



A target-cell limited model can reproduce influenza infection dynamics in hosts with differing immune responses

Rahil Sachak-Patwa ^{a,*}, Erin I. Lafferty ^b, Claude J. Schmit ^b, Robin N. Thompson ^{c,d}, Helen M. Byrne ^a

^a Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

^b Biosensors Beyond Borders Limited, 9 Bedford Square, London, WC1B 3RE, UK

^c Mathematics Institute, University of Warwick, Zeeman Building, Coventry, CV4 7AL, UK

^d Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry, CV4 7AL, UK

ARTICLE INFO

Dataset link: <https://github.com/rahilsp/withi-n-host-influenza>

Keywords:

Mathematical modelling
Within-host kinetics
Infectious disease
Viral dynamics
Theoretical immunology

ABSTRACT

We consider a hierarchy of ordinary differential equation models that describe the within-host viral kinetics of influenza infections: the IR model explicitly accounts for an immune response to the virus, while the simpler, target-cell limited TEIV and TV models do not. We show that when the IR model is fitted to pooled experimental murine data of the viral load, fraction of dead cells, and immune response levels, its parameters values can be determined. However, if, as is common, only viral load data are available, we can estimate parameters of the TEIV and TV models but not the IR model. These results are substantiated by a structural and practical identifiability analysis. We then use the IR model to generate synthetic data representing infections in hosts whose immune responses differ. We fit the TV model to these synthetic datasets and show that it can reproduce the characteristic exponential increase and decay of viral load generated by the IR model. Furthermore, the values of the fitted parameters of the TV model can be mapped from the immune response parameters in the IR model. We conclude that, if only viral load data are available, a simple target-cell limited model can reproduce influenza infection dynamics and distinguish between hosts with differing immune responses.

1. Introduction

Influenza is a short-lived infection: its viral kinetics in humans are characterised by rapid exponential growth over one to three days, to a peak viral load, followed by a decline over the subsequent week (Carrat et al., 2008). In uncomplicated cases, the virus infects epithelial cells in the respiratory tract of the host (Nicholson et al., 2003). Symptoms include high fevers, coughs, headaches, muscle and joint pains, sore throats, and runny noses. The majority of individuals recover within one week (World Health Organization, 2020), however in high-risk groups infection and symptoms may be prolonged by several days or even weeks (van der Vries et al., 2013). There are four types of influenza viruses (A, B, C, and D), with influenza A viruses being responsible for most human infections and severe illness (Centers for Disease Control and Prevention (CDC), 2020). Seasonal strains are estimated to cause approximately three to five million severe cases and result in 290,000 to 650,000 respiratory deaths annually worldwide. The majority of these deaths are in high-risk groups, such as young children, the elderly, and individuals with chronic health conditions (World Health Organization, 2020). These strains cause epidemics during the

winter months in temperate regions in the northern and southern hemispheres, while epidemics may occur at any time of the year near the equator. Considerably more lethal strains emerge sporadically, giving rise to global pandemics in 1918, 1957, and 1968 (Kilbourne, 2006).

The immune response to viral infections can be divided into the innate and adaptive responses (Sompayrac, 2019). The innate immune response is non-specific, limiting pathogen replication and spread in the body and priming the adaptive immune response. The secretion of proinflammatory cytokines, particularly type I interferon, induces an antiviral state in susceptible cells and further activates signalling pathways and cellular responses. Inflammatory mediators in the innate immune response cause systemic symptoms which may persist even after the virus has been primarily cleared and viral load levels are low (de Jong et al., 2006). The innate response also includes immune cells such as Natural Killer cells (Iwasaki and Pillai, 2014).

The adaptive immune response is specific to the virus (or, more generally, pathogen) causing the infection and can be characterised as either cellular or humoral. In humans, it takes four to seven days for the adaptive immune response to take effect following infection (Janeway

* Corresponding author.

E-mail address: rahil.sachak-patwa@maths.ox.ac.uk (R. Sachak-Patwa).

<https://doi.org/10.1016/j.jtbi.2023.111491>

Received 16 August 2022; Received in revised form 2 March 2023; Accepted 5 April 2023

Available online 11 April 2023

0022-5193/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

et al., 2001). Cellular adaptive immunity is mediated by T-cells: CD4⁺ T-cells ‘help’ other immune cells by releasing cytokines, and cytotoxic T-CD8⁺ cells clear infected cells via lysis, induce apoptosis, and produce pro-inflammatory cytokines (Grant et al., 2014). The humoral adaptive immune response controls infection via antibodies secreted by B-cells, which recognise and bind to specific epitopes on the virus. These antibodies (the most prominent being IgA, IgG, and IgM) bind to the hemagglutinin (HA) protein of the influenza virus, thereby inhibiting the virion’s attachment and entry into a target-cell, and effectively neutralising the virion (Nabel and Fauci, 2010). The immune response upon infection with influenza A is influenced by multiple factors including age (Haq and McElhaney, 2014; Lambert et al., 2012), and comorbidities including obesity (Honce and Schultz-Cherry, 2019) and heart disease (Hui et al., 2006).

Mathematical models have been used to investigate the mechanisms underlying influenza infection. They can be classified into two main types. First, target-cell limited models, in which the infection is limited by the availability of susceptible cells (Baccam et al., 2006; Dobrovolny et al., 2010; Hadjichrysanthou et al., 2016; Handel et al., 2007). Second, models can explicitly account for the immune response at varying levels of detail (Canini and Carrat, 2010; Cao and McCaw, 2017; Hernandez-Vargas et al., 2014; Lee et al., 2009; Saenz et al., 2010). Some models focus on just one aspect of the immune response (for instance, type I interferons Baccam et al., 2006; Saenz et al., 2010), while others detail many of the mechanisms involved in the immune response (Lee et al., 2009). Due to the availability of data, the quantity of interest is typically viral load, although symptom dynamics have also been considered (Canini and Carrat, 2010; Canini et al., 2014, 2016; Lukens et al., 2014; Manchanda et al., 2014). Mathematical modelling has also been used to assess the efficacy of antiviral treatments such as neuraminidase inhibitors and adamantanes (Canini et al., 2014; Handel et al., 2007; Kamal et al., 2015; Lee et al., 2009; Montaseri et al., 2018). Other models describe co-infection with influenza and another pathogen which may be another virus (Pinky and Dobrovolny, 2016; Pinky et al., 2019), or bacteria such as *Streptococcus pneumoniae* (Cheng et al., 2017; Smith, 2018a; Smith et al., 2013). These models usually comprise ordinary differential equations, although stochastic Markov chain models have also been considered (Bai et al., 2018; Handel et al., 2007; Pinky et al., 2019).

Model selection studies have been performed to compare models based on different immunological assumptions. For example, comparisons of target-cell limited and immune response models, similar to those explored in this study, have been made in the context of infection dynamics in mice of different ages (Hernandez-Vargas et al., 2014), drug treatment (Handel et al., 2007), and infection in immunocompromised and immunocompetent patients (Cao and McCaw, 2017). For a comprehensive review of within-host influenza models, see Beauchemin and Handel (2011), Handel et al. (2018), and Smith (2018a,b).

When constructing models to describe the within-host dynamics of a viral infection, a compromise has to be made between model complexity and parameter identifiability (i.e. the ability to determine parameter values from the available data). In many studies, point-estimates for model parameters are obtained (Baccam et al., 2006; Handel et al., 2009; Saenz et al., 2010), but it is not always clear whether these parameter values can be uniquely determined or whether other parameter values would yield equally good fits. Parameter identifiability can be an issue especially for complex models, as often, only viral load data are available (Miao et al., 2011; Nguyen et al., 2016). As we show in this study, when additional data on the immune response are available (for example, antibody or cytokine levels), values of model parameters may be identified with tighter credible intervals than when only viral load data are available (Dobrovolny et al., 2013).

In this study, we address the question of what type of model, with what level of epidemiological detail, should be used to model within-host influenza dynamics when fitting to host data. We investigate whether a target-cell limited model can reproduce and differentiate

between the viral load infection dynamics of different host groups. To address these questions, we consider a hierarchy of models which describe the kinetics of influenza infection: the IR (immune response) model is adapted from Handel et al. (2009), while the target-cell limited TEIV (target-cell, eclipse phase, infected phase, virus) model can be derived from the IR model by neglecting the immune response terms, and the TV (target-cell, virus) model comprises only one class of cells. We fit the IR model to time-series data from murine experiments. We pool data of the viral load, the fraction of dead cells, and the adaptive and innate immune responses from immunocompetent and immunosuppressed mice (they lack an adaptive immune response) (Iwasaki and Nozima, 1977), and show that the IR model can be parameterised using these data. However, commonly only data pertaining to the viral load are available, and when the IR model is fit to viral load data from the immunocompetent mice alone its parameters cannot be identified. This motivates us to consider using the simpler TEIV and TV models to describe viral load kinetics. To investigate this, we perform a structural identifiability analysis to determine whether the parameters of each model can be uniquely identified, and then a practical identifiability analysis to determine parameter identifiability given noise-free synthetic data. We fit the TV and TEIV models to the experimental murine viral load data, and then to synthetically generated viral load data, and show that, in many cases, both models can accurately reproduce the viral load data. The parameters of the TV model are also shown to depend on the immunological characteristics of the patient groups used to generate the data. Finally, by comparing the TV and IR models, we show that the TV model may be preferred when choosing a model to make forward projections of the viral load during an infection as it has fewer parameters and, hence, its parameter values can be reliably estimated, which is not the case for the IR model.

2. Methods

2.1. Models

In this study, we consider three models that describe the within-host dynamics of influenza infection. These models are termed the IR, TEIV, and TV models. Schematics of each model are shown in Fig. 1.

2.1.1. IR model

The immune response (IR) model comprises six ordinary differential equations (ODEs) and is adapted from Handel et al. (2009). Uninfected epithelial cells, T , that are infected by virions V at rate β , become latently infected cells, E . After a period of $1/\kappa$ days, latently infected cells become productively infected, I . Productively infected cells produce virions, at a maximum rate p (if the innate immune response is inactive). The lifespan of virions is $1/c$ days, while the lifespan of productively infected cells is $1/\delta$ days. Since the time-scale of the infection (one to two weeks) is much shorter than the average lifespan of uninfected epithelial cells (on the order of one year) (Rawlins and Hogan, 2008), we neglect their natural death and regrowth in our model.

