# Resolving Parameter Uncertainty in Outbreak Models Through Population-Level Serological Surveillance

Binod Pant[1,2], Matthew E. Levine[3,4], Anjalika Nande[5], Raúl Garrido García[1,6], George Dewey[1,2], Nicholas B. Link[1,7], and Mauricio Santillana[1,2,8]*

[1] *Machine Intelligence Group for the Betterment of Health and the Environment, Northeastern University, Boston, MA, 02115, USA*

[2] *Network Science Institute, Northeastern University, Boston, MA, 02115, USA*

[3] *Eric and Wendy Schmidt Center, Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA*

[4] *Basis Research Institute, Cambridge, MA, 02142, USA*

[5] *Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD, 21218, USA*

[6] *Department of Physics, Northeastern University, Boston, MA, 02115, USA*

[7] *Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA*

[8] *Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, 02115, USA*

**Abstract**

Outbreak modeling faces a critical blind spot: the vast majority of infectious disease cases go undetected, yet epidemiological models frequently ignore this reality. During major outbreaks—from H1N1 to COVID-19—surveillance systems capture only a fraction of true infections, with detection rates often below 10%. When compartmental models treat these incomplete data as ground truth, they produce seemingly accurate fits while generating catastrophically wrong estimates of transmission rates, epidemic size, and intervention effectiveness. We reveal the mathematical foundations of this crisis through two distinct but related problems. First, ignoring case underdetection entirely yields parameter errors exceeding 1000%, even when model fits appear visually perfect. Second, when models explicitly account for underdetection by including case ascertainment ratios as unknown parameters, an infinite spectrum of epidemiologically distinct scenarios can fit the data equally well due to structural unidentifiability. Using rigorous structural identifiability analysis, we prove that transmission rates and detection ratios cannot be uniquely identified from detected case data alone. Strikingly, incorporating just a single population-level seroprevalence measurement transforms model reliability by breaking this mathematical degeneracy. Serological surveys directly measure cumulative population exposure, unlike case surveillance which captures only ongoing infections imperfectly. This integration reduces parameter uncertainty by orders of magnitude and enables accurate inference of transmission rates, peak timing, and final outbreak size. The approach remains robust under realistic noise conditions—with up to 40% measurement noise in observed case counts and 10% in seroprevalence data. Our findings expose a systematic vulnerability in pandemic preparedness while providing an actionable remedy, showing that population-level serosurvey costs are orders of magnitude smaller than the healthcare expenditures prevented through improved outbreak modeling. Our uncertainty quantification framework provides a systematic approach for assessing when surveillance data are adequate for epidemic inference, making serological surveillance integration both a mathematical necessity and a strategic investment in pandemic preparedness.

**Keywords:** Uncertainty quantification, Parameter uncertainty, Seroprevalence data, Bayesian inference

*Corresponding author: Email: *msantill@g.harvard.edu*

# 1 Introduction

Across infectious diseases—from H1N1 influenza to dengue fever and COVID-19—surveillance systems consistently capture only a fraction of actual infections, with documented cases representing anywhere from 10% to less than 0.1% of true infection burden [1–8]. This underascertainment of infections produces a systematic bias which is further distorted by time-varying factors including testing fluctuations, reporting delays, and policy shifts. When mathematical models of infectious disease transmission are calibrated to directly match underascertained epidemiological case counts without acknowledging this underdetection problem [9–12], they generate severely biased—frequently by several orders of magnitude—parameter estimates that can differ dramatically from those likely occurring in the population during epidemic outbreaks. A model might suggest an outbreak is nearly contained when community transmission is actually accelerating, or recommend insufficient vaccine coverage when herd immunity thresholds are much higher than estimated.

The COVID-19 pandemic placed mathematical models at the forefront of critical public health decision-making affecting millions of lives and trillions in economic activity [13, 14]. A major challenge faced by modeling efforts was the sparsity and unreliability of early pandemic data—true infections worldwide were estimated to be 3 to 1000 times higher than those reported [15,16]. The combination of poor data quality with many models' limited ability to account for this massive underdetection led to seemingly accurate (model to data) fits but catastrophically biased estimates of transmission dynamics and intervention effectiveness that varied widely across models [17–19]. The epidemiological modeling community eventually recognized this limitation [20], which ultimately motivated modeling efforts such as the US COVID-19 Forecast Hub [21] to rethink the use of incidence case data as modeling targets.

The stakes of parameter misestimation are not merely theoretical. During the COVID-19 pandemic, forecasts from a variety of modeling approaches — including machine learning models and compartmental models — were used in state and national planning. However, retrospective evaluation of four widely used models revealed highly inaccurate point predictions and poor calibration of uncertainty (e.g. only 10% of predictions fell within 10% of the "ground truth," and most 95% prediction intervals failed to achieve nominal coverage) [22]. This case highlights how misestimation of key epidemiological parameters - such as ascertainment ratio, effective reproduction numbers, infection rates, or epidemic size - can mislead scientific conclusions and risk leading to misguided policy decisions.

The mathematical roots of this crisis involve two distinct but related problems. First, when models ignore case underascertainment entirely—treating detected cases as if they represent all infections—the resulting parameter estimates become systematically biased, with errors that worsen as detection ratio decline. Even when such models achieve visually convincing fits to observed case data, their estimates of transmission rates, cumulative infections, recovery rates, and reproduction numbers can be catastrophically uninformative [20].

Second, when models explicitly account for incomplete detection by including case ascertainment ratios as unknown parameters, they face well-documented structural identifiability problems [17, 23]. The transmission rate and detection ratio become mathematically confounded in the observed data, meaning that multiple epidemiologically distinct scenarios can fit the data equally well. Standard inference procedures cannot distinguish between high transmission with low detection versus low transmission with high detection, creating what we term an "identifiability crisis" where confident-looking parameter estimates rest on mathematically unstable foundations.

Current approaches to address case underascertainment fall short of resolving these problems. Some studies attempt to correct for underdetection by incorporating hospitalization or mortality data [24, 25], but these approaches require accurate estimates of case-hospitalization or case-fatality ratios—themselves subject to the same ascertainment biases [26, 27]—in order to properly work [28]. Others explicitly model case detection processes by including case ascertainment ratios as unknown parameters [8, 29], but without additional constraining data sources, this approach may even worsen identifiability problems by introducing new unidentifiable parameters that cannot resolve the underlying bias. The result is a modeling field where models achieve good fits to observed case data, and thus produce seemingly precise epidemiological estimates, while harboring fundamental uncertainties that propagate through all derived quantities of interest, from basic reproduction numbers to epidemic peak timing and cumulative infections.

Here, we demonstrate that simultaneously fitting models to observed case data and population-level seroprevalence data, in the context of historical reconstructions of an epidemic outbreak, provides a mathematically rigorous solution to this modeling crisis. Unlike case surveillance, which captures ongoing infections imperfectly, seroprevalence surveys can directly measure cumulative population exposure when properly designed and sampled [30–32]. For models that ignore underascertainment entirely, seroprevalence data reveal the true scale of population exposure, allowing us to reject these fundamentally biased models. For models that explicitly include detection ratio as unknown parameters, seroprevalence measurements break the mathematical confounding between transmission rates and ascertainment ratio by providing an independent constraint on cumulative infections. By integrating even a single well-timed seroprevalence measurement with detected case data, the identifiability crisis dissolves: parameter uncertainties drop by orders of magnitude, enabling reliable inference of critical epidemic quantities including transmission rates, peak timing, and final outbreak size.

Through comprehensive computational (synthetic) experiments using SIR models [33], we demonstrate that this approach transforms unreliable inference into robust parameter estimation. Our analysis reveals that the benefits persist under realistic values of noise—with up to 40% measurement noise in underreported case counts and 10% in seroprevalence measurements—making the approach practically viable. Most importantly, we quantify the economic case for implementation, showing that population-level serosurvey costs are orders of magnitude smaller than the healthcare expenditures prevented through improved outbreak modeling.

Overall, our findings expose a systematic vulnerability in pandemic preparedness while providing an actionable remedy. As the world prepares for future pandemic threats, integrating serological surveillance into routine epidemiological modeling represents both a mathematical necessity and a strategic investment in more reliable outbreak response. More broadly, our uncertainty quantification framework provides a systematic approach for identifying when surveillance data are adequate for their intended modeling purpose—a critical capability for evidence-based public health decision making.

