



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/232316/>

Version: Published Version

Article:

Hodgson, D., Hay, J., Jarju, S. et al. (2025) serojump: a Bayesian tool for inferring infection timing and antibody kinetics from longitudinal serological data. PLoS Computational Biology, 21 (9). e1013467. ISSN: 1553-734X

<https://doi.org/10.1371/journal.pcbi.1013467>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

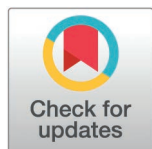
RESEARCH ARTICLE

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

David Hodgson^{1,2*}, James Hay³, Sheikh Jarju⁴, Dawda Jobe⁴, Rhys Wenlock⁴, Thushan I. de Silva^{4,5,6}, Adam J. Kucharski¹

1 Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **2** Charité Center for Global Health, Charité — Universitätsmedizin Berlin, Berlin, Germany, **3** Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom, **4** Vaccines and Immunity Theme, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, Fajara, The Gambia, **5** Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom, **6** Florey Institute of Infection and Sheffield NIHR Biomedical Research Centre, University of Sheffield, Sheffield, United Kingdom

* david.hodgson@charite.de



OPEN ACCESS

Citation: Hodgson D, Hay J, Jarju S, Jobe D, Wenlock R, de Silva TI, et al. (2025) *serojump*: A Bayesian tool for inferring infection timing and antibody kinetics from longitudinal serological data. PLoS Comput Biol 21(9): e1013467. <https://doi.org/10.1371/journal.pcbi.1013467>

Editor: Simon Cauchemez, Institut Pasteur, FRANCE

Received: March 6, 2025

Accepted: August 26, 2025

Published: September 8, 2025

Copyright: © 2025 Hodgson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All data and code required to reproduce the findings of this study are openly available. The data are provided within the manuscript and its [Supporting Information](#) files. The code, packaged in the open-source *serojump* R package, is available

Abstract

Understanding acute infectious disease dynamics at individual and population levels is critical for informing public health preparedness and response. Serological assays, which measure a range of biomarkers relating to humoral immunity, can provide a valuable window into immune responses generated by past infections and vaccinations. However, traditional methods for interpreting serological data, such as binary seropositivity and seroconversion thresholds, often rely on heuristics that fail to account for individual variability in antibody kinetics and timing of infection, potentially leading to biased estimates of infection rates and post-exposure immune responses. To address these limitations, we developed *serojump*, a novel probabilistic framework and software package that uses individual-level serological data to infer infection status, timing, and subsequent antibody kinetics. We validated *serojump* using simulated serological data and real-world SARS-CoV-2 datasets from The Gambia. In simulation studies, the model accurately recovered individual infection status, population-level antibody kinetics, and the relationship between biomarkers and immunity against infection, demonstrating robustness under observational noise. Benchmarking against standard serological heuristics in real-world data revealed that *serojump* achieves higher sensitivity in identifying infections, outperforming static threshold-based methods and precision in inferred infection timing. Application of *serojump* to longitudinal SARS-CoV-2 serological data taken during the Delta wave provided additional insights into i) missed infections based on sub-threshold rises in antibody level and ii) antibody responses to multiple biomarkers post-vaccination and infection. Our findings highlight the utility of *serojump* as a pathogen-agnostic,

at <https://github.com/seroanalytics/serojump>, with reproducible vignettes and documentation at <https://seroanalytics.org/serojump/articles/>.

Funding: TIdS was funded by a United Kingdom Research and Innovation Grant (No. MC_PC_19084). JH is supported by a Wellcome Trust Early Career Award (grant 225001/Z/22/Z). AJK and DH are supported by National Institutes of Health (1R01AI141534-01A1). AJK is supported by Wellcome (226142/Z/22/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

flexible tool for serological inference, enabling deeper insights into infection dynamics, immune responses, and correlates of protection. The open-source framework offers researchers a platform for extracting information from serological datasets, with potential applications across various infectious diseases and study designs.

Author summary

Tracking how infections spread and how our immune systems respond to them is essential for improving public health. One way to study this is by analysing blood samples to measure antibody levels, which can help estimate who has been infected, when they were infected, and how their immune response has developed over time. However, traditional methods for interpreting these antibody levels often use simple cutoffs that do not account for how much people's immune responses can vary, which can lead to inaccurate results. To solve this, we created a new tool in R called *serojump*. This tool uses advanced statistical methods to analyse changes in antibody levels over time, helping to more accurately identify who has been infected, when they were infected, and what happens to their antibodies afterwards. In our tests, *serojump* was better at detecting infections than traditional methods and provided more detailed insights about how people's immune systems responded to the virus and vaccines. It also revealed infections that were missed by standard testing methods. Our tool is flexible and can be used for many different diseases. By helping researchers and public health workers better understand infection patterns and immune responses, *serojump* can support efforts to control the spread of diseases and develop more effective treatments and vaccines.

1. Introduction

Serological samples can be analysed to detect the presence of biomarkers, such as antibodies, made in response to an infection long after the infection has cleared [1]. By combining validated biomarker assays with appropriate statistical inference techniques, it is increasingly possible to estimate incidence rates, antibody kinetics, and population-level susceptibility without ongoing syndromic surveillance [2]. Different types of serological assays contribute to such analyses, including morphological assays such as ELISA, which measure antibody concentrations, and functional assays such as neutralisation tests, which measure the ability of antibodies to inhibit infection in vitro. Analysing serological samples is important because syndromic surveillance systems typically only capture cases of disease, representing the upper parts of the reporting pyramid [3]. However, effective control and prediction of pathogen spread require understanding the full spectrum of the underlying epidemiology, including both symptomatic and asymptomatic infections, as asymptomatic or pauci-symptomatic individuals may also contribute to transmission dynamics and

population immunity. Therefore, analysing serological samples can enable researchers and healthcare professionals to infer crucial information about the epidemiology of a pathogen at the individual and population level, which case-based surveillance systems may otherwise miss [4]. Such analysis can in turn help the understanding of the immune system's ability to combat various pathogens, aid in developing new targeted intervention programmes and provide insights into patterns and drivers of the transmission dynamics of infectious diseases.

On the individual level, past infection with a specific pathogen has traditionally been inferred from measured antibodies using either i) an antibody threshold level (i.e., seropositivity) or ii) a threshold fold-rise between a pair of samples (i.e., seroconversion) [3]. Considerable research has focused on understanding how seropositivity and seroconversion rates change according to controlled host factors, such as age, geography, living conditions, sexual behaviour, etc [5–7]. On the population level, serological samples which are representative of a population (e.g., cross-sectional samples) can be used to estimate the prevalence of infectious diseases (seroprevalence) and determine how seroprevalence changes over time according to host factors [8–10].

Estimation of infection through analysis of seropositivity or seroconversion requires deriving an accurate absolute or relative threshold value, and these are often determined by rule-of-thumb heuristics (e.g., for influenza: 4-fold-rise for conversion, titre of 1:40 HAI for seropositivity) [11]. However, antibody responses vary greatly between individuals for many pathogens. Therefore, relying on these pre-determined heuristics to determine infections in serology studies can lead to both false positives and false negatives in inferred infection status, leading to biased estimates of prevalence [8,12,13]. Consequently, there has been a growth in new analytical methods to make better use of quantitative measurements derived from serological samples to inform infectious disease epidemiology and public health policy, a field of research termed 'serodynamics' [2].

In particular, efforts have been made to probabilistically infer individuals' infection status and timing in a serological cohort by analysing longitudinal changes in serological titre [14–16]. By modelling the expected antibody level over time following infection or vaccination, single or multiple measurements of an individual's antibody level can be used to back-calculate the presence of and likely timing of infection. These methods of inferring infection using an individual's changes in antibody levels are called "time-since-infection" or "titre-based" methods. However, the complexity of modelling these uncertain parameters—such as antibody kinetics, individual infection status, and the timing of infection—requires sophisticated statistical methods. Standard Bayesian methods such as Metropolis-Hastings and Hamiltonian Monte Carlo (HMC) require the assumption of a fixed-dimensional parameter space, where the number of parameters remains constant across posterior samples. However, serodynamic models often require inference of both the existence and timing of infection events, resulting in a transdimensional inference problem: an infection time can only be defined if an individual has a sampled infection event, therefore the number of inferred parameters varies across posterior samples. In these frameworks, the parameter space changes as the latent infection structure changes, introducing discontinuities in parameter space that violate key assumptions of traditional methods—particularly HMC, which depends on a smooth, differentiable log-likelihood surface.

One promising inference approach to overcome these issues is reversible jump MCMC [17], a transdimensional sampler that has been widely applied in general statistical problems involving model uncertainty, such as change-point detection [17–19] mixture models with an unknown number of components [20–22], and tests for overdispersion in hierarchical models [23]. Reversible-jump MCMC (RJ-MCMC) enable joint inference over both model structure (e.g., infection status) and associated parameters (e.g., infection time), using reversible proposals that traverse between models of differing dimensionality while preserving detailed balance via Jacobian corrections. This method has already provided valuable insights in specific studies, revealing patterns of infection and immunity for dengue and influenza that would otherwise remain hidden [15,16]. However, its application has been limited to narrowly defined problems due to the complexity of its underlying statistical methods and the need for problem-specific customisation of the sampler itself. This restriction prevents the broader adoption of reversible jump MCMC in epidemiological research despite its potential to infer accurate estimates of infection risk, correlates of protection, and antibody kinetics using only serological data.

Therefore, our study introduces *serojump*, a general-purpose statistical framework that applies reversible jump MCMC to infer infection status, timing, and antibody kinetics from serological data. By leveraging individual-level longitudinal changes in biomarker values, *serojump* provides a probabilistic approach to estimating epidemiological parameters, addressing the limitations of static serological heuristics. We validate this framework through simulation studies and real-world SARS-CoV-2 serological data from The Gambia, demonstrating its ability to recover infection histories, improve sensitivity in infection detection compared to traditional methods, and infer correlates of protection. By making these advanced inference techniques accessible via an R package, *serojump* enhances the utility of serological data for infectious disease surveillance and public health decision-making.

