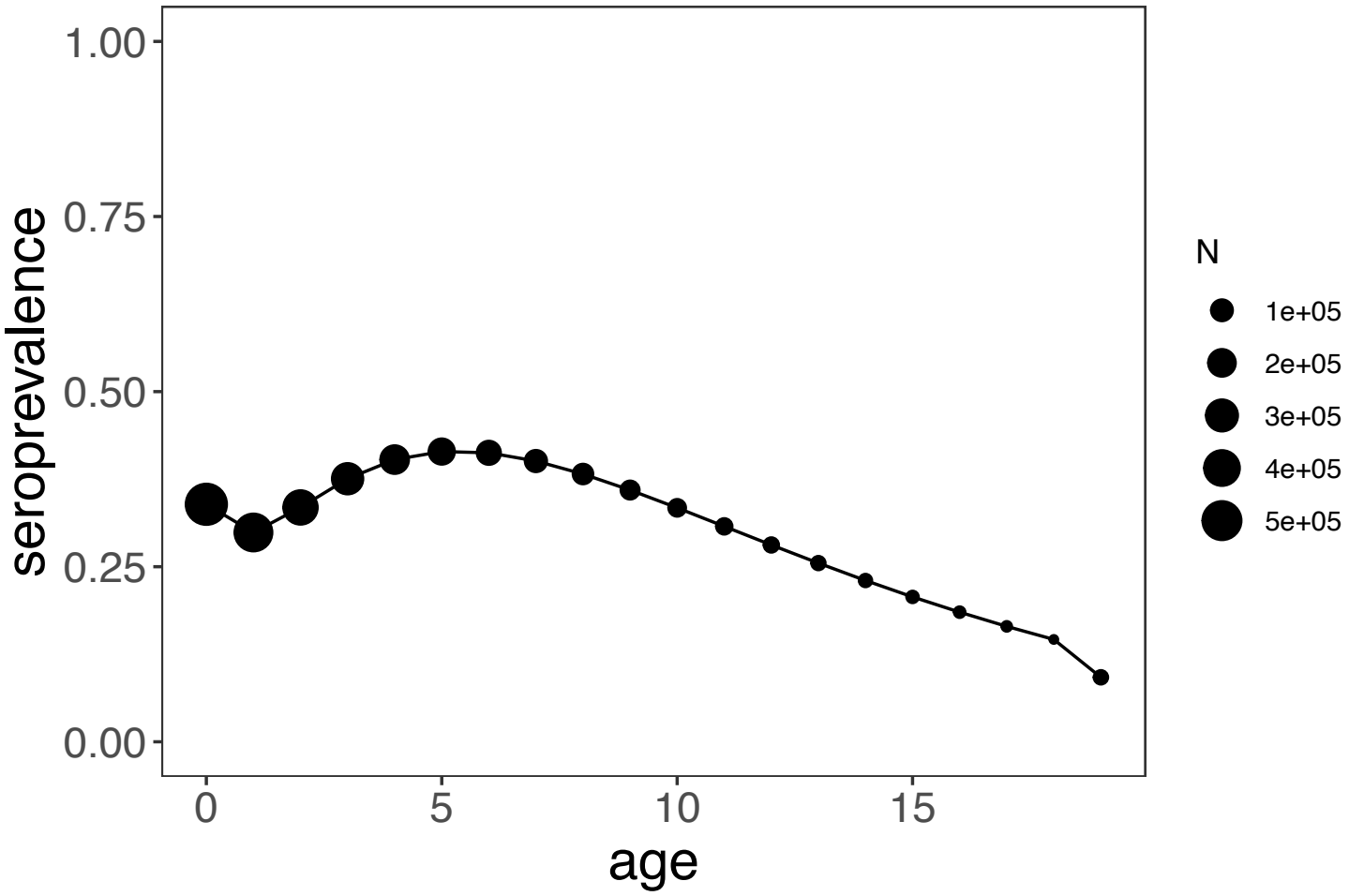


Fit could be better in **late age classes**....

