serojump: Inferring infection timing and antibody kinetics from longitudinal serology

Adam Kucharski (on behalf of David Hodgson)

Oct 2025 I ESWI 2025

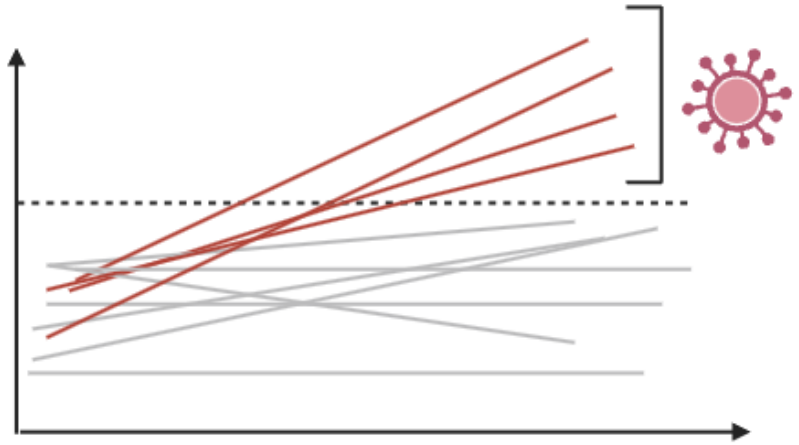
Charité — Universitätsmedizin Berlin

Charité Center
for Global Health

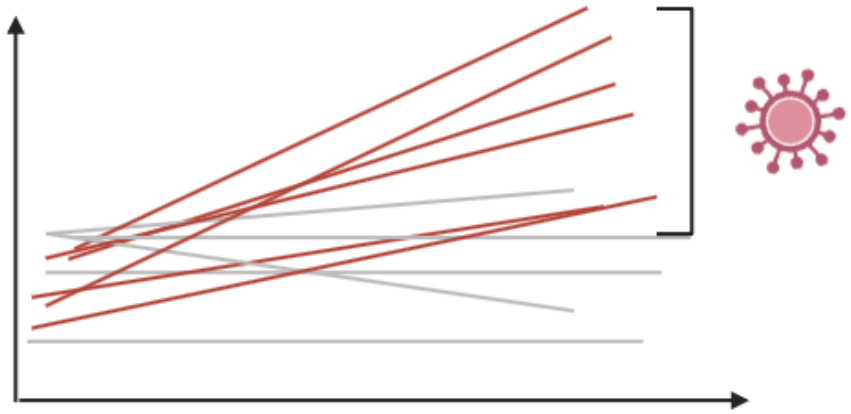
SERO-HEURISTICS

E.g.
Measles –
Flu – 1:40 HAI

Seropositivity thresholds



Seroconversion

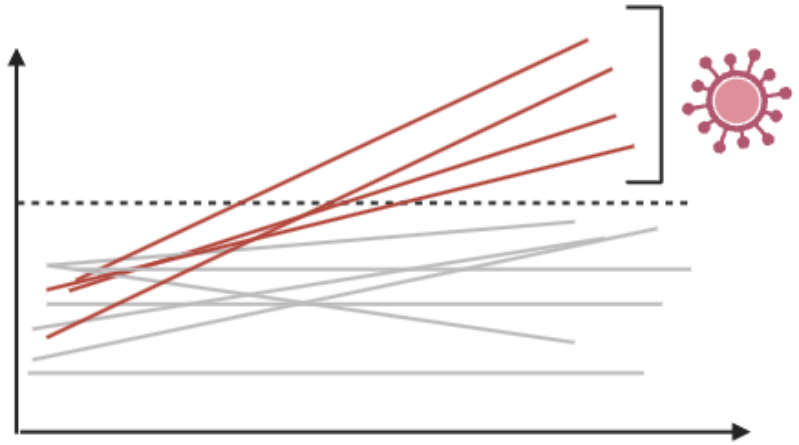


Measure:
Four-fold rise

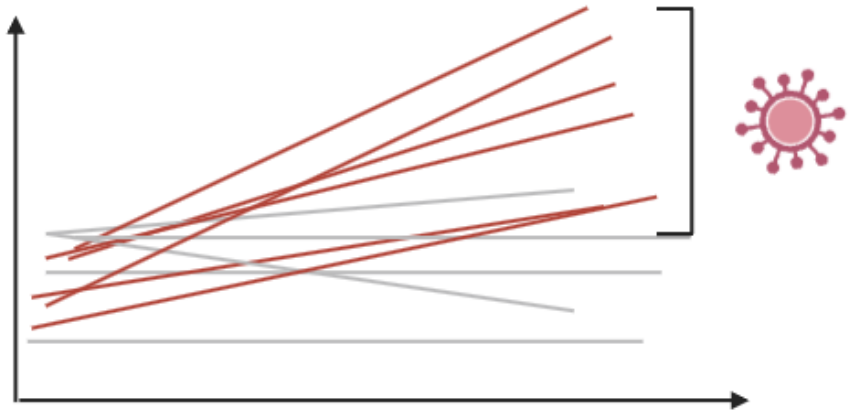
SERO-HEURISTICS

E.g.
Measles –
Flu – 1:40 HAI

Seropositivity thresholds



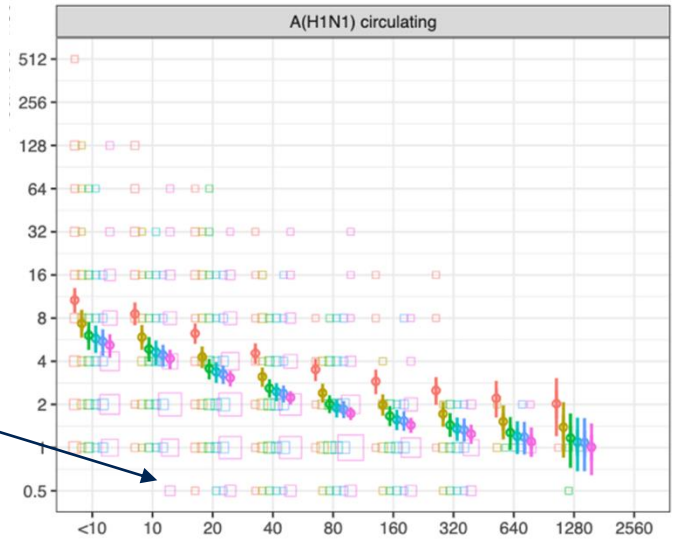
Seroconversion



Measure:
Four-fold rise

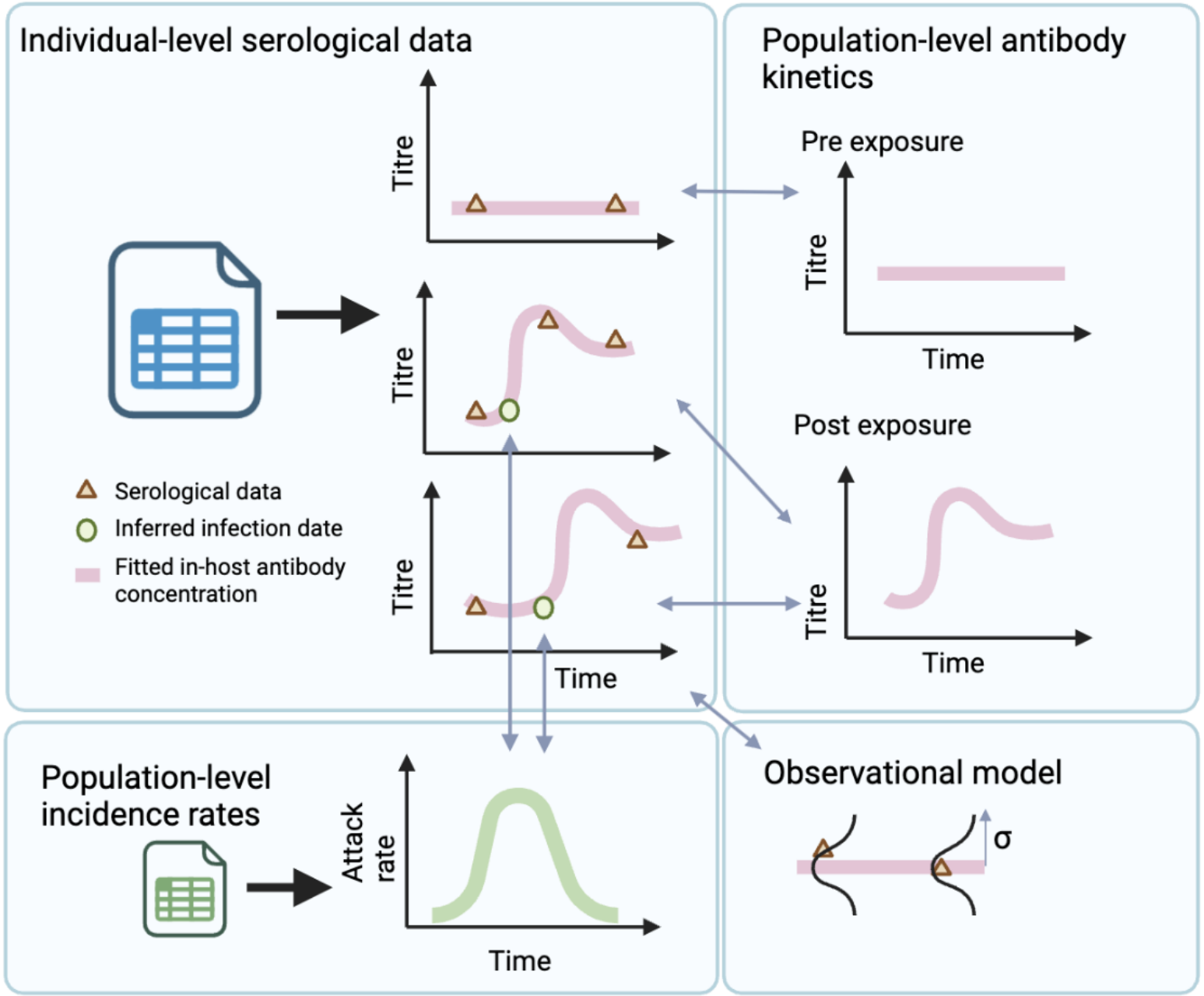
BUT—post-infection antibody kinetics depends on age, exposure history, and pre-exposure titre, so one rule not appropriate

Higher pre-vaccination titre = less boosting
Highly vaccinated = less boosting



Hodgson et al. 2024, *Vaccine*

BAYESIAN OVERVIEW



UNDERLYING MODEL

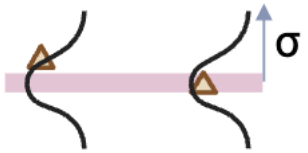