# 2 Mathematical Framework

## 2.1 The Parameter Identifiability Problem in Epidemiological Models

We demonstrate our computational approach using the canonical Susceptible-Infectious-Recovered (SIR) model [33], which serves as a prototype for a broad class of compartmental models facing similar identifiability challenges. The model categorizes the population into three compartments and evolves according to:

$$\begin{cases} \dfrac{dS}{dt} = -\beta \dfrac{SI}{N}, \\ \dfrac{dI}{dt} = \beta \dfrac{SI}{N} - \gamma I, \\ \dfrac{dR}{dt} = \gamma I, \end{cases} \tag{2.1}$$

where $\beta$ is the transmission rate, $\gamma$ is the recovery rate, and $N = S + I + R$ is the total population. While simple, this framework already showcases identifiability challenges that are present in more complex epidemiological models when the reality of undetected cases is considered.

The computational challenge emerges when surveillance systems observe only a fraction of true cases. If $I_{\text{new}}(t) = \beta S(t)I(t)/N$ represents true incidence, then detected incidence becomes:

$$I_{\text{new}}^{\text{obs}}(t) = \phi\beta \frac{S(t)I(t)}{N}, \tag{2.2}$$

where $\phi \in (0, 1]$ is the case ascertainment ratio. This creates a fundamental inverse problem: multiple parameter combinations $(\beta, \phi)$ can produce identical observed data, rendering standard parameter estimation approaches unreliable.

## 2.2 Structural Identifiability Theory and Computational Analysis

Structural identifiability analysis provides the mathematical foundation for understanding when parameters can be uniquely determined from observed data. A parameter is *uniquely (or globally) structurally identifiable* if, given perfect observations (defined as noise-free and continuous for all time), there exists a unique parameter value that can produce such observations. A model is structurally identifiable if all parameters are individually globally identifiable [34–37].

Our analysis highlights a critical mathematical result [38]: while the standard SIR model (2.1) with observations of either active cases ($I$) or true incidence ($\beta SI/N$) maintains global structural identifiability for both $\beta$ and $\gamma$, the extended system (2.1)—(2.2) becomes structurally unidentifiable under observations of detected incidence ($\phi\beta SI/N$). Specifically, $\gamma$ remains uniquely identifiable, but $\beta$ and $\phi$ cannot be separately determined—only their ratio $\beta/\phi$ is identifiable (see **Suppl. Materials S1**).

This identifiability crisis represents a fundamental computational limitation that affects a broad class of inverse problems in epidemiology and beyond, where scaling parameters confound the underlying system dynamics [39].

Despite the $\beta$ and $\phi$ confounding, the effective reproduction number $\mathcal{R}_e(t) = \frac{\beta}{\gamma}\frac{S(t)}{N}$ remains structurally identifiable from detected incidence (see **Suppl. Materials S1.3**).

## 2.3 Seroprevalence Integration as a Computational Solution

Seroprevalence data provide the mathematical constraint needed to resolve the identifiability crisis. Under ideal conditions (perfect sensitivity/specificity, representative sampling, persistent antibody detection), seroprevalence directly measures the fraction of the population that has moved from susceptible to infected or recovered status, equivalent to knowing $S(t)$ at survey time.

The computational power of this approach stems from complementary information: detected incidence captures temporal dynamics but suffers from unknown scaling (due to unknown $\phi$), while seroprevalence provides absolute scale with limited temporal resolution. Together, they break the mathematical degeneracy that renders $\beta$ and $\phi$ individually unidentifiable.

## 2.4 Computational Framework for Uncertainty Quantification

Our computational approach addresses parameter uncertainty through two complementary methods:

**Multi-start optimization** for noise-free scenarios identifies the manifold of parameter combinations that fit observed data perfectly, revealing the extent of identifiability issues.

**Bayesian inference with MCMC** quantifies uncertainty in realistic noisy conditions, providing full posterior distributions that capture both data limitations and prior knowledge.

The quantities of interest (QoIs) that drive epidemiological decision-making include: (i) state variables $S(t)$, $I(t)$, $R(t)$, true incidence $I_{\text{new}}(t)$, and detected incidence $I_{\text{new}}^{\text{obs}}(t)$ over time; (ii) case ascertainment ratio $\phi$; (iii) basic reproduction number $\mathcal{R}_0 = \beta/\gamma$; (iv) effective reproduction number $\mathcal{R}_e(t) = \beta/\gamma\, S(t)/N$; (v) peak infections $I_{\text{peak}} = \max_t I(t)$ and timing $t_{\text{peak}} = \arg\max_t I(t)$; (vi) final epidemic size $C_\infty = N - S_\infty$.

# 3 Computational Results

## 3.1 Systematic Bias from Ignoring Case Underascertainment

We first establish the mathematical foundation underlying parameter estimation failures in epidemic models. When compartmental models are fitted to detected case data without accounting for incomplete surveillance—treating observed cases as complete truth—the resulting parameter estimates exhibit systematic bias that worsens as detection rates decline. This represents a model-data mismatch.

**Figure 1** demonstrates this phenomenon computationally: fitting the SIR model to rescaled active case data (simulating various detection rates) produces visually acceptable fits while generating parameter errors exceeding 1000%. When only 1 in 50 cases is detected, the relative fitting error remains below 2%, yet relative errors in $\gamma$, $\beta$, and $\mathcal{R}_0$ reach approximately 1000%, 500%, and 50%, respectively. This mathematical result extends to detected incidence data (**Supplementary Figure S1**), confirming that good fits can mask catastrophically wrong parameter estimates.

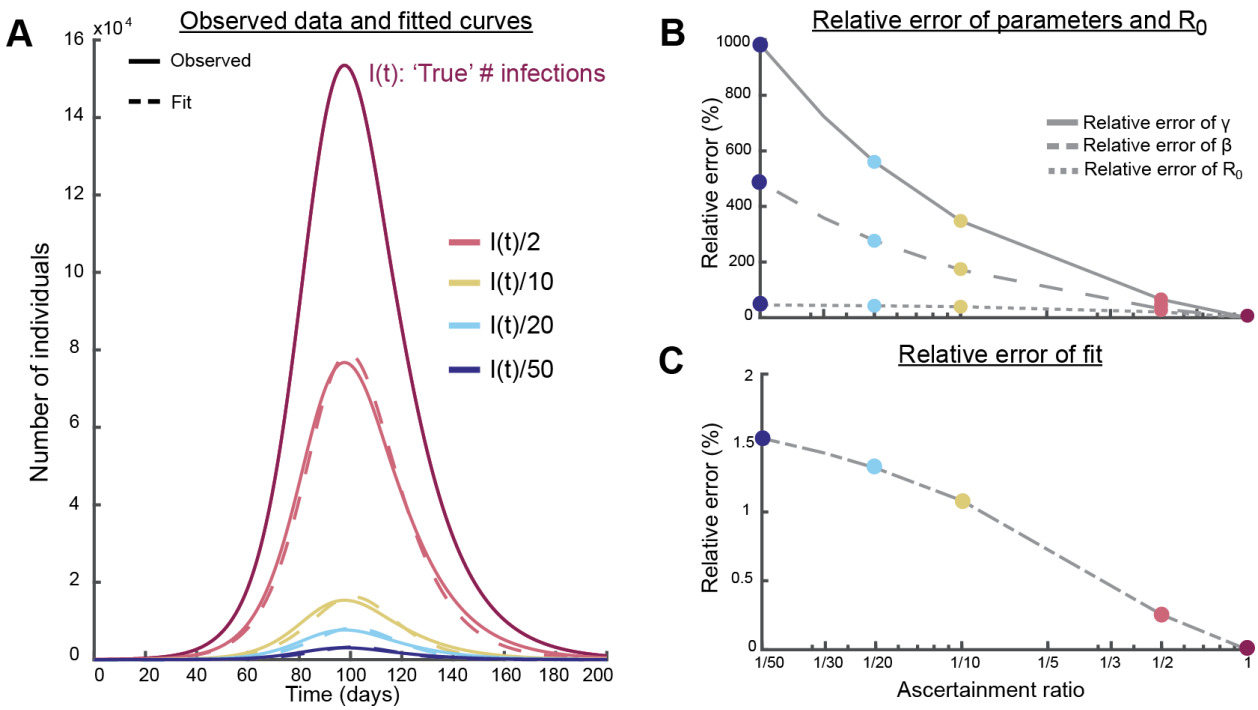


**Figure 1: Computational demonstration that ignoring case underascertainment yields severe parameter bias despite good fits. (A)** True infectious individuals $I(t)$ (magenta) with rescaled curves (solid lines) and fitted estimates (dashed lines) for ascertainment ratios 1, 1/2, 1/20, and 1/50. **(B)** Relative parameter errors in $\gamma$, $\beta$, and $\mathcal{R}_0$ as functions of ascertainment ratio. **(C)** Relative fitting errors for rescaled curves. Parameters: $\beta = 0.2$ day$^{-1}$, $\gamma = 0.1$ day$^{-1}$, $(S(0), I(0), R(0)) = (N - 50, 50, 0)$, $N = 10^6$.