Brook et al. 2019 *J of Animal Ecology*

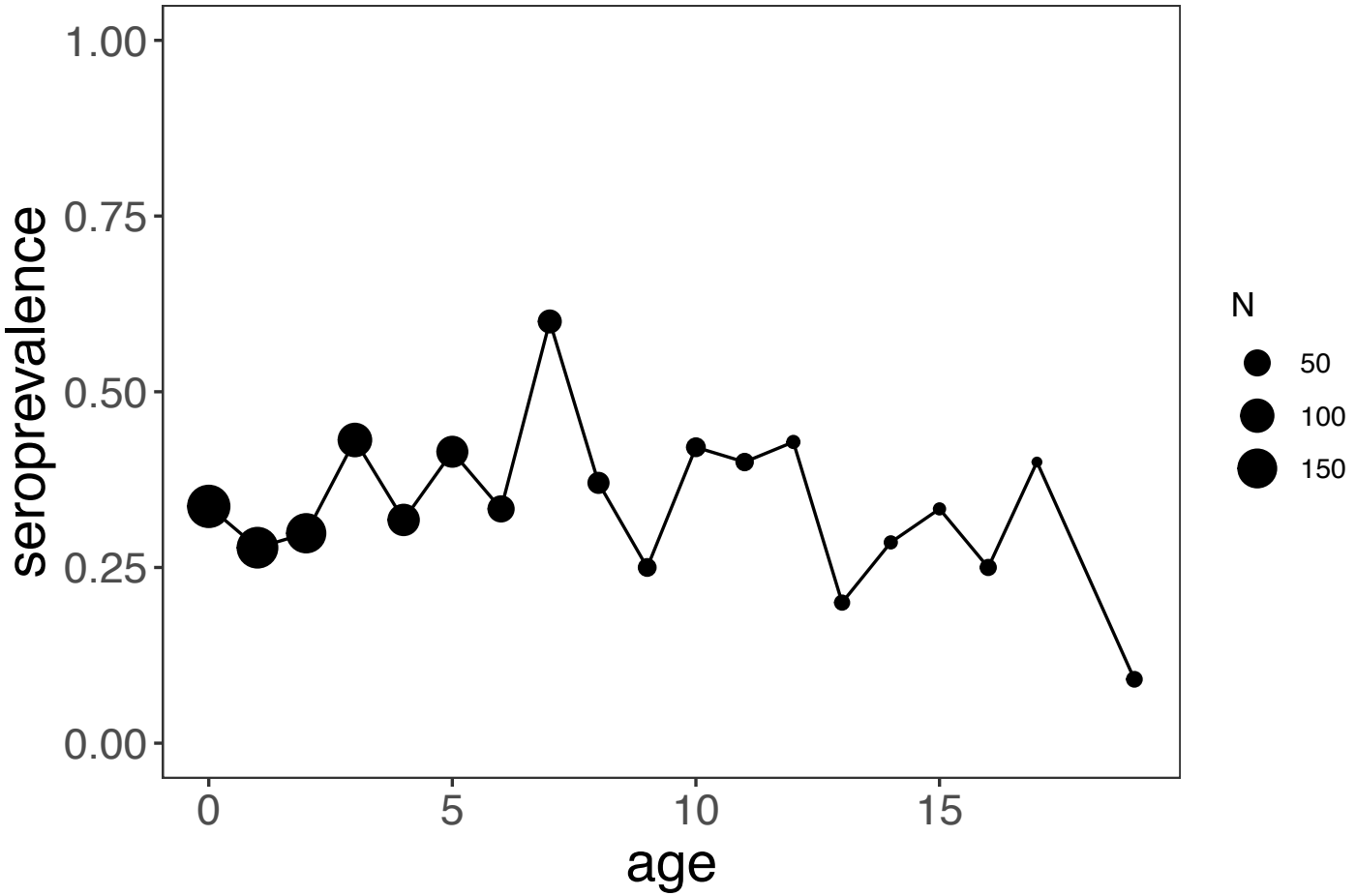
Simulation and evaluation for mechanistic models

