

RESEARCH ARTICLE

Bayesian uncertainty quantification to identify population level vaccine hesitancy behaviours

David J. Warne^{1,2*}, Abhishek Varghese², Aidan Brewster¹, Alexander P. Browning^{3,4}, Mario M. Krell⁵, Christopher Drovandi^{1,2}, Wenbiao Hu⁶, Antonietta Mira^{7,8}, Kerrie Mengersen^{1,2}, Adrienne L. Jenner^{1,2}

1 School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia, **2** Centre for Data Science, Queensland University of Technology, Brisbane, Queensland, Australia, **3** Mathematical Institute, University of Oxford, Oxford, United Kingdom, **4** School of Mathematics and Statistics, University of Melbourne, Melbourne, Victoria, Australia, **5** Graphcore Inc., Palo Alto, California, United States of America, **6** School of Public Health and Social Work, Queensland University of Technology, Brisbane, Queensland, Australia, **7** Faculty of Economics, Università della Svizzera Italiana, Lugano, Switzerland, **8** Department of Science and High Technology, Insubria University, Varese, Italy

* david.warne@qut.edu.au



OPEN ACCESS

Citation: Warne DJ, Varghese A, Brewster A, Browning AP, Krell MM, Drovandi C, et al. (2026) Bayesian uncertainty quantification to identify population level vaccine hesitancy behaviours. PLoS One 21(5): e0349499. <https://doi.org/10.1371/journal.pone.0349499>

Editor: Kristan Alexander Schneider, The University of New Mexico, UNITED STATES OF AMERICA

Received: June 26, 2025

Accepted: April 29, 2026

Published: May 26, 2026

Copyright: © 2026 Warne 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 relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: DJW, ALJ and AV were supported by a Queensland University of Technology Centre for Data Science First Byte Grant. DJW is

Abstract

When effective vaccines are available, vaccination programs are typically one of the best defences against the spread of an infectious disease. Such vaccination programs become particularly important during severe epidemics or pandemics to ensure sufficient vaccination coverage is achieved to increase protection or reduce transmission to a level that enables relaxation of non-pharmaceutical interventions, such as lockdowns, travel restrictions, or social distancing. Unfortunately, vaccination uptake in the community may be slow if there are substantial levels of vaccine hesitancy in the population. As a result, it is important to identify when these hesitancy behaviours are present in the community. Furthermore, understanding the main drivers of such behaviour can inform adjustments to public health strategies to improve community uptake. In this study, we consider the problem of identifying vaccination hesitancy behaviour during a vaccination roll-out that occurs in response to a severe epidemic. Specifically, our aim is to explore the extent to which mathematical modelling of reported case, death, and vaccination counts can be used to detect vaccine hesitancy and possible drivers. To do this, we develop a novel susceptible-exposed-infectious-recovered (SEIR) epidemiological model of disease transmission that incorporates changes in population behaviour relating to non-pharmaceutical interventions and vaccine uptake that are influenced by information reported through media or data dashboards about cases, deaths, and vaccination rates. We then use a Bayesian approach to analyse simulated data representing various hesitancy scenarios. Through this simulation study, our key findings are that individual parameters values related to drivers of vaccine hesitancy often cannot be identified. However, posterior correlation structures between these parameters enable the presence of

supported by an Australian Research Council (ARC) Discovery Early Career Researcher Award (DE250100396). CD is supported by an ARC Future Fellowship (FT210100260). AM received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016233. 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.

vaccine hesitancy in the community to be detected and provide some insight into the relative influence of key factors, such as vaccine safety concerns or complacency. While our simulation study is inspired by the public health response to the COVID-19 pandemic, our tools and techniques are general and could enable vaccination programs of various infectious diseases to be adapted rapidly in response to community behaviours in the future.

Introduction

To reduce transmission and mortality due to a vaccine-preventable infectious disease, a certain level of vaccination coverage is needed to ensure a sufficient level of protection or reduction in transmission such that non-pharmaceutical interventions (NPIs) can be relaxed. Depending on the specific disease, this vaccination coverage target may be related to achieving a certain level of protection via herd immunity, or in other cases, the vaccination target may be related to reduction in transmission as for endemic diseases with waning immunity and emerging variants. Regardless of the specific vaccination mechanism that enables relaxation of NPIs, both the supply and demand for the vaccine must be sufficiently high to maintain a large enough vaccination uptake to reach the required vaccination target. The factors contributing to the success of a vaccination program within a particular community are complex and must be assessed within the cultural and societal context of the program [1,2].

One particular concern within a vaccination program is that of *vaccine hesitancy*, that is, the delay in acceptance or refusal of vaccination despite availability of vaccination services [3–5]. The so-called 3C model [4] of vaccine hesitancy [2] identifies a variety of factors that influence vaccine hesitancy within the categories of *complacency*, *convenience*, and *confidence*. In this setting, complacency refers to the low perceived risk of a vaccine-preventable disease, convenience refers to the ease of getting vaccinated, and confidence refers to the trust in vaccine efficacy and safety, including trust in the healthcare system more broadly [4]. The different effects of vaccination are important to consider when it comes to vaccine confidence. For example, a vaccine may substantially reduce the severity of a disease while having little effect on transmission, and this may lead to reduced community confidence as reported cases continue to increase. While in many cases the majority accept vaccination, there can be context specific reasons why an understanding of vaccine hesitancy patterns within a sub-population is essential to ensure an effective vaccination program [4].

During a severe epidemic or pandemic in which NPIs are necessary control measures, understanding the impact of vaccine hesitancy is crucial to plan the relaxation of restrictions [5,6]. Assuming governments do take appropriate responsibility for decision making [7,8], then identifying vaccine hesitancy patterns in a population is crucial for affirmative action by governments, healthcare institutions and policy-makers to improve vaccine uptake. The goal of this work, is to demonstrate the feasibility of using reported case data to distinguish a slow vaccine uptake due to

vaccine hesitancy from limitations in vaccination rates due to supply and logistics. This identification of vaccine hesitancy behaviours in a population could be applied to assist in the design of targeted follow-up studies and policy amendments to address the main area affecting hesitancy.

A wealth of information regarding the driving forces behind vaccine hesitancy can be obtained through statistical machine learning approaches using survey data [9–13]. During a vaccine roll-out these approaches can be used to assess the willingness of individuals to get vaccinated [10] or account for effects of media and other information sources on attitudes to vaccination [9,11–14]. While these investigations analysing survey data are insightful in identifying predictors of vaccine hesitancy, there are limitations to the survey approach. In particular, survey-based approaches often rely on detailed data being available, thus limiting the applicability of the model to other countries or jurisdictions [15]. Furthermore, a survey cannot model the effect that disease dynamics has on vaccine hesitancy and *vice versa*. Unfortunately, there have been few mathematical modelling studies that consider this dynamic. Some examples explore the impacts of information on the behaviour related to vaccine hesitancy [16,17] and strategies to counter misinformation [18–20].

Following work by Buonomo [16], Warne et al., [21] and Le et al., [22], we develop mathematical modelling and statistical analysis tools to quantify vaccine hesitancy behaviour patterns during a vaccine roll-out while NPIs are eased. We use synthetically generated reported case numbers that are characteristic of real-world standardised measurements. We incorporate the effect of a single vaccine requiring two doses into a stochastic epidemiological model that accounts for changes in behaviour related to transmission (i.e., related to NPIs and community compliance) and vaccine uptake. Changes in population behaviour are driven by reported disease case data that result in feedback loops. We demonstrate that a variety of realistic dynamics can be produced with our model under different scenarios for NPI strategies, vaccination roll-outs, and hesitancy behaviours. To obtain insight into the identification of vaccine hesitancy behaviours during a real vaccination campaign, the model is calibrated to synthetic data to determine if reported case data are sufficiently informative in relation to vaccine hesitancy [23,24]. In this work, we address this important and often overlooked aspect of mathematical modelling by applying Bayesian uncertainty quantification. We demonstrate through a simulation study that different hesitancy behaviours, such as complacency and vaccine safety concerns can be distinguished from each other using reported case data and vaccination counts. This framework has the potential to assist public health policy makers in identifying the dominating causes of vaccine hesitancy rapidly to complement results of survey data analysis [15].

Materials and methods

In this section, we present a framework to analyse vaccine hesitancy behaviour patterns during an epidemic. We demonstrate our approach using synthetic data inspired by publicly available case data and vaccination counts that were actively tracked during the COVID-19 pandemic. Our approach consists of a novel stochastic epidemiological model that accounts for changes in transmission driven by behavioural changes related to NPIs and changes in vaccination uptake based on confidence and complacency types of hesitancy behaviour.

Data sources and structure

We consider a scenario in which time-series of cumulative reported cases, deaths and vaccination doses are actively maintained and monitored during the epidemic. There are a number of challenges in using this type of data source to model the epidemic evolution and understanding hesitancy. Firstly, the cumulative confirmed cases do not necessarily represent the actual number of infections that have occurred in the population. Primarily this will be due to undetected asymptomatic cases [25,26], but could also be impacted to a substantially lesser degree by lower sensitivity or specificity in diagnostic test results. These undetected cases need to be accounted for since asymptomatic cases can still effectively transmit disease [27]. Second, the cumulative case recoveries often need to be estimated as a proportion of cumulative cases that did not die and are not currently active. Finally, and most importantly for the study of vaccine hesitancy,

the cumulative case and death data is often not stratified by vaccination status. While the vaccination status of deaths is recorded by hospitals [28], it is typically not available in large scale aggregated data available from CDC dashboards [29,30]. Therefore the distribution of vaccination status across cases and deaths also needs to be inferred when working with dashboard data.

Since this work is designed to evaluate the potential of our methods to identify hesitancy behaviour patterns, we rely on simulated data using a model with known hesitancy patterns. However, our modelling and simulated data processes are constructed in such a way that they are consistent with real epidemiological data repositories such as those provided by Johns Hopkins University [29] or Our World in Data [30] during the COVID-19 pandemic. For our simulation study, we have selected key population and epidemiological parameters that are qualitatively similar to real scenarios that occurred during the COVID-19 pandemic, however, our various hesitancy scenarios are all hypothetical and no conclusions are claimed in relation to vaccine hesitancy during COVID-19 vaccination roll-outs.

Model development

Here, we describe our stochastic epidemiological model that incorporates both the effects on NPIs and the effect of vaccination hesitancy. Our approach builds upon the model employed by Warne et al., [21] and Le et al., [22] to quantify community behaviour in relation to NPIs and potential strategies for NPI relaxation. Nomenclature is provided in Tables 1 and 2 for reference throughout these sections. For full mathematical details of the models see S1 Appendix.

Our model is a stochastic susceptible-exposed-infectious-recovered (SEIR) compartmental model. In this framework, a population of size P is divided into four main epidemiological compartments: susceptible (S), exposed (E), infectious (I) and recovered (R). Since we cannot observe exactly the true number of active infectious individuals, these components are considered unobservable or latent states. We consider the introduction of a vaccination program including a single vaccine-brand with a two dose vaccination protocol (e.g., AstraZeneca or Pfizer only), that is, each compartment is further resolved by vaccination status: unvaccinated, vaccinated (1st dose), and fully vaccinated (2nd dose) (Fig 1).