Following infection, the innate immune response is activated when the virus is identified by immune cells (Takeuchi and Akira, 2009). We consider the role of interferons, F , which inhibit viral replication (Parkin and Cohen, 2001) and are secreted by infected cells. Rather than explicitly modelling interferon dynamics with a differential equation, we follow Handel et al. (2009) and Hernandez-Vargas et al. (2014), and account for their action via a prescribed function $F(t)$ which increases exponentially for $0 < t < t^*$ and decays exponentially for $t > t^*$ (see Eq. (7)). In Eq. (5), we assume that interferons reduce the rate at which productively infected epithelial cells produce virus particles, and introduce the parameter s to indicate the strength of this inhibition. As in Baccam et al. (2006) and Handel et al. (2009), a time delay τ is included to account for the time between interferon production and action. In practice, interferons initiate a signalling cascade that

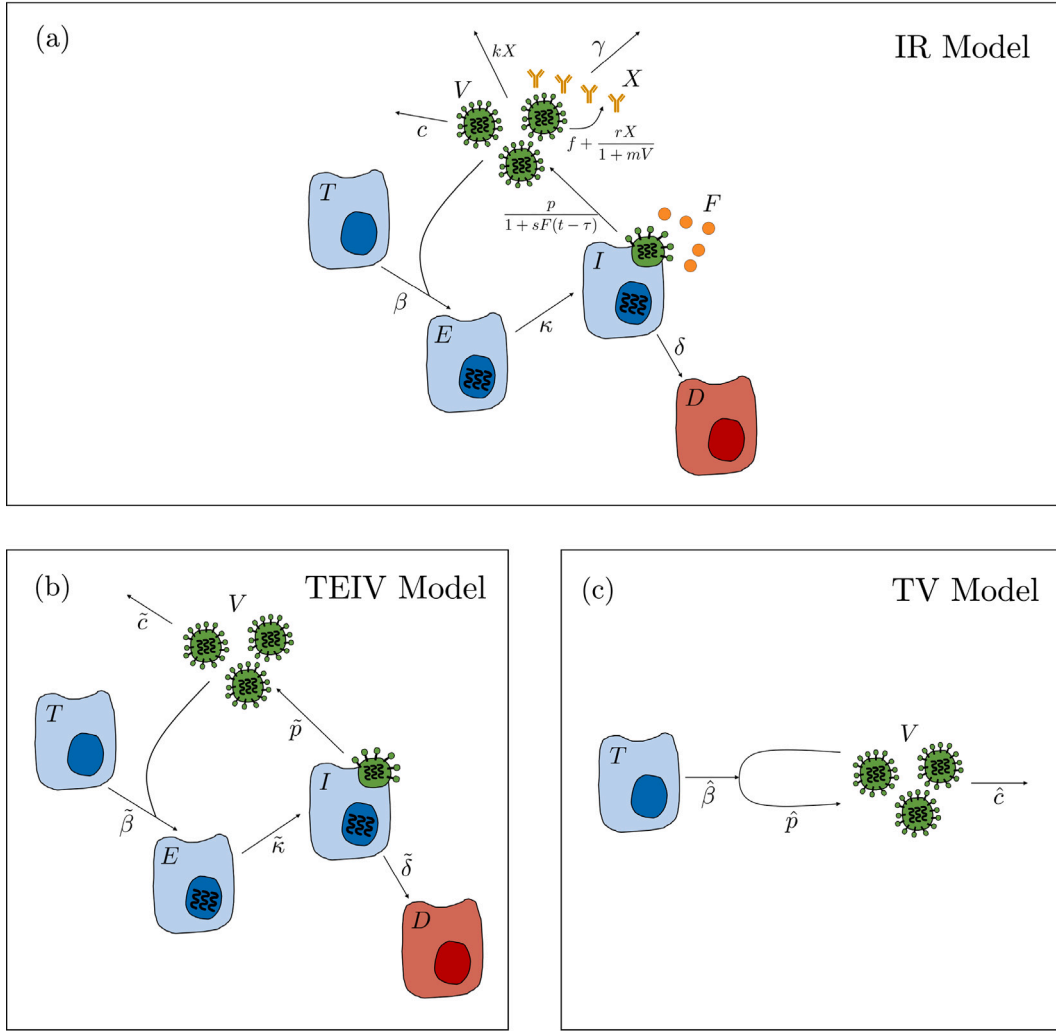


Fig. 1. Schematics of the within-host models of influenza infection. (a) IR model: Uninfected epithelial cells, T , are infected by free virions, V , at rate β , and then become latently infected, E , until they become productively infected, I , at rate κ . Productively infected cells die (becoming dead cells, D) at rate δ . They also produce virions at a maximum rate p . Following infection, the innate immune response, F , becomes active after a time delay τ and reduces virion production with a strength parameter s . The strength of the adaptive immune response, X , increases at a rate which depends on the numbers of free virions and productively infected cells, governed by parameters f , r , and m . It decays at rate γ . Free virions are cleared by the adaptive immune response at rate k and have a decay rate c . (b) TEIV model: Analogous to the IR model, but without the explicit inclusion of the immune response. It is obtained by neglecting the adaptive and immune response variables in the IR model ($F = X = 0$). (c) TV model: The simplest model of within-host infection dynamics in which free virions infect epithelial cells at rate $\hat{\beta}$ and \hat{p} new virions are produced per infected cell.

eventually induces an anti-viral state in cells, and thus a time delay between their production and action may be expected (García-Sastre, 2001).

The adaptive immune response is specific to the virus and takes longer to act than the innate immune response. Following Handel et al. (2009), we assume that the adaptive response, X , is activated at a rate proportional to the viral load V (rate constant f), and undergoes clonal expansion at rate r/m when viral densities are high, and at rate r when viral densities are low. In most cases, $V \gg 1$, and so the ratio r/m governs the exponential growth rate of the adaptive response. It has been shown that the main role of anti-influenza antibodies is to neutralise the virus by removing free virions from the extracellular space (Outlaw and Dimmock, 1991). Accordingly, we assume free virions are neutralised by antibodies at rate k . We further assume that antibodies have a natural ‘lifespan’ of $1/\gamma$ days. While Handel et al. (2009) assume that the reduction in viral load due to target-cell infection is significant, we follow Baccam et al. (2006), assuming, for simplicity, that virus uptake by target-cells can be neglected.

Combining the above assumptions, we arrive at the following system of equations which describe the time evolution of the infection:

$$\text{uninfected cells} \quad \frac{dT}{dt} = -\beta TV, \quad (1)$$

$$\text{latently infected cells} \quad \frac{dE}{dt} = \beta TV - \kappa E, \quad (2)$$

$$\text{productively infected cells} \quad \frac{dI}{dt} = \kappa E - \delta I, \quad (3)$$

$$\text{dead cells} \quad \frac{dD}{dt} = \delta I, \quad (4)$$

$$\text{free virions} \quad \frac{dV}{dt} = \frac{pI}{1 + sF(t - \tau)} - cV - kVX, \quad (5)$$

$$\text{adaptive IR} \quad \frac{dX}{dt} = fV + \frac{rXV}{1 + mV} - \gamma X. \quad (6)$$

In Eq. (5), the innate immune response is represented by the function $F(t)$ where

$$\text{innate IR} \quad F(t) = \begin{cases} 0, & \text{for } t \leq 0, \\ A_1 e^{\lambda_1 t} & \text{for } 0 < t \leq t^*, \\ A_2 e^{-\lambda_2 t} & \text{for } t > t^*. \end{cases} \quad (7)$$

As we neglect growth and natural death of uninfected cells, we can show that, by adding Eqs. (1)–(4), the total number of cells, $T + E + I + D$, is constant. Further, Eq. (4) decouples from the system. Throughout this study, and as in Handel et al. (2009) where the same murine data were analysed, we prescribe the functional form of $F(t)$ and

fix the parameters A_1 , A_2 , λ_1 , λ_2 , and t^* , that govern its dynamics. We justify this assumption by noting that the interferon dynamics in immunocompromised and immunocompetent mice are almost identical in the data we consider (see Fig. 2 and Iwasaki and Nozima, 1977).

We close Eqs. (1)–(7) by imposing the following initial conditions: $T(0) = T_0$, $E(0) = I(0) = D(0) = 0$, $V(0) = V_0$, $X(0) = 0$, where T_0 is the initial number of uninfected cells and V_0 is the initial viral load. In what follows, it will be convenient to introduce $C = D/T_0$, the fraction of dead cells, a term we will use when fitting the model to experimental data. Throughout this study, the viral titer V is measured in plaque forming units (PFU) and C is dimensionless. We shall denote the dimensions of X , the titer of neutralising antibodies, by u_X , and the dimensions of F , the titer of interferon, by u_F (see Iwasaki and Nozima, 1977 for further details on these units).

2.1.2. TEIV model

The TEIV model assumes that the infection is limited by the availability of uninfected target epithelial cells (Baccam et al., 2006). We derive the TEIV model by setting $F = X = 0$ in Eqs. (1)–(7), giving

$$\text{uninfected cells} \quad \frac{dT}{dt} = -\tilde{\beta}TV, \quad (8)$$

$$\text{latently infected cells} \quad \frac{dE}{dt} = \tilde{\beta}TV - \tilde{\kappa}E, \quad (9)$$

$$\text{productively infected cells} \quad \frac{dI}{dt} = \tilde{\kappa}E - \tilde{\delta}I, \quad (10)$$

$$\text{dead cells} \quad \frac{dD}{dt} = \tilde{\delta}I, \quad (11)$$

$$\text{free virions} \quad \frac{dV}{dt} = \tilde{p}I - \tilde{c}V, \quad (12)$$

where $\tilde{\beta}$, $1/\tilde{\kappa}$, $\tilde{\delta}$, \tilde{p} , and \tilde{c} are the infection rate, duration of the latent eclipse phase, infected cell clearance rate, virion production rate, and virion clearance rate, respectively. As before, Eq. (11) decouples. We close Eqs. (8)–(12) by imposing the following initial conditions: $T(0) = \tilde{T}_0$, $E(0) = I(0) = D(0) = 0$, and $V(0) = \tilde{V}_0$, where \tilde{T}_0 and \tilde{V}_0 are the initial number of uninfected cells and the initial viral load respectively.