simulated data at full population size



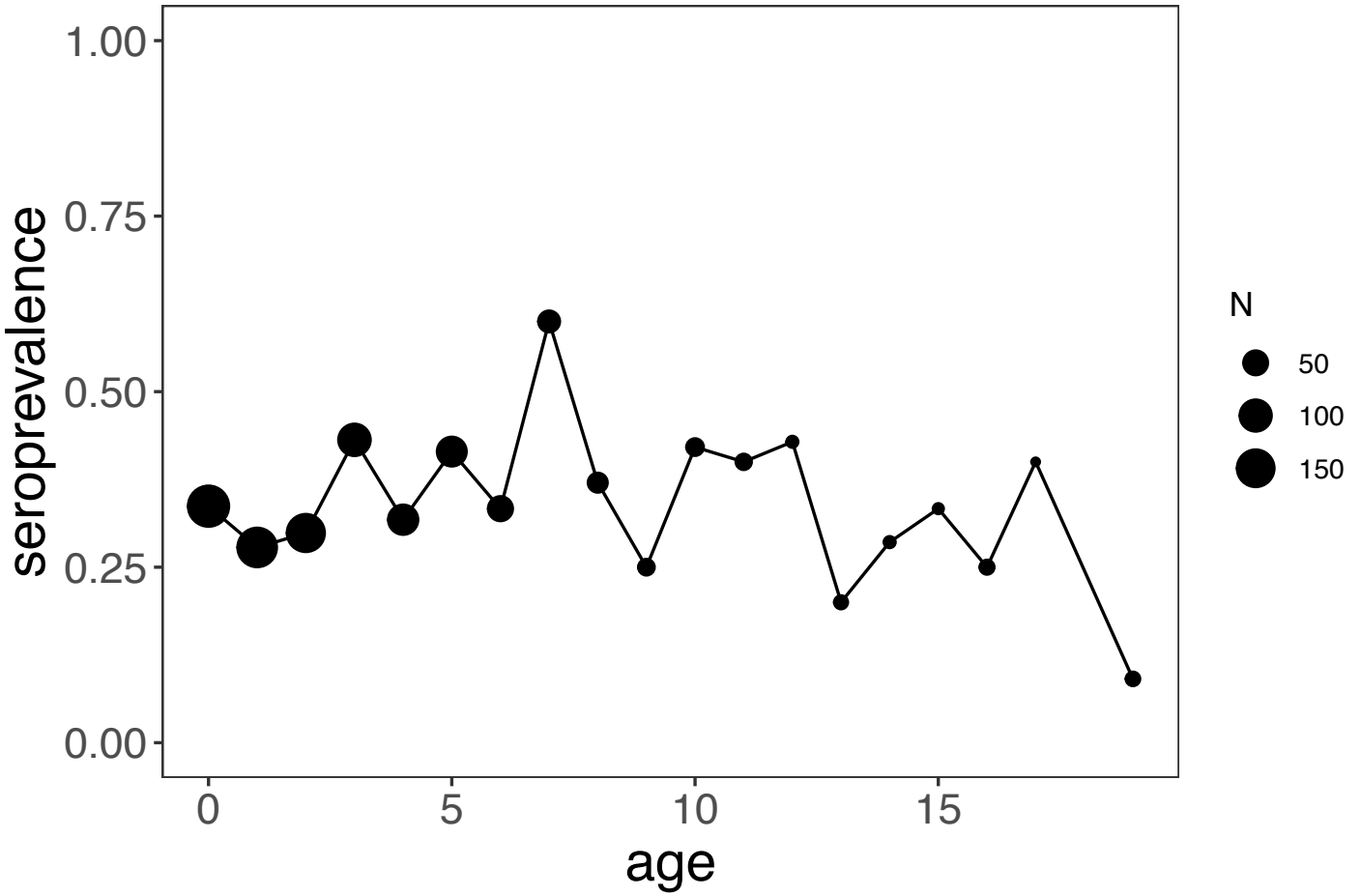
Simulation and evaluation for mechanistic models