$$\mathcal{L}(Y|Z,\tau,\theta) = \prod_{i=1}^{N} \prod_{b\in B} \prod_{t\in T_i} P_{obs}(Y_{i,t}^{b}|A_{i,t}^{b},\theta), \text{ where } A_{i,t}^{b} = f_a^{b}(Z_i,\tau_i,Y_{i,0}^{b},\theta) \qquad (1)$$

We are trying to sample from (Z,τ,θ) through the posterior

$$P(Z,\tau,\theta|Y) = \mathcal{L}(Y|Z,\tau,\theta)P(\tau|Z)P(Z)P(\theta), \text{ where } P(\tau|Z) = \prod_{\substack{i=1 \\ Z_i\neq 0}}^{N} P_{exp}(\tau_i) \qquad (2)$$

BUT, if $Z_i = 0$ then no value for τ_i, so dimensions of τ dynamically change

Observational model



UNDERLYING MODEL

$$\mathcal{L}(Y|Z,\tau,\theta) = \prod_{i=1}^{N}\prod_{b\in B}\prod_{t\in T_i} P_{obs}(Y_{i,t}^{b}|A_{i,t}^{b},\theta), \text{ where } A_{i,t}^{b} = f_a^{b}(Z_i,\tau_i,Y_{i,0}^{b},\theta) \qquad (1)$$

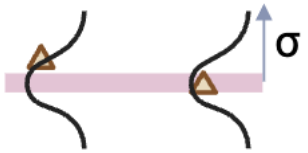
We are trying to sample from (Z,τ,θ) through the posterior

$$P(Z,\tau,\theta|Y) = \mathcal{L}(Y|Z,\tau,\theta)P(\tau|Z)P(Z)P(\theta), \text{ where } P(\tau|Z) = \prod_{\substack{i=1\\ Z_i\neq 0}}^{N} P_{exp}(\tau_i) \qquad (2)$$