## 3.2 Computational Implementation Framework

Our computational approach employs two complementary inference strategies tailored to different data scenarios and uncertainty quantification needs.

**Multi-start optimization** addresses the challenge of parameter identifiability in noise-free scenarios. We minimize the sum of squared errors between observations $\mathbf{Y} = [y_1(t_1), \ldots, y_n(t_n)]^T$ and model predictions $\mathcal{H}(t_i; \boldsymbol{\theta})$: $\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} [y_i(t_i) - \mathcal{H}(t_i; \boldsymbol{\theta})]^2$. To comprehensively explore the parameter landscape, we perform over 1000 optimization runs with randomized initial conditions, accepting solutions that achieve visually perfect fits (SSE $< 1000$). This approach reveals the full manifold of parameter combinations consistent with observed data, providing computational proof of identifiability issues.

**Bayesian MCMC inference** quantifies uncertainty under realistic noisy conditions. We implement the posterior distribution $p(\boldsymbol{\theta}|\mathbf{Y}) \propto p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ using the No-U-Turn Sampler (NUTS), an adaptive Hamiltonian Monte Carlo variant that efficiently explores high-dimensional posteriors. The likelihood incorporates multiplicative Gaussian noise:

$$I_{\text{new}}^{\text{noisy-obs}} = I_{\text{new}}^{\text{obs}} \cdot (1 + \epsilon_I), \quad \epsilon_I \sim \mathcal{N}(0, \sigma_I^2)$$
$$S^{\text{noisy}} = S \cdot (1 + \epsilon_S), \quad \epsilon_S \sim \mathcal{N}(0, \sigma_S^2).$$

This computational framework enables systematic exploration of how data quality and quantity affect parameter identifiability, providing rigorous uncertainty quantification for epidemiological decision-making under realistic surveillance conditions.

### 3.3 Computational Characterization of Structural Unidentifiability

When models explicitly account for incomplete detection by inferring the ascertainment ratio $\phi$ alongside transmission parameters, they face the structural identifiability problem proven in Section 2. Our computational analysis reveals the practical consequences of this mathematical limitation.

**Figure 2B** shows that despite visually perfect fits to detected incidence data, the inferred trajectories of true incidence and compartment dynamics exhibit enormous variability. This occurs because the optimization algorithm finds multiple parameter combinations $(\beta, \phi)$ that produce identical detected incidence curves while implying vastly different epidemic dynamics. The model cannot distinguish among a continuous spectrum of scenarios ranging from high transmission rates with low detection ratios to low transmission rates with high detection ratios—all scenarios fit the data equally well.

This computational result confirms our theoretical identifiability analysis: $\beta$ and $\phi$ are mathematically confounded, with only their ratio $\beta/\phi$ being uniquely determinable (**Supplementary Figure S2** and **Figure S3**). However, despite these underlying issues, fitted $\mathcal{R}_e(t)$ trajectories converge uniquely to true dynamics (**Supplementary Figure S4**), demonstrating its structural identifiability.

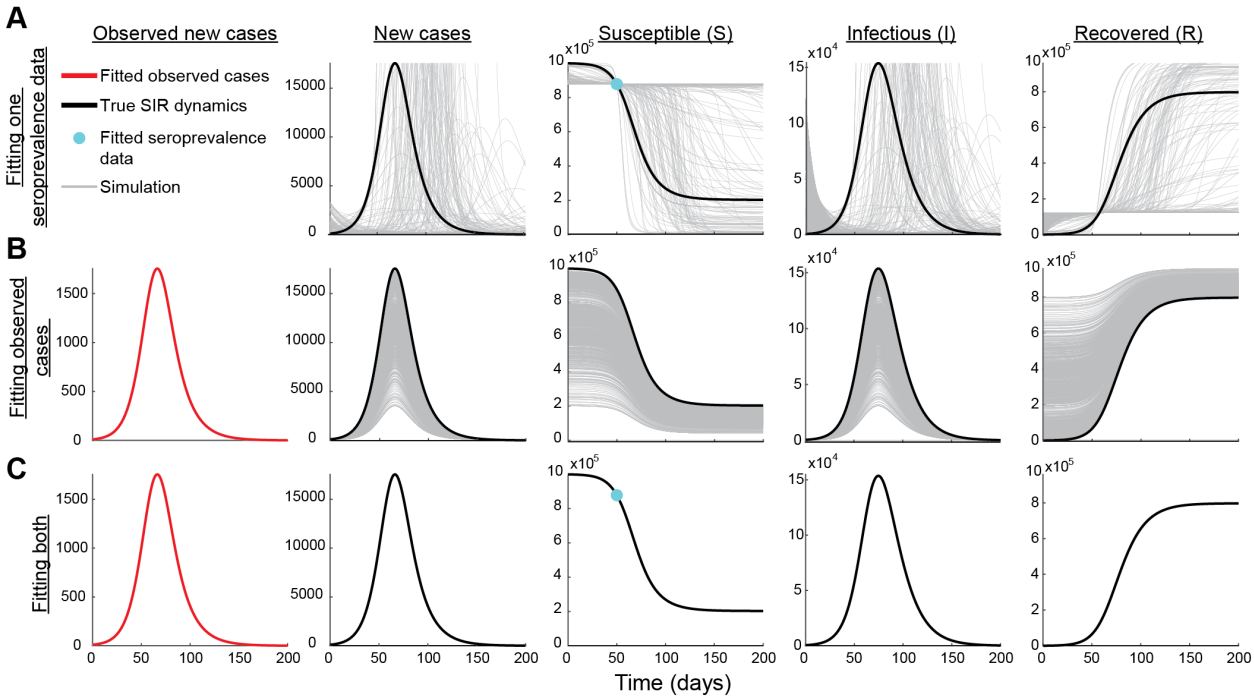


**Figure 2: Computational analysis of identifiability: seroprevalence data resolves parameter degeneracy.** Multi-start optimization results with parameters $\beta = 0.2$ day$^{-1}$, $\gamma = 0.1$ day$^{-1}$, $(S(0), I(0), R(0)) = (N - I(0), 500, 0)$, $N = 10^6$, $\phi = 0.1$. **(A)** Single seroprevalence point (cyan dot) provides insufficient constraint, yielding wide variability in solution trajectories (gray lines) versus true dynamics (black curves). **(B)** Detected incidence alone (red curve) produces multiple equivalent solutions despite perfect fits. **(C)** Combined seroprevalence and detected incidence constrains solutions near true dynamics. We perform over 1,000 multi-start optimization runs with randomized initial conditions sampled within bounds $\beta, \gamma, \phi \in (0, 1)$ and $S(0), I(0), R(0) \in [0, N]$ subject to $S(0) + I(0) + R(0) = N$.

## 3.4 Algorithmic Solution Through Seroprevalence Integration

The mathematical resolution to identifiability comes through integrating seroprevalence measurements. **Figure 2C** demonstrates that even a single seroprevalence data point, alongside detected incidence, dramatically constrains parameter estimates, breaking the high correlation between $\beta$ and $\phi$ (**Supplementary Figure S5**).

Our computational analysis reveals optimal data collection strategies: single seroprevalence measurements alongside detected incidence significantly improve inference, while three strategically timed seroprevalence measurements alone (early outbreak, epidemic peak, late stage) can reconstruct true dynamics with minimal uncertainty (**Supplementary Figure S6**).