2. Methods

2.1. Overview of inference framework

Using data on individual-level temporal changes in serological biomarker values, *serojump* can infer several key quantities within a single probabilistic framework: i) the mean population-level kinetics of the biomarker; ii) the individual-level probability of infection during a specific period; and iii) the distribution of timing of this infection. A core assumption of *serojump* is that each individual is infected at most once during the study period. This constraint facilitates transdimensional inference and reduces computational complexity, particularly in settings where serological data are sparse. This assumption is appropriate for many typical use cases, such as cohorts sampled over a single epidemic wave or transmission season, where reinfection is unlikely and the goal is to infer whether or when an individual was infected. Defining $Y = \{Y_{i,t}^b\}$ as the measured value at time, $1 \leq t \leq T$, to biomarker $b \in \{1, \dots, B\}$, for individual $i \in \{1, \dots, N\}$; $Z = \{Z_i\}$ as the latent binary vector representing the infection status of each individual (1: ever infected, 0: not yet infected), and $\tau = \{\tau_i\}$ is the infection time for those infected (where the dimension of the vector τ is equal to the number of infected individuals, $\|Z\|_1$), we define the likelihood of this algorithm as in Menezes et al [24]:

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

Where, $P_{obs}(Y_{i,t}^b | A_{i,t}^b, \theta)$ is the observational model, describing the likelihood of the serological data, $Y_{i,t}^b$, given the model predicted latent titre values, $A_{i,t}^b$, (defined by the function, f_a^b) and θ represents the model parameters being estimated.

The posterior distribution is therefore defined by:

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

Where $P_{exp}(\tau_i)$ is the prior distribution for the timing of infection for individual i , $P(Z)$ is the prior distribution on the infection status binary vector, and $P(\theta)$ is the prior distribution on the parameters describing the observational model and antibody kinetics model. To sample from Equation 2, we used a reversible-jump MCMC algorithm, which allows for efficient exploration over the fixed dimensions of (Z, θ) and the dynamically changing dimensions of τ . The full algorithm and its derivation are given in **Sections Aa–Ac in S1 Methods**.

2.1.1 Antibody kinetics model. In the most general case, the antibody kinetics model, for an individual i , we define k immunological stimuli, $\{e_1, e_2, \dots, e_{k-1}, e_k\}$ which occur at times $\{t_{e_1}, t_{e_2}, \dots, t_{e_{k-1}}, t_{e_k}\}$, such that $t_{e_i} < t_{e_{i+1}}$ for all i . These immunological stimuli could represent exposure to an infection or vaccination. If the functions $\{f_{e_1}^b, f_{e_2}^b, \dots, f_{e_{k-1}}^b, f_{e_k}^b\}$

represent the antibody kinetics following immunological stimulus e_i for biomarker b , then, given a serological sample at time $t_{e_k} < t^*$, we estimate the biomarker value (e.g., log neutralising titre) at this time, A_{i,t^*}^b , in the model as:

$$A_{i,t^*}^b = f_a^b(Z_i, \tau_i, Y_{i,0}^b, \theta) = \begin{cases} A_{i,t_{e_1}}^b = Y_{i,0}^b \\ A_{i,t_{e_2}}^b = f_{e_1}^b(t_{e_2} - t_{e_1}, A_{i,t_{e_1}}^b | \theta_{e_1}) \\ \dots \\ A_{i,t_{e_k}}^b = f_{e_{k-1}}^b(t_{e_k} - t_{e_{k-1}}, A_{i,t_{e_{k-1}}}^b | \theta_{e_{k-1}}) \\ A_{i,t^*}^b = f_{e_k}^b(t^* - t_{e_k}, A_{i,t_{e_k}}^b | \theta_{e_k}) \end{cases}, \quad (3)$$

That is, the antibody titre is governed by an individual's most recent exposure and the function for this response is conditional on the time since this exposure, the biomarker level at that time, and any exposure-specific parameters. Given that our analysis focuses on inferring infections, we assume that for a subset of individuals, the immunological stimuli, e_i , results in an infection, and the timing of this infection $\tau_i = t_{e_i}$. We refer to this framework as an antibody kinetics model as, in this study, we are measuring titre values to antibody levels, but note the same method can be applied to any measured immunological biomarker (e.g., memory B cell concentrations).

2.1.2 Observational model. The observation model defines the likelihood function that links the model-estimated biomarker value at the sampled time to the measured value at that time. In many applications, particularly when serological measurements are log-transformed, it is common to assume a normally distributed error structure for expected log assay observations, alongside censoring where measurements are discrete and bounded (e.g., dilution thresholds). This assumption can be appropriate for assays such as ELISA, where measured values are continuous, bounded away from zero, and approximately homoscedastic on the log scale. However, this is not universally applicable: for example, PRNT and HAI assays often exhibit scale-dependent variance and discretisation effects due to dilution steps. Users of *serojump* are therefore encouraged to tailor the likelihood function to their specific assay, as the R package allows full flexibility to define custom likelihoods (e.g., interval-censored, overdispersed, or non-normal). In our analysis, we used a normal likelihood, $P_{obs}(Y_i, tb|A_i, tb, \theta) = F_N(Y_i, tb|A_i, tb, \theta)$ on log-transformed ELISA values, which we found to be well calibrated and consistent with the distributional properties of the underlying measurements [25].

2.1.3 Infection model. The prior for the infection time, $P_{exp}(\tau_i)$, for each individual can be estimated on the population-level from existing serological surveys or surveillance data to identify likely periods of infection. For example, surveillance data can still indicate the distribution of infection burden over time, if not the true cumulative number of infections. A key feature of our framework is that $P_{exp}(\tau_i)$ can be specified to reflect a shared force of infection (FOI) across individuals, ensuring that infection times are conditionally dependent given the population-level FOI. This avoids the naive assumption that $P_{exp}(\tau_i)$ values are fully independent for each individual. If this prior is not specified in the model, or there is no known data to infer the epidemic curve, the model will assume a uniform risk of infection over the study period (**Section Ad in S1 Methods**).

2.1.4 Prior distributions. The prior distribution on the parameters describing the observational model and antibody kinetics model, $P(\theta)$, can be defined by the user and are assumed to be independent. We note that, the prior distributions chosen for the simulated and empirical models described below reflect practical considerations rooted in both biological plausibility and computational stability. For instance, the decay rate parameter, $b \sim N(0.3, 0.05)$, reflects empirically observed values for SARS-CoV-2 and other viral infections, while the observational noise prior $\sigma \sim U(0.0, 4.0)$ is calibrated to the expected scale of log-transformed ELISA values. While these priors may appear narrow, they are not required by the RJ-MCMC algorithm per se; rather, they help stabilise inference under transdimensional sampling, where high-dimensional or weakly identified models may suffer from poor mixing or identifiability issues. The prior distribution on the infection status binary vector, $P(Z)$, as with similar inference problems, [25] has an implicit prior resulting from the definition of the reversible jump algorithm. This is Bin(0.5, N), or 50% of the population infected on average. To remove

this implicit prior and instead assign a uniform distribution to the total number of infected individuals during the epidemic, we choose the combinatorial prior where $\|Z\|_1$ is the number of infected individuals:

$$P(Z) = \frac{(N - \|Z\|_1)! \|Z\|_1!}{N!}, \quad (4)$$

If prior knowledge on the number of missed infected individuals is known (e.g., from random community testing regardless of symptoms), these can be added to this prior (**Sections Ad-Ae in S1 Methods**). To support model robustness and transparency, we conduct prior predictive checks for both the antibody kinetics functions and the observational models to ensure they produce biologically plausible outputs before fitting to data. These checks are described in **Section Af in S1 Methods**. We strongly encourage users of *serojump* to follow similar workflows and adapt the priors to reflect assay characteristics, study populations, and research questions. The R package supports fully custom prior specifications for all components of the model, including hierarchical or unbounded alternatives.

2.1.5. Post-processing of posterior distributions and correlates of risk. After fitting the model with the *serojump* algorithm, we obtain posterior distributions for the set $(\hat{\tau}, \hat{Z}, \hat{\theta})$. We calculate the posterior distribution of the post-exposure antibody trajectories for exposure e and biomarker b , $\hat{p}_e^b(t, \hat{\theta})$. To assess the posterior distribution of the number of infected individuals across the population, we calculate $\|\hat{Z}\|_1$ (the number of infected individual) for each sample of the posterior distribution. We also calculate for each posterior sample, the “reference titre value” for each individual and biomarker. The reference titre value for those infected is the titre at which individuals become infected (τ_i) for each individual and biomarker $s_i^b = A_{i,\tau_i}^b$. For those not infected, we sample a random value from their titre trajectories across the study weighted by the prior probability of exposure timing $s_i^b = A_{i,T}^b$, where $T \sim P_{exp}$. Using the reference titre value, we assess the relationship between the biomarker value (e.g., antibody titre) and infection risk, to estimate a *correlate of risk (COR)*. A COR represents either a direct mediator of immunity or a proxy for immune protection. To estimate this relationship from *serojump*, we fit a logistic curve with parameters k (gradient), x_0 (mid-point) and L (asymptote, representing the exposure rate), to describe how reference biomarker values relate to the infection risk, given by:

$$R(x) = L \left(1 - \frac{1}{1 + \exp(-k(x - x_0))} \right)$$

From this, we define the *conditional protection curve (CPC)* as the fitted probability curve, representing the direct relationship between biomarker levels and protection from infection. It can be derived from the fitted $R(x)$ as:

$$CPC(x) = \frac{1}{1 + \exp(-k(x - x_0))}$$

The *relative risk (RR)*, is a rescaled version of the CPC that normalises protection relative to a reference titre level (in this case the lower limit of detection. It can also be derived from fitted $R(x)$ as

$$RR(x) = \frac{R(x)}{R[\min(x)]}$$

These estimates allow us to quantify the protective effect of immune markers and assess how biomarker levels modulate infection risk across a population (**Sections Af-Ag in S1 Methods**).

2.2 Application to simulated data

2.2.1 Description of the simulated data. To test whether it is possible to correctly recover epidemiological and immunological dynamics from common serological cohort structures, we first simulate a serological dataset using the *serosim* R package [24] to test the *serojump* framework. We simulate continuous epidemic serosurveillance (CES) cohort data, which represents a serostudy in which individuals are followed over an epidemic wave and sampled at multiple random time points throughout. The simulated data includes $N=200$ individuals with serological samples taken within the first seven days of the study's starting and a sample within the last seven days of the study's ending. These individuals also had three samples taken randomly throughout the study during an 120-day epidemic wave.