Since the population counts within each of these compartments are not directly observable in reality, these states are treated as unobservable latent states and we include an observation process based on diagnostic testing in the infectious population. Within the same vaccination status i , infectious individuals (I_i) will be identified as active confirmed cases

Table 1. List of compartments in the full SEIR model including the case reporting process of a two-dose vaccination program. A state is considered observable if it is available in a publicly available repository. States labelled with an asterisk "*" indicate data used in reporting, but aggregated before public release.

Symbol	Description	Observable
P	Total population of a region	yes
S_i	Susceptible vaccinated dose $i \in \{0, 1, 2\}$	no
E_i	Exposed vaccinated dose $i \in \{0, 1, 2\}$	no
I_i	Infectious vaccinated dose $i \in \{0, 1, 2\}$	no
R_i	Recovered vaccinated dose $i \in \{0, 1, 2\}$	no
A_i^*	Active confirmed case vaccinated dose $i \in \{0, 1, 2\}$	no
R_i^*	Case recovery vaccinated dose $i \in \{0, 1, 2\}$	no
D_i^*	Case fatality vaccinated dose $i \in \{0, 1, 2\}$	no
A^*	Total active confirmed cases	yes
R^*	Total case recoveries	yes
D^*	Total case fatalities	yes
V_i^*	Total vaccinated population dose $i \in \{1, 2\}$	yes

<https://doi.org/10.1371/journal.pone.0349499.t001>

Table 2. Model parameters fixed for synthetic dataset generation. Parameters related to vaccination rates and hesitancy are estimated via Bayesian inference. All other parameter values are obtained from the literature [21,31].

Symbol	Description	Units	Value
α_0	Residual transmission rate	[1/days]	0.05
α	Transmission rate	[1/days]	0.4
$1/\beta$	mean incubation period	[days]	14
γ	Case identification rate	[1/days]	0.05
ρ	Case recovery rate	[1/days]	0.05
δ	Case death rate	[1/days]	0.03
η	Unobserved recovery rate	[1/days]	0.1
κ	initial ratio infectious to confirmed active cases	–	10
ζ	initial proportion of living cases that are active	–	0.1527
$1/\omega$	mean duration between doses	[days]	21
α_v^1	Transmission reduction with dose 1	–	0.693
α_v^2	Transmission reduction with dose 2	–	0.048
δ_v^1	Death rate reduction with dose 1	–	0.08
δ_v^2	Death rate reduction with dose 2	–	0.047
T_d	the day that restrictions are lifted	[days]	92
T_v	the day that vaccine roll-out begins	[days]	92
n	response function $g(\cdot)$ slope parameter	–	10
n_v	hesitancy function $h(\cdot)$ slope parameter	–	2
w_A	response function active case weight	–	3×10^{-4}

<https://doi.org/10.1371/journal.pone.0349499.t002>

(A_i^*) at rate $\gamma > 0$. These active confirmed cases will transition to reported deaths (D_i^*) or recoveries (R_i^*) at rates $\delta_i > 0$ and $\rho > 0$ respectively. If we accumulate these observable case numbers over vaccination status, then we obtain three observables: the total confirmed active cases $A^* = A_0^* + A_1^* + A_2^*$, total case deaths, $D^* = D_0^* + D_1^* + D_2^*$, total case recoveries, $R^* = R_0^* + R_1^* + R_2^*$, total vaccinations (1st dose), $V_1^* = S_1 + E_1 + I_1 + R_1 + A_1^* + D_1^* + R_1^*$, and total fully vaccinated $V_2^* = S_2 + E_2 + I_2 + R_2 + A_2^* + D_2^* + R_2^*$. These observables correspond to reported data available from online dashboards and repositories during the COVID-19 pandemic [29,30]. It should be noted that most data sources only reliably record the cumulative case numbers, $C^* = A^* + R^* + D^*$, and reported deaths, D^* . In addition, the reported C^* will be an underestimate of the true number of cumulative cases since infectious individuals can also recover without being reported as a confirmed case. In our model this transition occurs with rate $\eta > 0$.

We assume vaccination only affects the disease transmission and mortality rate. At each vaccination level $i \in \{0, 1, 2\}$, modifiers $\alpha_{v,i} \in (0, 1]$ and $\delta_{v,i} \in (0, 1]$ are applied to the baseline transmission rates $\alpha > 0$ and mortality rate $\delta > 0$, respectively. For example, at vaccination level i the effective mortality rate is $\delta'_i \delta$. We set and assume $\alpha_v^2 < \alpha_v^1 < 1$ and $\delta_v^2 < \delta_v^1 < 1$. Estimates for these modifier parameters would typically be obtained from the literature resulting from vaccination clinical trials [28,31–33]. For the purposes of our simulation study, we derived rough estimates based on literature on the Oxford-AstraZeneca vaccine (ChAdOx1) in relation to the B.1.617.2 (Delta) variant of SARS-CoV-2 virus [31]. There are also vaccination rate parameters, $\nu > 0$ and $\omega > 0$. Here, ν represents the rate at which individuals get vaccinated in the absence of any vaccine hesitancy, and ω is the rate at which individuals who received their first dose $1/\omega$ days ago will proceed to second dose as per protocol, this is implemented using a delay term (See S1 Appendix). We assume that those currently infected, that is exposed and infectious individuals, do not get vaccinated until after recovery. This simplifies the model so that the only possible transitions between vaccination levels are for the susceptible and recovered compartments, S_i , R_i , and R_i^* .

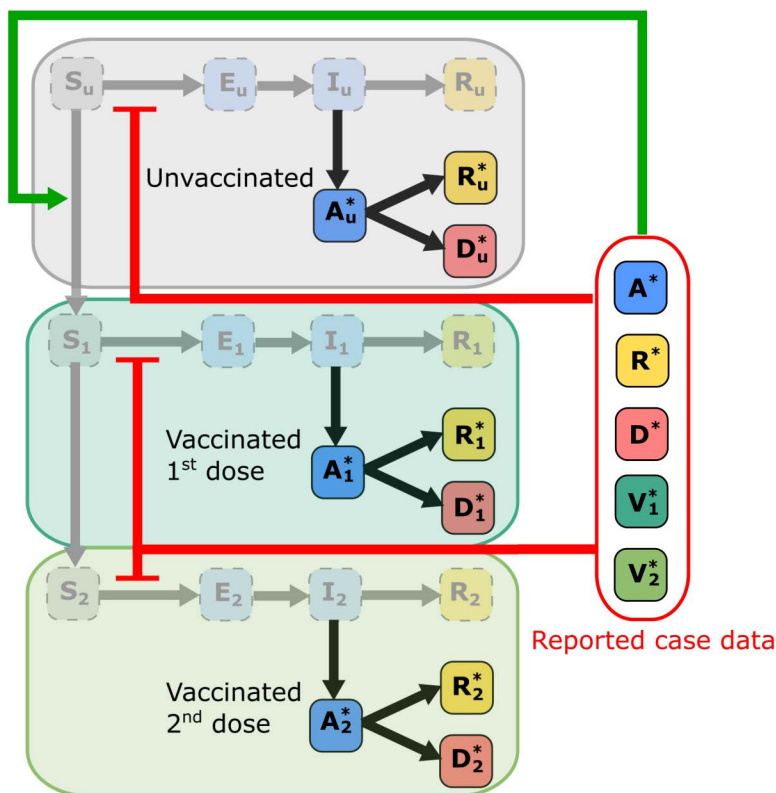


Fig 1. Schematic of the epidemiological model including two-dose vaccination program. States are marked as labelled boxes and arrows indicate state transitions. The model consists of three vaccinations stages, each with its own SEIR model with the reported case data arising through the transition from I_u to A_u^* (resp. I_1 to A_1^* and I_2 to A_2^*). The transmission is inhibited by a response function that depends on the aggregated observed reported case data. Similarly, vaccination is promoted through the hesitancy effect function that depends on the aggregated observed reported case data.

<https://doi.org/10.1371/journal.pone.0349499.g001>

Given the large well-mixed populations that we consider here (Fig 2), it is true that deterministic models may provide good approximations for the mean, however there are a number of key reasons we focus on the stochastic setting in this work. Firstly, even though the observable states are large, this does not mean that extinction events or very small populations cannot occur within any single component; ignoring this possibility can have a non-negligible effect on the system dynamics due to the two non-linear feedback mechanisms. Secondly, even when the simulation error is small, this can have an unpredictable effect on parameter inference and uncertainty quantification [34,35]. Finally, we wish to provide a computational framework that is widely applicable to other settings that include smaller populations [21]. While working with a stochastic model comes at a computational cost, we mitigate this computational burden using efficient approximate stochastic simulation and Bayesian inference scheme.

In our model, the virus transmission rate and the vaccination rate are modulated by two feedback mechanisms representing how a population changes their behaviour in response to information related to the observable states C^* , D^* , V_1^* and V_2^* . The first is a negative feedback *response function* $g(\cdot) \in [0, 1]$ that affects the transmission rate and represents the implementation of voluntary or mandated NPIs including the compliance thereof (See S1 Appendix for details). Such mechanisms have been explored for a variety of infectious disease to identify media influences [21,22,36,37]. A variety of forms for this response function can be implemented [21], here we employ an NPI implementation strategy based on a trigger threshold,