BUT, if $Z_i = 0$ then no value for τ_i, so dimensions of τ dynamically change

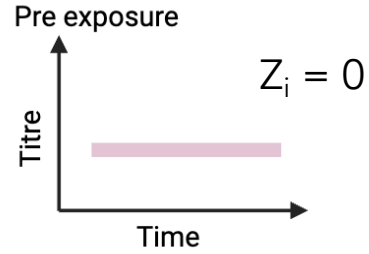
UNDERLYING MODEL

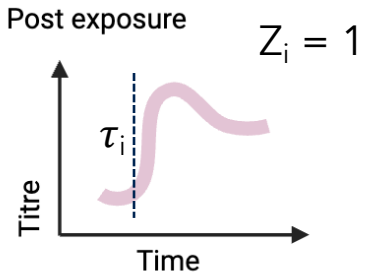
Observational model



$$\mathcal{L}(Y|Z,\tau,\theta) = \prod_{i=1}^{N}\prod_{b\in B}\prod_{t\in T_i} P_{obs}(Y_{i,t}^{b}|A_{i,t}^{b},\theta), \text{ where } A_{i,t}^{b} = f_a^{b}(Z_i,\tau_i,Y_{i,0}^{b},\theta) \quad (1)$$

Ab kinetics model





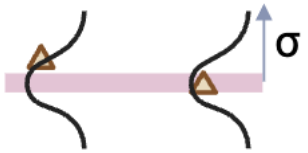
We are trying to sample from (Z,τ,θ) through the posterior

$$P(Z,\tau,\theta|Y) = \mathcal{L}(Y|Z,\tau,\theta)P(\tau|Z)P(Z)P(\theta), \text{ where } P(\tau|Z) = \prod_{\substack{i=1 \\ Z_i\neq 0}}^{N} P_{exp}(\tau_i) \quad (2)$$

BUT, if $Z_i = 0$ then no value for τ_i, so dimensions of τ dynamically change

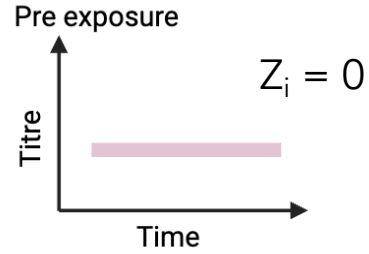
UNDERLYING MODEL

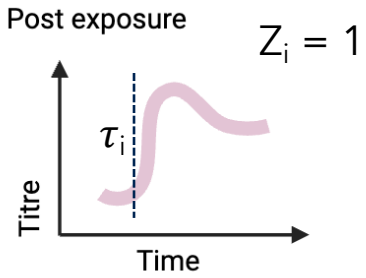
Observational model



$$\mathcal{L}(Y|Z,\tau,\theta) = \prod_{i=1}^{N}\prod_{b\in B}\prod_{t\in T_i} P_{obs}(Y^b_{i,t}|A^b_{i,t},\theta), \text{ where } A^b_{i,t} = f^b_a(Z_i,\tau_i,Y^b_{i,0},\theta) \quad (1)$$

Ab kinetics model

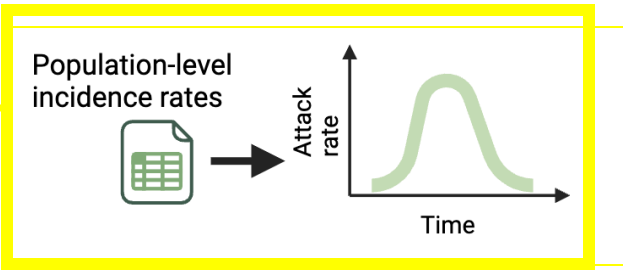




We are trying to sample from (Z,τ,θ) through the posterior

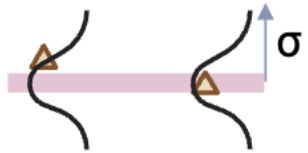
$$P(Z,\tau,\theta|Y) = \mathcal{L}(Y|Z,\tau,\theta)P(\tau|Z)P(Z)P(\theta), \text{ where } P(\tau|Z) = \prod_{\substack{i=1 \\ Z_i\neq 0}}^{N} P_{exp}(\tau_i) \quad (2)$$

Correction factor for implicit priors from Binomial dist



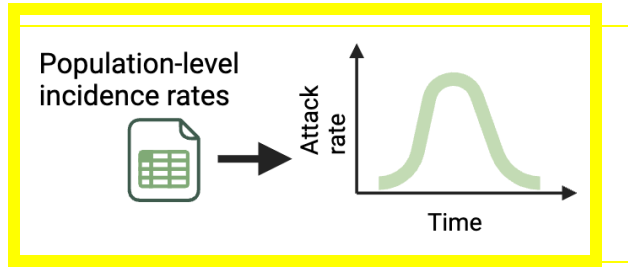
BUT, if $Z_i = 0$ then no value for τ_i, so dimensions of τ dynamically change

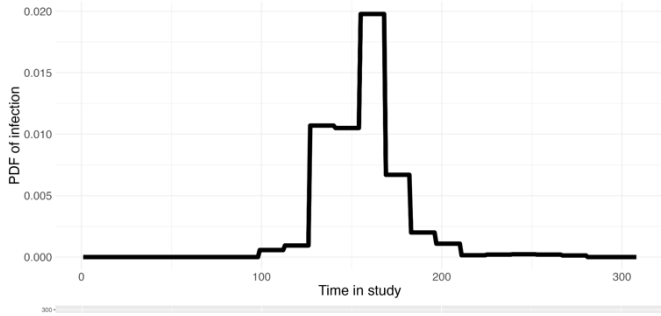
Observational model



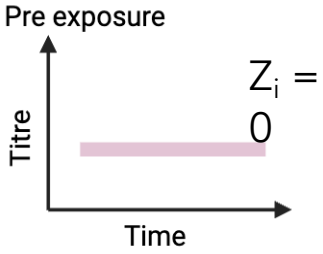
Normal distribution, $Y_{i,t}^{b} \sim N\left(A_{i,t}^{b}, \sigma\right)$,