To address concerns about the generalizability of our findings and ensure that the observed patterns are not artifacts of specific parameter choices, we conducted sensitivity analyses examining both parameter variations and measurement considerations. The qualitative features described in this section remain consistent across these variations. We show that the result of Figure 2C holds even when the ideal seroprevalence measurement is taken at different timepoints during an outbreak: incorporating even a single seroprevalence measurement dramatically reduces uncertainty in epidemiological quantities of interest regardless of when during the outbreak it is collected (**Supplementary Figure S7**). We assessed the impact of seroprevalence estimation errors and found that underestimating true seroprevalence by 10—30% leads to incorrect mapping from detected to true incidence, causing inferred quantities to deviate from their true values with increasing error in seroprevalence (**Supplementary Figure S8**). Varying transmission rates (corresponding to $\mathcal{R}_0$: 2–5) and ascertainment ratio independently show that uncertainties in QoIs persist when only fitting to detected incidence (see **Supplementary Figure S9 and Figure S10**, respectively). Unlike Figure 2C, where trajectories generated from fitted parameters are systematically biased to one side of the true dynamics, varying the initial number of recovered individuals produces trajectories that fall on both sides of the true dynamics (**Supplementary Figure S11**), demonstrating that the observed bias in Figure 2C is an artifact of the ground truth data being generated under the assumption of no prior immunity in the population (i.e., $R(0) = 0$).

## 3.5 Bayesian Computational Framework for Noisy Data

Real epidemiological data contain measurement noise, requiring sophisticated uncertainty quantification methods. We implement Bayesian inference with MCMC to fully characterize parameter uncertainty under realistic conditions.

**Figure 3** demonstrates that the identifiability solution remains robust under noise. Even with 40% noise in detected incidence and 10% noise in seroprevalence data, simultaneous fitting dramatically improves parameter estimation compared to using detected incidence alone.

**Figure 4** provides comprehensive quantification of uncertainty reduction across all epidemiological quantities of interest. The combined approach (orange violins) yields posterior distributions tightly concentrated around true values with dramatically reduced uncertainty compared to detected incidence alone (green violins), even at high noise levels. The distributions of the priors of the initial conditions and parameters for both noise-levels (10% and 40%) considered in this section, as well as their corresponding trace plots, are given in **Supplementary Section S11**.

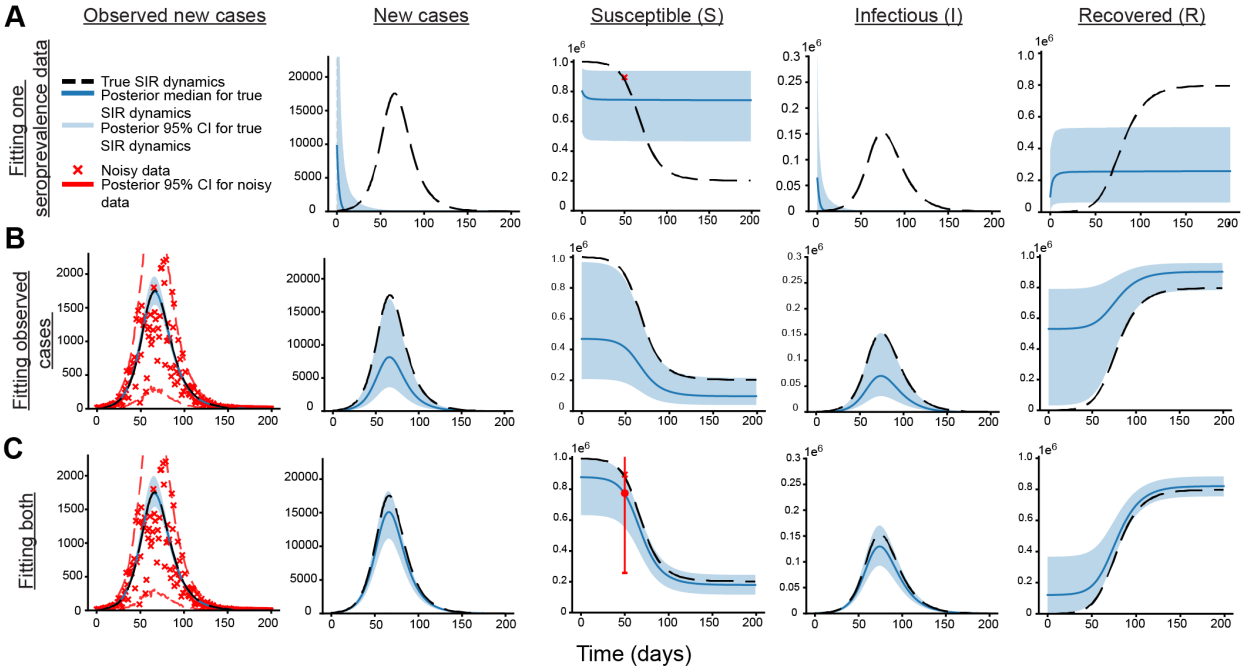


**Figure 3: Bayesian computational analysis demonstrates robustness under realistic noise conditions.** MCMC inference with $\beta = 0.2$ day$^{-1}$, $\gamma = 0.1$ day$^{-1}$, $(S(0), I(0), R(0)) = (N - I(0) - R(0), 500, 0.1N)$, $N = 10^6$, $\phi = 0.1$. **(A)** Single noisy seroprevalence measurement yields wide 95% credible intervals (blue regions) deviating from true dynamics (dashed black). **(B)** Detected incidence with 40% noise (red crosses) produces wide uncertainty bands. **(C)** Combined noisy data substantially constrains posterior distributions.

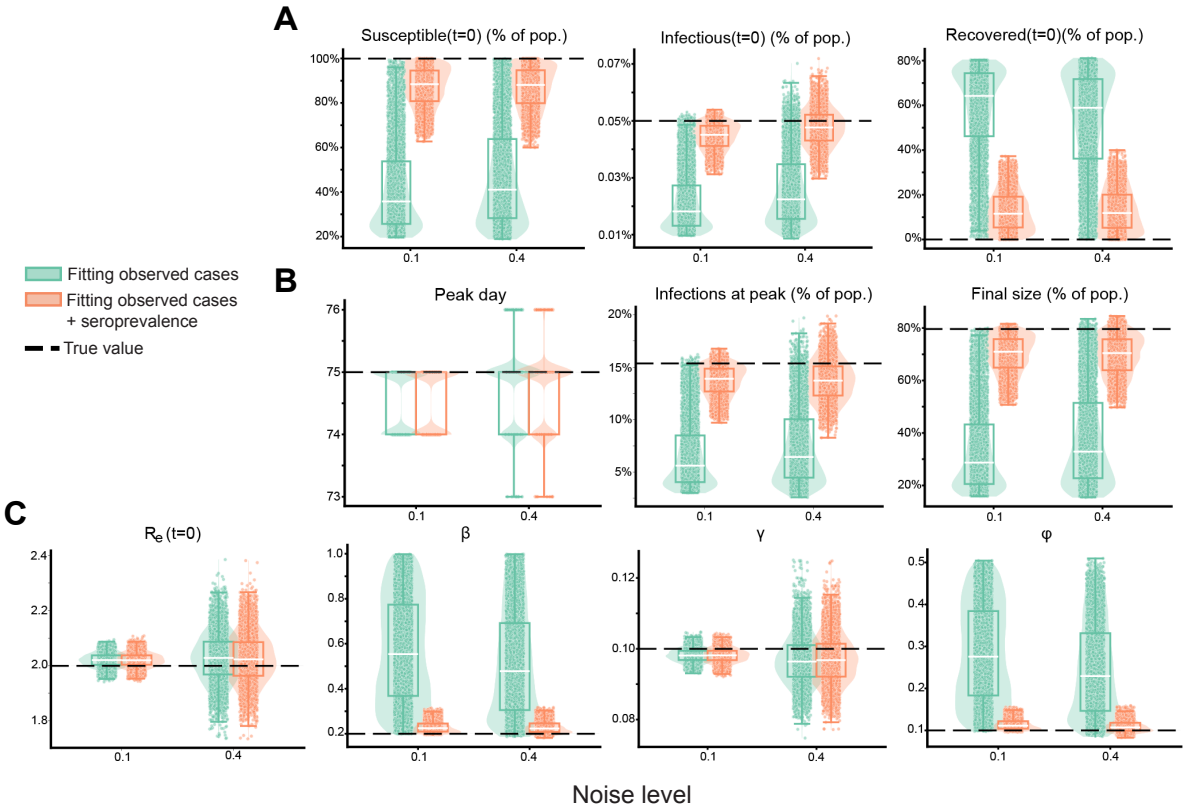


**Figure 4: Comprehensive uncertainty quantification demonstrates systematic improvement across all epidemiological quantities.** Posterior distributions comparing detected incidence alone (green violins: 10% noise; light green: included seroprevalence cases) versus combined data (orange violins: 40% incidence noise, 10% seroprevalence noise). Black dashed lines indicate true values. Seroprevalence integration consistently reduces uncertainty and improves accuracy across all QoIs. Parameters as in **Figure 3**.