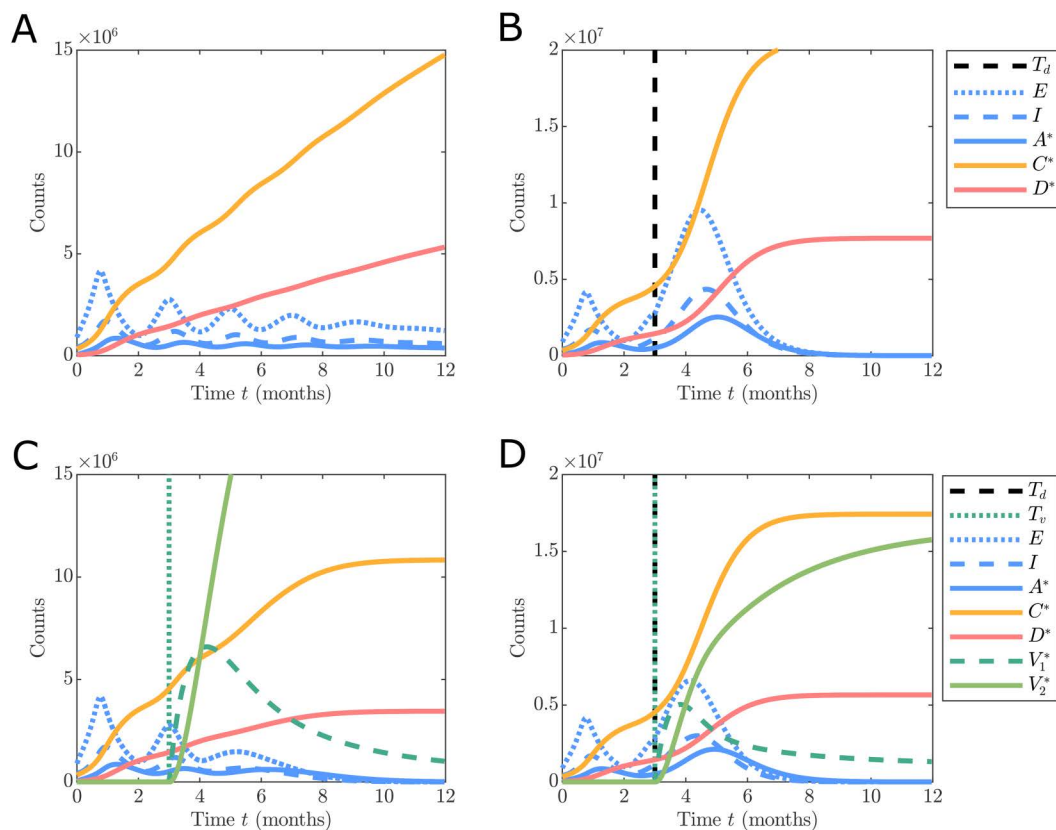


Fig 2. Comparison of the effect of an AstraZeneca-like vaccination program under different configurations of response function and vaccination inclusion (assuming no vaccine hesitancy). Simulations are initialised based on the UK data on 1st September 2020 with $C_0^* = 337,798$, $D_0^* = 41,551$. (A) and (C): corresponds to ongoing small lockdowns, (B) and (D): corresponds to complete easing of restrictions after $T \geq 92$ (roughly the 2nd of December 2020). A single stochastic realisation is shown in (A)–(D) using an approximate stochastic simulation scheme (See S1 Appendix).

<https://doi.org/10.1371/journal.pone.0349499.g002>

$$g(A^*, t) = \frac{1}{1 + (w_A A^*)^n} \mathbb{1}_{(0, \tau_d)}(t) + \mathbb{1}_{(\tau_d, \infty)}(t), \quad (1)$$

where $T_d > 0$ is the point in time that NPIs cease and w_A is a weight parameter setting the number of reported active cases at which NPI efficacy reaches 50%. Note that for $t > T_d$ the $g(A^*, t) = 1$. See S1 Appendix for additional details.

Fig 2 demonstrates the effect that vaccination has on the virus spread and severity using example stochastic realisations from our model parameterised by Fig 2 under different scenarios of response behaviour with the assumption no hesitancy behaviour (that is, perfect vaccination rate of $\nu > 0$ for $t > T_v > 0$ where T_v is the start time of the vaccination roll-out). Fig 2(A) shows a continued lockdown using Eq. (1) with $T_d \rightarrow \infty$ and no vaccination roll-out $T_v \rightarrow \infty$. Fig 2(B) shows a similar scenario, however with NPIs ceasing at $T_d = 3$ (months). Comparing Fig 2(A) and 2(B) demonstrates the effect of ceasing NPIs in the absence of any effective vaccine, that is, an order of magnitude more reported cases and deaths. Fig 2(C) and 2(D) show the same lockdown scenarios as Fig 2(A) and 2(B), however, introducing a vaccination roll-out commencing at $T_v = 3$ (months). We observe, as expected, a reduction in the total case deaths due to vaccination (Compare Fig 2(A) with 2(C) and Fig 2(B) with 2(D)). We also note a chance of elimination occurring when restrictions continue along with the vaccination program (Fig 2(C)), such elimination behaviours can only be captured using a stochastic model.

The second feedback loop relates to the vaccine hesitancy. We denote $h(\cdot) \in [0, 1]$ as the *hesitancy effect* function resulting in an effective vaccination rate of $\nu h(\cdot)$. We assume that vaccine hesitancy behaviour only affects the probability of an individual receiving the first dose with every individual that gets vaccinated continuing to obtain a second dose to be “fully vaccinated”. We note that this assumption may not always be appropriate, especially for diseases requiring boosters or adverse reactions to the first dose reducing the likelihood of an individual returning for the second dose (see Discussion). Just as with the response function, the hesitancy effect function only depends on the observable states and time. We treat $h(\cdot)$ as an increasing function with $h(\cdot) = 0$ corresponding to a situation where individuals refuse vaccination entirely and $h(\cdot) = 1$ leads to no hesitancy effect with the maximum rate ν achieved. This function is intended to model the way in which reported case and vaccination data may increase the probability of an individual seeking vaccination.

We can construct a hesitancy effect function to capture the complacency and confidence components of the 3C model of vaccine hesitancy [4]. We do not consider the convenience component due to the well-mixed assumption (See Discussion). If we consider a population that is only affected by complacency, then it is reasonable to assume the incidence of a disease in a community will tend to increase the likelihood of an individual to seek vaccination [16,18] as the perceived risk increases to overcome complacency. This complacency pattern could be modelled with the function,

$$h(C^*, D^*, t) = \frac{(w_C C^* + w_D D^*)^{n_v}}{1 + (w_C C^* + w_D D^*)^{n_v}} \mathbb{1}_{[T_v, \infty)}(t), \tag{2}$$

where $w_C, w_D > 0$ are weights reflecting the influence of data reports on population behaviour and $n_v > 0$ is a slope parameter that governs the sensitivity of the behaviour change to increased incidence, and T_v is that time the vaccine roll-out starts with $\mathbb{1}_{[T_v, \infty)}(t) = 0$ for $t < T_v$. The relative importance of case numbers and deaths in influencing an individuals decision to get vaccinated can be assessed by considering the case where $C^* = (w_D/w_C)D^*$, that is each new death has as much influence as w_D/w_C new cases. Of course, motivation could be completely dominated by cases (resp. deaths), in which case $w_D = 0$ (resp. $w_C = 0$). In practice, due to the strong correlation between total confirmed cases and deaths, it is perfectly reasonable to set $w_C = 0$.

Another cause of vaccine hesitancy behaviour are vaccine safety concerns, related to confidence in the 3C model [4]. Such safety concerns were prevalent in the COVID-19 pandemic due to rapid development of vaccines [38] and some widely publicised, though rare, side effects [19,20]. If only vaccination safety is a concern, then critical mass of fully vaccinated individuals may be necessary to alleviate these concerns. This can be reflected in the function,

$$h(V_2^*, t) = \frac{(w_V V_2^*)^{n_v}}{1 + (w_V V_2^*)^{n_v}} \mathbb{1}_{[T_v, \infty)}(t), \tag{3}$$

with weight parameter $w_V > 0$. Note that in this case, at time T_v some initial non-zero vaccinated population must be present, otherwise the vaccinated population will remain at $V_2^* = 0$. The initial population could represent participants in relevant clinical trials, or an initial vaccination mandate for particular occupations (e.g., healthcare workers). In reality, other factors beyond vaccine safety could affect confidence especially in situations where the vaccine reduces disease severity without impacting transmission. In this setting, Eq. (3) could be extended to enable continued increase in reported cases to reduce confidence, which would work against the increasing case numbers in the complacency function Eq (2).

Finally, we can consider a population hesitancy behaviour that is due to safety concerns and complacency,

$$h(C^*, D^*, V_2^*, t) = \frac{(w_C C^* + w_D D^* + w_V V_2^*)^{n_v}}{1 + (w_C C^* + w_D D^* + w_V V_2^*)^{n_v}} \mathbb{1}_{[T_v, \infty)}(t). \tag{4}$$

Once again the weights $w_C, w_D, w_V \geq 0$ provide a measure of the relative influence each new case, death or vaccination has on the probability of an individual getting vaccinated. In reality, these parameters will be unknown, however if standard

COVID-19 case data can provide insight into the values of these parameters [21], then we can start to assess the trends in hesitancy behaviours for a given population of interest.

Simulation study

Given that reported data are aggregated, it is unclear if the data are informative enough to identify what type of hesitancy effect is occurring in reality. For example, three different scenarios are shown in Fig 3 corresponding to no hesitancy (Fig 3(A)), complacency only (Fig 3(B)), and complacency and vaccine safety concerns (Fig 3(C)). While each evolution is different qualitatively, it is completely unclear if the scenarios including different hesitancy effects can be distinguished from a scenario with no hesitancy effect but slower vaccination rate. For example, such is the difference in the hesitancy behaviour between Fig 3(C) and the no hesitancy case of Fig 3(A), yet the differences in the dynamics are subtle. Here Fig 3(C) looks like Fig 3(A) with a slower vaccination rate, however, in reality both have the same rate and the effect is entirely due to hesitancy.

The behaviour observed in Fig 3 provides motivation to explore the informativity of aggregated reported COVID-19 data as provided by online repositories synthetically for the purpose of analysing vaccine hesitancy behaviour. To explore this, we consider a simulation study and develop Bayesian analysis techniques.

We generate multiple simulated datasets under three different vaccine uptake behaviours. Specifically, we consider: complacency-based hesitancy ($w_V=0$, $w_C>0$, and $w_D>0$), complacency and vaccine safety ($w_V>0$, $w_C>0$, and $w_D>0$) hesitancy, and no hesitancy ($h(\cdot) = 1$). In each of the simulation settings we vary the weights and the maximum vaccination rate, the remaining parameters are fixed based on the literature [21,28,31–33] as listed in Table 2.

Bayesian analysis

Given a simulated dataset, $\mathcal{D} = \{(C_t^*, D_t^*, V_{1,t}^*, V_{2,t}^*)\}_{t=0}^{t=450}$, we infer the parameters related to the vaccination uptake $\theta = [\nu, n_V, w_C, w_D, w_V]$ through sampling the Bayesian posterior distribution with density,

$$p(\theta | \mathcal{D}) \propto \mathcal{L}(\theta; \mathcal{D})p(\theta),$$

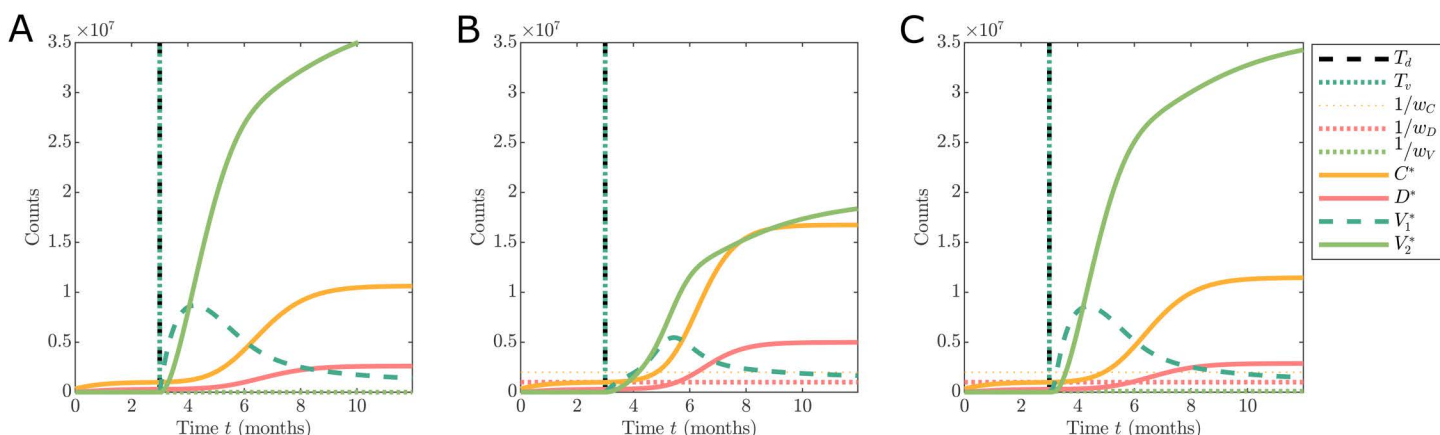


Fig 3. Comparison of different hesitancy effect functions show the dynamics of the observable states. The scenarios are: (A) no hesitancy behaviour with $h(\cdot, t) = \mathbb{1}_{[T_v, \infty)}(t)$; (B) hesitancy effect driven by complacency only with $h(\cdot)$ defined in Eq. (2); and (C) hesitancy effect driven by mild safety concerns and some complacency only with $h(\cdot)$ defined in Eq. (4). Simulations are initialised based on the UK data on 1st September 2020 with $C_0^* = 337,798$, $D_0^* = 41,551$, and model parameters are given in Table 2 with $\nu = 0.01$, $n_V = 4$, $w_C = 2 \times 10^{-6}$, $w_D = 1 \times 10^{-6}$, and $w_V = 1 \times 10^{-5}$.