simulated data after sub-sampling to 1000 bats



Simulation and evaluation for mechanistic models

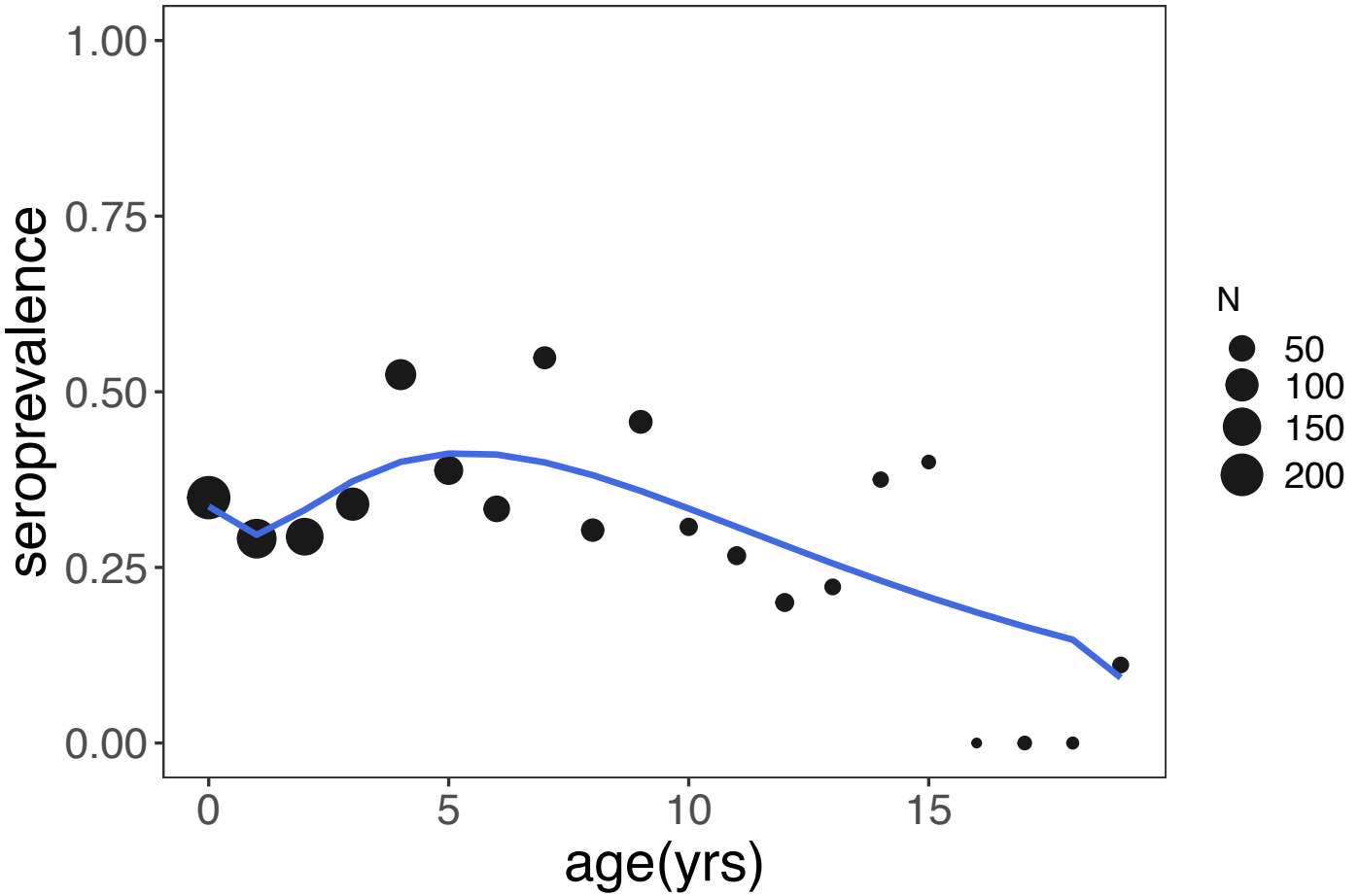
simulated data after sub-sampling to 1000 bats



Would this plot look the same if we sub-sampled a second time?