The computational framework reveals that recovery rate estimates ($\gamma$) remain nearly identical regardless of data combination, consistent with our theoretical identifiability analysis. This parameter's identifiability from temporal patterns makes it insensitive to seroprevalence addition. Similarly, the effective reproduction number ($\mathcal{R}_e(t)$) maintains minimal uncertainty even from detected incidence alone (**Supplementary Figure S23**), consistent with its structural identifiability. In contrast, transmission-related parameters $\beta$ and $\phi$ as well as initial conditions of the state variables show show dramatic improvement with seroprevalence integration.

# 4 Discussion

This study demonstrates that the integration of population-level seroprevalence data provides a mathematically rigorous and practically robust solution to the critical problem of parameter uncertainty in epidemiological models caused by case underascertainment. Our computational framework reveals two fundamental failures of models that rely solely on being fitted to detected case data. First, when models ignore underdetection, they produce catastrophically biased estimates of transmission rates and epidemic size, even when achieving deceptively reasonable fits to the observed data. Second, when models explicitly account for underdetection by including an ascertainment ratio as a parameter, they suffer from a structural unidentifiability crisis, where a wide spectrum of epidemiologically distinct scenarios can explain the data equally well. Our central finding is that incorporating even a single, well-timed seroprevalence measurement resolves both of these issues by providing an independent constraint on the cumulative scale of the epidemic, thereby

breaking the mathematical degeneracy between the transmission rate and the ascertainment ratio.

An important nuance in our findings concerns the effective reproduction number, a critical metric for real-time outbreak assessment. Unlike the transmission rate and detection ratio, the effective reproduction number is identifiable from detected incidence alone. This has significant implications for the adequacy-for-purpose view of modeling, which evaluates models based on their fitness for specific intended uses rather than general accuracy [40–42]. A model fitted to detected cases may be adequate-for-purpose when the goal is determining whether an outbreak is growing or declining—the primary question guiding intervention timing. The effective reproduction number's insensitivity to case ascertainment means that detected incidence data, despite being incomplete, can reliably inform this specific decision. However, the same model-data pairing becomes inadequate-for-purpose when inferential goals shift to estimating true infection rates, cumulative attack rates, or final epidemic size—quantities that directly depend on the unidentifiable individual parameters. This underscores how the adequacy of a data source must always be judged relative to the specific inferential task at hand.

The implications of these findings for public health and pandemic preparedness are substantial. Accurate estimation of infection rates, and epidemic size is critical because these quantities not only guide intervention strategies, vaccination targets, and resource allocation, but also shape our understanding of how diseases spread, evolve, and compare across pathogens and strains. We have computationally confirmed that relying on models fitted exclusively to reported case counts is an inherently flawed strategy that can lead to severely misguided policy decisions. By exposing this systematic vulnerability, our work underscores the necessity of moving beyond a single source of surveillance data. While several modeling studies include other data sources such as hospitalizations, deaths, or wastewater-based viral counts [25, 43, 44], these are also only proxies with uncertain scaling to true infections. In contrast, serology uniquely overcomes this limitation by directly measuring cumulative exposure. The integration of serological data transforms outbreak modeling from a potentially misleading exercise into a reliable tool for quantifying critical metrics such as the true infection rate, the effective reproduction number, and the final size of an epidemic. This enhanced accuracy is essential for effective resource allocation, intervention planning, and building resilient public health systems capable of responding to future pandemic threats with confidence.

The challenge of parameter unidentifiability, while highlighted here in an epidemiological context, represents a broader class of inverse problems prevalent across computational science [36, 37, 45–47]. In any system where the true underlying dynamics are observed through a process that involves unknown scaling factors or biases, standard inference methods can face substantial challenges. Our proposed solution—constraining the system with a complementary data source that provides information on absolute scale rather than temporal dynamics—offers a generalizable strategy for resolving such parameter degeneracies [48]. The computational framework, which combines multi-start optimization to map the landscape of possible solutions under a noise-free setting and Bayesian MCMC to quantify uncertainty under realistic noise, provides a transferable methodology for diagnosing and solving similar identifiability crises.

Our work highlights how synthetic experiments provide a powerful framework for assessing uncertainties in parameters, state variables, and derived quantities, offering a comprehensive picture of which aspects of system dynamics can be reliably inferred. When analytical identifiability becomes intractable due to model complexity, fitting to noise-free data remains computationally feasible and can reveal practical limitations of the model-data pairing in estimation and reconstruction of quantities of interest. Moreover, by systematically introducing observational noise into synthetic data, we can bridge the gap between ideal noiseless conditions and real-world complexity, isolating the impact of measurement error from structural model limitations. Real-world applications also introduce additional challenges through model-data mismatch, where simplifying model assumptions may fail to capture true data-generating processes [49]. Understanding the model's ability to infer dynamical system behavior under synthetic noise-free and noisy conditions thus provides an essential foundation for interpreting discrepancies with real data and guiding model refinement, and should be undertaken as a first step for any new model as a matter of best practice, prior to fitting with real-world data.

A key strength of our findings is the demonstrated robustness of the solution under reasonable values of errors present in current surveillance systems. Our Bayesian inference framework showed that the dramatic

reduction in parameter uncertainty holds even in the presence of substantial measurement noise, including up to 40% noise in detected incidence data and 10% in seroprevalence estimates. Furthermore, our sensitivity analyses showed that the benefits of seroprevalence integration are consistent across different transmission scenarios and are not highly dependent on the precise timing of the serological survey, strengthening the case for its broad implementation in routine surveillance. Both of these findings demonstrate that our approach is not merely a theoretical curiosity but a practically viable strategy for real-world application.

While this framework provides a clear path forward, we acknowledge several limitations that suggest avenues for future research. Our analysis relies on a simplified SIR model with a constant case ascertainment ratio, whereas real-world detection efforts and testing capacity fluctuate over time. Future work should extend this framework to models that incorporate such complexities. We also assume idealized seroprevalence measurements; further investigation is needed to quantify the impact of real-world complexities such as sampling biases, antibody waning [50–52], and imperfect test sensitivity and specificity [53–56]. Moreoever, while we demonstrate the powerful constraining effect of idealized serosurveys, the optimal design of real seroprevalence surveys, including determining the ideal timing, frequency, and sample size to maximize information gain, remains an important open question for future computational and epidemiological research. Finally, a limitation of our approach is that we have established that we can successfully reconstruct the dynamics (and infer the associated epidemiological parameters) of a fully observed disease outbreak. Future work should explore the extent to which population-level serological data can improve parameter inference in real-time, with access to incomplete data, such as the early stages of an outbreak and before the peak has been observed.

Our work provides not only a mathematical and computational solution to a long-standing problem in outbreak modeling but also a compelling economic and practical argument for implementing serosurveys in disease surveillance. With as few as 1,000 samples, reliable seroprevalence estimates can be obtained for populations of 1 million, at a total cost of about $69,000 ($69 per test [57]), provided true prevalence exceeds 2% [32, 58]. By enabling more accurate characterization of outbreak dynamics, this targeted data collection can prevent costly interventions and reduce the burden on healthcare systems, yielding a return on investment of several orders of magnitude while also informing policies that help safeguard lives. Therefore, integrating serological surveillance into routine epidemiological modeling should be viewed as both a mathematical necessity and a strategic investment in robust, evidence-based pandemic preparedness.