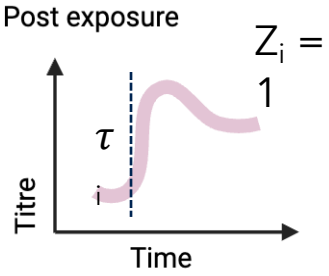
$\sigma \sim \text{Exponential}(1)$





Ab kinetics model

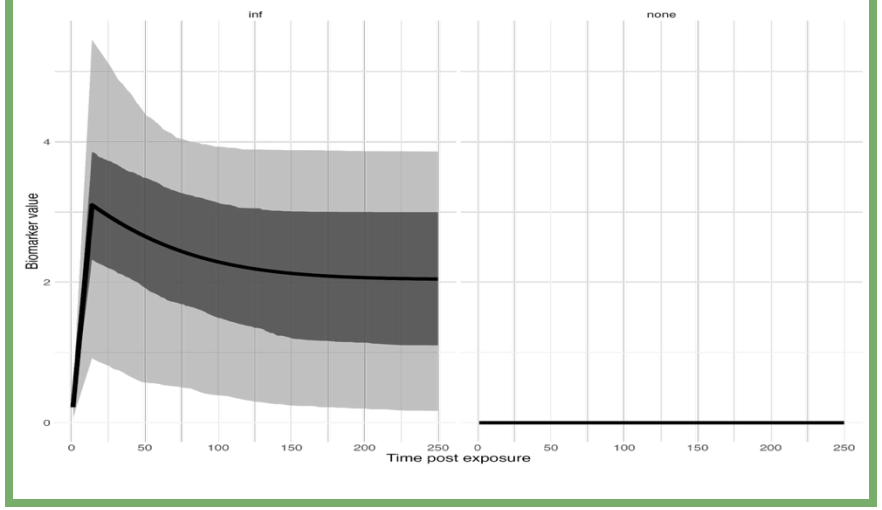




$$f_{e_{pre}}\left(t,\ Y_{i,0}\right) = Y_{i,0} - tw,$$
$$w \sim U(0,1)$$

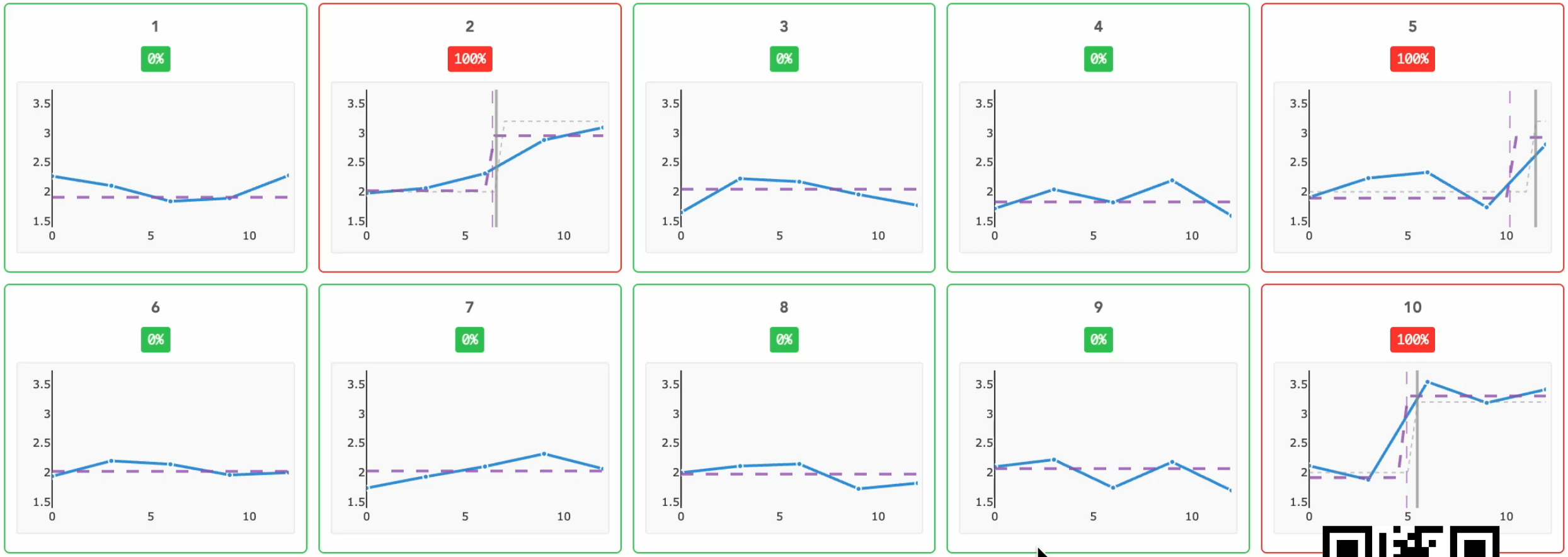
$$f_{sri}^{b}(t|\theta) = \begin{cases} \log(\exp(a) + \exp(c))t/14 \\ \log(\exp(a)\exp(-(b/14)(t-14)) + \exp(c)) \end{cases}$$