We model the epidemic as a stochastic transmission process, where individuals can be either susceptible, exposed, or infected. The probability of exposure is set at 60% per individual, meaning that individuals coming into contact with an infectious person have this probability of becoming exposed, assuming no prior immunity. Among those exposed, the probability of progressing to infection is 30%, meaning that some individuals remain uninfected despite exposure. This corresponds to a scenario where a pathogen spreads within a population that starts with partial immunity. Over the study timeframe we assume that individuals can have a maximum of one exposure. To model a symmetric epidemic peak, we simulate the exposure time from a normal distribution, $N(60, 20)$ days. To assess the impact of biomarker levels on infection risk, we simulate two different relationships between the measured biomarker and the probability of infection given exposure: (1) No COP model (Model A), where infection occurs with a fixed probability of 50% regardless of biomarker titre, representing a non-titre-dependent correlate of protection; and (2) With COP model (Model B), where infection probability follows a logistic function of biomarker titre, a commonly used correlate of protection model [26,27]. Explicitly, the simulated probability of infection, given a titre value, x , is given by

$$f(x) = \frac{1}{1 + \exp(-(-4 + 2x))}$$

Following infection, the antibody kinetics are assumed to follow a linear rise to a peak at 14 days, followed by an exponential decay to a set-point value [28]. The formula for this biphasic trajectory is given by

$$f_{sri}^b(t|\theta) = \begin{cases} \log(\exp(a) + \exp(c)) \cdot \frac{t}{14}, & \text{if } 0 \leq t < 14 \\ \log(\exp(a)\exp(-(\frac{b}{14})(t-14)) + \exp(c)), & \text{if } t \geq 14, \end{cases} \quad (5)$$

Biologically, Equation 5 models a biphasic antibody response characterised by a rapid linear rise to a peak (typically within ~2 weeks post-infection), followed by an exponential decay to a stable plateau. This reflects typical antibody dynamics observed after acute viral infections such as influenza, SARS-CoV-2, and dengue, where antibodies first expand quickly (activation phase), then contract and stabilise (memory phase). In this parameterisation, a controls the peak magnitude of the response, b the rate of exponential decay, and c the long-term plateau value. This simple yet biologically interpretable form has been shown to effectively reproduce observed kinetics in controlled infection studies [28]. The simulated values we use are $a=1.5$, $b=0.3$, $c=1$ and t , is the number of days post-infection. If an individual is not infected over the period, their titre remains unchanged and, therefore, equal to their start titre throughout. We assume all individuals have the same antibody kinetic response following infection but we add observational noise into our model, assuming a normal distribution with standard deviation σ . In the base case, we assume $\sigma = 0.1$. We tested the robustness of the simulation recovery to noise by fitting the *serojump* algorithm to simulated data with increasing observational error. Specifically, we simulated using *serosim*, serological data with standard deviations on the normally distributed error for 11 values, $(0.01, 0.05, 0.1, 0.15, \dots, 0.45, 0.5)$, and evaluated the capacity for *serojump* to recover i) the infection status of individuals, ii) the timing of the epidemic and iii) the recovered population-level antibody kinetics.

2.2.2 Model specification for serojump. We assume each individual's serological profile is characterised by the IgG biomarker ($b = \{\text{IgG}\}$) and that this biomarker is measured using ELISA, with the units given by IU/mL. Two immunological events are considered: e_{pre} , the pre-infection state and e_{inf} , the infected state. The structural assumptions on the kinetics of the pre-infection and infection state are; for the pre-infection state, we assume the prior on the antibody levels Y_i are modeled as a linear wane

$$f_{e_{pre}}(t, Y_{i,0}) = Y_{i,0} - wt$$

with a waning rate following the uniform distribution with a prior $w \sim U(0, 0.1)$. For the post-infection dynamics, we assume that antibody-kinetics boost according to the functional definition defined as $f_{e_{inf}}^{IgG} = f_{sri}^{IgG}(t | a, b, c)$ (Equation 5) with the priors, $a \sim N(2, 2)$, $b \sim N(0.3, 0.05)$, $c \sim U(0, 4)$.

For the observational model, we assume that antibody measurements follow the normal distribution probability density function $[F_N(\mathbf{X} | \mu, \sigma)]$;

$$P_{obs}(Y_{i,t}^b | A_{i,t}^b) = F_N(Y_{i,t}^b | A_{i,t}^b, \sigma)$$

with the prior $\sigma \sim U(0, 4)$. For the prior on the risk of infection over time, we assume this is uniform across the whole time period $P_{exp}(\tau_i) = U(1, 120)$ for all i . Finally, the distribution for the number of infection events, $P(Z)$, is constrained by a combinatorial factor reflecting the arrangement of events across the population as described in Equation 4. A summary for the method specification for *serojump* can be found in the **Section B** in [S1 Methods](#).

2.3 Application to empirical data

2.3.1 Description of the real-world serological data on SARS-CoV-2 from The Gambia. To test *serojump* on real-world serological data, we used longitudinal serological data collected before and after the Delta wave [29] in a household cohort study of 349 individuals in The Gambia. Quantitative ancestral anti-spike and nucleocapsid (NCP) IgG data were generated using previously validated ELISA assays [10], calibrated to the WHO International Standard for anti-SARS-CoV-2 immunoglobulin (cat no NIBSC 20/136). In addition to serological data, SARS-CoV-2 PCR testing was conducted weekly regardless of symptoms and additionally for those who presented with influenza-like illness symptoms during this period. In addition to PCR-positive diagnoses, dates of SARS-CoV-2 vaccination and PCR-positive pre-Delta infections were also known. At the start of the study, which was just prior to the Delta variant circulating, 56.7% of people were seropositive for SARS-CoV-2 [29]. Over the study period, the Delta variant predominantly circulated with 99 PCR-confirmed cases, with an additional 21 PCR confirmed SARS-CoV-2 infections with pre-Delta variants, and 49 individuals who received a dose of vaccine. In this analysis, we explicitly model vaccination as a separate immunological stimulus with a known date of administration, which is provided as an input to the model. This ensures that vaccine-induced boosts in antibody levels—particularly to the spike protein—are not misclassified as infections. When vaccination dates are unknown or uncertain, such boosts could in principle be misattributed; however, the inclusion of multiple biomarkers (spike and nucleocapsid IgG) allows the model to disambiguate immunological responses. A description of the timing of the sample collection, and infections for each individual is given in **Section C** in [S1 Methods](#).

2.3.2 Model specification for serojump. Each individual's serological profile is characterised by two biomarkers: IgG to SARS-CoV-2 ancestral spike and NCP protein ($b = \{\text{spike}, \text{NCP}\}$). We model four immunological events: e_{pre} , the pre-infection state, $e_{pre-delta}$, the pre-Delta infection state, e_{delta} , the Delta-infected state, and e_{vax} the vaccination state. For the pre-infection state, we assume the prior on the antibody levels Y_i for biomarker b are modeled as a linear wane:

$$f_{e_{pre}}(t, Y_{i,0}) = Y_{i,0} - wt$$

with a waning rate, w , with a prior $w \sim U(0, 0.1)$, estimated separately for each biomarker. For post-infection (pre-Delta and Delta) and vaccination, we assume that antibody kinetics follow Equation 6 (below), which describes a heterogeneous antibody response incorporating variation in response magnitude, time-to-peak, and decay rate. This functional form is adapted from Teunis et al. (2009) [30], who demonstrated its utility in seroepidemiological inference across multiple pathogens. For SARS-CoV-2, recent work (e.g., Srivastava et al., 2024 [28]) has shown that antibody responses often peak around 14–21 days and then decline heterogeneously, with substantial inter-individual variation. Thus, parameters y_1 , and t_1 describe the peak titre and time to peak, respectively, r governs the decay rate, and σ the observational uncertainty. This flexible form allows us to model the diverse immune responses to both infection and vaccination with SARS-CoV-2, and is well-suited to longitudinal data with multiple exposures and biomarkers. In addition, titre-dependent boosting of SARS-CoV-2 has been observed in previous studies [28,31], and we defined α as the parameter which drives the degree to which titre-value at exposure attenuates the boosting effect. Therefore, we assume that antibody kinetics follow a boost of $\log(I_{teu}^b) * I_{dep}^b$, defined by:

$$I_{teu}^b(t|\theta) = \begin{cases} \exp(\mu t), & \text{if } t < t_1 \\ y_1(1 + (r-1)y_1^{-1}v(t-t_1)^{-\frac{1}{r-1}}), & \text{if } t \geq t_1 \end{cases}, \quad (6)$$

and

$$I_{dep}^b(A_0^b, \alpha) = \max(0, 1 - \alpha A_0^b), \quad (7)$$

where $\mu = y_1(1/t_1)$, and the priors governing the boosting response are given by: $r \sim U(0, 1)$, $y_1 \sim U(0, 6)$, $t_1 \sim U(7, 21)$, and $\alpha \sim U(0, 1)$. The value of v is fixed to 0.001, which corresponds to a slow, long-term waning of the antibodies coming from the long-lasting memory B cell response. Thus, in the posterior distribution, each immunological event (pre-Delta infection, Delta infection, vaccination) and each biomarker has distinct parameter values describing its kinetics, but they share the same prior distributions. For the observational model, we assume that antibody measurements follow the normal distribution probability density [$F_N(\mathbf{X}|\mu, \sigma)$];

$$P_{obs}(Y_{i,t}^b | A_{i,t}^b) = F_N(Y_{i,t}^b | A_{i,t}^b, \sigma),$$

with prior $\sigma \sim \text{Exponential}$ [1]. We choose an empirical distribution based on the PCR-positive data for the prior on the risk of infection over time. Finally, the distribution of the number of infection events, $P(Z)$, is constrained by a combinatorial factor reflecting the arrangement of events across the population as described in Equation 4. A summary for the method specification for *serojump* can be found in the **Section C in S1 Methods**.

2.3.3 Benchmarking against standard serological heuristics. We evaluate how well *serojump* and other serological approaches can infer true infection status (as defined by PCR-positivity) using serological data alone, without incorporating PCR results into the inference process. Specifically, we compare five different infection classification strategies based on serology: *serojump*, a four-fold rise in spike or NCP, and seropositivity thresholds for spike and NCP (IC50 values of 6.39 and 3.00) [29]. To assess the performance of these classification approaches, we compute and compare their sensitivity in detecting PCR-confirmed infections.