# 5 Methods

## 5.1 Synthetic Data Generation Framework

We generate controlled synthetic datasets using the SIR model (2.1) to evaluate our computational framework under known ground truth conditions. This approach enables rigorous assessment of identifiability issues and uncertainty quantification methods without the confounding effects of model misspecification or unknown data-generating processes.

### 5.1.1 Noise-free synthetic data

Synthetic epidemic trajectories are generated by solving the SIR system (2.1) using MATLAB's `ode45` solver, an explicit Runge-Kutta (4,5) implementation based on the Dormand-Prince method.

Detected case data are generated by rescaling true trajectories according to the case ascertainment ratio $\phi$. For detected active cases, we compute $I^{\text{obs}}(t) = \phi \times I(t)$, with initial conditions adjusted accordingly: $I^{\text{obs}}(0) = \phi \times I(0)$ and $S(0) = N - (I^{\text{obs}}(0) + R(0))$, with $R(0) = 0$. For detected incidence experiments, the ascertainment ratio is incorporated through equation (2.2), with initial infectious population inferred during parameter estimation.

### 5.1.2 Noisy synthetic data

To assess robustness under realistic surveillance conditions, we add multiplicative Gaussian noise to both detected incidence and seroprevalence data:

$$
\begin{aligned}
I_{\text{new}}^{\text{noisy-obs}} &= I_{\text{new}}^{\text{obs}} \cdot (1+\epsilon_I) \quad \text{where } \epsilon_I \sim \mathcal{N}(0, \sigma_I^2) \\
S^{\text{noisy}} &= S \cdot (1+\epsilon_S) \quad \text{where } \epsilon_S \sim \mathcal{N}(0, \sigma_S^2).
\end{aligned} \tag{5.1}
$$

The noise parameters $\sigma_I$ and $\sigma_S$ represent multiplicative standard deviations. For example, $\sigma_I = 0.1$ corresponding to 10% noise where approximately 68% of observations fall within ±10% of true values. We refer to $\sigma_X$ with $X = \{I, S\}$ as $\sigma$% "percent noise" on $X$.

We set seroprevalence noise at 10% to reflect typical uncertainties in COVID-19 serological assays, where sensitivity ranges 85-95% and specificity ranges 95-99% [53,54]. For detected incidence, we explore low noise (10%) and moderate noise (40%) scenarios, representing variability from test performance, data aggregation artifacts, and surveillance system inconsistencies—distinct from the systematic underascertainment modeled through $\phi$.

## 5.2 Computational Inference Methods

### 5.2.1 Multi-start optimization for noise-free analysis

Our multi-start optimization approach systematically explores parameter spaces to identify identifiability issues. We minimize the sum of squared errors between observations $\mathbf{Y} = [y_1(t_1), y_2(t_2), \ldots, y_n(t_n)]^T$ and model predictions $\mathcal{H}(t_i; \boldsymbol{\theta})$:

$$
\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} \left[ y_i(t_i) - \mathcal{H}(t_i; \boldsymbol{\theta}) \right]^2
$$

Implementation uses MATLAB's `fminsearch`, `lsqcurvefit`, and `fmincon` optimizers. To comprehensively map the solution space, we perform over 1,000 optimization runs with randomized initial parameter guesses. Solutions achieving SSE $< 1000$ (corresponding to visually perfect fits) are retained for analysis. This approach reveals the complete manifold of parameter combinations consistent with observed data, providing computational evidence for theoretical identifiability results.

We highlight that in the case of fitting detected active case (Figure 1) or detected incidence data (Figure S1) to the SIR model without accounting for the reality that not all cases are observed, we use a multi-start approach to find parameter sets that minimize SSE the most.

In the context of multiple optimization runs with various initial guesses, the resulting distribution of parameter estimates does not represent true statistical uncertainty. This distribution merely reflects the sensitivity to initial guesses and the optimization algorithm's convergence patterns rather than actual parameter uncertainty in the data. As a result, we do not display the histogram distribution of parameters using this method, nor do we use this method for the noise-present data fitting exercise (where a Bayesian method will be used).

### 5.2.2 Bayesian inference for uncertainty quantification

For realistic noisy conditions, we implement Bayesian inference to quantify parameter uncertainty. The posterior distribution $p(\boldsymbol{\theta}|\mathbf{Y}) \propto p(\mathbf{Y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ combines the likelihood defined by the noise model (5.1) with uniform priors on parameters and initial conditions.

Prior specifications include $\beta, \gamma, \phi \sim \mathcal{U}([0.01, 1])$ for model parameters and $S(0), I(0), R(0) \sim \mathcal{U}(0, N)$ subject to $S(0) + I(0) + R(0) = N$. Initial condition priors are implemented using Dirichlet distributions with concentration parameters $\alpha = [1, 1, 1]$ then rescaled by $N$, providing equivalent uniform constraints with improved computational efficiency.

Noise parameters $\sigma_I, \sigma_S \sim \mathcal{U}([0.001, 0.5])$ are jointly estimated, with full posterior distributions provided in Supplementary Materials.

#### 5.2.3 MCMC implementation

Sampling uses the No-U-Turn Sampler (NUTS), an adaptive Hamiltonian Monte Carlo variant implemented in NumPyro [59]. Our sampling protocol includes:

- Preliminary warmup: 2,000 iterations (step size 1e-3, target acceptance 0.4)
- Burn-in: 5,000 iterations (discarded)
- Final warmup: 2,000 iterations (step size 1.0, target acceptance 0.9)
- Posterior sampling: 5,000 iterations for analysis

This custom implementation ensures robust convergence in scenarios with complex identifiability structures. Simpler default schemes suffice when both seroprevalence and case data are available, but the enhanced protocol handles challenging unidentifiable cases. Convergence assessment uses visual trace plot inspection, with full diagnostics in Supplementary Materials.

### 5.3 Structural Identifiability Analysis

Structural identifiability analysis uses two complementary symbolic computation tools: DAISY [60] and Julia's `StructuralIdentifiability.jl` [61]. We use these tools to determine parameter identifiability as theoretical properties independent of data quality or estimation algorithms, while assuming unknown initial conditions during structural identifiability analysis.

### 5.4 Parameter Ranges and Computational Settings

All experiments use population size $N = 10^6$ individuals. Baseline parameters are $\beta = 0.2$ day$^{-1}$, $\gamma = 0.1$ day$^{-1}$ (corresponding to $\mathcal{R}_0 = 2$), with case ascertainment ratio of $\phi = 0.1$. Sensitivity analyses explore transmission rates $\beta \in [0.2, 0.5]$ day$^{-1}$ ($\mathcal{R}_0 \in [2, 5]$) with case ascertainment ratios ranging from $\phi = 0.1$ to $\phi = 1$. Initial conditions and parameters vary by experiment, with specific values provided in figure captions. The computational framework is designed to handle diverse parameter regimes while maintaining theoretical rigor and practical applicability.

## 6 Acknowledgements

AN was supported by the CDC-CFA award for the Atlantic Coast Center for Infectious Disease Dynamics and Analytics (ACCIDDA): NU38FT000012-01-00 (UNC).

# References


[1] C. Reed, F. J. Angulo, D. L. Swerdlow, M. Lipsitch, M. I. Meltzer, D. Jernigan, and L. Finelli, "Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009," *Emerging infectious diseases*, vol. 15, no. 12, p. 2004, 2009.

[2] S. S. Shrestha, D. L. Swerdlow, R. H. Borse, V. S. Prabhu, L. Finelli, C. Y. Atkins, K. Owusu-Edusei, B. Bell, P. S. Mead, M. Biggerstaff, *et al.*, "Estimating the burden of 2009 pandemic influenza a (H1N1) in the United States (April 2009–April 2010)," *Clinical Infectious Diseases*, vol. 52, no. suppl_1, pp. S75–S82, 2011.

[3] K. Chong, H. Fong, and C. Zee, "Estimating the incidence reporting rates of new influenza pandemics at an early stage using travel data from the source country," *Epidemiology & Infection*, vol. 142, no. 5, pp. 955–963, 2014.