$$a \sim N(2,2), b \sim N(0.3, 0.05), c \sim N(0,4)$$



10 Individuals

SEROJUMP WIDGET



Link: https://tinyurl.com/bde2kw9j

APPLICATION TO DATASETS

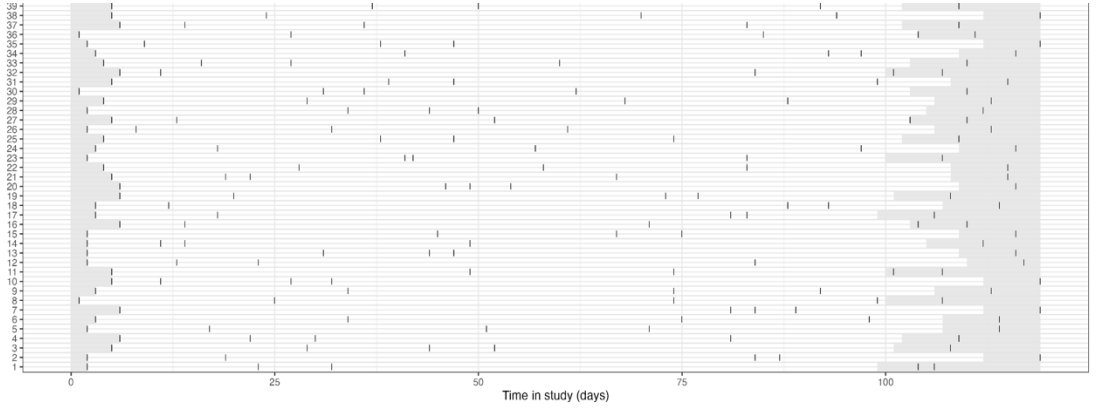
Simulated data

N = 200 people

Over 120 days

5 bleeds person

1 biomarker



Empirical data, TRANSVIR

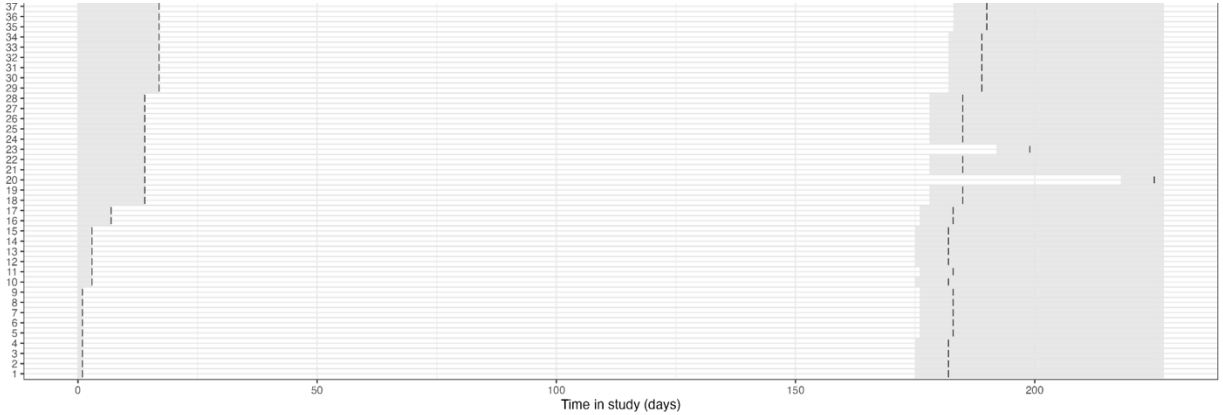
Gambia cohort to SARS-COV-2

N = 256 people

Over 308 days

2 bleeds person

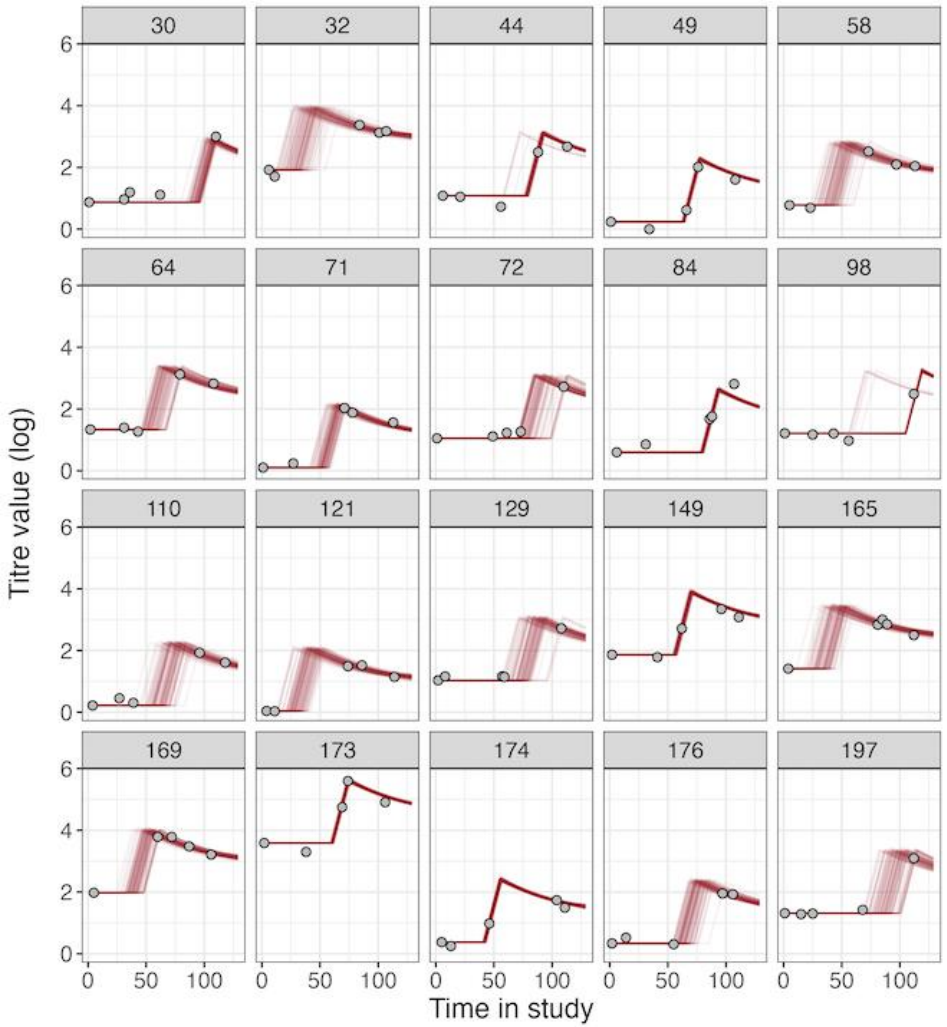
2 biomarkers: spike and NCP



CHARITÉ

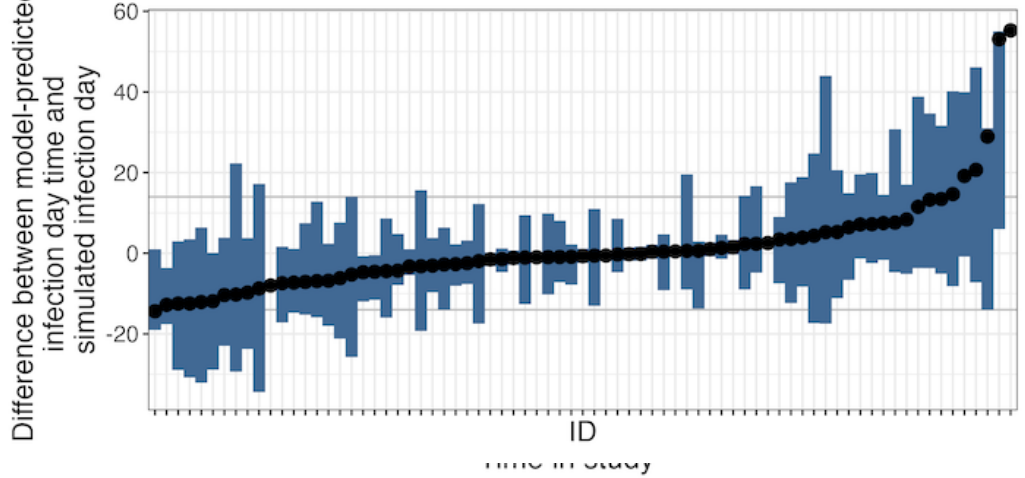
SIMULATED DATA RESULTS:

A

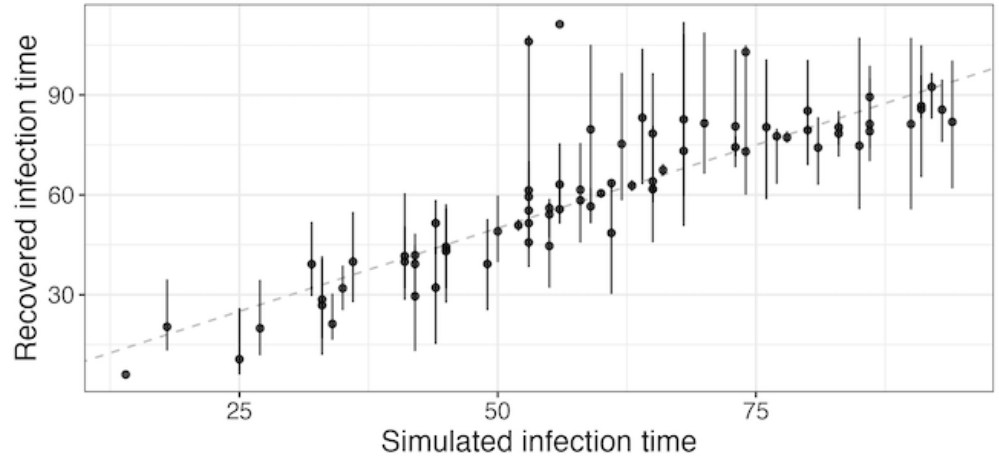


B

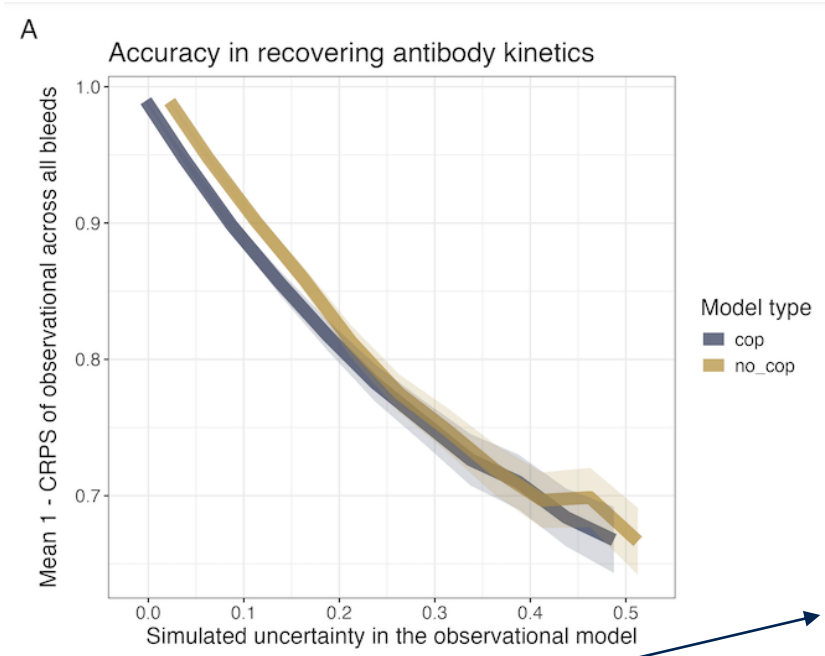
C

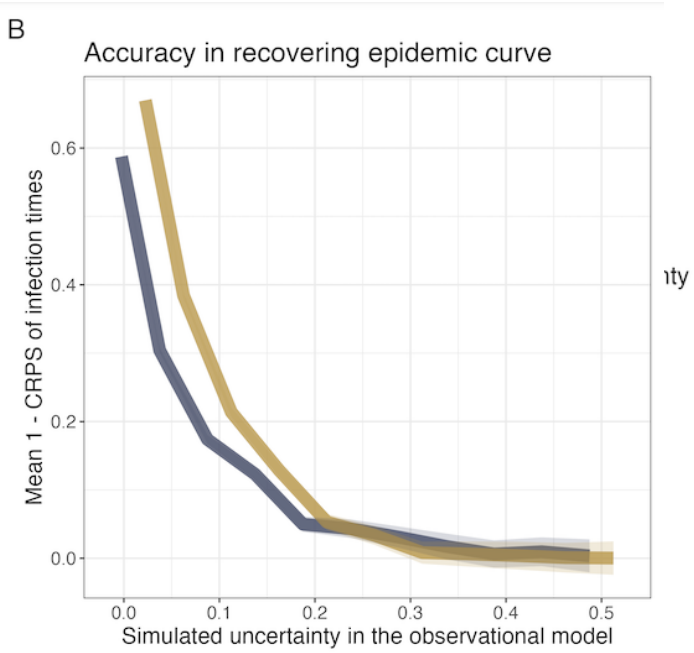


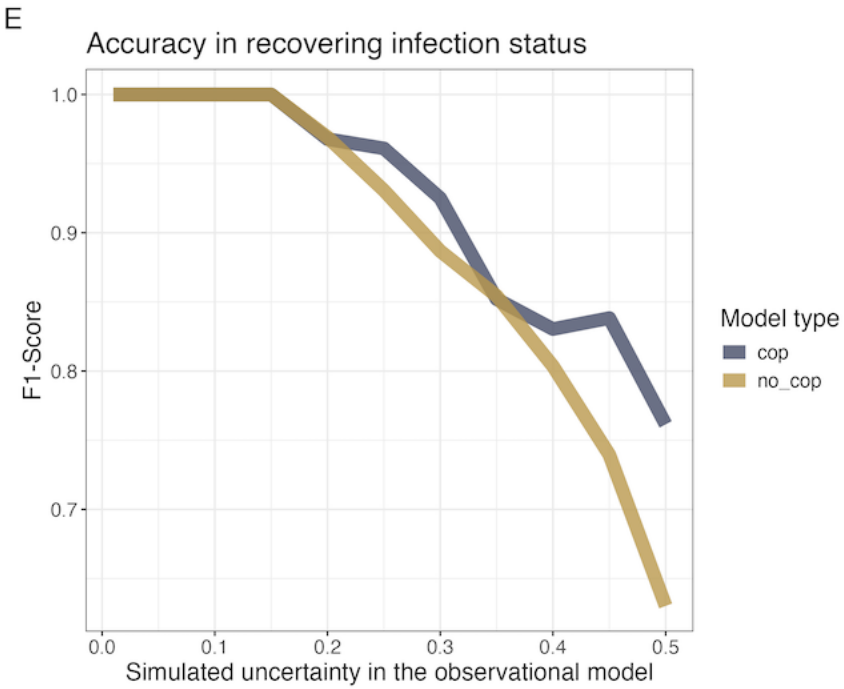
E



SIMULATED DATA RESULTS: STABILITY UNDER UNCERTAINTY