<https://doi.org/10.1371/journal.pone.0349499.g003>

where $\mathcal{L}(\theta; \mathcal{D})$ is the likelihood function and $p(\theta)$ is the prior probability density. Due to the stochastic nature of the model and the large number of latent states, the likelihood function is intractable, and we rely upon approximate Bayesian computation (ABC) [39,40] to sample the approximate posterior,

$$p(\theta \mid \rho(\mathcal{D}, \mathcal{D}_s) \leq \epsilon) \propto \mathbb{P}(\rho(\mathcal{D}, \mathcal{D}_s) \leq \epsilon \mid \theta)p(\theta), \tag{5}$$

where \mathcal{D}_s is a simulated dataset generated from the model with given parameter values, θ , $\rho(\mathcal{D}, \mathcal{D}_s)$ is a discrepancy metric between the data and simulations, and ϵ is a target discrepancy threshold. For this work, our discrepancy metric between two datasets is the Frobenius norm of the matrix of differences between counts in each dataset (S2 Appendix). To achieve accurate inferences we apply an adaptive sequential Monte Carlo sampler for ABC (SMC-ABC) [41] that produces N samples with as small a discrepancy as practically possible. We repeat this inference process for all M synthetic datasets, $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_M$, that span the different hesitancy scenarios. Thereby arriving at M sets of posterior samples $\{\theta_{1j=1}^j, \theta_{2j=1}^j, \dots, \theta_{Mj=1}^j\}$. The structure of the resulting posterior samples can be used to investigate if there is information about vaccine hesitancy available in COVID-19 case data and vaccination counts. In all cases the priors are $\nu \sim \mathcal{U}(0, 0.02)$, $n_v \sim \mathcal{U}(0, 10)$, $-\log w_C \sim \mathcal{U}(3, 8)$, $-\log w_D \sim \mathcal{U}(3, 8)$, and $-\log w_V \sim \mathcal{U}(3, 8)$.

Model fit is assessed through sampling the within-sample posterior predictive distribution with density,

$$p(\mathcal{D}_s \mid \mathcal{D}) = \int_{\Theta} \mathcal{L}(\theta; \mathcal{D}_s)p(\theta \mid \mathcal{D}) d\theta. \tag{6}$$

If the model fits well, then we expect the data to lie within the 95% Credible Intervals (95% CrI). Since the models we use to generate synthetic data are all special cases of the full model that is based on (4) and (5), there is no model misspecification error to account for in the inference. However, the posterior predictive check is useful to assess the accuracy of the SMC-ABC sampler with respect to the discrepancy threshold. In practice, model misspecification is important to consider in real applications and we discuss this in the Discussion section.

Identifiability analysis

In the setting of deterministic models, a parameter is defined as *structurally identifiable* if the mapping from parameter value to a model solution curve is an invertible function. Various methods to analyse structural identifiability for deterministic dynamical systems are available in the literature [42–45]. In contrast, very few methods are available for structural identifiability for stochastic models [24,46]. For simplicity, we derive an equivalent deterministic ordinary differential equation (ODE) approximation (S1 Appendix) to our stochastic model and apply generating series approaches to investigate identifiability [45]. This is appropriate in this setting as the ODE can be considered as an approximation of the mean of the stochastic process as the variance is comparatively negligible due to the large populations in each compartment [24].

In addition to structural identifiability, Bayesian posterior samples provide a way to explore practical identifiability. While frequentist approaches exist for exploring practical identifiability, such as profile likelihoods [47–49], they are challenging to apply for models with intractable likelihood functions [50]. Therefore we adopt a Bayesian approach where practical identifiability is primarily explored via the marginal posterior distributions typically using Markov chain Monte Carlo [24,51] and posterior predictive sampling. Since we are sampling from the joint posterior distribution we automatically have access to samples from the marginal posterior distributions. When marginals themselves demonstrate practical non-identifiability, exploring correlation structures can provide additional insight. In this work, we visually inspect the bivariate marginal structures for each simulation case and Spearman’s rank correlation between pairs of parameters.

Results

In this section, we present the results of our simulation study. We highlight the key simulation scenarios explored and appropriateness of the model calibration process. We then demonstrate that while individual vaccine uptake parameters may not be uniquely identified from COVID-19 case data, the posterior distributions contain sufficient information to determine the vaccination scenario occurring in the data.

Simulation scenarios and model calibration

Synthetic datasets were constructed using the model under each vaccination hesitancy behaviour scenario considered in the Methods section. Example synthetic datasets for these three cases are shown in Fig 4.

For each dataset, \mathcal{D} , we apply adaptive SMC-ABC (S2 Appendix) targeting the joint posterior $p(\nu, n_v, w_C, w_D, w_V | \mathcal{D})$. The result is $N=1000$ samples from the approximate posterior for each synthetic dataset. These samples are used for Monte Carlo integration of posterior probability densities, practical identifiability analysis, and visualisation of parameter correlation structures to investigate the use of COVID-19 case data for identification of vaccine hesitancy patterns.

The model fitness is evaluated for each synthetic dataset using within-sample posterior predictive checks. Since the model used for each synthetic dataset is nested within our full model we do not need to account for model misspecification. See Discussion for considerations when misspecification is relevant. The purpose of the posterior predictive check is to ensure the sampler has adapted to a sufficiently small discrepancy threshold for inference. Fig 5 provides example posterior predictive simulations and demonstrates typical model fits obtained from the Bayesian calibration using equivalent convergence criteria for the posterior sampler.

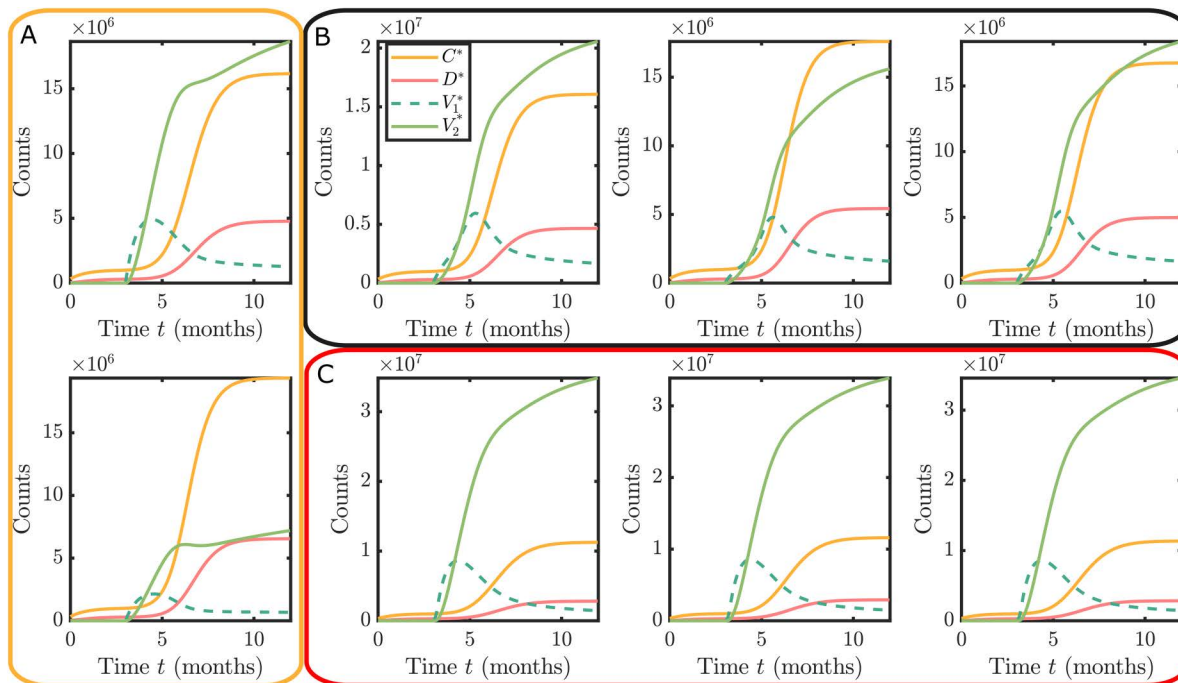


Fig 4. Example synthetic datasets for the three hesitancy scenarios: (A) no hesitancy, but smaller maximum vaccination rate than in (B) and (C); (B) complacency only with $w_c > 0$, $w_p > 0$ and $w_v = 0$; and (C) complacency and vaccine safety concerns with $w_c > 0$, $w_p > 0$ and $w_v > 0$.