Simulation and evaluation for mechanistic models

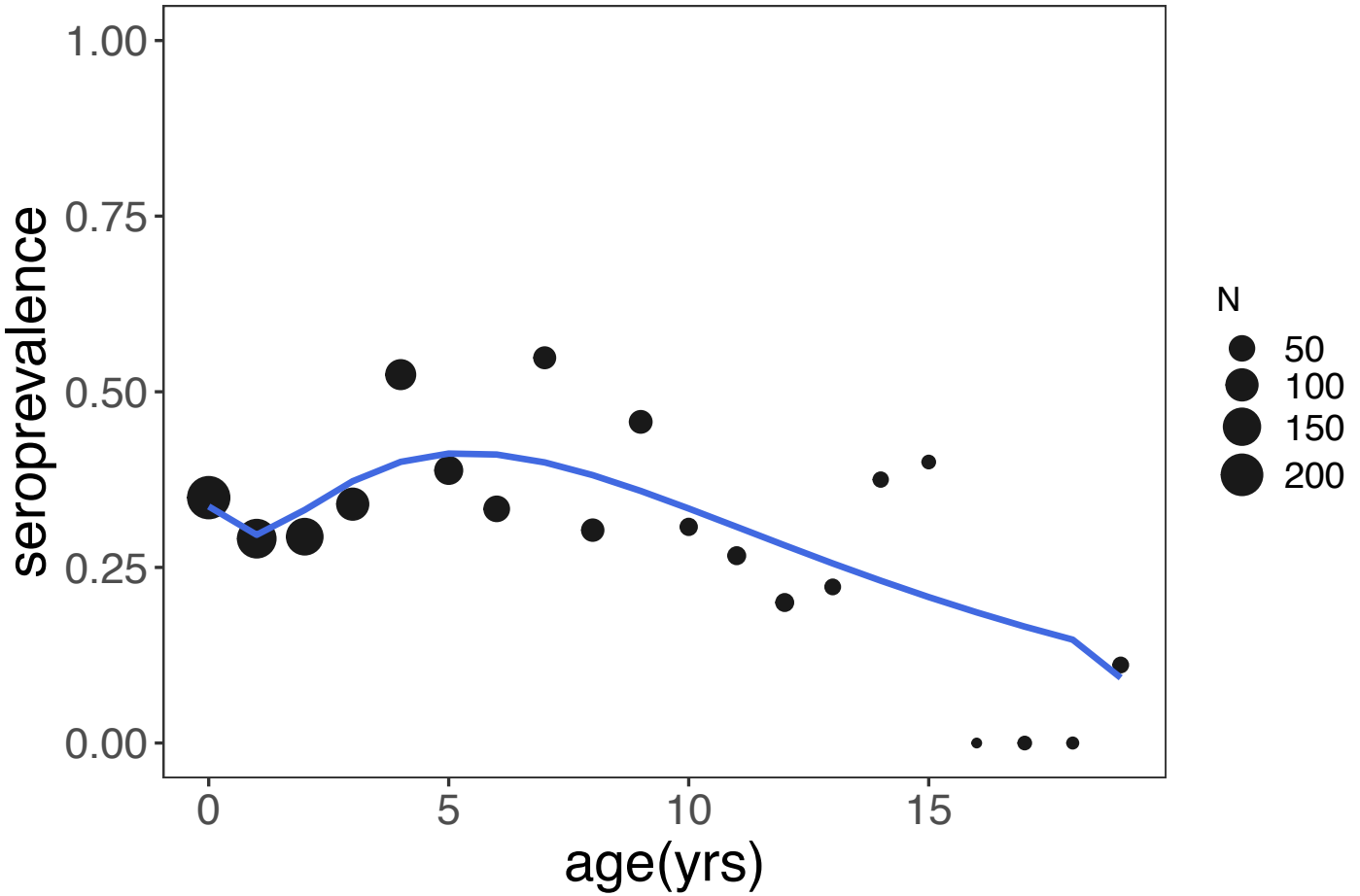
fitted lifelong immunity model (hyp1) to data subsample