2.4. Implementation

The code is written in R (v.4.4.1) [32] and Rcpp (v1.0.14) [33], open source and packaged at <https://github.com/seroanalytics/serojump>. The package is flexible, allowing the user to input their functional forms and associated prior distributions for the antibody kinetics model (I_{e}^b) for each immunological stimulus and biomarker, or the user can select built-in functions

from previous papers such as those mentioned in this study. Further, the user can specify their functional form for the observational model likelihood $[P_{obs}(Y_{i,t}^b | A_{i,t}^b)]$ and the prior distributions on the arguments. Finally, the user can customise the prior on the risk of infection over the time period, $P_{exp}(\tau_i)$, and the prior on the distribution of infection events $P(Z)$. The inputs for *serojump* to run the model on the simulated data and empirical data are summarised in **Sections B and C** in [S1 Methods](#). The package *serojump* has vignettes to reproduce this study, which can be found at <https://seroanalytics.org/serojump/articles/>. To ensure that posterior samples adequately represent the target distribution, we recommend users assess convergence using both standard MCMC diagnostics and transdimensional-specific tools. For fixed-dimensional parameters, we report Gelman-Rubin R, effective sample size (ESS), and trace plots, which are useful for evaluating mixing, between-chain consistency, and sampling efficiency. We interpret R-hat values exceeding 1.1 as indicative of potential non-convergence, consistent with standard practice. Values below this threshold are considered evidence of adequate mixing and convergence. For trans-dimensional exploration we consider transdimensional convergence metrics such as the Structural Model Index (SMI) [34]. These diagnostics are implemented via standard R packages and are directly compatible with *serojump* output. The assessment of convergence statistics and model fits for the case studies presented here are summarised in [S1–S9 Figs](#) and information about the computational performance of the framework is provided in **Section D** in [S1 Methods](#).

3. Results

3.1 Case 1: simulated data

3.1.1. Simulation recovery for the base case simulated dataset. For the default simulated dataset with low observational noise ($\sigma = 0.1$), the individual-level antibody kinetics show good agreement between simulated and recovered kinetics under the With COP and No COP simulated data ([Fig 1A–B](#)) with the observed data points (grey dots) close to the median posterior distributions of the model-fitted individual-level antibody trajectories (red lines). We recover the infection status of each individual with 100% sensitivity and specificity, and the differences between simulated infection days and model-predicted infection days ([Fig 1C–D](#)) are small (within 14 days) for almost all individuals. This results in a close alignment between the simulated and the recovered epidemic curves ([Fig 1E–F](#)). These observations suggest that for the default simulated dataset, both the With COP and No COP data are well recovered, as evidenced by closer alignment of posterior medians to observed data, with minor errors in infection time predictions and an accurate reconstruction of the infection time distribution.

3.1.2. Stability of simulation recovery for increasing observational noise. We assess predictive performance by measuring the Continuous Ranked Probability Score (CRPS, [35]) a metric that evaluates the accuracy of probabilistic predictions by comparing the predicted cumulative distribution function to the observed value. We find the functional form of the antibody kinetics for both No COP and With COP models is well recovered (CRPS < 0.35) across all levels of observational noise; however, decreasing in accuracy as observational noise increases ([Fig 2A](#)). In contrast, the recovery of the epidemic curve is adequate at low levels of uncertainty but quickly becomes inaccurate (CRPS > 0.7) at a moderate level of noise ([Fig 2B](#)). This suggests that for datasets with a high degree of observational error, the recovery of the epidemic timing may be inaccurate when a uniform prior on infection timing is used in *serojump*.

For recovery infection status, we find the model perfectly recovers the infection status of each individual (F1-score = 1) up until an observational error of 0.15, after which there is a decrease in F1-score as observational error increases, suggesting the sensitivity and specificity of the recovery of infection status is decreasing ([Fig 2C–E](#)). This decrease is driven by the increase in the false positive rate, with the true positive rate remaining consistent across all observational noise, suggesting that the model correctly identifies all infections but may incorrectly detect infections at high levels of noise. We find that the functional form of the COP is well recovered in both the *With COP* and *No COP* datasets when observational

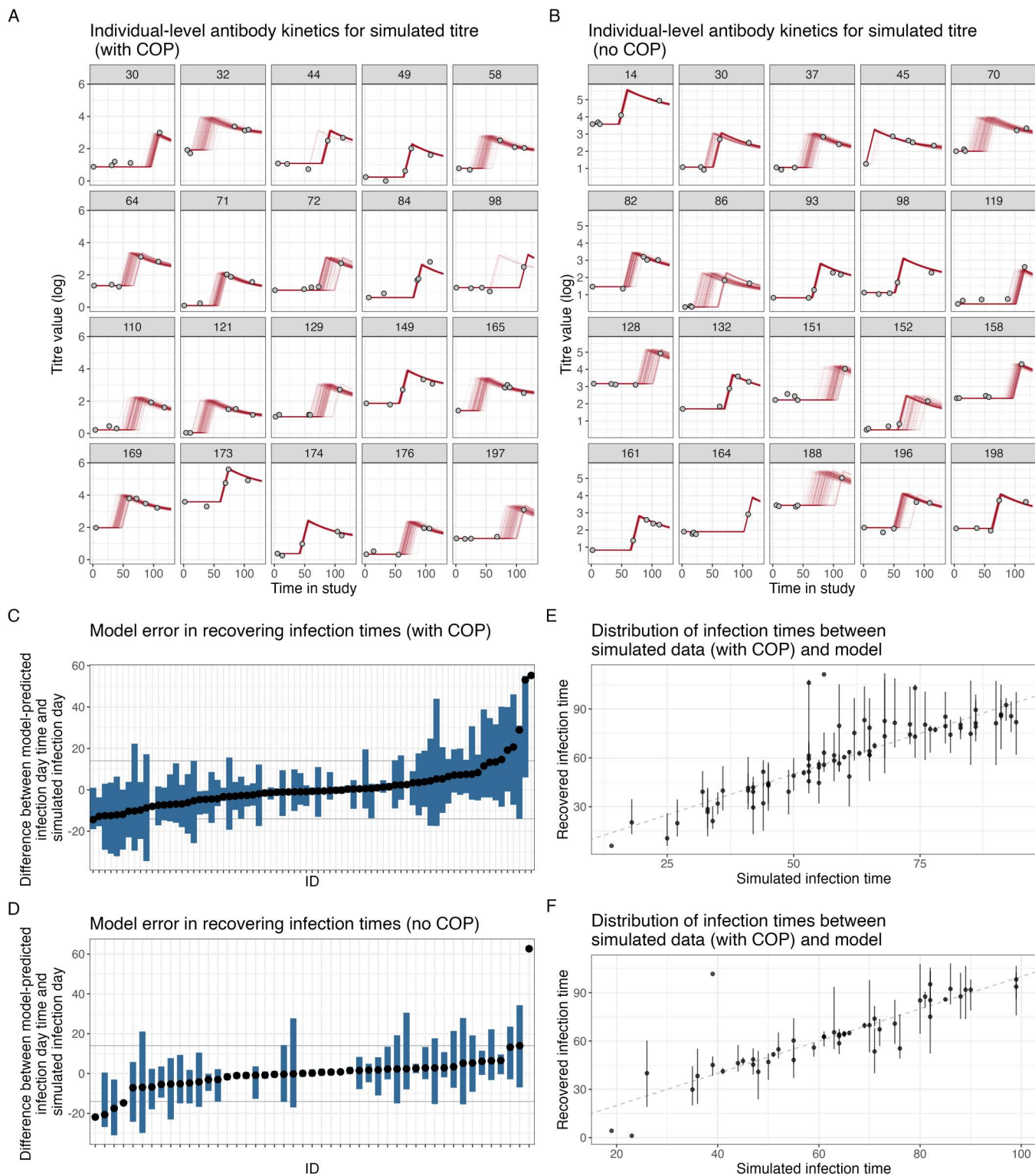


Fig 1. Comparison of simulation recovery for With COP and No COP models with observational error (0.1). (A, B) Individual-level antibody kinetics for a subset of individuals, showing observed data points (grey dots), and individual-level posterior medians (red lines) for With COP data and No COP data. (C, D) Model error (mean with 95% credible intervals) in recovering infection times, represented as the difference between simulated and model-predicted infection days for infected individuals With COP data and No COP data. (E, F) Distribution of infection times, comparing simulated infection times with recovered infection times.

<https://doi.org/10.1371/journal.pcbi.1013467.g001>

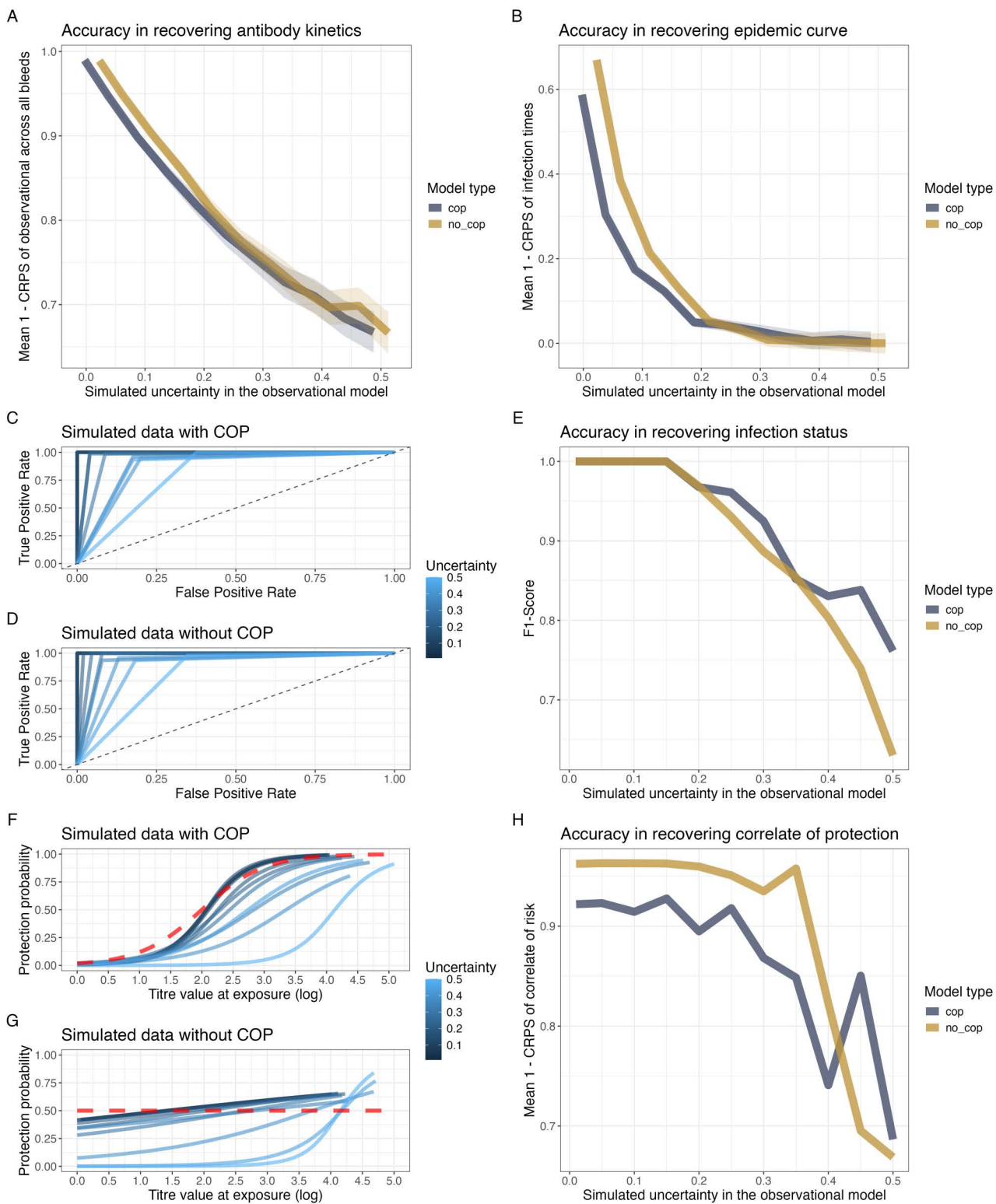


Fig 2. Evaluation of model performance under varying levels of observational uncertainty. (A) Accuracy in recovering antibody kinetics over various simulated observational errors, measured as the mean CRPS across all individuals of the observational model (B) Accuracy in recovering the epidemic curve over various simulated observational errors, assessed using the mean CRPS across all infection times (C, D) ROC curves for infection