2.1.3. TV model

The TV model does not distinguish between different types of infected cells, and can be derived from the TEIV model by assuming that there is no eclipse phase and that the viral dynamics occur over a shorter timescale than the infected cell dynamics (Hadjichrysanthou et al., 2016). Free virions are assumed to infect epithelial cells at rate $\hat{\beta}$, and \hat{p} determines how many virions are produced by each infected cell. The model equations are given by

$$\text{uninfected cells} \quad \frac{dT}{dt} = -\hat{\beta}TV, \quad (13)$$

$$\text{free virions} \quad \frac{dV}{dt} = \hat{p}\hat{\beta}TV - \hat{c}V. \quad (14)$$

We note that although the three models contain similar functional forms, their interpretations differ. In particular, in the IR model, we fix δ and c , the inverse of the lifespan of infected cells and free virions, respectively. However, in the TEIV model, we view $\tilde{\delta}$ and \tilde{c} as fitting parameters. Similarly, in the TV model, \hat{c} is a fitting parameter. Descriptions and estimated values of the parameters for all three models are given in Table 1.

2.2. Experimental data

Following Handel et al. (2009), we use data from an experimental study by Iwasaki and Nozima (1977) in which naive C3H/He mice were infected with the H1N1 influenza strain A/PR8. The data are from the tracheobronchial washings of immunocompetent mice, and immunosuppressed mice with suppressed antibody production. The data were collected over 12 days and consist of viral load titer, measured in plaque forming units (PFU), the fraction of dead cells as measured by lung lesions, combined type I and II interferon (IFN) levels, and IgA,

IgM, and IgG antibody titers. The antibody data represent the combined levels of IgA and IgM (IgG is neglected since it was not found to influence the infection dynamics Handel et al., 2009; Iwasaki and Nozima, 1977). Further details of the experimental procedure are given in the original study (Iwasaki and Nozima, 1977). Since the original data are unavailable, they were extracted from Iwasaki and Nozima (1977) using the data extraction tool <https://automeris.io/WebPlotDigitizer/>.

2.3. Structural identifiability analysis

We perform a structural identifiability analysis to determine whether it is possible to uniquely determine the parameters of a particular model, given perfect data for one or more model outputs (Bellman and Åström, 1970). In general terms, we wish to establish the structural identifiability of a model with unknown parameters θ , observable quantities $y(t; \theta) = [y_1(t; \theta), \dots, y_N(t; \theta)]$ and known initial conditions $y(0) = [y_1(0), \dots, y_N(0)]$. We employ the Taylor series approach (Chis et al., 2011), which assumes that the observable functions $y_j(t; \theta)$ for $j = 1, \dots, N$ and their time derivatives are unique analytic functions of time. Hence, it is possible to represent the observable functions via Taylor series expansions in a neighbourhood of an initial state. Uniqueness of this representation in terms of the model parameters and initial conditions guarantees structural identifiability. The Taylor series expansion of the observable functions in a neighbourhood of $t = 0$, is given by

$$y_j(t, \theta) = y_j(0; \theta) + y_j^{(1)}(0; \theta)t + y_j^{(2)}(0; \theta)\frac{t^2}{2!} + \dots + y_j^{(i)}(0; \theta)\frac{t^i}{i!} + \dots, \quad (15)$$

for $j = 1, \dots, N$, where

$$y_j^{(i)}(0, \theta) = \frac{d^i}{dt^i} y_j(0; \theta). \quad (16)$$

Let us define

$$a_j^i(\theta) = \lim_{t \rightarrow 0} y_j^{(i)}(t; \theta). \quad (17)$$

Then a sufficient condition for global structural identifiability is

$$a_j^i(\theta) = a_j^i(\theta^*) \text{ for } i = 1, \dots, i_{\max} \text{ and } j = 1, \dots, N \implies \theta = \theta^*, \quad (18)$$

where i_{\max} is the smallest integer for which the parameters θ can be determined in terms of the Taylor series coefficients $a_j^i(\theta)$ (Chis et al., 2011). Hence, the problem reduces to establishing whether a particular system of algebraic equations for the parameters θ has a unique solution. A parameter is termed globally identifiable if a unique value of the parameter can be identified, and locally identifiable if the value of the parameter can be uniquely identified in a local region of parameter space. Otherwise, the parameter is deemed non-identifiable.

2.4. Practical identifiability analysis

Structural identifiability depends on model structure and is independent of experimental data. As such it is a prerequisite for practical identifiability (Raue et al., 2009). A parameter is deemed practically identifiable if it can be uniquely determined in a local region of parameter space by fitting the model to data (either real-world or synthetic). We use a profile likelihood-based approach to determine practical identifiability in which differences between the model output and data are calculated as parameters vary. This approach is widely used in biological models with large numbers of fitting parameters and does not impose any restrictions on the form of the model equations (Raue et al., 2009, 2014).

Using the likelihood function (S1) given in Supplementary Information Section S.3, we define $\chi^2(\theta)$ by

$$\chi^2(\theta) = J(\theta; y) = \sum_{i=1}^M \sum_{j=1}^{N_i} \left(\frac{-[z_i(t_j; \theta) - y_{ij}]^2}{2\sigma_i^2} \right), \quad (19)$$

where $z_i(t_j, \theta)$ represent model outputs generated with parameters θ , y_{ij} represent data, and σ_i are noise scaling parameters. The subscripts

Table 1

Descriptions and parameter values of the IR, TEIV, and TV models.

(a): IR model — Eqs. (1)–(7).			
Parameter	Description	Values (Units)	Source
β	Infection rate	Fitting parameter (PFU ⁻¹ days ⁻¹)	
p	Virion production rate	Fitting parameter (PFU days ⁻¹ cells ⁻¹)	
$1/\delta$	Lifespan of infected cells	2 days	Handel et al. (2009)
$1/c$	Lifespan of free virions	10 days	Handel et al. (2009)
$1/\kappa$	Duration of latent eclipse phase	4 days	Handel et al. (2009)
τ	Innate IR action time delay	Fitting parameter (days)	
s	Strength of innate IR	Fitting parameter (u_F^{-1})	
r	Expansion rate of adaptive IR	Fitting parameter (PFU ⁻¹ days ⁻¹)	
f	Recruitment rate of adaptive IR	Fitting parameter (u_X PFU ⁻¹ days ⁻¹)	
m	Adaptive IR growth parameter	1	
γ	Decay rate of adaptive IR	1.159 days ⁻¹	Dietzen (2018)
k	Kill rate of adaptive IR	Fitting parameter (u_X^{-1} days ⁻¹)	
A_1	Innate IR dynamics parameter	0.8236 u_F	Handel et al. (2009)
A_2	Innate IR dynamics parameter	2.128×10^6 u_F	Handel et al. (2009)
λ_1	Innate IR dynamics parameter	1.241 days ⁻¹	Handel et al. (2009)
λ_2	Innate IR dynamics parameter	1.712 days ⁻¹	Handel et al. (2009)
r^*	Innate IR dynamics parameter	5 days	Handel et al. (2009)
T_0	Initial no. of uninfected cells	7×10^9 cells	Handel et al. (2009)
V_0	Initial viral load	Fitting parameter (PFU)	
(b): TEIV model — Eqs. (8)–(12).			
Parameter	Description	Values (Units)	Source
$\hat{\beta}$	Infection rate	Fitting parameter (PFU ⁻¹ days ⁻¹)	
\hat{p}	Virion production rate	Fitting parameter (PFU days ⁻¹ cells ⁻¹)	
$\hat{\delta}$	Clearance rate of infected cells	Fitting parameter (days ⁻¹)	
\hat{c}	Clearance rate of virions	Fitting parameter (days ⁻¹)	
$1/\hat{\kappa}$	Duration of latent eclipse phase	4 days	Handel et al. (2009)
\hat{T}_0	Initial no. of uninfected cells	7×10^9 cells	Handel et al. (2009)
\hat{V}_0	Initial viral load	Fitting parameter (PFU)	
(c): TV model — Eqs. (13)–(14).			
Parameter	Description	Values (Units)	Source
$\hat{\beta}$	Infection rate	Fitting parameter (PFU ⁻¹ days ⁻¹)	
\hat{p}	Virion production coefficient	Fitting parameter (PFU cells ⁻¹)	
\hat{c}	Clearance rate of virions	Fitting parameter (days ⁻¹)	
\hat{T}_0	Initial no. of uninfected cells	7×10^9 cells	Handel et al. (2009)
\hat{V}_0	Initial viral load	Fitting parameter (PFU)	

$i = 1, \dots, M$ relate to the observables (e.g. viral titer, fraction of dead cells, etc.) and the subscripts $j = 1, \dots, N_i$ relate to the output time-points. The maximum likelihood estimated value of θ corresponds to

$$\hat{\theta} = \arg \min_{\theta \in \Theta} J(\theta; y). \quad (20)$$

Likelihood-based confidence intervals can be defined from a confidence region

$$\{\theta | \chi^2(\theta) - \chi^2(\hat{\theta}) < \Delta_\alpha\} \quad \text{with} \quad \Delta_\alpha = \chi^2(\alpha, df), \quad (21)$$

the boundaries of which represent confidence intervals (Meeker and Escobar, 1995), where Δ_α is the quantile of the χ^2 distribution with df degrees of freedom and confidence level α . As in Raue et al. (2009), we use $\Delta_{\alpha=0.68}$. We also calculate the profile likelihood (PL) χ_{PL}^2

$$\chi_{PL}^2(\theta_i) = \min_{\theta_{j \neq i}} [\chi^2(\theta_i)], \quad (22)$$

meaning the re-optimisation of $\chi^2(\theta_i)$ with respect to all other fitting parameters $\theta_{j \neq i}$ (i.e. for each value of θ_i , the value of $\chi^2(\theta_i)$ is minimised using $\theta_{j \neq i}$ as fitting parameters). Following Raue et al. (2009), a parameter estimate $\hat{\theta}_i$ is defined to be practically identifiable if the likelihood-based confidence region does not extend infinitely in increasing and/or decreasing directions of θ_i (i.e. $\chi_{PL}^2(\theta_i)$ crosses the threshold Δ_α for a desired confidence level, α , to the left and right of $\hat{\theta}_i$).