<https://doi.org/10.1371/journal.pone.0349499.g004>

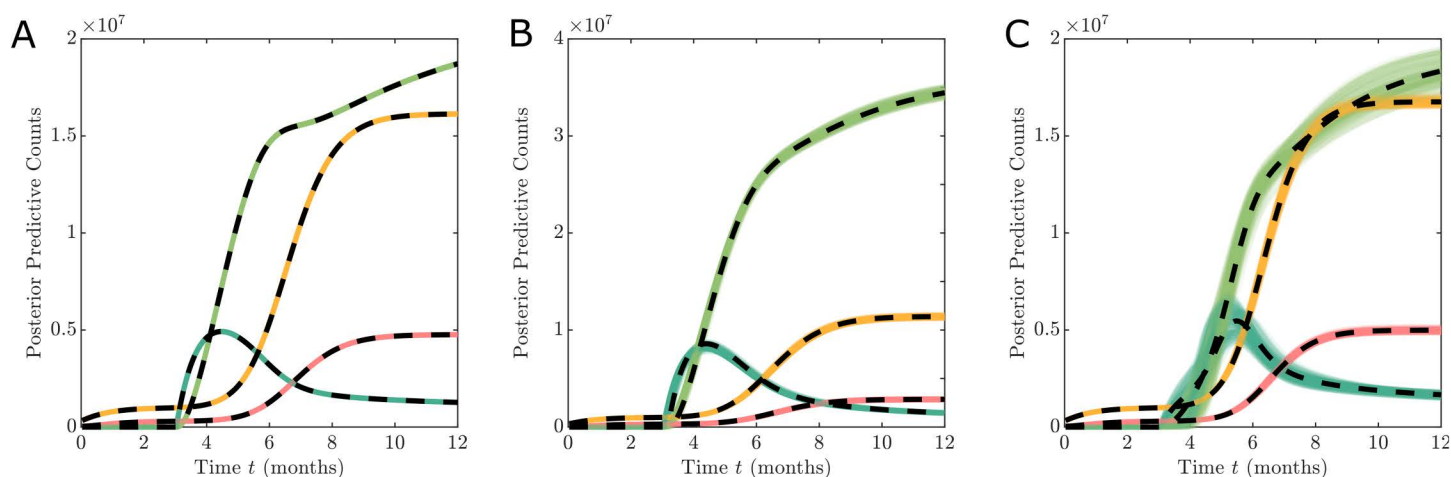


Fig 5. Example posterior predictive simulations (shaded regions) demonstrating model fit for different synthetic datasets (black dashed lines). (A) no hesitancy (true parameters $\theta = (0.005, 4, 1, 1, 1)$); (B) complacency only (true parameters $\theta = (0.01, 4, 5 \times 10^{-7}, 1 \times 10^{-6}, 0)$); and (C) Complacency and vaccine safety concerns (true parameters $\theta = (0.01, 4, 5 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5})$).

<https://doi.org/10.1371/journal.pone.0349499.g005>

Identifiability of hesitancy function parameters

We use the `GenSSI 2.0` toolkit [43,45] to analyse the structural identifiability of the parameters related to vaccination uptake $\theta = (\nu, n_v, w_C, w_R, w_V)$. This approach is applied to a deterministic version of the model. We consider the case when $T_v = T_d$, that is, NPI based restrictions are lifted at the same time as the vaccine roll-out, and the case when $T_v \ll T_d$, that is, NPI restrictions continue throughout the vaccination roll-out. In both cases we observe only the total reported case numbers, deaths, and vaccinations.

In both cases, all of the vaccine uptake parameters are shown to be locally structurally identifiable. This means that model parameters can be uniquely determined within a subset of parameter space. While this structural identifiability analysis is not directly applicable to the stochastic model, for very large case numbers, the stochastic effects are negligible. Therefore, we conclude that the observation process, if it was continuous, is informative enough to investigate hesitancy behaviours.

In addition, since our real observation process only records daily counts, we still need to explore practical identifiability as the data are not continuous noise-free observations. The posterior marginal distributions demonstrate practical non-identifiability. This can be observed in the example marginal posterior density plots (Fig 6). Regardless of the synthetic data scenario, the hesitancy function parameters w_C, w_D , and w_V shows substantial uncertainty. The only parameter that is well identified is the maximum vaccination rate ν . While the shape of the marginals (particularly the skewness) provide some hints toward recovering the true hesitancy scenario for the respective synthetic data, the differences between the scenarios are subtle.

Hesitancy inferred through dependency structures

The marginal posterior distributions indicate that key parameters related to vaccination rates and hesitancy behaviours are practically non-identifiable using COVID-19 case data. However, practical non-identifiability does not preclude the study of the posterior structure to provide insight into the hesitancy behaviours. In particular, the joint posterior distributions contain information about the correlation structure between these parameters. The qualitative form of these correlation structures can be used to uncover the true hesitancy behaviour, albeit without the ability to quantify the actual parameter values reliably. Fig 7 show characteristic example bivariate plot matrices for each of the different vaccination hesitancy scenarios

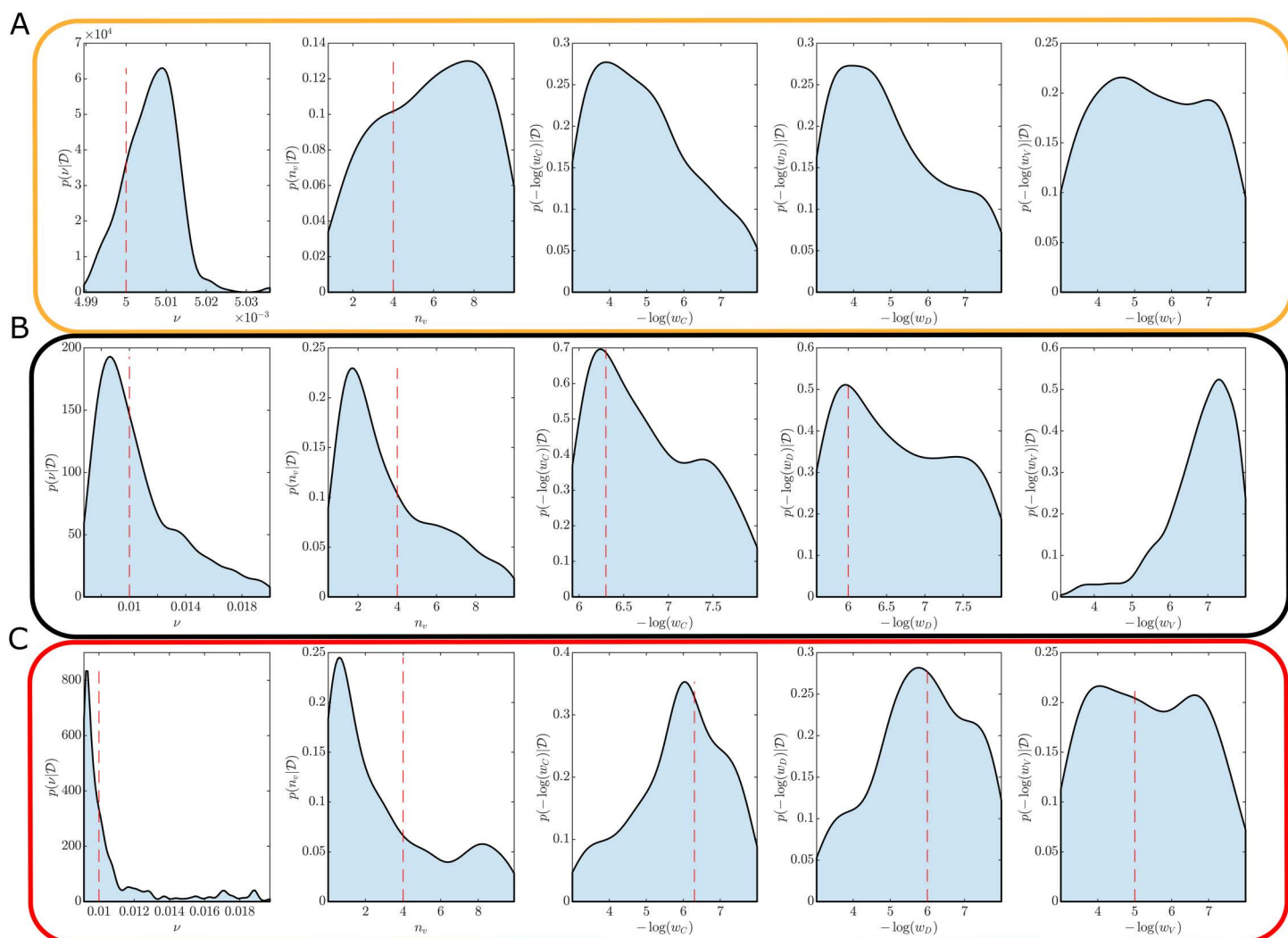


Fig 6. Example marginal posterior densities obtained using synthetic datasets for the three hesitancy scenarios: (A) no hesitancy; (B) complacency only; (C) complacency and vaccine safety concerns. True parameter values are indicated (red dashed lines). Substantial uncertainties in the hesitancy function parameters n_v , w_C , w_D , and w_V indicate practical unidentifiability problems. Cases without a dashed red line indicate cases when the true parameter corresponds to: (A) $w_C, w_D, w_V \rightarrow \infty$; and (B) $w_V \rightarrow 0$.

<https://doi.org/10.1371/journal.pone.0349499.g006>

(See [S3 Appendix](#) for full bivariate distributions). Throughout, all statistical hypothesis tests are performed at the 0.01 significance level, and correlations refer to Spearman's rank correlation.

Firstly, by comparing [Fig 7\(A\)](#), that corresponds to synthetic data generated with no hesitancy, with [Fig 7\(B\)](#) and [7\(C\)](#), that correspond to different hesitancy scenarios, we observe qualitative and quantitative differences in the bivariate posteriors. In the absence of hesitancy in the data, there is no evidence for correlation between ν and the negative log weights, or with the slope parameter n_v as all correlations are statistically insignificant ([Fig 7\(A\)](#)). In contrast, there are statistically significant negative correlations between ν and n_v , and statistically significant positive correlations between ν and negative log weights $-\log(w_C)$, $-\log(w_D)$, $-\log(w_V)$ ([Fig 7\(B\)](#) and [7\(C\)](#)). This demonstrates that correlation between the maximum vaccination rate ν with the negative log weight parameters $-\log(w_C)$, $-\log(w_D)$, and $-\log(w_V)$ provide a