status predictions under simulated data with COP (C) and without COP (D) under different observational error (colour) (E) F1-scores for recovering infection status across different observational errors and model types (No COP and With COP). (F, G) Posterior means of the COP curves across various observational uncertainty values (blue colours) compared to the simulated function (dashed red line) (H) Accuracy in recovering the functional form for the COP, measured by Mean CRPS, across various levels of observational uncertainty for both model types (With COP and No COP).

<https://doi.org/10.1371/journal.pcbi.1013467.g002>

noise is below 0.3, with CRPS values less than 0.05 (Fig 2F–H). In high-noise settings, we observe that underestimation of the correlate of protection (COP) curve can lead to elevated false positive rates, as individuals with relatively high antibody titres may still be incorrectly inferred as infected. This interplay between inferred protection and infection classification underscores the difficulty of simultaneously recovering both infection dynamics and immunity correlates, particularly when data are noisy or sparse.

3.2 Case study 2: serological inference for SARS-CoV-2 in the Gambia during the delta wave

3.2.1. Testing sensitivity of serological detection methods. We evaluate the sensitivity of various serological detection heuristics in identifying infection status. Using each of the five infection classification strategies based on serology (*serojump*, a four-fold rise in spike or NCP, and seropositivity thresholds for spike and NCP), we classified each of the PCR-confirmed individuals by their derived infection status and then calculated the sensitivity of each method (Fig 3). The *serojump* method demonstrates the highest sensitivity, outperforming the other four heuristics. The four-fold spike rise achieves the second-highest sensitivity, closely followed by the four-fold NCP rise. Both seropositive threshold methods (spike and NCP) perform less well, highlighting the limitations of using simple thresholds for infection detection. These observations highlight that leveraging dynamic changes in antibody kinetics can outperform static thresholds for

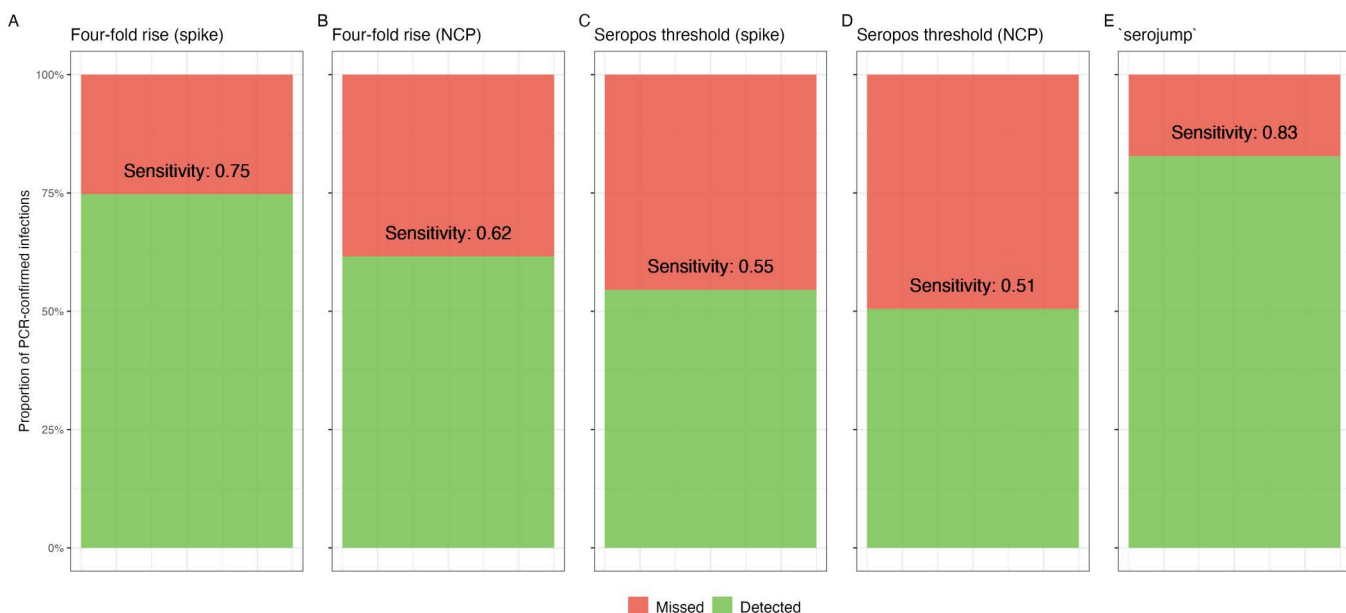


Fig 3. Sensitivity of serological detection heuristics in identifying infection status for 99 PCR-positive individuals using serological data only. Posterior probabilities of recovery for individuals using five detection methods: (A) Four-fold spike rise, (B) Four-fold NCP rise, (C) Seropositive threshold spike, (D) Seropositive threshold NCP, and (E) *serojump*. Green bars represent individuals inferred as true positives, while red points indicate false negatives.

<https://doi.org/10.1371/journal.pcbi.1013467.g003>

serological detection of infection, with the *serojump* method demonstrating superior sensitivity, indicating its potential as a preferred approach for infection inference.

3.2.2. Combining PCR and serological data to infer epidemiological dynamics of the Delta wave. We estimate the dynamics of antibody responses, epidemic patterns, and infection risk in the context of SARS-CoV-2 infection and vaccination during the Delta wave in The Gambia. This uses the same serological dataset described in Section 3.1 but includes PCR-confirmed infections, anchoring the infection status and infection time for 99 individuals. We infer antibody kinetics trajectories for the biomarkers, spike and NCP, for individuals infected prior to the Delta variant, during the Delta wave, and following vaccination (Fig 4A). We find that infections with Delta and infections with variants before Delta stimulate both spike and NCP responses, with spike responses seeing more sustained responses compared to NCP responses (AUC: 395 for spike and 257 for NCP for Delta infection). The spike responses remain above a four-fold rise up until day 46 for infection with the Delta variant. Following vaccination, we find robust spike responses but very low NCP responses (AUC: 607 vs 67 for spike and NCP, respectively). These observations are consistent with the spike-based composition of the vaccine used. We also summarise the total number of infections inferred from the *serojump* framework and compare this to the known PCR-confirmed infections (Fig 4B and 4C). Considering only PCR-confirmed infections, we find that the attack rate is 30%, increasing to 52% when we use *serojump*, implying analysing changes in serological kinetics finds many infections which are missed through virological surveillance alone. Using our informed prior ensures that the inferred timings of these missed infections agree with the known epidemic wave. Finally, we compared the relationship between titre levels at infection and the individual-level posterior probability of infection, serving as a correlate of risk (COR) (Fig 4D). For NCP and spike biomarkers, a negative relationship is observed between titre levels at infection and the probability of infection, suggesting that higher antibody titres are associated with reduced infection risk. When we consider the CPC after normalising the titre scales, spike has a slightly steeper gradient compared to NCP, suggesting a stronger correlate of protection (Fig 4E). In addition, high titre values providing only 60% protection suggest the Ancestral spike and NCP proteins correlate with protection despite being antigenically distant from the infecting variant (Delta), suggesting cross-immunity with the analysed proteins. The RR shows that having a log spike titre of 3.2 provides 50% more protection against infection compared to those with no measurable titre (Fig 4F).

4. Discussion

In this study, we present an open-source flexible framework for serological inference, *serojump*, which uses a reversible-jump MCMC algorithm to probabilistically estimate whether an individual has been infected or not. Specifically, it can infer individual-level infections, antibody kinetics, and population-level dynamics using longitudinal serological data. Our approach uses mechanistic modelling of antibody kinetics to estimate latent infection states and provides an alternative to conventional serological heuristics, which are often based on crude static thresholds. Our results demonstrate that *serojump* can achieve high sensitivity and specificity in recovering infection status, infection timing, and population-level epidemic patterns from simulated and real-world datasets. The framework effectively captures antibody kinetics post-infection and vaccination, allowing us to differentiate between infections during different epidemic waves and post-vaccination responses. Specifically, the inferred kinetics show that vaccination induces robust spike antibody responses but negligible NCP responses, consistent with the composition of SARS-CoV-2 vaccines and previous observations.[36,37] Our results also show how static heuristics, such as four-fold rises in antibody titres or seropositivity thresholds, underperform compared to the *serojump* framework. This underscores the need for dynamic, data-driven methods to accurately infer infections, especially in scenarios where virological data are incomplete or unavailable. For example, using real-world data from The Gambia, *serojump* inferred an attack rate significantly higher than that estimated by PCR-confirmed cases alone, suggesting under-detection by virological sampling methods, potentially due to a relatively short period of peak viral shedding [4].