2.5. Parameter estimation

We use two methods to estimate model parameters: Markov chain Monte Carlo (MCMC), with the Metropolis–Hastings algorithm (Hastings, 1970) (Figs. 2, 3, and 7), and maximum likelihood estimation (MLE) (Figs. 3, 4, 5, and 6). For each MCMC simulation, we perform 5×10^5 sampling iterations, discarding the first 10^4 iterations as the ‘burn-in’ period and recording every 100 iterations thereafter to reduce autocorrelation. See Supplementary Information Section S.3 for further details on parameter estimation.

We use the Bayesian information criterion (BIC) (Schwarz, 1978) and the corrected Akaike Information Criterion (AIC) (Burnham and Anderson, 2004) to compare different model projections.

3. Results

3.1. Fitting the IR model to pooled experimental data

As in Handel et al. (2009), we fitted the IR model to pooled data of the viral load, fraction of dead cells, and adaptive immune response in immunosuppressed and immunocompetent mice (Fig. 2 and Table S1). We used MCMC parameter estimation and, therefore, determined the distributions of unknown parameter values rather than point-estimates. The same parameter distributions were assumed to describe the infection dynamics in immunocompetent and immunosuppressed mice, except that no adaptive immune response was elicited in the immunosuppressed mice (i.e. $f = r = 0$).

As shown in Fig. 2, the model captures the infection dynamics of both sets of mice. Initially the infection dynamics are similar for the

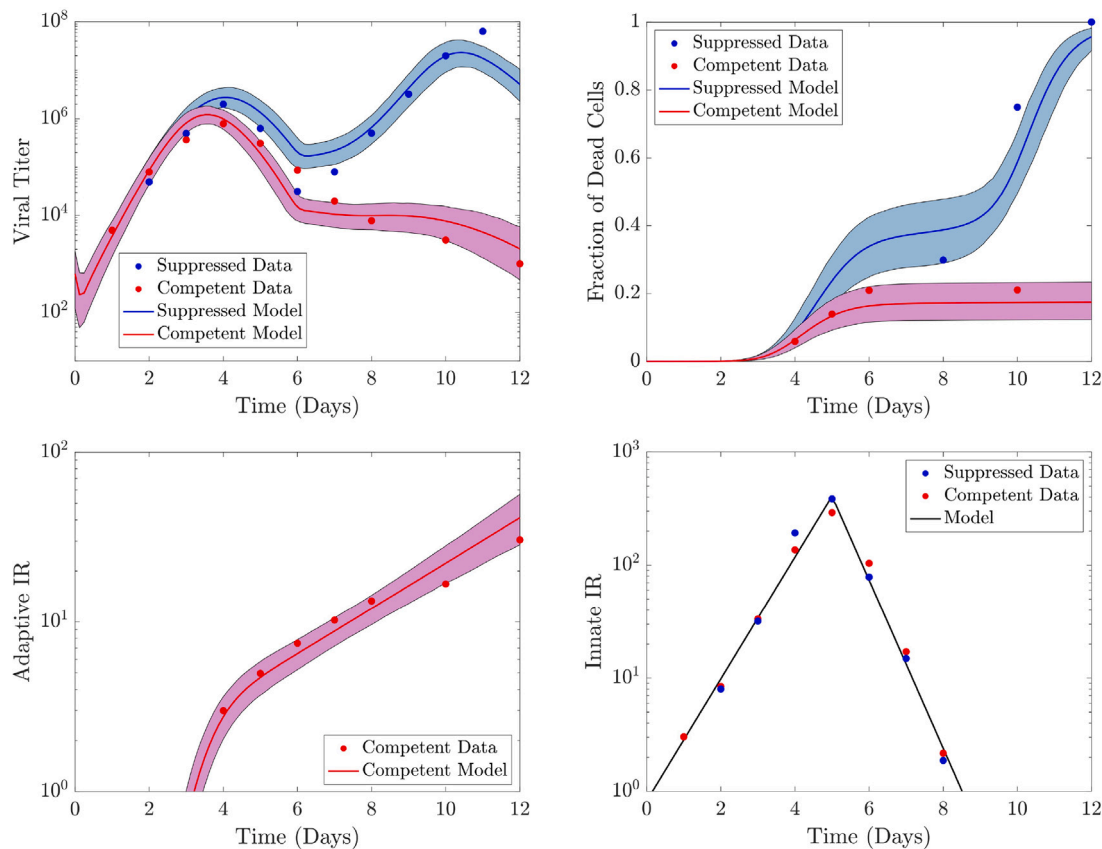


Fig. 2. The IR model fitted to pooled experimental data on the viral load, fraction of dead cells, and adaptive immune response in immunosuppressed and immunocompetent mice. Solid lines and shaded regions indicate the mean and 95% CrIs of model trajectories based on the posterior distributions of parameters. The innate immune response is assumed to be known precisely and is described by the same function as in [Handel et al. \(2009\)](#). Fixed and estimated parameter values are stated in [Tables 1](#) and [S1](#) respectively. Posterior distributions of fitted parameters given in [Figure S1](#).

two cases. However, after approximately 4 days, the immunocompetent mice elicit an adaptive immune response, which clears the virus and limits the number of cells which die. By contrast in the immunosuppressed mice, the viral load resurges after reaching an initial peak because the innate immune response decays after approximately 5 days and they do not elicit an adaptive immune response. Consequently, viral production increases, causing all remaining cells in the immunosuppressed mice to become infected and then to die. As shown in [Figure S1](#), the distributions of all fitting parameters are unimodal, with credible intervals spanning approximately one order of magnitude, except for V_0 which has a larger credible interval (estimated parameter values are given in [Table S1](#)). We used the IR model, with these parameter estimates, to generate synthetic data later in this study.

We investigated how the dynamics of the IR model change as parameters s , τ , k , and r , associated with the immune response, are varied. We considered three model outputs: (i) the area under the viral load curve (AUC), (ii) the duration of symptomatic infection, and (iii) the fraction of dead cells at the end of the infection (see [Supplementary Information Section S.4](#) for definitions). In general, as the innate and adaptive immune response parameters are increased, the infection outcome of the infection is improved as the AUC, the final fraction of dead cells, and the infection duration all decrease. We also performed a global parameter sensitivity analysis using the eFAST and Sobol methods (see [Supplementary Information Section S.5](#)). The parameters β , p , and T_0 were found to have the largest influence on the area under the viral load curve, the duration of symptomatic infection and the fraction of dead cells at the end of the infection. However, the immune response parameters s , τ , r , and k , which may differ between hosts based on their immunological characteristics, significantly influence the dynamics of the IR model; we explore this further in the following sections.

3.2. Fitting the models to experimental viral load data

In [Section 3.1](#) we showed that we can parameterise the IR model using pooled data. In practice, however, often, only viral load data are available. Therefore we then fitted the IR model to only data of the viral load kinetics in immunocompetent mice ([Fig. 3](#)). The shaded regions represent the 95% credible intervals (CrIs) of model trajectories based on the posterior distributions of fitting parameters calculated using MCMC. We also show the maximum likelihood model trajectory (dashed red line) in each case. For the IR model, the posterior distributions of the fitting parameters span many orders of magnitude ([Figure S2 \(a\)](#)), showing that they are not identifiable given these data ([Siekmann et al., 2012](#)).

As the IR model parameters are not identifiable, this motivates us to use the simpler TEIV and TV models to fit the data. When fitting these models to the same murine viral load data, we find that the models capture the exponential increase and decrease in viral load, with only small differences. [Figure S2 \(b\)](#) shows that the posterior distributions of the TEIV model parameters are unimodal, although they also span several orders of magnitude. By contrast, the TV model parameters have tighter credible intervals ([Figure S2 \(c\)](#)) and their distributions are bounded above and below. Estimated parameter values for each model are stated in [Table S2](#). As the BIC values of the TEIV and TV models are lower than that those of the IR model ([Table S3](#)), the TEIV and TV models should be preferred over the IR model when describing this particular viral load data. We also note that although the MLE fit of the IR model is more accurate than for the TEIV and TV models, as shown by the lower sum of squared residuals (SSR), the IR model's AIC value is greater because there are a larger number of fitting parameters ([Table S3](#)), again confirming that the TEIV and TV models should be preferred.

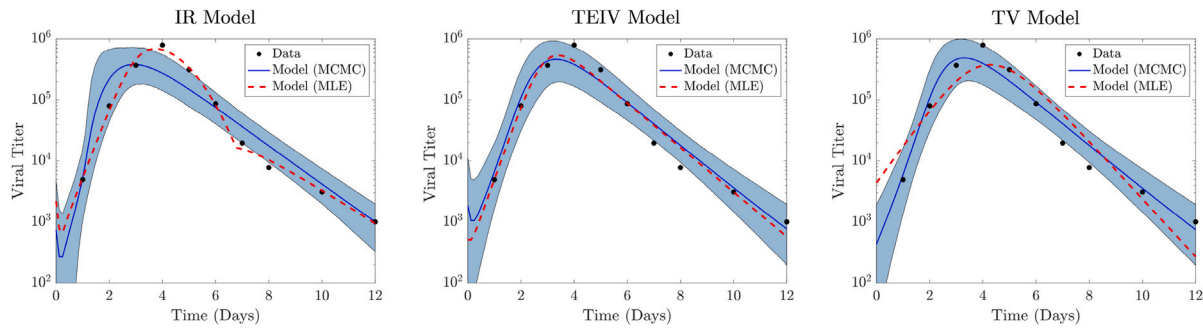


Fig. 3. The IR, TEIV, and TV models fitted to experimental data on the viral load of infected immunocompetent mice. Solid lines and shaded regions indicate the mean and 95% CrIs of model trajectories based on the posterior distributions of parameters. Red dashed lines indicate model trajectories relating to the maximum likelihood estimated parameter set. Fixed and estimated parameter values are stated in Tables 1 and S2 respectively. Posterior distributions of fitted parameters given in Figure S2.