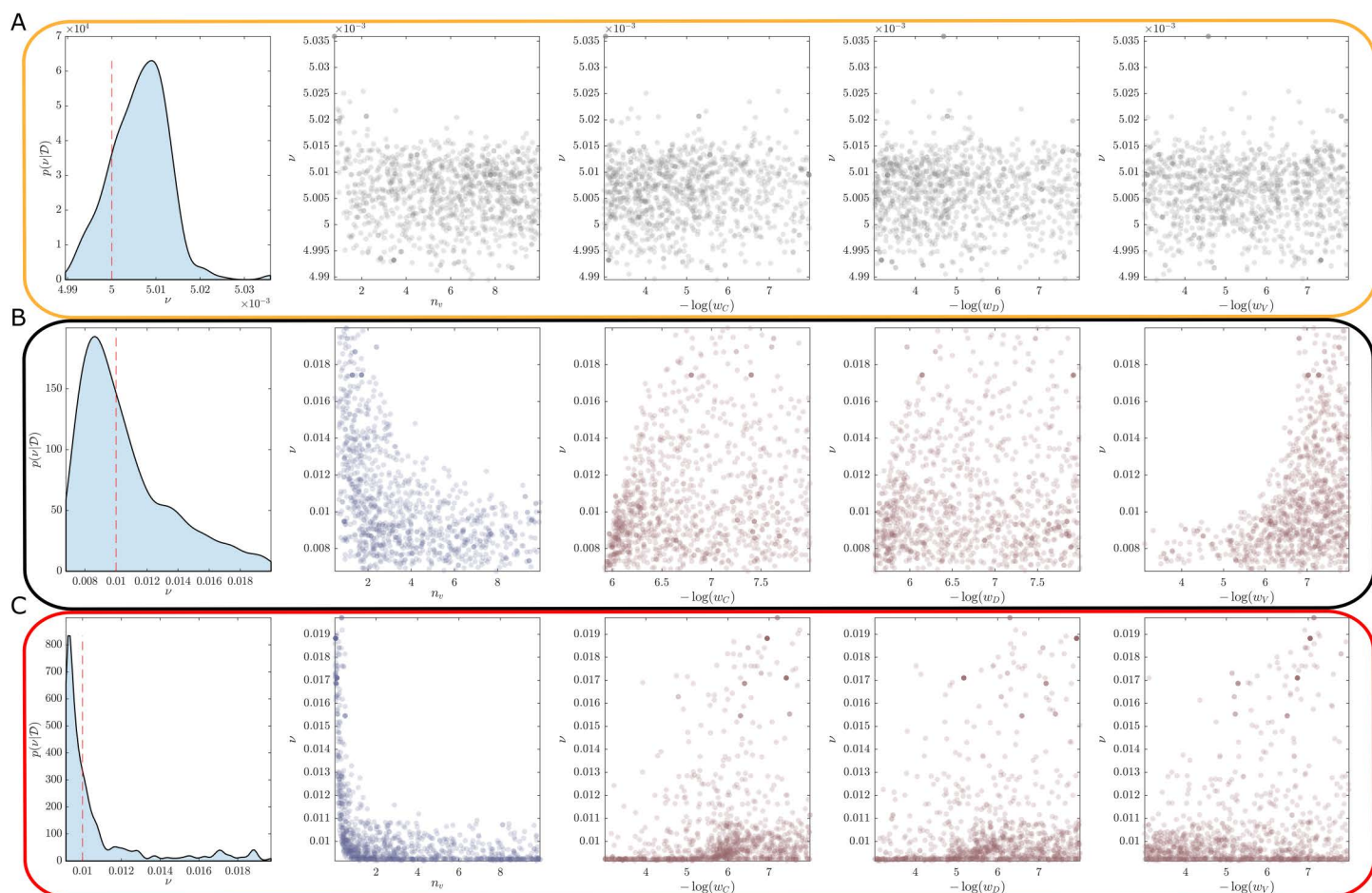


Fig 7. Univariate and bivariate marginal distributions obtained using synthetic datasets for the three hesitancy scenarios: (A) no hesitancy; (B) complacency only; and (C) complacency and vaccine safety concerns. The univariate marginal posterior densities for the maximum vaccination rate, ν , is show with dashed red lines indicating the true values. The scatter plots show the bivariate posterior samples involving ν , from which the pairwise correlations are computed. Plots with significantly positive, significantly negative and statistically insignificant correlations are shown in red, blue and grey respectively.

<https://doi.org/10.1371/journal.pone.0349499.g007>

strong indicator for the presence of hesitancy behaviours, despite univariate marginal parameter estimates being practically non-identifiable.

Second, comparing correlation structures for different cases of hesitancy can inform the components that are dominating behaviour. For example, [Fig 7\(B\)](#) shows the results for a synthetic dataset with only complacency behaviour. Here, the strength of the correlation between ν and $-\log(w_V)$ is substantially stronger than that of ν with $-\log(w_C)$ and $-\log(w_D)$. This indicates the vaccine counts have a much smaller effect on the vaccine uptake than the confirmed cases and deaths. By comparison, a scenario that includes both complacency and vaccine safety concerns ([Fig 7\(C\)](#)) has weaker correlation strengths that are on similar scales. The interpretation here is that increased cases, deaths and vaccine counts all increase the probability of individuals get vaccinated.

Finally, the correlations between the slope n_v parameters also provide some insight, but this information mainly complements the insights obtained through correlations with ν and the other parameters ([S3 Appendix](#)). These correlations

seem to allow some identification of the dominating complacency drivers between total case numbers or death numbers, however, more work is required to explore this in more detail.

Summary

Through simulated synthetic dataset scenarios and Bayesian analysis, we demonstrate the ability to identify the presence of vaccine hesitancy at a population level. This is possible despite practical non-identifiability of individual parameters. We achieve this through inspection of the correlation structures in the joint posterior distribution for the parameters relevant to vaccine hesitancy. Further, once vaccine hesitancy behaviours have been identified, it is possible to isolate the specific factors that are driving the hesitancy behaviour. We explore the potential implications and limitations of these results in the Discussion section.

Discussion

In this work, we explore the feasibility of isolating vaccine hesitancy behaviour and key drivers of this hesitancy using reported case data, such as those available from COVID-19 online dashboards. Using a stochastic epidemic model that accounts for the effects of NPIs and vaccinations over time, we generate synthetic datasets for various hesitancy scenarios with model parameterisation inspired by the early AstraZeneca COVID-19 vaccination roll-out in the United Kingdom. While hesitancy parameters are structurally identifiable, Bayesian analysis reveals that there are practical identifiability challenges for the stochastic model. Despite this, our analysis framework can distinguish the different hesitancy behaviours through inspection of the joint posterior densities and correlation structures. This is potentially a useful tool for future vaccination roll-outs for many diseases aside from COVID-19.

This analysis implies that, in principle, protocols could be developed to quantify the population level effects of hesitancy behaviour during a vaccine roll-out. This could supplement traditional approaches based on survey data to develop information and education campaigns to target hesitancy causes that will have the greatest impact on vaccine uptake. This could be particularly valuable in supplementing survey data when the epidemic situation is rapidly changing as sudden changes in attitudes are difficult to capture with survey data [15]. Furthermore, if reported data is appropriately subdivided, differences in hesitancy behaviour could be assessed for different jurisdictions or vaccine eligibility groups. This could have been applicable during the COVID-19 vaccine roll-out of AstraZeneca when various regulators adjusted the minimum age recommendations in response to concerns around thrombosis and thrombocytopenia syndrome [19,52].

There are several key assumptions within our framework and it is important to highlight them here and discuss the applicability of our methods for real data. Firstly, we consider a spatially homogeneous population in which all individuals have an equal probability of interaction. While we have presented our framework for a large population size, in many realistic settings this assumption will be mainly applicable to smaller spatial scales, such as cities rather than countries. Second, we do not consider an age-structured population nor a staged vaccination roll-out. This does not preclude the utility of our method to be applied to staged programs, provided the reported data is available for the various ages structures and eligibility criterion for each vaccine stage. Third, we assume there is sufficient information on the epidemic and vaccine efficacies to obtain values for all parameters except the vaccination rate and hesitancy parameters. This is not necessarily a major constraint, since substantial epidemiological research will typically proceed vaccination programs and the setting of vaccination roll-out targets, such as vaccine coverage, typically rely on estimates of epidemiological parameters [53–55]. Importantly, as we consider a Bayesian setting, any uncertainties in epidemic parameters or vaccine efficacy can be propagated through our framework through prior specification. Fourth, we only consider a single vaccine with a two dose regime with no hesitancy effects on the second dose. However, the overall approach would still be applicable to multiple vaccine brands by either extending the model to incorporate more vaccination types, or treat vaccination efficacy parameters as an averaged effect. As many data sources do not provide detailed breakdowns of vaccination counts by brand, the latter approach may be preferred. Finally, we assume for simplicity that there is no model misspecification. A

key motivation for this was to obtain “best-case” scenarios on what can theoretically be identified about vaccine hesitancy behaviour during a vaccine roll-out. While this assumption will almost certainly not hold in practice, our likelihood-free approach to Bayesian inference (See [S2 Appendix](#)) has well-established properties for natural robustness to misspecification and accounts for model uncertainty [56–58].

In this work we have focused on the identification of hesitancy behaviours for a single well-mixed population. As we have stated in our assumptions, this does not limit the use of our approach to more complex and realistic populations. However, future development of our methods could dramatically expand the range of possibilities. Various extensions that account for spatial heterogeneity [59], age-structure [60], and seasonality [61] would improve the broad applicability of our methods to more realistic settings. Some extensions, such as spatial heterogeneity, would be challenging to implement using deterministic models, however, these are feasible and completely valid within a stochastic framework. As COVID-19 becomes endemic to a population effects of waning immunity and vaccination booster programs become important [62]. Due to our focus on the initial roll-out of a new vaccine, we have not accounted for these effects here, however, future research should extend our approaches to this case. Finally, vaccine hesitancy could be different for each vaccine brand, therefore a multiple vaccine system could be developed, though it is unclear if the individual hesitancy behaviours will be identifiable in this case.

To implement any of the above mentioned improvements to our modelling framework requires either an extension of the compartment space or the inclusion of hierarchical effects. For models with intractable likelihood functions, as our stochastic model is, such extensions have the potential to dramatically increase the computational cost and inhibit the feasibility of including additional model complexity. However, advanced algorithms for likelihood-free Bayesian computation have the potential to resolve some of these issues. For example, in the hierarchical modelling setting, new computational schemes have been developed in a likelihood-free setting to deal with waning immunity in epidemiological modelling [63]. Furthermore various approaches that exploit surrogate or approximate models have been shown to dramatically improve performance whilst sacrificing minimal accuracy [34,35,64–70]. Finally, state-of-the-art massively parallel computing hardware has potential to complement algorithmic advances to enable real-time analysis [69,71–78].

Conclusion

Our modelling and computational framework provides a new approach to monitoring vaccine hesitancy using reported case data and vaccination counts that have been widely available during the COVID-19 pandemic. We show that vaccine hesitancy behaviours can be identified from these data provided sufficient information on the epidemic exists preceding the vaccine roll-out. A full Bayesian analysis is performed to be able to identify correlation structure in parameters. It is likely that vaccine hesitancy will continue to be a barrier for vaccine uptake and a major concern for governments during the roll-out of new vaccines in future epidemics. Therefore, our tools that enable the rapid assessment of trends in vaccine hesitancy behaviours within a population can greatly assist public health policy makers and practitioners in addressing public concerns.

Supporting information

S1 Appendix. Stochastic model details.

(PDF)

S2 Appendix. Approximate Bayesian Analysis.

(PDF)

S3 Appendix. Additional Results.

(PDF)

S1 Code. Matlab implementations of models and analysis.

(ZIP)

Acknowledgments

We thank the two anonymous reviewers and associate editor for the valuable comments on our work. Computational resources were provided by the eResearch Office, Queensland University of Technology.

Author contributions

Conceptualization: David J Warne, Kerrie Mengersen, Adrienne L Jenner.

Data curation: Abhishek Varghese.

Formal analysis: David J Warne, Alexander P Browning.

Funding acquisition: David J Warne, Christopher Drovandi, Antonietta Mira, Adrienne L Jenner.

Investigation: Abhishek Varghese, Adrienne L Jenner.

Methodology: David J Warne, Abhishek Varghese, Alexander P Browning, Mario M Krell, Christopher Drovandi, Adrienne L Jenner.

Project administration: David J Warne.

Resources: Mario M Krell.

Software: David J Warne, Abhishek Varghese, Aidan Brewster, Mario M Krell, Christopher Drovandi.

Supervision: David J Warne, Adrienne L Jenner.