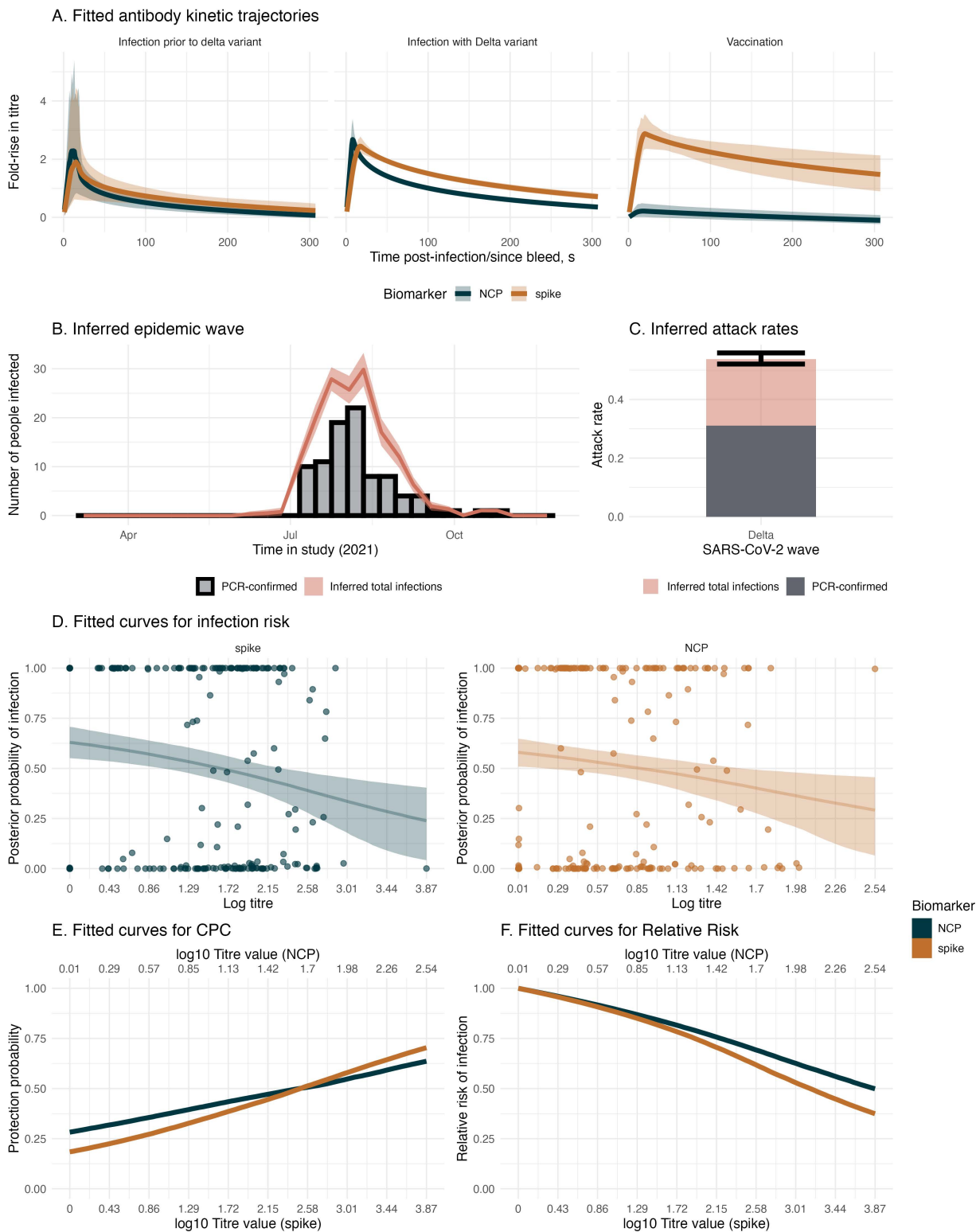


Fig 4. Antibody kinetics, recovery of infection timings, and correlates of protection from empirical data with PCR information. (A) Fitted antibody kinetic trajectories for individuals with infection prior to the Delta variant (left), infection during the Delta variant wave (center), and following vaccination (right), showing fold-rise in antibody titre over time for NCP (blue) and spike (orange) biomarkers. Shaded regions represent 95% credible intervals. (B) Inferred epidemic wave during the Delta wave, illustrating the temporal distribution of PCR-confirmed cases (black bars) and total inferred

infections (pink shaded curve), with shaded uncertainty. (C) The estimated attack rate during the Delta wave, partitioned into PCR-confirmed cases (dark grey) and total inferred infections (pink). Error bars indicate 95% credible intervals. (D) The fitted logistic curve to the infection risk, showing the posterior probability of infection as a function of antibody titre at infection for NCP (blue) and spike (orange) biomarkers, with shaded regions representing 95% credible intervals. (E) The CPC from the fitted infection risk curve and (F) shows the relative risk for the same biomarkers.

<https://doi.org/10.1371/journal.pcbi.1013467.g004>

Our framework addresses several challenges associated with serological studies, including variability in antibody responses and the need for flexible, biologically informed models. By integrating individual- and population-level data into a unified probabilistic framework, *serojump* offers a powerful tool for assessing pathogen transmission dynamics, evaluating vaccine efficacy, and identifying correlates of protection. These capabilities are critical for informing public health interventions, especially in resource-limited settings where virological sampling may be sparse. Furthermore, the negative relationship between antibody titres at infection and the individual-level probability of infection observed in our analysis highlights the potential for *serojump* to identify correlates of risk and protection. These insights can guide the development of immunological benchmarks for vaccine efficacy and inform strategies for boosting population immunity. We highlight that the framework is currently pathogen agnostic and thus is adaptable to multiple pathogens with different serological profiles, highlighting the robustness of the model in handling diverse datasets, which could include varying sample sizes, intervals, and noise levels.

Inferring latent infection status and timings from serological data is a highly complex inference challenge. This task requires joint modelling of epidemic dynamics, variation in individual infection timing and status, and complex antibody kinetics that depend on the timing, nature, and antigens of infections. While it would be ideal to have a single framework capable of addressing all these mechanisms simultaneously, such an approach is challenging and demands extensive high-quality data. As a result, different software tools have been developed, each targeting specific dimensions of this complexity and tailored to particular types of serological data and research questions. For instance, *Rsero* [38] and *serofoi* [39] use serocatalytic models to estimate age- and time-specific forces of infection from seroprevalence data. *serosolver* [25] is designed to reconstruct individual-level infection histories across multiple epidemic waves, often involving multiple antigens and using both cross-sectional and longitudinal data in discrete time steps. In contrast, *serojump* is optimised for studies with dense temporal sampling over shorter periods, focusing on the precise inference of single infection events. Although *serojump* and *serosolver* share the same underlying likelihood model [24], they are designed for different epidemiological contexts. *serosolver* is suited to long-term dynamics and complex infection histories, including multiple infections and antigenic relationships. *serojump*, by contrast, prioritises high temporal resolution in short-term outbreaks, supports continuous-time sampling, and links biomarkers to infection probability.

A key design trade-off in *serojump* lies between model flexibility and computational feasibility. It assumes shared, parametric antibody kinetics across individuals and excludes random effects, as incorporating individual-level heterogeneity proved infeasible within the reversible-jump MCMC framework during development. Instead, it relies on parsimonious, biologically grounded models that balance realism and robustness—particularly valuable when working with sparse or noisy data. More flexible approaches, such as splines or piecewise functions, are intentionally avoided due to potential identifiability issues when disentangling infection timing from antibody dynamics. Ultimately, *serosolver* and *serojump* offer complementary approaches to serological inference. By tailoring each tool to distinct epidemiological settings—whether long-term, antigenically complex dynamics or short-term, high-resolution outbreak analysis—researchers are better equipped to address the diverse and nuanced challenges of interpreting serological data.

Previous studies have used Reversible-Jump MCMC methods to infer individual-level infection times and statuses from serological data [15,16]. These approaches model functional forms of antibody kinetics to estimate infection times and/or the force of infection, and subsequently establish relationships between estimated antibody titres at the time of infection and protection against disease. While these studies share similarities, such as the ability to infer multiple infections for an individual, their models and codebases are often tailored specifically to the datasets or research questions

they address. As a result, they typically lack publicly available, user-friendly interfaces for broader applications to other datasets. In contrast, the methods implemented in *serojump* focus on a streamlined design that simplifies inference by only allowing a single infection to be inferred. This simplification enables a flexible and modular approach, allowing users to adapt and extend the model framework with user-written code. Unlike many existing methods, *serojump* provides a plug-and-play interface, offering general usability without requiring the re-coding of bespoke model structures, thus making advanced inference techniques more accessible to a broader range of users. While the implementation uses fixed or adaptively-tuned proposal distributions optimised for general use, future versions may allow user-defined proposal, offering advanced users greater control over inference behaviour in novel settings. This flexibility represents a significant advance in accessibility and usability, facilitating broader adoption and application across different datasets.

Our study has limitations. The accuracy of the *serojump* framework depends on the quality of input data, including the timing, frequency, and noise levels of serological samples. However, simulation studies like the ones presented above can be used to assess the likely performance of the method with a given data structure. In particular, systematic assessment of model misspecification—such as evaluating performance when fitting incorrect kinetic forms—remains an important area for future development, both for *serojump* and other serological inference frameworks. Our framework does not currently account for individual-level variation in kinetics (e.g., random effects) or the influence of covariates on antibody kinetics. This choice was motivated by computational constraints: estimating individual-specific parameters would greatly expand the parameter space requiring a prohibitively large number of MCMC iterations and complicating convergence. While more efficient reparameterisations (e.g., non-centered parameterisation) may eventually allow such extensions, our current goal was to provide a practical, general-purpose tool prioritising flexibility, usability, and convergence for typical use cases. Incorporating individual heterogeneity remains a key area for future development. Scalability and runtime also pose computational challenges; while effective for smaller datasets, further optimisation would be beneficial for applications exceeding 1,000 individuals. Potential extensions to the framework include inferring multiple infections per individual and incorporating multiple biomarkers for antigenically varied pathogens. Conceptually, this would require expanding the reversible-jump MCMC framework to accommodate multiple sets of latent infection times and corresponding antibody boosts per individual, with careful handling of identifiability, particularly in cases of overlapping or weakly-separated responses. These additions could enable the exploration of longer-term immunological phenomena (e.g., antigenic seniority, [40]) and improve inference for pathogens like influenza and SARS-CoV-2. However, such extensions would significantly increase the parameter space and computational demands. Strategies such as optimising samplers or integrating population-level MCMC algorithms (e.g., parallel tempering) could mitigate these challenges and allow for more complex frameworks to be evaluated in the future.

Therefore, we developed a general-purpose statistical framework that leverages reversible jump MCMC to systematically infer antibody kinetics, infections and their timing using individual-level antibody kinetics for an epidemic outbreak. This tool provides a robust, flexible approach to better inform epidemiological parameters from serological data. By making these advanced methods more accessible and adaptable in the form of an R package *serojump*, we hope this framework will substantially enhance our ability to track and control infectious diseases in diverse settings.