Although, as stated earlier, the interpretation of the parameters in the IR and TEIV models differ, it is worth noting that if δ and c were used as fitting parameters in the IR model, it would be even more challenging to identify precise parameter estimates for that model compared to in this study.

3.3. Structural identifiability analysis

To further investigate the parameter identifiability of the IR, TEIV, and TV models given different data, we performed structural identifiability analyses; the results are summarised in Table 2. Through nondimensionalisation (see Supplementary Information Section S.6), we can show that, for each model, only the grouping pT_0 can be identified, as we do not observe data for the numbers of uninfected and/or infected cells (Baccam et al., 2006; Hadjichrysanthou et al., 2016; Handel et al., 2009; Saenz et al., 2010). Furthermore, if X is not observed in the IR model, then the grouping fk is identifiable but not each parameter f and k independently.

For the IR model, if V and X , or C and X , are observed (see Table 2, Cases 3 and 4), then the parameters β , p , τ , s , r , f , k , and V_0 are at least locally identifiable if there is no time delay ($\tau = 0$ in Eq. (5)). If C and V are observed until some time t where $t > t^* + \tau$, then all unknown parameters including τ are globally identifiable (Case 1a). Results proving identifiability of certain parameter groups if only one or two sets of data are observed are also presented. For Cases 2–7 we assume certain parameter values, in addition to θ , are known (e.g. we assume the value of k is known for Case 2) such that we could determine identifiability; with only θ known, it was computationally infeasible to prove structural identifiability of all other parameters due to the high order of derivatives needed for symbolic computation.

For the TEIV and TV models, all unknown parameters are identifiable (Cases 8 and 9), and can be estimated from viral load data. From this analysis, we deduce that it is, in theory, possible to determine values of model parameters given the available data from Iwasaki and Nozima (1977). See Supplementary Information Section S.6 for details of the structural identifiability calculations for the IR model, given data for V , C , and X .

3.4. Practical identifiability analysis

Structural identifiability does not guarantee practical identifiability. Indeed, even if a parameter is structurally identifiable, it may not be possible to determine its value by fitting the model to data. Profile likelihood curves for Cases 1, 8, and 9 are presented in Figure S5 (see Supplementary Information Section S.7 for Cases 2–7). Parameters are deemed to be practically identifiable if their profile likelihoods intersect the thresholds (black horizontal dotted lines) from above and below, creating a ‘U’ shape. When V , C , and X are observed then all unknown parameters in the IR model are identifiable (Case 1, Figure S5 (a)) However, if only one or two variables are observed, then most of the

parameters are not practically identifiable. We note further that if X is not observed, then the profile likelihood curves for f and k are identically zero for all values of f and k ; this result is consistent with our earlier observation that f and k are not structurally identifiable.

All four parameters of the TV model are practically identifiable when V is observed (Case 9, Figure S5 (c)). For the TEIV model, when V is observed, the parameters β , \tilde{c} , and \tilde{V}_0 are identifiable, while \tilde{p} and $\tilde{\delta}$ are not (Case 8, Figure S5 (b)). In what follows, as in Baccam et al. (2006), we continue to estimate all five parameters of the TEIV model for comparison with the TV model which has four unknown parameters.

The results of our structural and practical identifiability analyses substantiate the findings of fitting the models to the experimental data. We conclude that parameter values for the IR model can be identified if that model is fitted to pooled data including that of the immune response. However, if only viral load data are available, then the TEIV and TV models should be preferred (see Table 3).

3.5. Using the TEIV and TV models to reproduce viral load dynamics of the IR model

We now investigate whether the TEIV and TV models can reproduce the dynamics of the IR model and, in particular, the viral load dynamics. In practice, the immune response may vary between hosts. Therefore, we generated synthetic data from the IR model for different values of the parameters associated with the immune response (i.e. r , k , s , and τ in Eqs. (1)–(7)). To generate the data, one parameter was varied at a time while the others were held fixed. We view $\tilde{\beta}$, \tilde{p} , $\tilde{\delta}$, and \tilde{c} as fitting parameters for the TEIV model, and $\hat{\beta}$, \hat{p} , and \hat{c} as fitting parameters for the TV model (see Table 1). We fix $\tilde{V}_0 = \hat{V}_0 = 100$ PFU (a relatively small starting value) as using the initial viral load as an extra fitting parameter did not improve model fits (results not shown).

For each parameter in the IR model that we vary, there is a region of parameter space in which the TEIV and TV models accurately replicate the viral load dynamics (with a sum of squared residuals (SSR) of less than 15). In each case, the best-fit values of the parameters for the TEIV and TV models depend in a predictable way on the corresponding parameter of the IR model. With only three fitting parameters, these variations are clearer in the TV model. For example, as the value of s , the strength of the innate immune response, is increased, the best-fit values of $\hat{\beta}$ in the TV model increase, and the values of \hat{p} and \hat{c} decrease (Fig. 5(h)). For illustrative purposes, we also show the synthetic data and best-fit models (Figs. 4 and 5 right columns) for parameter values that are 0.1, 0.5, 2, and 10 times the values given in Table S1. In general, the TEIV and TV models can accurately replicate dynamics for which the viral load increases exponentially to a peak value and then decreases exponentially (see, for example, the bottom left panels in Figs. 4(f) and 5(f)). However, neither model can reproduce dynamics for which the viral load exhibits a resurgence and second peak (see, for example, the top left panels in Figs. 4(l) and 5(l)).

Table 2

Results of the structural identifiability analysis of the IR, TEIV, and TV models given different observable data. For the IR model (Case 1a) V^* and C^* refer to data which are observed until some time $t > t^* + \tau$. The set of parameters that is assumed to be known for the IR model is $\Theta = [A_1, A_2, \lambda_1, \lambda_2, t^*, \kappa, \delta, c, m, \gamma, T_0]$.

Model	Case	Observable data	Globally identifiable	Locally identifiable	Known
IR	1a	V^*, C^*, X	$\beta, p, \tau, s, r, f, k, V_0$		Θ
	1b	V, C, X	$\beta, p, s, r, f, k, V_0$		$\Theta, \tau = 0$
	2	V, C	β, f, V_0	p, s, r	$\Theta, \tau = 0, k$
	3	V, X	r, f, k, V_0	β, p, s	$\Theta, \tau = 0$
	4	C, X		$\beta, p, s, r, f, k, V_0$	$\Theta, \tau = 0$
	5	V	V_0, f	β, p, s	$\Theta, \tau = 0, k, r$
	6	C	β, p, f, s		$\Theta, \tau = 0, k, r, V_0$
	7	X		$\beta, p, s, r, f, k, V_0$	$\Theta, \tau = 0$
TEIV	8	V	\tilde{V}_0, \tilde{c}	$\tilde{\beta}, \tilde{p}, \tilde{\delta}$	$\tilde{T}_0, \tilde{\kappa}$
TV	9	V	$\hat{V}_0, \hat{\beta}, \hat{p}, \hat{c}$		\hat{T}_0

Table 3

Results of the practical identifiability analysis of the IR, TEIV, and TV models given different observable data. Synthetic data were generated by simulating the relevant models using the parameter values given in Table 1, MLE values in Table S1 for the IR model, and MLE values in Table S2 for the TEIV and TV models, over a 12 day period. The set of parameters that is assumed to be known for the IR model is $\Theta = [A_1, A_2, \lambda_1, \lambda_2, t^*, \kappa, \delta, c, m, \gamma, T_0]$. Profile likelihood curves relating to Cases 1, 8, and 9 are shown in Figure S5, while curves for Cases 2–7 are given in Supplementary Information Section S.7.

Model	Case	Observable data	Identifiable	Non-identifiable	Known
IR	1	V, C, X	$\beta, p, \tau, s, r, f, k, V_0$		Θ
	2	V, C	β, p, τ	s, r, f, k, V_0	Θ
	3	V, X	τ, s, r, f, V_0	β, p, k	Θ
	4	C, X	r	$\beta, p, \tau, s, f, k, V_0$	Θ
	5	V		$\beta, p, \tau, s, r, f, k, V_0$	Θ
	6	C		$\beta, p, \tau, s, r, f, k, V_0$	Θ
	7	X		$\beta, p, \tau, s, r, f, k, V_0$	Θ
TEIV	8	V	$\tilde{\beta}, \tilde{c}, \tilde{V}_0$	$\tilde{p}, \tilde{\delta}$	$\tilde{T}_0, \tilde{\kappa}$
TV	9	V	$\hat{\beta}, \hat{p}, \hat{c}, \hat{V}_0$		\hat{T}_0

When comparing the TEIV and TV model fits, we observe that the SSRs are similar (Figs. 4 and 5, left columns). To compare the model fits, we computed the mean SSR for each varied value of the IR parameter (see Table S4). In each case, the TV model has a lower mean SSR, suggesting that the extra fitting parameter in the TEIV model does not improve the model fitting. We conclude that the TV model should be used to describe synthetic viral load data generated from the IR model rather than the TEIV model, as the TV model has fewer unknown parameters and, therefore, model fitting is computationally simpler.

3.6. Using the TV model to reproduce viral load dynamics in hosts with different immunological characteristics

In the previous section we showed that the TEIV and TV models could be fit to viral load data generated from the IR model when individual parameters were varied. In practice, the strength of the innate and adaptive immune responses can vary greatly between hosts; it depends on multiple factors, including age and the presence of comorbidities, such as obesity (Haq and McElhaney, 2014; Honce and Schultz-Cherry, 2019; Hui et al., 2006; Lambert et al., 2012). We now determine whether the TV model can reproduce viral load dynamics when multiple immune response parameters of the IR model are varied simultaneously. We consider four hypothetical groups of hosts; ‘baseline’, ‘elderly’, ‘obese’, and ‘enhanced immunity’, and use the IR model to generate synthetic viral load data relating to hosts in each group. Although the IR model is parameterised with murine data, we use this parameterisation to explore the qualitative differences that might exist in humans. We assume that the immune response parameters s, τ , and k vary between and within groups. Accordingly, as is common in mathematical biology (Sachak-Patwa et al., 2021), we prescribe gamma distributions for these three parameters (while holding all other model parameters fixed) to characterise the variations between hosts within each group. The baseline group refers to healthy hosts (i.e. hosts who

are not considered to be elderly, obese, or have enhanced immunity, and whose parameters have mean values as in Table S1).