Validation: David J Warne, Aidan Brewster, Alexander P Browning, Mario M Krell, Christopher Drovandi, Wenbiao Hu, Antonietta Mira, Kerrie Mengersen, Adrienne L Jenner.

Visualization: David J Warne, Abhishek Varghese, Aidan Brewster.

Writing – original draft: David J Warne, Adrienne L Jenner.

Writing – review & editing: David J Warne, Alexander P Browning, Mario M Krell, Christopher Drovandi, Wenbiao Hu, Antonietta Mira, Kerrie Mengersen, Adrienne L Jenner.

References

1. Dubé E, Gagnon D, Nickels E, Jeram S, Schuster M. Mapping vaccine hesitancy—country-specific characteristics of a global phenomenon. *Vaccine*. 2014;32(49):6649–54. <https://doi.org/10.1016/j.vaccine.2014.09.039> PMID: [25280436](https://pubmed.ncbi.nlm.nih.gov/25280436/)
2. World Health Organization. Behavioural and social drivers of vaccination: tools and practical guidance for achieving high uptake. World Health Organization; 2022.
3. Larson HJ, Jarrett C, Eckersberger E, Smith DMD, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: A systematic review of published literature, 2007–2012. *Vaccine*. 2014;32(19):2150–9. <https://doi.org/10.1016/j.vaccine.2014.01.081>
4. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: Definition, scope and determinants. *Vaccine*. 2015;33(34):4161–4. <https://doi.org/10.1016/j.vaccine.2015.04.036> PMID: [25896383](https://pubmed.ncbi.nlm.nih.gov/25896383/)
5. Olivares A, Staffetti E. Uncertainty quantification of a mathematical model of COVID-19 transmission dynamics with mass vaccination strategy. *Chaos Solitons Fractals*. 2021;146:110895. <https://doi.org/10.1016/j.chaos.2021.110895> PMID: [33814733](https://pubmed.ncbi.nlm.nih.gov/33814733/)
6. Hogan AB, Winskill P, Watson OJ, Walker PGT, Whittaker C, Baguelin M, et al. Within-country age-based prioritisation, global allocation, and public health impact of a vaccine against SARS-CoV-2: A mathematical modelling analysis. *Vaccine*. 2021;39(22):2995–3006. <https://doi.org/10.1016/j.vaccine.2021.04.002> PMID: [33933313](https://pubmed.ncbi.nlm.nih.gov/33933313/)
7. Betsch C, Schmid P, Heinemeier D, Korn L, Holtmann C, Böhm R. Beyond confidence: Development of a measure assessing the 5C psychological antecedents of vaccination. *PLoS One*. 2018;13(12):e0208601. <https://doi.org/10.1371/journal.pone.0208601> PMID: [30532274](https://pubmed.ncbi.nlm.nih.gov/30532274/)
8. Attwell K, Hannah A, Leask J. COVID-19: talk of “vaccine hesitancy” lets governments off the hook. *Nature*. 2022;602(7898):574–7. <https://doi.org/10.1038/d41586-022-00495-8> PMID: [35194212](https://pubmed.ncbi.nlm.nih.gov/35194212/)

9. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. *Lancet Public Health*. 2021;6(4):e210–21. [https://doi.org/10.1016/S2468-2667\(21\)00012-8](https://doi.org/10.1016/S2468-2667(21)00012-8) PMID: [33556325](https://pubmed.ncbi.nlm.nih.gov/33556325/)
10. Lincoln TM, Schlier B, Strakeljahn F, Gaudio BA, So SH, Kingston J, et al. Taking a machine learning approach to optimize prediction of vaccine hesitancy in high income countries. *Sci Rep*. 2022;12(1):2055. <https://doi.org/10.1038/s41598-022-05915-3> PMID: [35136120](https://pubmed.ncbi.nlm.nih.gov/35136120/)
11. Gupta H, Verma OP. Vaccine hesitancy in the post-vaccination COVID-19 era: a machine learning and statistical analysis driven study. *Evol Intell*. 2023;16(3):739–57. <https://doi.org/10.1007/s12065-022-00704-3> PMID: [35281291](https://pubmed.ncbi.nlm.nih.gov/35281291/)
12. Bass SB, Kelly PJ, Hoadley A, Arroyo Lloret A, Organtini T. Mapping Perceptual Differences to Understand COVID-19 Beliefs in Those with Vaccine Hesitancy. *J Health Commun*. 2022;27(1):49–61. <https://doi.org/10.1080/10810730.2022.2042627> PMID: [35199628](https://pubmed.ncbi.nlm.nih.gov/35199628/)
13. Walsh JC, Comar M, Folan J, Williams S, Kola-Palmer S. The psychological and behavioural correlates of COVID-19 vaccine hesitancy and resistance in Ireland and the UK. *Acta Psychol (Amst)*. 2022;225:103550. <https://doi.org/10.1016/j.actpsy.2022.103550> PMID: [35259642](https://pubmed.ncbi.nlm.nih.gov/35259642/)
14. Reno C, Maietti E, Di Valerio Z, Montalti M, Fantini MP, Gori D. Vaccine Hesitancy towards COVID-19 Vaccination: Investigating the Role of Information Sources through a Mediation Analysis. *Infect Dis Rep*. 2021;13(3):712–23. <https://doi.org/10.3390/idr13030066> PMID: [34449654](https://pubmed.ncbi.nlm.nih.gov/34449654/)
15. Lindvall J, Rönnerstrand B. Challenges for public-service delivery: the case of Covid-19 vaccine hesitancy. *J Eur Public Policy*. 2022;30(12):2601–22. <https://doi.org/10.1080/13501763.2022.2123024>
16. Buonomo B. Effects of information-dependent vaccination behavior on coronavirus outbreak: insights from a SIRI model. *Ricerche Mat*. 2020;69(2):483–99. <https://doi.org/10.1007/s11587-020-00506-8>
17. Oduro B, Miloua A, Apenteng OO, Osei PP. COVID-19 Vaccination Hesitancy Model: The Impact of Vaccine Education on Controlling the Outbreak. *Math Model Appl*. 2021;6(4):81. <https://doi.org/10.11648/j.mma.20210604.11>
18. Buonomo B, Della Marca R, d'Onofrio A, Groppi M. A behavioural modelling approach to assess the impact of COVID-19 vaccine hesitancy. *J Theor Biol*. 2022;534:110973. <https://doi.org/10.1016/j.jtbi.2021.110973> PMID: [34896166](https://pubmed.ncbi.nlm.nih.gov/34896166/)
19. Lau CL, Mayfield HJ, Sinclair JE, Brown SJ, Waller M, Enjeti AK, et al. Risk-benefit analysis of the AstraZeneca COVID-19 vaccine in Australia using a Bayesian network modelling framework. *Vaccine*. 2021;39(51):7429–40. <https://doi.org/10.1016/j.vaccine.2021.10.079> PMID: [34810000](https://pubmed.ncbi.nlm.nih.gov/34810000/)
20. Sinclair JE, Mayfield HJ, Short KR, Brown SJ, Puranik R, Mengersen K, et al. A Bayesian network analysis quantifying risks versus benefits of the Pfizer COVID-19 vaccine in Australia. *NPJ Vaccines*. 2022;7(1):93. <https://doi.org/10.1038/s41541-022-00517-6> PMID: [35953502](https://pubmed.ncbi.nlm.nih.gov/35953502/)
21. Warne DJ, Ebert A, Drovandi C, Hu W, Mira A, Mengersen K. Hindsight is 2020 vision: a characterisation of the global response to the COVID-19 pandemic. *BMC Public Health*. 2020;20(1):1868. <https://doi.org/10.1186/s12889-020-09972-z> PMID: [33287789](https://pubmed.ncbi.nlm.nih.gov/33287789/)
22. Le T-M, Raynal L, Talbot O, Hambridge H, Drovandi C, Mira A, et al. Framework for assessing and easing global COVID-19 travel restrictions. *Sci Rep*. 2022;12(1):6985. <https://doi.org/10.1038/s41598-022-10678-y> PMID: [35484268](https://pubmed.ncbi.nlm.nih.gov/35484268/)
23. Kretzschmar ME, Ashby B, Fearon E, Overton CE, Panovska-Griffiths J, Pellis L, et al. Challenges for modelling interventions for future pandemics. *Epidemics*. 2022;38:100546. <https://doi.org/10.1016/j.epidem.2022.100546> PMID: [35183834](https://pubmed.ncbi.nlm.nih.gov/35183834/)
24. Browning AP, Warne DJ, Burrage K, Baker RE, Simpson MJ. Identifiability analysis for stochastic differential equation models in systems biology. *J R Soc Interface*. 2020;17(173):20200652. <https://doi.org/10.1098/rsif.2020.0652> PMID: [33323054](https://pubmed.ncbi.nlm.nih.gov/33323054/)
25. Ma Q, Liu J, Liu Q, Kang L, Liu R, Jing W, et al. Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis. *JAMA Netw Open*. 2021;4(12):e2137257. <https://doi.org/10.1001/jamanetworkopen.2021.37257>
26. Subramanian R, He Q, Pascual M. Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *Proc Natl Acad Sci U S A*. 2021;118(9):e2019716118. <https://doi.org/10.1073/pnas.2019716118> PMID: [33571106](https://pubmed.ncbi.nlm.nih.gov/33571106/)
27. Long Q-X, Tang X-J, Shi Q-L, Li Q, Deng H-J, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med*. 2020;26(8):1200–4. <https://doi.org/10.1038/s41591-020-0965-6> PMID: [32555424](https://pubmed.ncbi.nlm.nih.gov/32555424/)
28. Rosenberg ES, Holtgrave DR, Dorabawila V, Conroy M, Greene D, Lutterloh E, et al. New COVID-19 Cases and Hospitalizations Among Adults, by Vaccination Status — New York, May 3–July 25, 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70(37):1306–11. <https://doi.org/10.15585/mmwr.mm7037a7>
29. Johns Hopkins University. Coronavirus resource centre. 2020. Accessed 5 December 2022. Available from: <https://coronavirus.jhu.edu/>
30. Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus Pandemic (COVID-19). 2020. Accessed 5 December 2022. Available from: <https://ourworldindata.org/coronavirus>
31. Bernal JL, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N Engl J Med*. 2021;385(7):585–94. <https://doi.org/10.1056/nejmoa2108891>
32. Lopez Bernal J, Andrews N, Gower C, Robertson C, Stowe J, Tessier E, et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ*. 2021;n1088. <https://doi.org/10.1136/bmj.n1088>
33. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet*. 2021;397(10269):99–111. [https://doi.org/10.1016/S0140-6736\(20\)32661-1](https://doi.org/10.1016/S0140-6736(20)32661-1) PMID: [33306989](https://pubmed.ncbi.nlm.nih.gov/33306989/)