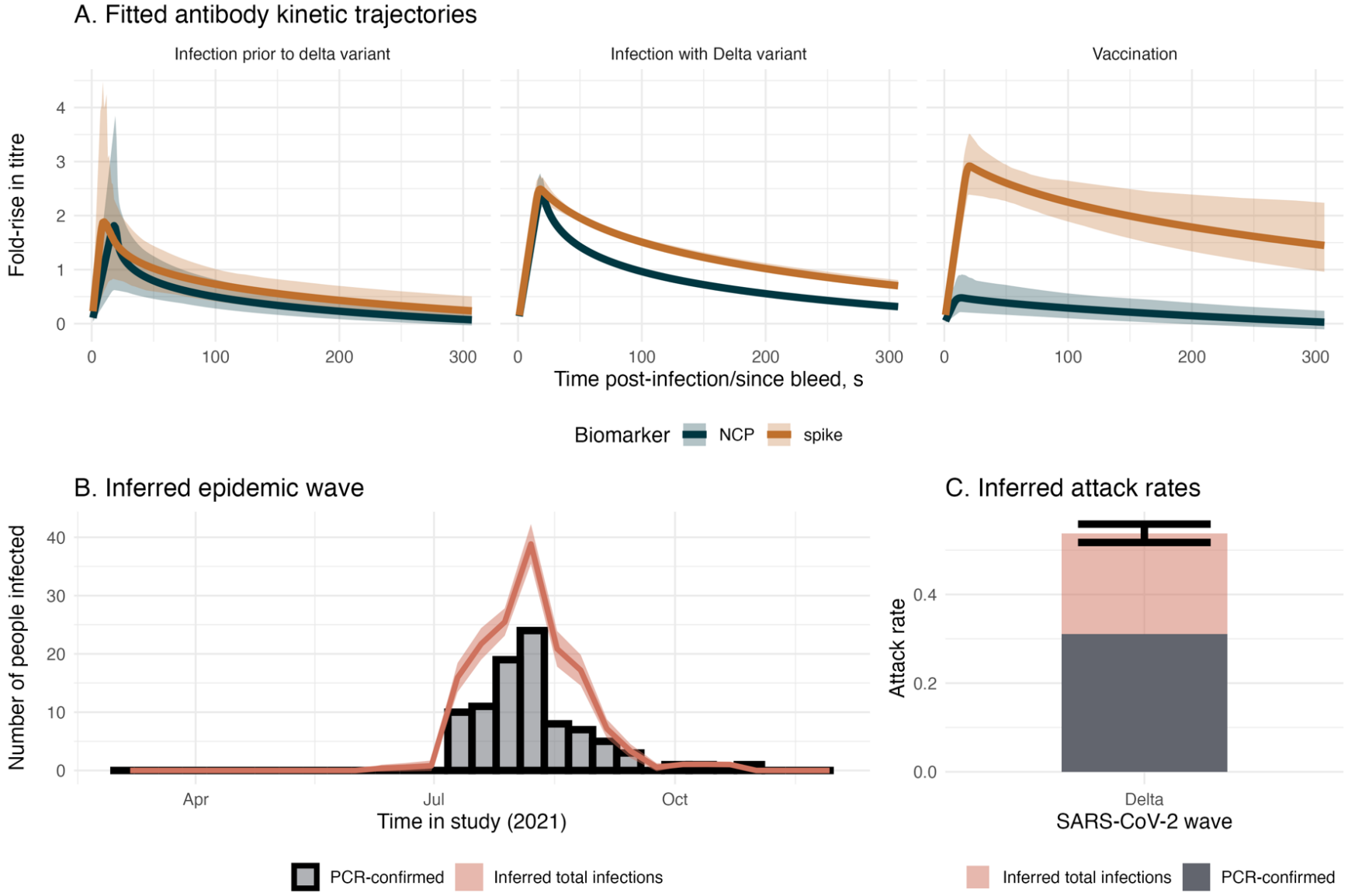


Uniform prior on infection time so recovery poor

Also found that this had better sensitivity than seropositive and seroconversion metrics!