Influenza severity is strongly linked with age due to the gradual deterioration of the immune system during natural aging (Haq and McElhaney, 2014). Both the innate and adaptive immune responses are altered during viral infections in the elderly. For example, it has been shown that influenza A infected monocytes from older hosts have impaired interferon production. Indeed, values of IFN- β levels in the monocytes of elderly hosts were half those in similarly infected monocytes of young hosts (Pillai et al., 2016). The antibody response also changes with age, with virus-specific antibody titers shown to be smaller in elderly than in young hosts (Blomberg and Frasca, 2011). Based on these results, we shall assume that in elderly hosts values of the adaptive IR kill rate k and the strength of the innate IR s are half the baseline values given in Table S1.

Similarly, obesity is linked with influenza severity as it impairs the host’s innate and adaptive immune responses (Mancuso, 2013). For example, when obese hosts are infected with influenza, their expression of type I interferons is delayed and reduced compared to non-obese hosts (Honce and Schultz-Cherry, 2019). Based on these findings, we characterise obese hosts as having values of s and τ which are half and double their respective baseline values. We also consider hosts who have enhanced immunity. We assume that their values of k and s are double the baseline values. The mean values of s, τ , and k for each group are summarised in Table S5.

We found that the TV model accurately replicates the infection dynamics of individuals in each group (Fig. 6(a)). We also generated 10^5 viral load datasets for each group and fit the TV model to each dataset using MLE. Noting that previously an SSR value below 15 corresponded to a good fit, Fig. 6(b) shows that the TV model describes the synthetic data accurately for each group. In Fig. 6(c)–(k) we show how the distributions of the IR model parameters s, τ , and κ for each group map to the TV model parameters $\hat{\beta}, \hat{p}$, and \hat{c} . The histograms and scatter plots of the TV model parameters show that for each group, the fitted parameter values cluster in different regions of parameter

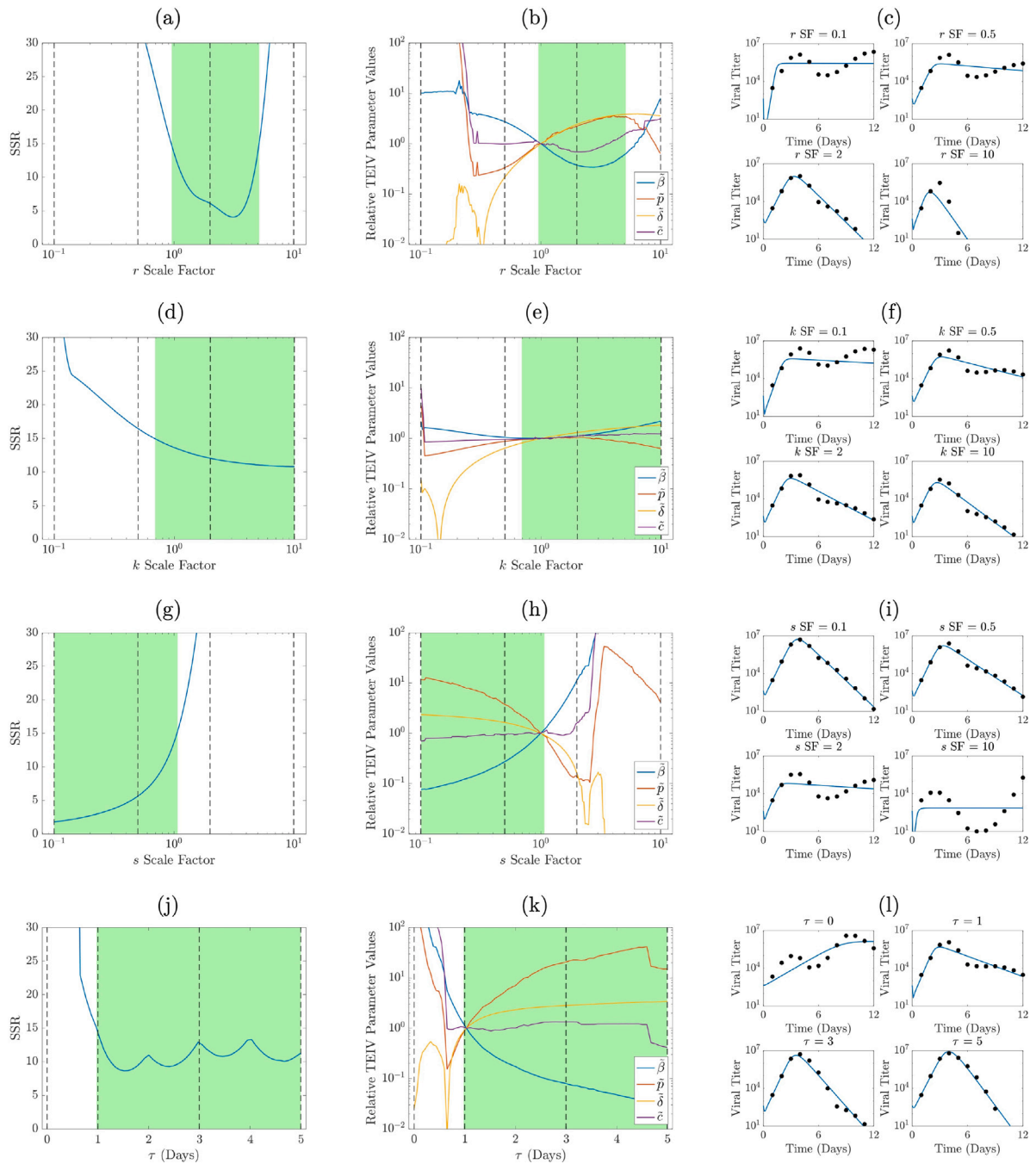


Fig. 4. The TEIV model fitted to synthetic data generated from the IR model as the immune response parameters are varied between hosts. Left column ((a), (d), (g), (j)): The sum of squared residuals (SSR) of the best-fit TEIV model fitted to the synthetic data generated from the IR model. Central column ((b), (e), (h), (k)): Best-fit parameter values when the TEIV model is fitted to the synthetic data generated from the IR model. The relative TEIV parameter values refer to parameter values normalised by those given in Table S2. Right column ((c), (f), (i), (l)): The synthetic data (black dots) and best-fit TEIV model (blue lines) for different values of the IR model parameters. Green shaded regions represent regions of parameter space where $SSR < 15$ as this indicates a good fit. The parameter values of the IR model are defined in Tables 1 and S1 unless otherwise stated. Scale factors (SF) refer to IR parameter values normalised by those given in Table S1. Vertical dashed lines in the left and central columns refer to scale factors of 0.1, 0.5, 2, and 10 respectively as used to generate synthetic data in the right column. The TEIV model is fitted using MLE, with $\tilde{V}_0 = 100$ and $\tilde{T}_0 = 7 \times 10^9$.

space. We observe that the infection rate $\hat{\beta}$ is smaller for elderly and obese hosts than for the baseline group, whereas the virion production coefficient \hat{p} is larger than the baseline for these two groups. The virion clearance rate $\hat{\epsilon}$ is greater than the baseline in the enhanced immunity and obese groups. We conclude that the TV model can accurately reproduce the viral load dynamics of influenza infection in hosts with different immunological characteristics, with different model parameter values for each group.

3.7. Projections of influenza viral load dynamics

We have used synthetic data to show that the TV model can reproduce the viral load dynamics in different host groups. We now present a practical application of a scenario where the TV model would be preferred to the more complex IR model. Specifically, we consider making real-time projections of patient-specific viral load dynamics due to influenza infection.