AIC = 1277

Simulation and evaluation for mechanistic models

fitted lifelong immunity model (hyp1) to data subsample

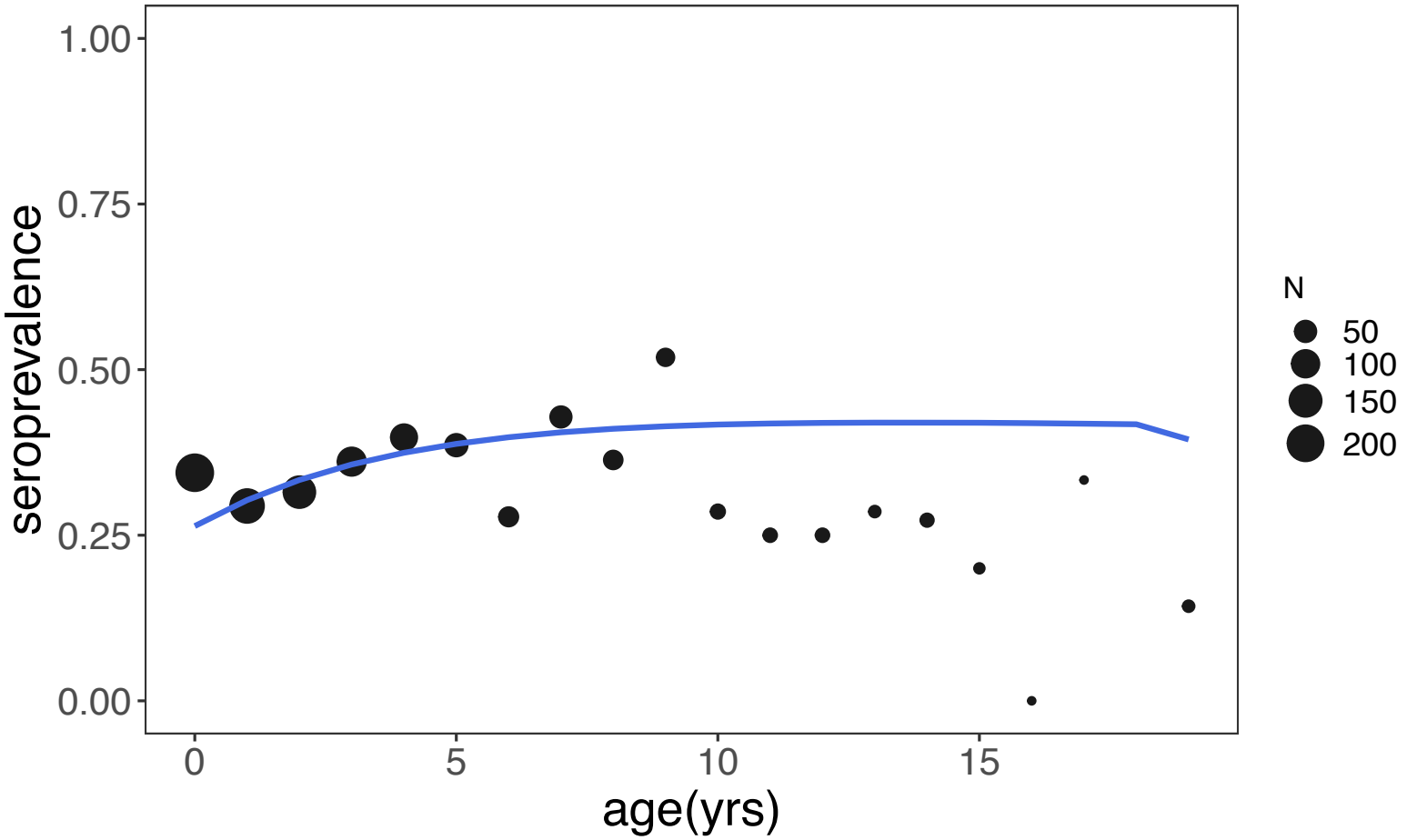


AIC = 1277

What would change if we sub-sampled more data?

Simulation and evaluation for mechanistic models

fitted waning immunity model (hyp2) to data subsample



AIC = 1291