34. Bon JJ, Warne DJ, Nott DJ, Drovandi C. Bayesian score calibration for approximate models. *J Mach Learn Res.* 2025;26(301):1–40. <http://jmlr.org/papers/v26/24-1179.html>
35. Warne DJ, Baker RE, Simpson MJ. Rapid Bayesian Inference for Expensive Stochastic Models. *J Comput Graph Stat.* 2021;31(2):512–28. <https://doi.org/10.1080/10618600.2021.2000419>
36. Collinson S, Heffernan JM. Modelling the effects of media during an influenza epidemic. *BMC Public Health.* 2014;14:376. <https://doi.org/10.1186/1471-2458-14-376> PMID: 24742139
37. Teng TRY, Lara-Tuprio EPD, Macalalag JMR. An HIV/AIDS epidemic model with media coverage, vertical transmission and time delays. In: AIP Conference Proceedings. AIP Publishing; 2019.
38. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigran A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur J Epidemiol.* 2020;35(8):775–9. <https://doi.org/10.1007/s10654-020-00671-y> PMID: 32785815
39. Sisson SA, Fan Y, Beaumont MA. *Handbook of Approximate Bayesian Computation.* Chapman and Hall/CRC; 2018.
40. Sunnåker M, Busetto AG, Numminen E, Corander J, Foll M, Dessimoz C. Approximate Bayesian computation. *PLoS Comput Biol.* 2013;9(1):e1002803. <https://doi.org/10.1371/journal.pcbi.1002803> PMID: 23341757
41. Drovandi CC, Pettitt AN. Estimation of parameters for macroparasite population evolution using approximate bayesian computation. *Biometrics.* 2010;67(1):225–33. <https://doi.org/10.1111/j.1541-0420.2010.01410.x> PMID: 20345496
42. Bellu G, Saccomani MP, Audoly S, D'Angiò L. DAISY: a new software tool to test global identifiability of biological and physiological systems. *Comput Methods Programs Biomed.* 2007;88(1):52–61. <https://doi.org/10.1016/j.cmpb.2007.07.002> PMID: 17707944
43. Chiş O-T, Banga JR, Balsa-Canto E. Structural identifiability of systems biology models: a critical comparison of methods. *PLoS One.* 2011;6(11):e27755. <https://doi.org/10.1371/journal.pone.0027755> PMID: 22132135
44. Chiş O, Banga JR, Balsa-Canto E. GenSSI: a software toolbox for structural identifiability analysis of biological models. *Bioinformatics.* 2011;27(18):2610–1. <https://doi.org/10.1093/bioinformatics/btr431> PMID: 21784792
45. Ligon TS, Fröhlich F, Chiş OT, Banga JR, Balsa-Canto E, Hasenauer J. GenSSI 2.0: multi-experiment structural identifiability analysis of SBML models. *Bioinformatics.* 2017;34(8):1421–3. <https://doi.org/10.1093/bioinformatics/btx735>
46. Pieschner S, Hasenauer J, Fuchs C. Identifiability analysis for models of the translation kinetics after mRNA transfection. *J Math Biol.* 2022;84(7):56. <https://doi.org/10.1007/s00285-022-01739-x> PMID: 35577967
47. Pawitan Y. *In all likelihood: statistical modelling and inference using likelihood.* Oxford University Press; 2001.
48. Murphy RJ, Maclaren OJ, Calabrese AR, Thomas PB, Warne DJ, Williams ED, et al. Computationally efficient framework for diagnosing, understanding and predicting biphasic population growth. *J R Soc Interface.* 2022;19(197):20220560. <https://doi.org/10.1098/rsif.2022.0560> PMID: 36475389
49. Simpson MJ, Browning AP, Warne DJ, Maclaren OJ, Baker RE. Parameter identifiability and model selection for sigmoid population growth models. *J Theor Biol.* 2022;535:110998. <https://doi.org/10.1016/j.jtbi.2021.110998> PMID: 34973274
50. Warne DJ, Maclaren OJ, Carr EJ, Simpson MJ, Drovandi C. Generalised likelihood profiles for models with intractable likelihoods. *Stat Comput.* 2023;34(1). <https://doi.org/10.1007/s11222-023-10361-w>
51. Hines KE, Middendorf TR, Aldrich RW. Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach. *J Gen Physiol.* 2014;143(3):401–16. <https://doi.org/10.1085/jgp.201311116> PMID: 24516188
52. Leask J, Carlson SJ, Attwell K, Clark KK, Kaufman J, Hughes C, et al. Communicating with patients and the public about COVID-19 vaccine safety: recommendations from the Collaboration on Social Science and Immunisation. *Med J Aust.* 2021;215(1):9–12.e1. <https://doi.org/10.5694/mja2.51136> PMID: 34137034
53. Elsaid M, Nasef MA, Huy NT. R0 of COVID-19 and its impact on vaccination coverage: compared with previous outbreaks. *Hum Vaccin Immunother.* 2021;17(11):3850–4. <https://doi.org/10.1080/21645515.2020.1865046> PMID: 34612165
54. Kayser V, Ramzan I. Vaccines and vaccination: history and emerging issues. *Hum Vaccin Immunother.* 2021;17(12):5255–68. <https://doi.org/10.1080/21645515.2021.1977057> PMID: 34582315
55. Zhong L, Diagne M, Wang Q, Gao J. Vaccination and three non-pharmaceutical interventions determine the dynamics of COVID-19 in the US. *Humanit Soc Sci Commun.* 2022;9(1). <https://doi.org/10.1057/s41599-022-01142-3>
56. Frazier DT, Robert CP, Rousseau J. Model Misspecification in Approximate Bayesian Computation: Consequences and Diagnostics. *J R Stat Soc Ser B Stat Methodol.* 2020;82(2):421–44. <https://doi.org/10.1111/rssb.12356>
57. Schälte Y, Hasenauer J. Efficient exact inference for dynamical systems with noisy measurements using sequential approximate Bayesian computation. *Bioinformatics.* 2020;36(Suppl_1):i551–9. <https://doi.org/10.1093/bioinformatics/btaa397> PMID: 32657404
58. Wilkinson RD. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Stat Appl Genet Mol Biol.* 2013;12(2):129–41. <https://doi.org/10.1515/sagmb-2013-0010> PMID: 23652634
59. Dutta I, Basu T, Das A. Spatial analysis of COVID-19 incidence and its determinants using spatial modeling: A study on India. *Environ Chall (Amst).* 2021;4:100096. <https://doi.org/10.1016/j.envc.2021.100096> PMID: 38620946
60. Blyuss KB, Kyrychko YN. Effects of latency and age structure on the dynamics and containment of COVID-19. *J Theor Biol.* 2021;513:110587. <https://doi.org/10.1016/j.jtbi.2021.110587> PMID: 33450286

61. Merow C, Urban MC. Seasonality and uncertainty in global COVID-19 growth rates. *Proc Natl Acad Sci U S A*. 2020;117(44):27456–64. <https://doi.org/10.1073/pnas.2008590117> PMID: 33051302
62. Menni C, May A, Polidori L, Louca P, Wolf J, Capdevila J, et al. COVID-19 vaccine waning and effectiveness and side-effects of boosters: a prospective community study from the ZOE COVID Study. *Lancet Infect Dis*. 2022;22(7):1002–10. [https://doi.org/10.1016/S1473-3099\(22\)00146-3](https://doi.org/10.1016/S1473-3099(22)00146-3) PMID: 35405090
63. Alahakoon P, McCaw JM, Taylor PG. Estimation of the probability of epidemic fade-out from multiple outbreak data. *Epidemics*. 2022;38:100539. <https://doi.org/10.1016/j.epidem.2022.100539> PMID: 35093850
64. Everitt RG, Rowińska PA. Delayed Acceptance ABC-SMC. *J Comput Graph Stat*. 2020;30(1):55–66. <https://doi.org/10.1080/10618600.2020.1775617>
65. Prangle D. Lazy ABC. *Stat Comput*. 2014;26(1–2):171–85. <https://doi.org/10.1007/s11222-014-9544-3>
66. Prescott TP, Baker RE. Multifidelity Approximate Bayesian Computation. *SIAM/ASA J Uncertain Quantif*. 2020;8(1):114–38. <https://doi.org/10.1137/18m1229742>
67. Prescott TP, Baker RE. Multifidelity Approximate Bayesian Computation with Sequential Monte Carlo Parameter Sampling. *SIAM/ASA J Uncertain Quantif*. 2021;9(2):788–817. <https://doi.org/10.1137/20m1316160>
68. Jasra A, Jo S, Nott D, Shoemaker C, Tempone R. Multilevel Monte Carlo in approximate Bayesian computation. *Stoch Anal Appl*. 2019;37(3):346–60. <https://doi.org/10.1080/07362994.2019.1566006>
69. Warne DJ, Sisson SA, Drovandi C. Vector Operations for Accelerating Expensive Bayesian Computations – A Tutorial Guide. *Bayesian Anal*. 2022;17(2). <https://doi.org/10.1214/21-ba1265>
70. Warne DJ, Baker RE, Simpson MJ. Multilevel rejection sampling for approximate Bayesian computation. *Comput Stat Data Anal*. 2018;124:71–86. <https://doi.org/10.1016/j.csda.2018.02.009>
71. Hurn S, Lindsay K, Warne DJ. A heterogeneous computing approach to maximum likelihood parameter estimation for the Heston stochastic volatility model. *ANZIAM J*. 2016;57:364. <https://doi.org/10.21914/anziamj.v57i0.10425>
72. Kulkarni S, Tsyplikhin A, Krell MM, Moritz CA. Accelerating Simulation-based Inference with Emerging AI Hardware. In: 2020 International Conference on Rebooting Computing (ICRC). IEEE; 2020.
73. Kulkarni S, Krell MM, Nabarro S, Moritz CA. Hardware-accelerated Simulation-based Inference of Stochastic Epidemiology Models for COVID-19. *J Emerg Technol Comput Syst*. 2022;18(2):1–24. <https://doi.org/10.1145/3471188>
74. Lee A, Yau C, Giles MB, Doucet A, Holmes CC. On the utility of graphics cards to perform massively parallel simulation of advanced Monte Carlo methods. *J Comput Graph Stat*. 2010;19(4):769–89. <https://doi.org/10.1198/jcgs.2010.10039> PMID: 22003276
75. Mahani AS, Sharabiani MTA. SIMD parallel MCMC sampling with applications for big-data Bayesian analytics. *Comput Stat Data Anal*. 2015;88:75–99. <https://doi.org/10.1016/j.csda.2015.02.010>
76. Lueckmann JM, Gonçalves PJ, Bassetto G, Öcal K, Nonnenmacher M, Macke JH. Flexible statistical inference for mechanistic models of neural dynamics. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. NIPS'17. Red Hook (NY): Curran Associates Inc.; 2017. p. 1289–1299.
77. Papamakarios G, Sterratt D, Murray I. Sequential Neural Likelihood: Fast Likelihood-free Inference with Autoregressive Flows. In: Chaudhuri K, Sugiyama M, editors. Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics. vol. 89 of Proceedings of Machine Learning Research. PMLR; 2019. p. 837–48.
78. Cranmer K, Brehmer J, Louppe G. The frontier of simulation-based inference. *Proc Natl Acad Sci U S A*. 2020;117(48):30055–62. <https://doi.org/10.1073/pnas.1912789117> PMID: 32471948