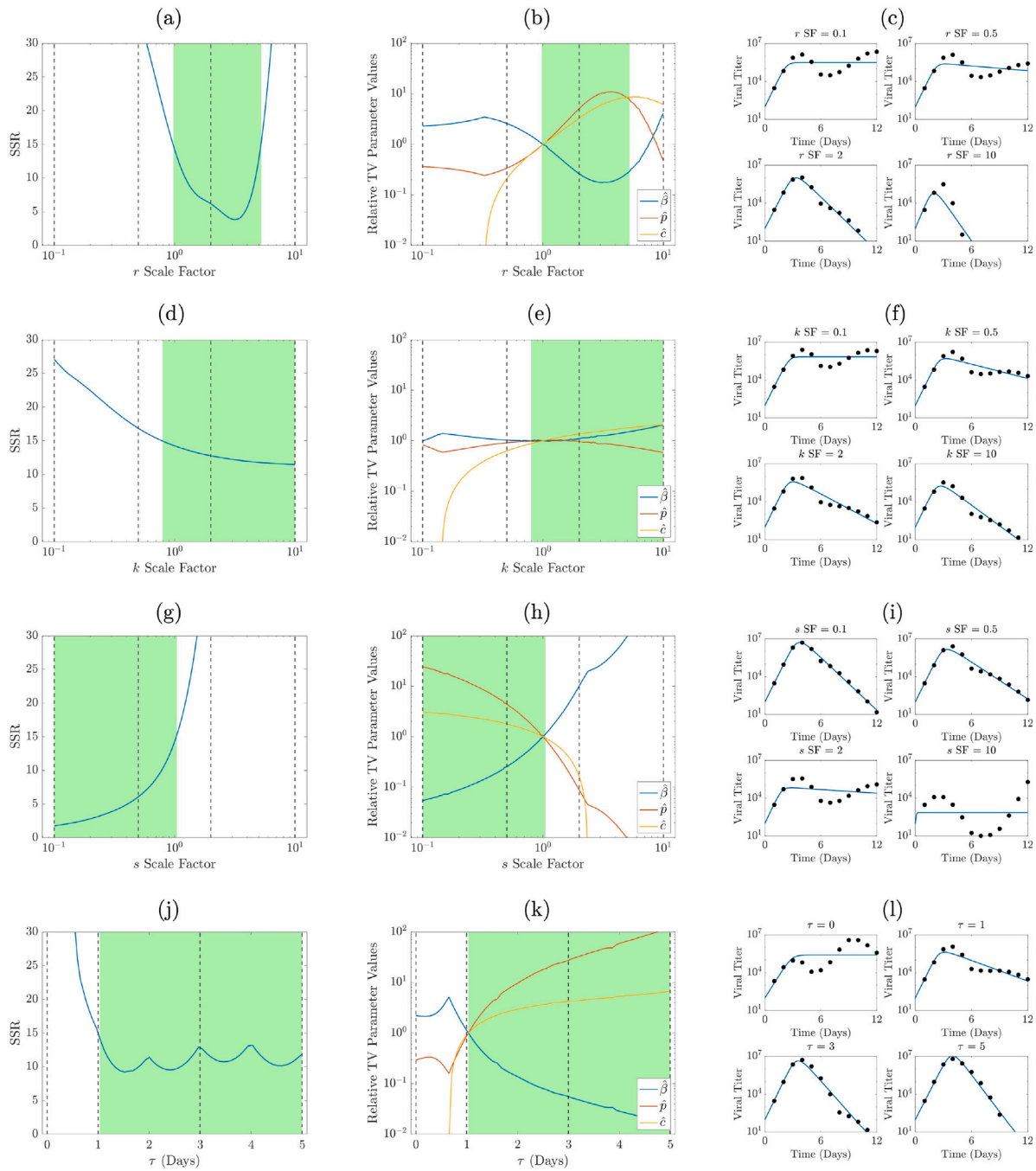


Fig. 5. The TEIV model fitted to synthetic data generated from the IR model as the immune response parameters are varied between hosts. Left column ((a), (d), (g), (j)): The sum of squared residuals (SSR) of the best-fit TV model fitted to the synthetic data generated from the IR model. Central column ((b), (e), (h), (k)): Best-fit parameter values when the TV model is fitted to the synthetic data generated from the IR model. The relative TV parameter values refer to parameter values normalised by those given in Table S2. Right column ((c), (f), (i), (l)): The synthetic data (black dots) and best-fit TV model (blue lines) for different values of the IR model parameters. Green shaded regions represent regions of parameter space where $SSR < 15$ as this indicates a good fit. The parameter values of the IR model are defined in Tables 1 and S1, unless otherwise stated. Scale factors (SF) refer to IR parameter values normalised by those given in Table S1. Vertical dashed lines in the left and central columns refer to scale factors of 0.1, 0.5, 2, and 10 respectively as used to generate synthetic data in the right column. The TV model is fitted using MLE, with $\hat{p}_0 = 100$ and $\hat{r}_0 = 7 \times 10^9$.

In Fig. 7, we present projections of viral load progression generated from the TV and IR models for a single simulated patient. The prior model parameter distributions used in Fig. 7 were obtained by fitting each model to a training dataset comprised of viral load data from six other simulated patients. This provided distributional estimates of the values of the parameters of each model that do not account for the specific patient data considered in Fig. 7. When the projections are made in advance (Fig. 7(a)–(b)), the models are simulated with parameter values sampled repeatedly from these distributions (there is

no additional model fitting as the projections are made in advance). When the projections are instead made at day six (Fig. 7(c)–(d)), the distributions are used as priors in the fitting process; posteriors are obtained by then fitting the models to the viral load data up to and including day six from the specific patient under consideration. The training data, along with the test data, were generated from the IR model, using parameter values relating to obese hosts as this group gave the lowest SSR when fitted to the TV model (Fig. 6(b)).

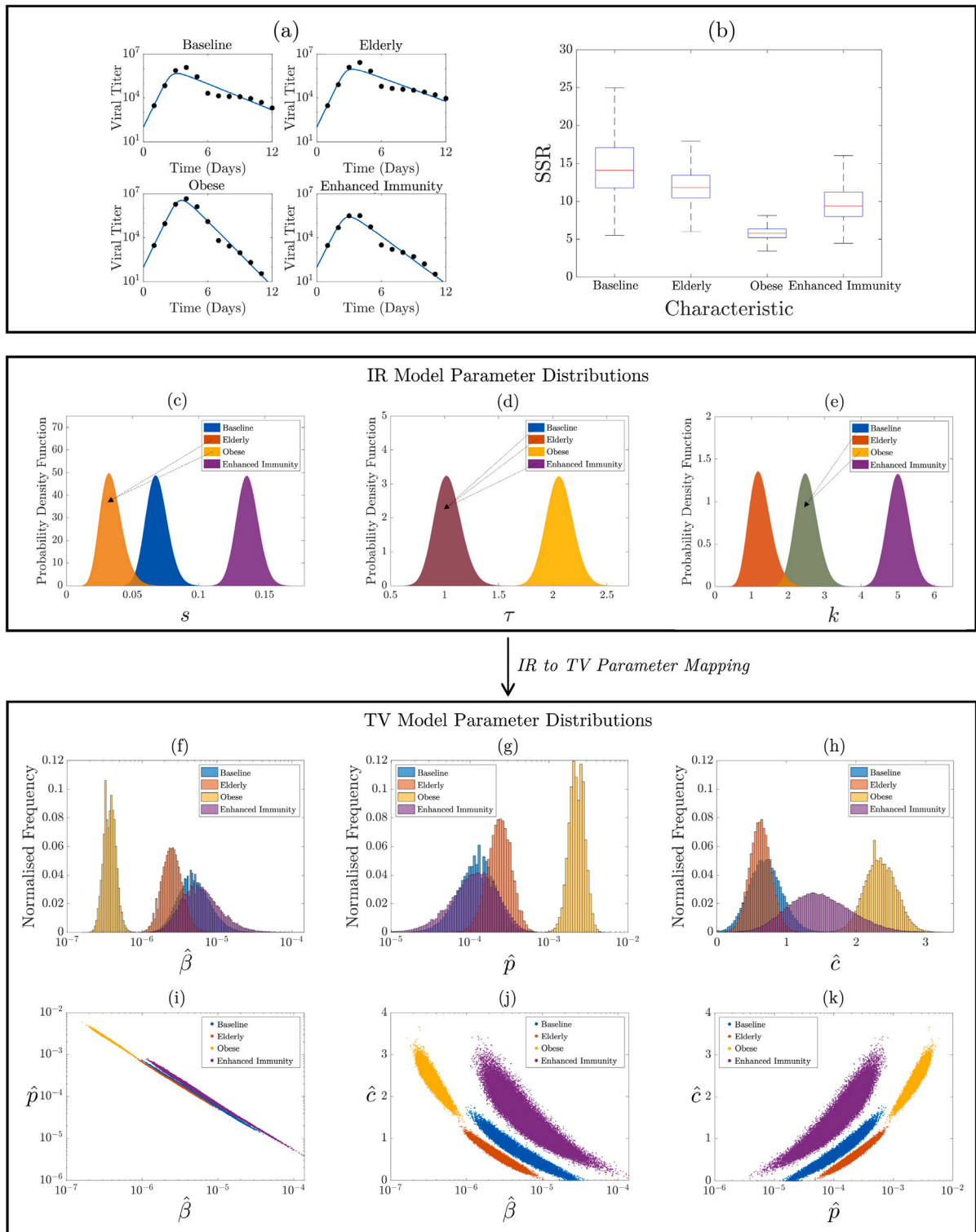


Fig. 6. The TV model fitted to synthetic data relating to different patient groups. (a): TV model fitted to viral load data relating to the four groups using MLE, with mean values of s , τ and k used to generate the data. (b): Boxplots of the SSR of the TV model fitted to synthetic data relating to hosts from each group. (c)–(e): Distributions of s , τ , and k for the different groups in the IR model. (f)–(k): Histograms and scatter plots of the fitted TV model parameters for each group. For each group, 10^5 synthetic datasets were generated. The TV model was fitted to the synthetic data using MLE, with $V_0 = 100$ and $T_0 = 7 \times 10^9$. The synthetic data were generated using the IR model with the default parameter values given in Tables 1 and S1.

For projections made in advance, prior to infection, with parameters calibrated using training data, the 95% credible intervals for both the TV and IR model projections encompass the test data (Fig. 7(a)–(b)). However, when projections are made 6 days after infection, using the new test data up until that time (Fig. 7(c)–(d)), the TV model

projection is significantly more accurate; the TV model yields a BIC value of 993 which is over three times smaller than the value for the IR model. This is because eight parameters need to be estimated in the IR model, and it is infeasible to infer them accurately. This is in addition to the technical difficulty of exploring parameter space