Supporting information

S1 Methods. Supporting documentation for the methods section.
(DOCX)

S1 Fig. Convergence diagnostics for simulated data with COP and 0.1 uncertainty in observational error. (A)

Trace plots for fitted parameters (σ , wane, a , b , c) across four Markov chains, illustrating the mixing and convergence of the parameters. **(B)** Trace plots for the log posterior across the four chains, showing the variability and stabilization of the log posterior over iterations. **(C)** Convergence diagnostics for fitted parameters, including effective sample size

(`ess_bulk`, `ess_tail`) and R^{hat} , which assess the adequacy of sampling and convergence for each parameter. **(D)** Trace plots for transdimensional convergence of the model dimension, with histogram counts of model dimensions sampled across the chains. **(E)** Trace plots for transdimensional convergence for the SMI (Structural Model Index) and histogram counts of the log-transformed SMI values across chains. **(F)** Convergence diagnostics for transdimensional parameters, including effective sample size (`ess_bulk`, `ess_tail`) and R^{hat} , summarizing the adequacy of sampling and convergence for the transdimensional space.

(PNG)

S2 Fig. Convergence diagnostics of infection timing for simulated data with COP and 0.1 uncertainty in observational error. **(A)** Trace plots for the timing of infection for individuals with posterior $P(Z) > 0.5$ display estimates across four Markov chains. Each point and its uncertainty interval reflect the sampled infection timing for each individual over iterations. **(B)** Convergence diagnostics for the timing of infection for individuals with posterior $P(Z) > 0.5$, showing R^{hat} values for each individual. The red dashed line indicates the threshold for $R^{\text{hat}} = 1.1$, which marks convergence.

(PNG)

S3 Fig. Convergence diagnostics for simulated data No COP and 0.1 uncertainty in observational error. **(A)** Trace plots for fitted parameters (`sigma`, `wane`, `a`, `b`, `c`) across four Markov chains, illustrating the mixing and convergence of the parameters. **(B)** Trace plots for the log posterior across the four chains, showing the variability and stabilization of the log posterior over iterations. **(C)** Convergence diagnostics for fitted parameters, including effective sample size (`ess_bulk`, `ess_tail`) and R^{hat} , which assess the adequacy of sampling and convergence for each parameter. **(D)** Trace plots for transdimensional convergence of the model dimension, with histogram counts of model dimensions sampled across the chains. **(E)** Trace plots for transdimensional convergence for the SMI (Structural Model Index) and histogram counts of the log-transformed SMI values across chains. **(F)** Convergence diagnostics for transdimensional parameters, including effective sample size (`ess_bulk`, `ess_tail`) and R^{hat} , summarizing the adequacy of sampling and convergence for the transdimensional space.

(PNG)

S4 Fig. Convergence diagnostics of infection timing for simulated data No COP and 0.1 uncertainty in observational error. **(A)** Trace plots for the timing of infection for individuals with posterior $P(Z) > 0.5$ display estimates across four Markov chains. Each point and its uncertainty interval reflect the sampled infection timing for each individual over iterations. **(B)** Convergence diagnostics for the timing of infection for individuals with posterior $P(Z) > 0.5$, showing R^{hat} values for each individual. The red dashed line indicates the threshold for $R^{\text{hat}} = 1.1$, which marks convergence.

(PNG)

S5 Fig. Convergence diagnostics for empirical data without PCR information **(A)** Trace plots for fitted parameters (`sigma`, `wane`, `a`, `b`, `c`) across four Markov chains, illustrating the mixing and convergence of the parameters. **(B)** Trace plots for the log posterior across the four chains, showing the variability and stabilization of the log posterior over iterations. **(C)** Convergence diagnostics for fitted parameters, including effective sample size (`ess_bulk`, `ess_tail`) and R^{hat} , which assess the adequacy of sampling and convergence for each parameter. **(D)** Trace plots for transdimensional convergence of the model dimension, with histogram counts of model dimensions sampled across the chains. **(E)** Trace plots for transdimensional convergence for the SMI (Structural Model Index) and histogram counts of the log-transformed SMI values across chains. **(F)** Convergence diagnostics for transdimensional parameters, including effective sample size (`ess_bulk`, `ess_tail`) and R^{hat} , summarizing the adequacy of sampling and convergence for the transdimensional space.

(PNG)

S6 Fig. Convergence diagnostics for empirical data without PCR information. **(A)** Trace plots for the timing of infection for individuals with posterior $P(Z) > 0.5$ display estimates across four Markov chains. Each point and its uncertainty

interval reflect the sampled infection timing for each individual over iterations. **(B)** Convergence diagnostics for the timing of infection for individuals with posterior $P(Z) > 0.5$, showing R^{hat} values for each individual. The red dashed line indicates the threshold for $R^{\text{hat}} = 1.1$, which marks convergence.

(PNG)

S7 Fig: Convergence diagnostics for empirical data with PCR information **(A)** Trace plots for fitted parameters (σ , w , a , b , c) across four Markov chains, illustrating the mixing and convergence of the parameters. **(B)** Trace plots for the log posterior across the four chains, showing the variability and stabilization of the log posterior over iterations. **(C)** Convergence diagnostics for fitted parameters, including effective sample size (ess_bulk , ess_tail) and R^{hat} , which assess the adequacy of sampling and convergence for each parameter. **(D)** Trace plots for transdimensional convergence of the model dimension, with histogram counts of model dimensions sampled across the chains. **(E)** Trace plots for transdimensional convergence for the SMI (Structural Model Index) and histogram counts of the log-transformed SMI values across chains. **(F)** Convergence diagnostics for transdimensional parameters, including effective sample size (ess_bulk , ess_tail) and R^{hat} , summarizing the adequacy of sampling and convergence for the transdimensional space.

(PNG)

S8 Fig. Convergence diagnostics for empirical data with PCR. **(A)** Trace plots for the timing of infection for individuals with posterior $P(Z) > 0.5$ display estimates across four Markov chains. Each point and its uncertainty interval reflect the sampled infection timing for each individual over iterations. **(B)** Convergence diagnostics for the timing of infection for individuals with posterior $P(Z) > 0.5$, showing R^{hat} values for each individual. The red dashed line indicates the threshold for $R^{\text{hat}} = 1.1$, which marks convergence.

(PNG)

S9 Fig. Antibody kinetics, recovery of infection timings, and correlates of protection from empirical data without PCR information. **(A)** Fitted antibody kinetic trajectories for individuals with infection prior to the Delta variant (left), infection during the Delta variant wave (center), and following vaccination (right), showing fold-rise in antibody titre over time for NCP (blue) and spike (orange) biomarkers. Shaded regions represent 95% credible intervals. **(B)** Inferred epidemic wave during the Delta wave, illustrating the temporal distribution of PCR-confirmed cases (black bars) and total inferred infections (pink shaded curve), with shaded uncertainty. **(C)** The estimated attack rate during the Delta wave, partitioned into PCR-confirmed cases (dark grey) and total inferred infections (pink). Error bars indicate 95% credible intervals. **(D)** The fitted logistic curve to the infection risk, showing the posterior probability of infection as a function of antibody titre at infection for NCP (blue) and spike (orange) biomarkers, with shaded regions representing 95% credible intervals. **(E)** The CPC from the fitted infection risk curve and **(F)** shows the RR for the same biomarkers.

(PNG)

Author contributions

Conceptualization: David Hodgson, Thushan I. de Silva, Adam J. Kucharski.

Data curation: Sheikh Jarju, Dawda Jobe, Rhys Wenlock.

Formal analysis: David Hodgson.

Funding acquisition: Thushan I. de Silva, Adam J. Kucharski.

Investigation: Sheikh Jarju, Dawda Jobe.

Methodology: David Hodgson, James Hay, Thushan I. de Silva, Adam J. Kucharski.

Project administration: Thushan I. de Silva, Adam J. Kucharski.

Resources: Sheikh Jarju, Dawda Jobe.

Software: David Hodgson, James Hay.

Supervision: Thushan I. de Silva, Adam J. Kucharski.

Validation: David Hodgson, Sheikh Jarju, Dawda Jobe.

Visualization: David Hodgson, James Hay.

Writing – original draft: David Hodgson, Adam J. Kucharski.

Writing – review & editing: David Hodgson, James Hay, Sheikh Jarju, Dawda Jobe, Rhys Wenlock, Thushan I. de Silva, Adam J. Kucharski.