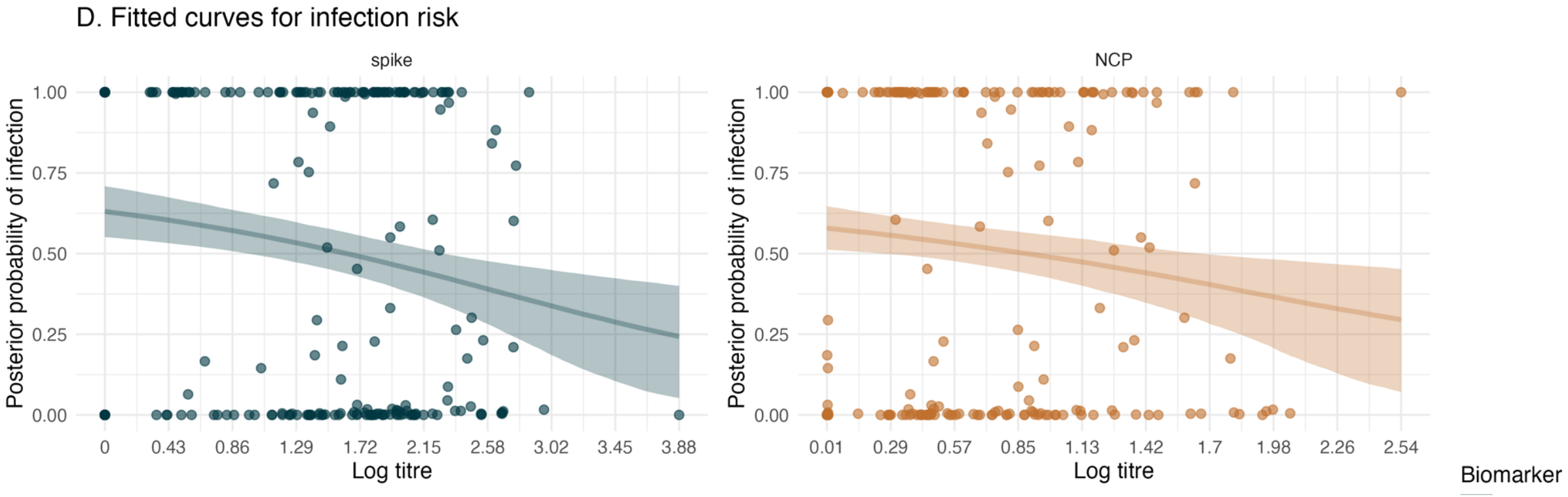
If uncertainty bigger than boost, can't recover infections well

EMPIRICAL DATA RESULTS: DELTA WAVE



EMPIRICAL DATA RESULTS: DELTA WAVE

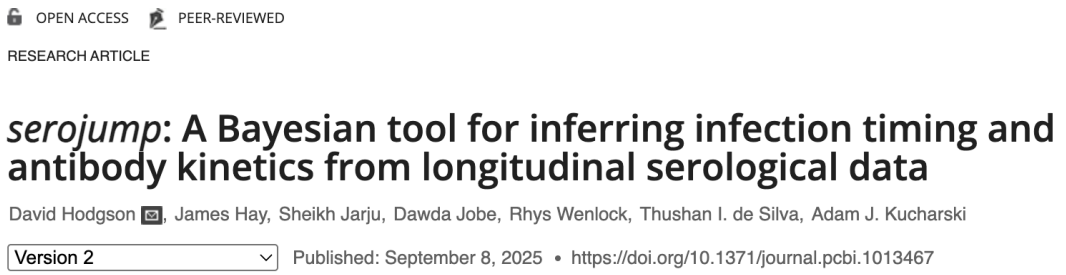
Infection risk by titre



SUMMARY

serojump: pathogen and biomarker agnostic way of detecting infections and kinetics for an outbreak

- Hope it will improve *seroheuristic* approaches and give a more probabilistic and holistic approach to serological inference
- Better inference on immunological kinetics
- Understand nuances in immunology
- Better estimates for COP and establish thresholds specific to individuals
- Packaged up and encouraging people to use



OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

***serojump*: A Bayesian tool for inferring infection timing and antibody kinetics from longitudinal serological data**

David Hodgson, James Hay, Sheikh Jarju, Dawda Jobe, Rhys Wenlock, Thushan I. de Silva, Adam J. Kucharski

Version 2 Published: September 8, 2025 • https://doi.org/10.1371/journal.pcbi.1013467

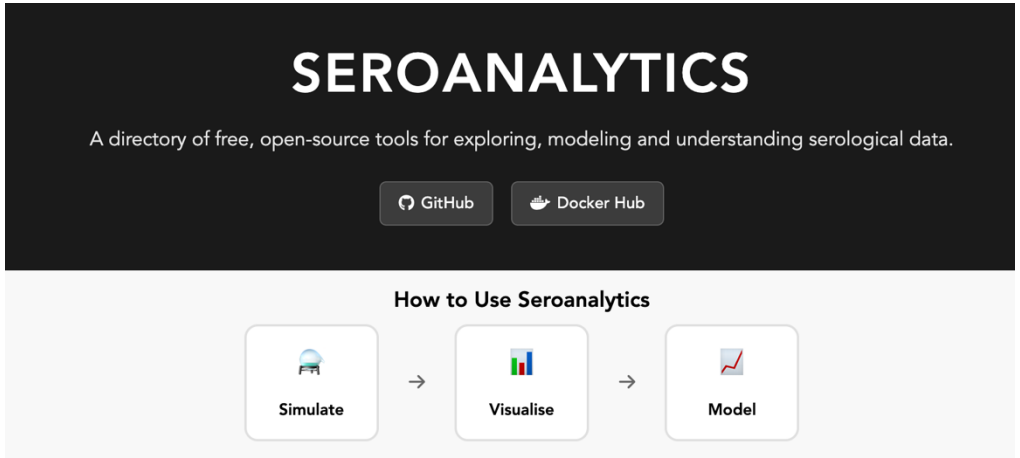
LINKS:

R PACKAGE: https://seroanalytics.org/serojump/

PAPER: https://doi.org/10.1371/journal.pcbi.1013467

ACKNOWLEDGEMENTS

Dr. James Hay
Dr. Sheikh Jarju
Dr.Dawda Jobe
Dr. Rhys Wenlock
Dr. Thushan I de Silva
Prof Adam J Kucharski



FOLLOW FOR MORE INFO!
david.hodgson@charite.de
LinkedIn: https://www.linkedin.com/in/dchodgson/
Bluesky: dchodge.bsky.social

CHARITÉ | Center for Global Health.

> seroanalytics.org