[4] S. Runge-Ranzinger, P. J. McCall, A. Kroeger, and O. Horstick, "Dengue disease surveillance: an updated systematic literature review," *Tropical Medicine & International Health*, vol. 19, no. 9, pp. 1116–1160, 2014.

[5] E. Colman, G. A. Puspitarani, J. Enright, and R. R. Kao, "Ascertainment rate of sars-cov-2 infections from healthcare and community testing in the uk," *Journal of Theoretical Biology*, vol. 558, p. 111333, 2023.

[6] T. W. Russell, N. Golding, J. Hellewell, S. Abbott, L. Wright, C. A. B. Pearson, K. van Zandvoort, C. I. Jarvis, H. Gibbs, Y. Liu, R. M. Eggo, W. J. Edmunds, A. J. Kucharski, A. K. Deol, C. J. Villabona-Arenas, T. Jombart, K. O'Reilly, J. D. Munday, S. R. Meakin, R. Lowe, A. Gimma, A. Endo, E. S. Nightingale, G. Medley, A. M. Foss, G. M. Knight, K. Prem, S. Hué, C. Diamond, J. W. Rudge, K. E. Atkins, M. Auzenbergs, S. Flasche, R. M. G. J. Houben, B. J. Quilty, P. Klepac, M. Quaife, S. Funk, Q. J. Leclerc, J. C. Emery, M. Jit, D. Simons, N. I. Bosse, S. R. Procter, F. Y. Sun, S. Clifford, K. Sherratt, A. Rosello, N. G. Davies, O. Brady, D. C. Tully, G. R. Gore-Langton, and C. C.-w. group, "Reconstructing the early global dynamics of under-ascertained covid-19 cases and infections," *BMC Medicine*, vol. 18, no. 1, p. 332, 2020.

[7] N. E. Kogan, S. Gantt, D. Swerdlow, C. Viboud, M. Semakula, M. Lipsitch, and M. Santillana, "Leveraging serosurveillance and postmortem surveillance to quantify the impact of coronavirus disease 2019 in Africa," *Clinical Infectious Diseases*, vol. 76, no. 3, pp. 424–432, 2023.

[8] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman, "Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)," *Science*, vol. 368, no. 6490, pp. 489–493, 2020.

[9] B. Wang, Y. Hu, Z. Cen, J. Huang, T. He, and A. Xu, "A SIR model with incomplete data for the analysis of influenza a spread in ningbo," *Discrete Dynamics in Nature and Society*, vol. 2024, no. 1, p. 7694770, 2024.

[10] V. N. Leonenko and S. V. Ivanov, "Fitting the SEIR model of seasonal influenza outbreak to the incidence data for russian cities," *Russian Journal of Numerical Analysis and Mathematical Modelling*, vol. 31, no. 5, pp. 267–279, 2016.

[11] A. Pandey, A. Mubayi, and J. Medlock, "Comparing vector–host and SIR models for dengue transmission," *Mathematical Biosciences*, vol. 246, no. 2, pp. 252–259, 2013.

[12] L. Kong, J. Wang, Z. Li, S. Lai, Q. Liu, H. Wu, and W. Yang, "Modeling the heterogeneity of dengue transmission in a city," *International Journal of Environmental Research and Public Health*, vol. 15, no. 6, p. 1128, 2018.

[13] S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, *et al.*, "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe," *Nature*, vol. 584, no. 7820, pp. 257–261, 2020.

[14] S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, and A. B. Gumel, "To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19pandemic," *Infectious Disease Modelling*, vol. 5, pp. 293–308, 2020.

[15] F. S. Lu, A. T. Nguyen, N. B. Link, M. Molina, J. T. Davis, M. Chinazzi, X. Xiong, A. Vespignani, M. Lipsitch, and M. Santillana, "Estimating the cumulative incidence of COVID-19 in the United States using influenza surveillance, virologic testing, and mortality data: Four complementary approaches," *PLOS Computational Biology*, vol. 17, no. 6, p. e1008994, 2021.

[16] S. G. Krantz and A. S. R. S. Rao, "Level of underreporting including underdiagnosis before the first peak of COVID-19 in various countries: Preliminary retrospective results based on wavelets and deterministic modeling," *Infection Control & Hospital Epidemiology*, vol. 41, pp. 857–859, July 2020.

[17] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li, "Why is it difficult to accurately predict the COVID-19 epidemic?," *Infectious Disease Modelling*, vol. 5, pp. 271–281, 2020.

[18] M. T. Meehan, D. P. Rojas, A. I. Adekunle, O. A. Adegboye, J. M. Caldwell, E. Turek, B. M. Williams, B. J. Marais, J. M. Trauer, and E. S. McBryde, "Modelling insights into the COVID-19 pandemic," *Paediatric Respiratory Reviews*, vol. 35, pp. 64–69, Sept. 2020.

[19] Y. Xiang, Y. Jia, L. Chen, L. Guo, B. Shu, and E. Long, "COVID-19 epidemic prediction and the impact of public health interventions: A review of COVID-19 epidemic models," *Infectious Disease Modelling*, vol. 6, pp. 324–342, Jan. 2021.

[20] C. J. E. Metcalf, C. Viboud, D. J. Spiro, and B. T. Grenfell, "Using serology with models to clarify the trajectory of the SARS-CoV-2 emerging outbreak," *Trends in Immunology*, vol. 41, no. 10, pp. 849–851, 2020.

[21] V. K. Lopez, E. Y. Cramer, R. Pagano, J. M. Drake, E. B. O'Dea, M. Adee, T. Ayer, J. Chhatwal, O. O. Dalgic, M. A. Ladd, B. P. Linas, P. P. Mueller, J. Xiao, J. Bracher, A. J. C. Rivadeneira, A. Gerding, T. Gneiting, Y. Huang, D. Jayawardena, A. H. Kanji, K. Le, A. Mühlemann, J. Niemi, E. L. Ray, A. Stark, Y. Wang, N. Wattanachit, M. W. Zorn, S. Pei, J. Shaman, T. K. Yamana, S. R. Tarasewicz, D. J. Wilson, S. Baccam, H. Gurung, S. Stage, B. Suchoski, L. Gao, Z. Gu, M. Kim, X. Li, G. Wang, L. Wang, Y. Wang, S. Yu, L. Gardner, S. Jindal, M. Marshall, K. Nixon, J. Dent, A. L. Hill, J. Kaminsky, E. C. Lee, J. C. Lemaitre, J. Lessler, C. P. Smith, S. Truelove, M. Kinsey, L. C. Mullany, K. Rainwater-Lovett, L. Shin, K. Tallaksen, S. Wilson, D. Karlen, L. Castro, G. Fairchild, I. Michaud, D. Osthus, J. Bian, W. Cao, Z. Gao, J. L. Ferres, C. Li, T.-Y. Liu, X. Xie, S. Zhang, S. Zheng, M. Chinazzi, J. T. Davis, K. Mu, A. P. y. Piontti, A. Vespignani, X. Xiong, R. Walraven, J. Chen, Q. Gu, L. Wang, P. Xu, W. Zhang, D. Zou, G. C. Gibson, D. Sheldon, A. Srivastava, A. Adiga, B. Hurt, G. Kaur, B. Lewis, M. Marathe, A. S. Peddireddy, P. Porebski, S. Venkatramanan, L. Wang, P. V. Prasad, J. W. Walker, A. E. Webber, R. B. Slayton, M. Biggerstaff, N. G. Reich, and M. A. Johansson, "Challenges of COVID-19 Case Forecasting in the US, 2020–2021," *PLOS Computational Biology*, vol. 20, p. e1011200, May 2024.

[22] V. Chin, N. I. Samia, R. Marchant, O. Rosen, J. P. Ioannidis, M. A. Tanner, and S. Cripps, "A case study in model failure? COVID-19 daily deaths and icu bed utilisation predictions in new york state: V. chin et al.," *European Journal of Epidemiology*, vol. 35, no. 8, pp. 733–742, 2020.

[23] L. Gallo, M. Frasca, V. Latora, and G. Russo, "Lack of practical identifiability may hamper reliable predictions in COVID-19 epidemic models," *Science Advances*, vol. 8, no. 3, p. eabg5234, 2022.

[24] B. Pant, S. Safdar, M. Santillana, and A. B. Gumel, "Mathematical assessment of the role of human behavior changes on SARS-CoV-2 transmission dynamics in the United States," *Bulletin of Mathematical Biology*, vol. 86, no. 8, p. 92, 2024.