References

1. Cutts FT, Hanson M. Seroepidemiology: an underused tool for designing and monitoring vaccination programmes in low- and middle-income countries. *Trop Med Int Health*. 2016;21(9):1086–98. <https://doi.org/10.1111/tmi.12737> PMID: [27300255](https://pubmed.ncbi.nlm.nih.gov/27300255/)
2. Hay JA, Routledge I, Takahashi S. Serodynamics: A primer and synthetic review of methods for epidemiological inference using serological data. *Epidemics*. 2024;49(100806):100806.
3. Haselbeck AH, Im J, Prifti K, Marks F, Holm M, Zellweger RM. Serology as a Tool to Assess Infectious Disease Landscapes and Guide Public Health Policy. *Pathogens*. 2022;11(7):732. <https://doi.org/10.3390/pathogens11070732> PMID: [35889978](https://pubmed.ncbi.nlm.nih.gov/35889978/)
4. Hellewell J, Russell TW, SAFER Investigators and Field Study Team, Crick COVID-19 Consortium, CMMID COVID-19 working group, Beale R, et al. Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections. *BMC Med*. 2021;19(1):106. <https://doi.org/10.1186/s12916-021-01982-x> PMID: [33902581](https://pubmed.ncbi.nlm.nih.gov/33902581/)
5. Crawford DH, Macsween KF, Higgins CD, Thomas R, McAulay K, Williams H, et al. A cohort study among university students: identification of risk factors for Epstein-Barr virus seroconversion and infectious mononucleosis. *Clin Infect Dis*. 2006;43(3):276–82. <https://doi.org/10.1086/505400> PMID: [16804839](https://pubmed.ncbi.nlm.nih.gov/16804839/)
6. Wansom T, Muangnoicharoen S, Nitayaphan S, Kitsiripornchai S, Crowell TA, Francisco L, et al. Risk Factors for HIV sero-conversion in a high incidence cohort of men who have sex with men and transgender women in Bangkok, Thailand. *EClinicalMedicine*. 2021;38:101033. <https://doi.org/10.1016/j.eclinm.2021.101033> PMID: [34505031](https://pubmed.ncbi.nlm.nih.gov/34505031/)
7. Dhar-Chowdhury P, Paul KK, Haque CE, Hossain S, Lindsay LR, Dibernardo A, et al. Dengue seroprevalence, seroconversion and risk factors in Dhaka, Bangladesh. *PLoS Negl Trop Dis*. 2017;11(3):e0005475. <https://doi.org/10.1371/journal.pntd.0005475> PMID: [28333935](https://pubmed.ncbi.nlm.nih.gov/28333935/)
8. Chan Y, Fornace K, Wu L, Arnold BF, Priest JW, Martin DL, et al. Determining seropositivity—A review of approaches to define population seroprevalence when using multiplex bead assays to assess burden of tropical diseases. *PLoS Negl Trop Dis*. 2021;15(6):e0009457. <https://doi.org/10.1371/journal.pntd.0009457> PMID: [34181665](https://pubmed.ncbi.nlm.nih.gov/34181665/)
9. van den Berg OE, Stanoeva KR, Zonneveld R, Hoek-van Deursen D, van der Klis FR, van de Kasstelee J, et al. Seroprevalence of *Toxoplasma gondii* and associated risk factors for infection in the Netherlands: third cross-sectional national study. *Epidemiol Infect*. 2023;151:e136. <https://doi.org/10.1017/S095026882300122X> PMID: [37503608](https://pubmed.ncbi.nlm.nih.gov/37503608/)
10. Colton H, Hodgson D, Hornsby H, Brown R, Mckenzie J, Bradley KL, et al. Risk factors for SARS-CoV-2 seroprevalence following the first pandemic wave in UK healthcare workers in a large NHS Foundation Trust. *Wellcome Open Res*. 2022;6:220. <https://doi.org/10.12688/wellcomeopenres.17143.3> PMID: [35600250](https://pubmed.ncbi.nlm.nih.gov/35600250/)
11. Xu C, Liu L, Ren B, Dong L, Zou S, Huang W, et al. Incidence of influenza virus infections confirmed by serology in children and adult in a suburb community, northern China, 2018–2019 influenza season. *Influenza Other Respir Viruses*. 2021;15(2):262–9. <https://doi.org/10.1111/irv.12805> PMID: [32978902](https://pubmed.ncbi.nlm.nih.gov/32978902/)
12. Cauchemez S, Horby P, Fox A, Mai LQ, Thanh LT, Thai PQ, et al. Influenza infection rates, measurement errors and the interpretation of paired serology. *PLoS Pathog*. 2012;8(12):e1003061. <https://doi.org/10.1371/journal.ppat.1003061> PMID: [23271967](https://pubmed.ncbi.nlm.nih.gov/23271967/)
13. Borremans B, Hens N, Beutels P, Leirs H, Reijnders J. Estimating Time of Infection Using Prior Serological and Individual Information Can Greatly Improve Incidence Estimation of Human and Wildlife Infections. *PLoS Comput Biol*. 2016;12(5):e1004882. <https://doi.org/10.1371/journal.pcbi.1004882> PMID: [27177244](https://pubmed.ncbi.nlm.nih.gov/27177244/)
14. Pepin KM, Kay SL, Golas BD, Shriner SS, Gilbert AT, Miller RS, et al. Inferring infection hazard in wildlife populations by linking data across individual and population scales. *Ecol Lett*. 2017;20(3):275–92. <https://doi.org/10.1111/ele.12732> PMID: [28090753](https://pubmed.ncbi.nlm.nih.gov/28090753/)
15. Salje H, Cummings DAT, Rodriguez-Barraquer I, Katzelnick LC, Lessler J, Klungthong C, et al. Reconstruction of antibody dynamics and infection histories to evaluate dengue risk. *Nature*. 2018;557(7707):719–23. <https://doi.org/10.1038/s41586-018-0157-4> PMID: [29795354](https://pubmed.ncbi.nlm.nih.gov/29795354/)
16. Tsang TK, Perera RAPM, Fang VJ, Wong JY, Shiu EY, So HC, et al. Reconstructing antibody dynamics to estimate the risk of influenza virus infection. *Nat Commun*. 2022;13(1):1557. <https://doi.org/10.1038/s41467-022-29310-8> PMID: [35322048](https://pubmed.ncbi.nlm.nih.gov/35322048/)
17. Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*. 1995;82(4):711–32. <https://doi.org/10.1093/biomet/82.4.711>

18. Gibson GJ. Markov Chain Monte Carlo Methods for Fitting Spatiotemporal Stochastic Models in Plant Epidemiology. *Journal of the Royal Statistical Society Series C: Applied Statistics*. 1997;46(2):215–33. <https://doi.org/10.1111/1467-9876.00061>
19. Lyyjynen H. Reversible Jump Markov Chain Monte Carlo: Some Theory and Applications [Master's Thesis]. [Bergen]: University of Bergen; 2014.
20. Dellaportas P, Papageorgiou I. Multivariate mixtures of normals with unknown number of components. *Stat Comput*. 2006;16(1):57–68. <https://doi.org/10.1007/s11222-006-5338-6>
21. Rantini D, Iriawan N, Irhamah. On the Reversible Jump Markov Chain Monte Carlo (RJMCMC) Algorithm for Extreme Value Mixture Distribution as a Location-Scale Transformation of the Weibull Distribution. *Applied Sciences*. 2021;11(16):7343. <https://doi.org/10.3390/app11167343>
22. Richardson S, Green PJ. On Bayesian analysis of mixtures with an unknown number of components. *J R Stat Soc B*. 1997;59(4):731–92.
23. Hastie DJ, Green PJ. Model choice using reversible jump Markov chain Monte Carlo. *Statistica Neerlandica*. 2012;66(3):309–38. <https://doi.org/10.1111/j.1467-9574.2012.00516.x>
24. Menezes A, Takahashi S, Routledge I, Metcalf CJE, Graham AL, Hay JA. serosim: An R package for simulating serological data arising from vaccination, epidemiological and antibody kinetics processes. *PLoS Comput Biol*. 2023;19(8):e1011384. <https://doi.org/10.1371/journal.pcbi.1011384> PMID: [37578985](https://pubmed.ncbi.nlm.nih.gov/37578985/)
25. Hay JA, Minter A, Ainslie KEC, Lessler J, Yang B, Cummings DAT, et al. An open source tool to infer epidemiological and immunological dynamics from serological data: serosolver. *PLoS Comput Biol*. 2020;16(5):e1007840. <https://doi.org/10.1371/journal.pcbi.1007840> PMID: [32365062](https://pubmed.ncbi.nlm.nih.gov/32365062/)
26. Hobson D, Curry RL, Beare AS, Ward-Gardner A. The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses. *J Hyg (Lond)*. 1972;70(4):767–77. <https://doi.org/10.1017/s0022172400022610> PMID: [4509641](https://pubmed.ncbi.nlm.nih.gov/4509641/)
27. Khoury DS, Cromer D, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat Med*. 2021;27(7):1205–11. <https://doi.org/10.1038/s41591-021-01377-8> PMID: [34002089](https://pubmed.ncbi.nlm.nih.gov/34002089/)
28. Srivastava K, Carreño JM, Gleason C, Monahan B, Singh G, Abbad A, et al. SARS-CoV-2-infection- and vaccine-induced antibody responses are long lasting with an initial waning phase followed by a stabilization phase. *Immunity*. 2024;57(3):587–599.e4. <https://doi.org/10.1016/j.immuni.2024.01.017> PMID: [38395697](https://pubmed.ncbi.nlm.nih.gov/38395697/)
29. Jarju S, Wenlock RD, Danso M, Jobe D, Jagne YJ, Darboe A, et al. High SARS-CoV-2 incidence and asymptomatic fraction during Delta and Omicron BA.1 waves in The Gambia. *Nat Commun*. 2024;15(1):3814. <https://doi.org/10.1038/s41467-024-48098-3> PMID: [38714680](https://pubmed.ncbi.nlm.nih.gov/38714680/)
30. Teunis PFM, van Eijkeren JCH, de Graaf WF, Marinović AB, Kretzschmar MEE. Linking the seroresponse to infection to within-host heterogeneity in antibody production. *Epidemics*. 2016;16:33–9. <https://doi.org/10.1016/j.epidem.2016.04.001> PMID: [27663789](https://pubmed.ncbi.nlm.nih.gov/27663789/)
31. Nilles EJ, Roberts K, Aubin M de S, Mayfield H, Restrepo AC, Garnier S, et al. Convergence of SARS-CoV-2 spike antibody levels to a population immune setpoint. *eBioMedicine [Internet]*. 2024 Oct 1 [cited 2025 Jul 3];108. Available from: https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964%2824%2900355-4/fulltext?utm_source=chatgpt.com
32. R Core Team. R: A Language and Environment for Statistical Computing. 2016; Available from: <http://www.r-project.org/>
33. Seamless R and C++ Integration [R package Rcpp version 1.0.14]. 2025 Jan 12 [cited 2025 Mar 4]; Available from: <https://CRAN.R-project.org/package=Rcpp>
34. Somogyvári M, Reich S. Convergence Tests for Transdimensional Markov Chains in Geoscience Imaging. *Math Geosci*. 2019;52(5):651–68. <https://doi.org/10.1007/s11004-019-09811-x>
35. Zamo M, Naveau P. Estimation of the Continuous Ranked Probability Score with Limited Information and Applications to Ensemble Weather Forecasts. *Math Geosci*. 2017;50(2):209–34. <https://doi.org/10.1007/s11004-017-9709-7>
36. Bayart J-L, Morimont L, Closset M, Wieërs G, Roy T, Gerin V, et al. Confounding Factors Influencing the Kinetics and Magnitude of Serological Response Following Administration of BNT162b2. *Microorganisms*. 2021;9(6):1340. <https://doi.org/10.3390/microorganisms9061340> PMID: [34205564](https://pubmed.ncbi.nlm.nih.gov/34205564/)
37. Błaszczuk A, Michalski A, Malm M, Drop B, Polz-Dacewicz M. Antibodies to NCP, RBD and S2 SARS-CoV-2 in Vaccinated and Unvaccinated Healthcare Workers. *Vaccines (Basel)*. 2022;10(8):1169. <https://doi.org/10.3390/vaccines10081169> PMID: [35893818](https://pubmed.ncbi.nlm.nih.gov/35893818/)
38. nathoze. nathoze/Rsero [Internet]. 2025 [cited 2025 Jul 3]. Available from: <https://github.com/nathoze/Rsero>
39. Cucunubá ZM, T. Domínguez N, Lambert B, Nouvellet P. serofoi: Bayesian Estimation of the Force of Infection from Serological Data [Internet]. 2025 [cited 2025 Jul 3]. Available from: <https://github.com/epiverse-trace/serofoi>
40. Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, et al. Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China. *PLoS Pathog*. 2012;8(7):e1002802. <https://doi.org/10.1371/journal.ppat.1002802> PMID: [22829765](https://pubmed.ncbi.nlm.nih.gov/22829765/)