[25] L. LeJeune, N. Ghaffarzadegan, L. M. Childs, and O. Saucedo, "Mathematical analysis of simple behavioral epidemic models," *Mathematical Biosciences*, vol. 375, p. 109250, 2024.

[26] M. P. Smith, "Role of ascertainment bias in determining case fatality rate of COVID-19," *Journal of epidemiology and global health*, vol. 11, no. 2, pp. 143–145, 2021.

[27] Y. Suhail, J. Afzal, and Kshitiz, "Incorporating and addressing testing bias within estimates of epidemic dynamics for SARS-CoV-2," *BMC medical research methodology*, vol. 21, no. 1, p. 11, 2021.

[28] K. V. Parag, C. A. Donnelly, and A. E. Zarebski, "Quantifying the information in noisy epidemic curves," *Nature Computational Science*, vol. 2, no. 9, pp. 584–594, 2022.

[29] R. Subramanian, Q. He, and M. Pascual, "Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity," *Proceedings of the National Academy of Sciences*, vol. 118, no. 9, p. e2019716118, 2021.

[30] C. J. E. Metcalf, J. Farrar, F. T. Cutts, N. E. Basta, A. L. Graham, J. Lessler, N. M. Ferguson, D. S. Burke, and B. T. Grenfell, "Use of serological surveys to generate key insights into the changing global landscape of infectious disease," *The Lancet*, vol. 388, no. 10045, pp. 728–730, 2016.

[31] A. K. Winter and S. T. Hegde, "The important role of serology for COVID-19 control," *The Lancet Infectious Diseases*, vol. 20, no. 7, pp. 758–759, 2020.

[32] D. B. Larremore, B. K. Fosdick, K. M. Bubar, S. Zhang, S. M. Kissler, C. J. E. Metcalf, C. O. Buckee, and Y. H. Grad, "Estimating sars-cov-2 seroprevalence and epidemiological parameters with uncertainty from serological surveys," *Elife*, vol. 10, p. e64206, 2021.

[33] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.

[34] M. C. Eisenberg and M. A. Hayashi, "Determining identifiable parameter combinations using subset profiling," *Mathematical Biosciences*, vol. 256, pp. 116–126, 2014.

[35] N. Tuncer and T. T. Le, "Structural and practical identifiability analysis of outbreak models," *Mathematical Biosciences*, vol. 299, pp. 1–18, 2018.

[36] R. Bellman and K. J. Åström, "On structural identifiability," *Mathematical Biosciences*, vol. 7, no. 3-4, pp. 329–339, 1970.

[37] C. Cobelli and J. J. Distefano 3rd, "Parameter and structural identifiability concepts and ambiguities: a critical review and analysis," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 239, no. 1, pp. R7–R24, 1980.

[38] F. Bergström, M. Favero, and T. Britton, "Identifiability in epidemic models with prior immunity and under-reporting," June 2025.

[39] G. Chowell, S. Dahal, Y. R. Liyanage, A. Tariq, and N. Tuncer, "Structural identifiability analysis of epidemic models based on differential equations: a tutorial-based primer," *Journal of Mathematical Biology*, vol. 87, no. 6, p. 79, 2023.

[40] W. S. Parker, "Model evaluation: An adequacy-for-purpose view," *Philosophy of Science*, vol. 87, no. 3, pp. 457–477, 2020.

[41] A. Bokulich and W. Parker, "Data models, representation and adequacy-for-purpose," *European Journal for Philosophy of Science*, vol. 11, no. 1, p. 31, 2021.

[42] A. Bokulich and W. S. Parker, "Understanding data uncertainty," *Studies in History and Philosophy of Science*, vol. 112, pp. 90–101, 2025.

[43] A. Oveson, M. Girvan, and A. B. Gumel, "Modeling the impact of hospitalization-induced behavioral changes on the spread of COVID-19 in new york city," *Infectious Disease Modelling*, 2025.

[44] B. Pant, S. Safdar, C. N. Ngonghala, and A. B. Gumel, "Mathematical assessment of wastewater-based epidemiology to predict SARS-CoV-2 cases and hospitalizations in Miami-Dade County," *Acta Biotheoretica*, vol. 73, no. 1, p. 2, 2025.

[45] E. Walter and L. Pronzato, "Identification of parametric models: from experimental data," *(No Title)*, 1997.

[46] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, 2009.

[47] A. F. Villaverde, A. Barreiro, and A. Papachristodoulou, "Structural identifiability of dynamic systems biology models," *PLoS computational biology*, vol. 12, no. 10, p. e1005153, 2016.

[48] L. Arriola and J. M. Hyman, "Sensitivity analysis for uncertainty quantification in mathematical models," in *Mathematical and statistical estimation approaches in epidemiology*, pp. 195–247, Springer, 2009.

[49] J. W. Glasser and Z. Feng, "Mechanistic models are hypotheses: A perspective," *Mathematical Biosciences*, vol. 383, p. 109419, 2025.

[50] A. J. Pollard and E. M. Bijker, "A guide to vaccinology: from basic principles to new developments," *Nature Reviews Immunology*, vol. 21, pp. 83–100, Feb. 2021.

[51] I. Wahl and H. Wardemann, "Sterilizing immunity: Understanding COVID-19," *Immunity*, vol. 55, pp. 2231–2235, Dec. 2022.

[52] E. K. Accorsi, X. Qiu, E. Rumpler, L. Kennedy-Shaffer, R. Kahn, K. Joshi, E. Goldstein, M. J. Stensrud, R. Niehus, M. Cevik, *et al.*, "How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19," *European journal of epidemiology*, vol. 36, no. 2, pp. 179–196, 2021.

[53] K. M. Taylor, K. M. Ricks, P. A. Kuehnert, A. A. Eick-Cost, M. R. Scheckelhoff, A. R. Wiesen, T. L. Clements, Z. Hu, S. E. Zak, S. P. Olschner, *et al.*, "Seroprevalence as an indicator of undercounting of covid-19 cases in a large well-described cohort," *AJPM focus*, vol. 2, no. 4, p. 100141, 2023.

[54] X. Ma, Z. Li, M. G. Whelan, D. Kim, C. Cao, M. Yanes-Lane, T. Yan, T. Jaenisch, M. Chu, D. A. Clifton, *et al.*, "Serology assays used in sars-cov-2 seroprevalence surveys worldwide: a systematic review and meta-analysis of assay features, testing algorithms, and performance," *Vaccines*, vol. 10, no. 12, p. 2000, 2022.

[55] D. McConnell, C. Hickey, N. Bargary, L. Trela-Larsen, C. Walsh, M. Barry, and R. Adams, "Understanding the challenges and uncertainties of seroprevalence studies for SARS-CoV-2," *International Journal of Environmental Research and Public Health*, vol. 18, no. 9, p. 4640, 2021.

[56] R. M. Boyce, B. E. Shook-Sa, and A. E. Aiello, "A tale of 2 studies: study design and our understanding of severe acute respiratory syndrome coronavirus 2 seroprevalence," *Clinical Infectious Diseases*, vol. 73, no. 9, pp. e3124–e3126, 2021.

[57] LabCorp, "Coronavirus disease (covid-19) individual faqs." `https://www.labcorp.com/coronavirus-disease-covid-19/individuals/faqs`, 2024. Accessed June 30, 2025.

[58] A. Nikiforuk, I. Sekirov, and A. Jassem, "Simple approximation of sample size for precise estimates of SARS-CoV-2 infection from point-seroprevalence studies," *Public Health*, vol. 212, pp. 7–9, 2022.

[59] NumPyro, *Markov Chain Monte Carlo (MCMC)*, 2024. Accessed June 17, 2025.

[60] G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò, "DAISY: A new software tool to test global identifiability of biological and physiological systems," *Computer Methods and Programs in Biomedicine*, vol. 88, no. 1, pp. 52–61, 2007.

[61] R. Dong, C. Goodbrake, H. Harrington, and P. G., "Differential elimination for dynamical models via projections with applications to structural identifiability," *SIAM Journal on Applied Algebra and Geometry*, vol. 7, no. 1, pp. 194–235, 2023.