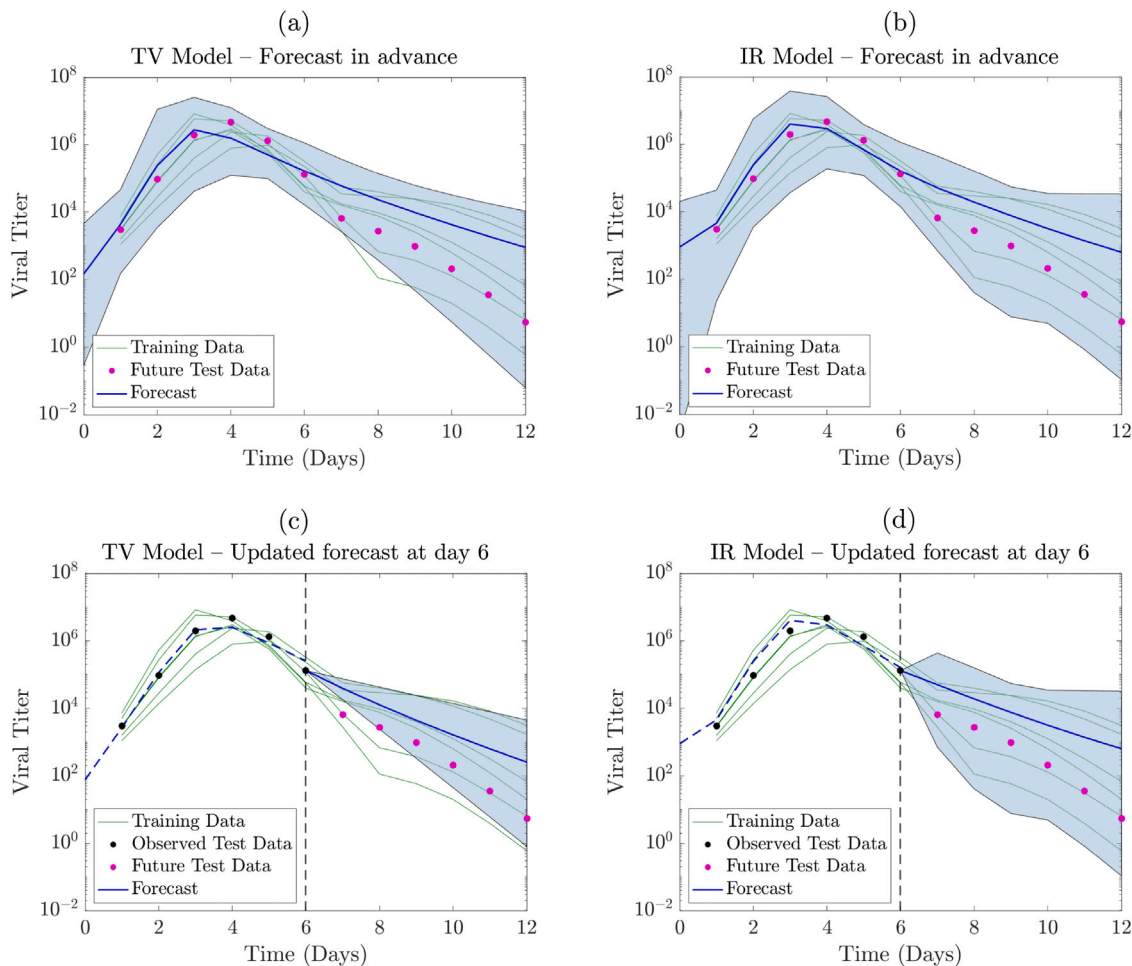


Fig. 7. Comparison of viral load projections obtained using the TV and IR models. Training data from six hosts were used to calibrate the models. Dashed blue lines indicate the mean of the model trajectories during the calibration period. Solid blue lines and shaded regions indicate the mean and 95% CrIs of the projections, based on the posterior parameter distributions. Synthetic training and test data were generated using the IR model by sampling parameter values from the distributions for obese patients (Table S5) and fixing the remaining parameters at the default values given in Tables 1 and S1. BIC values are: (a) 4965, (b) 4253, (c) 993, and (d) 3265.

efficiently in the MCMC estimation when there are so many different parameters requiring estimation. However, as the TV model contains only three fitting parameters and they are all identifiable, it can be used to update projections in real-time to improve projection accuracy (Fig. 7(c)). Furthermore, computing projections with the TV model is less computationally expensive than using the IR model. Although our analysis was based on synthetic data from a single patient rather than a comprehensive investigation, this illustrative example shows how the simpler TV model may be used to make predictions when only viral load data are accessible, even if alternative conclusions could be reached with different data.

4. Discussion

Influenza infection within an individual host is a complex epidemiological process, involving multiple interactions between the immune response, epithelial cells, and the virus (Dou et al., 2018). Consequently, when developing mathematical models to understand the within-host mechanisms driving viral infections, compromises must be made between epidemiological realism and analytical tractability.

In this study, we have considered a hierarchy of ordinary differential equation models which describe influenza infection: the IR model explicitly models the innate and adaptive immune responses, whereas the simpler TEIV and TV models do not. We first fit the IR model to pooled data from infected immunocompetent and immunosuppressed mice that detail viral load, fraction of dead cells, and immune response

levels (Iwasaki and Nozima, 1977), showing that the model's parameter values can be identified using these data. However, when only viral load data from immunocompetent mice are available, the unknown parameter values of the IR model cannot be determined unambiguously. This motivates the use of the simpler TEIV and TV models to describe viral load dynamics. We show, through model fitting, that the parameters of the TV and TEIV models can be estimated given dynamic viral load data from immunocompetent mice. These findings were then substantiated by structural and practical identifiability analyses, showing that, given perfect data, the unknown parameters of the IR model can be identified given dynamic data on the innate and adaptive immune responses, while the TV model parameters (and some of the TEIV model parameters) can be identified using viral load data alone.

We then generated synthetic viral load data using the IR model and showed that the TV model can reproduce the data accurately in cases where the data rise to a single peak and then decay. We also showed that when considering such synthetic data, the TV model should be preferred over the TEIV since it described the data more accurately. Finally, we presented an example of viral load projections for which the simpler TV model may be preferred over the IR model. Although we considered influenza, our results may also generalise to other viral infections which can be modelled using target-cell limited and immune response models, such as SARS-CoV-2 (Fadai et al., 2021), West Nile virus (Banerjee et al., 2016), Adenovirus (Li and Handel, 2014), Respiratory Syncytial Virus (González-Parra and Dobrovolny,

2015), yellow fever virus (Moore et al., 2017), and Zika virus (Best and Perelson, 2018) among others.

The way in which we modelled the immune response in this study was driven by the available data. Thus, to advance our understanding of the immunological processes that affect infection dynamics, it is key that modellers and experimentalists work together in a cycle of data-driven mathematical modelling and model-driven experimental design (Smith, 2018a,b). There are numerous ways in which the IR model could be extended. In this study, the roles of interferons in the innate immune response and antibodies in the adaptive immune response were included in the IR model because experimental data on interferon and antibody levels were available. These mechanisms were incorporated such that interferons reduce viral load, and antibodies clear free virions (Handel et al., 2009). However, other mechanisms may be at play and could be incorporated into the IR model. For example, interferons may promote resistance to infection in target epithelial cells, by moving them to a refractory class where they are no longer susceptible (Hernandez-Vargas et al., 2014). Furthermore, although interferon levels may be used as a proxy for the entire innate immune response (Handel et al., 2007), cellular components such as neutrophils and macrophages may also cause infected cell death and/or clear free virions (Maines et al., 2008). Regarding the adaptive immune response, T-cells are also known to kill infected cells (Bender et al., 1992). Different expansion and activation mechanisms of the innate and adaptive immune responses could also be considered. In particular, a mechanistic differential equation model which describes the innate immune response could replace Eq. (7).

Due to the availability of data, we have used viral load as the predominant infection-related quantity of interest. However, influenza pathogenesis is linked to both viral load and elevated cytokine levels (de Jong et al., 2006), a prime example of the latter being the 1918 influenza pandemic in which it is hypothesised that ‘cytokine storms’ led to severe disease and mortality (Peiris et al., 2010). Metrics related to cytokine and chemokine expression may be better proxies for disease severity (Hayden et al., 1998; Liu et al., 2015). Thus future work could address the problem of modelling influenza severity and pathogenesis more directly (Canini and Carrat, 2010), for example by modelling a host’s ‘symptom score’ which is based on the number of symptoms an infected host is suffering from. Although the TV model replicated the viral load dynamics well in many of the cases we considered, it could not reproduce more complex dynamics, such as when the viral load rose to a second peak. In order to describe more complex viral kinetics, a model incorporating the immune response may be needed, along with sufficient data to parameterise it. Different sampling intervals could also be considered, which may affect the ability of the TV model (or any other model) to describe the data. Furthermore, data could also be generated from a stochastic model to account for the inherent randomness in the infection process (Bai et al., 2018; Handel et al., 2007).

We generated synthetic viral load data pertaining to different hypothetical groups of hosts; baseline, elderly, obese, and enhanced immunity. By fitting the TV model to these data, we were able to cluster hosts into different groups. If real-world data on different patient groups were available, future work could examine whether it is possible to classify patients into different groups based on the values of the TV model parameters. The goal of this study though was not to model the specific infection dynamics of any group in particular, but instead show more generally that a simple model can be used to distinguish the infection dynamics in groups whose immune responses function differently. It is worth noting that the interpretation of the TV model parameters may not be immunological. As a result, if the TV model were applied to real-world data, it may not be possible to distinguish between groups based solely on differences in their immune response or non-immunological factors, such as the baseline production rate of the virus. However, as we have shown, the TV model does provide a

way to easily differentiate between groups of individuals with distinct within-host infection dynamics.

Model selection should be guided by the problem being addressed. If the aim is to make accurate projections given past data without infection-mechanism interpretability, then machine-learning (Wang et al., 2014) or phenomenological models which do not attempt to describe the epidemiology (Daniels and Nemenman, 2015) may be sufficient. If the aim is to understand the mechanisms affecting the infection, then more complex mathematical models may be more appropriate. For instance, it was shown by Cao and McCaw (2017) that the assumed mechanism of viral control (either target-cell limitation or immune-mediated control) markedly altered the viral kinetics of infection under antiviral treatment and, as such, an immune response model was more suitable for predicting infection dynamics and the effect of antiviral treatment in immunocompetent hosts.

In summary, we have explored which models should be used to describe within-host influenza infection dynamics. Parameter identifiability is often a challenge in within-host modelling studies (Miao et al., 2011; Nguyen et al., 2016). Given sufficient data relating to the immune response, we showed that the IR model could be parameterised. However, these data are not routinely collected (especially for infections in humans), and thus only dynamic data pertaining to the viral load are typically available. In such cases, we found that simple target-cell limited models such as the TV model should be used to describe the kinetics of infection. Moreover, we showed that the TV model can differentiate between the infection dynamics in hosts which have differently functioning immune responses. Hence, although target-cell limited models do not fully describe the underlying epidemiology of infection, they maintain immunological interpretability and their simplicity and parameter identifiability make them a useful tool for describing and predicting infection dynamics.

CRedit authorship contribution statement

Rahil Sachak-Patwa: Conceived the study, Conducted the analysis, Wrote and revised the manuscript. **Erin I. Lafferty:** Conceived the study, Supervised the research, Wrote and revised the manuscript. **Claude J. Schmit:** Conceived the study, Supervised the research, Wrote and revised the manuscript. **Robin N. Thompson:** Conceived the study, Supervised the research, Wrote and revised the manuscript. **Helen M. Byrne:** Conceived the study, Supervised the research, Wrote and revised the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and MATLAB code underlying this article are available at <https://github.com/rahilsp/within-host-influenza>.

Acknowledgements

We thank members of the Wolfson Centre for Mathematical Biology (University of Oxford) for helpful discussions about this research. All authors approved the final version of the manuscript.

Funding

This publication is based on work supported by the EPSRC Centre For Doctoral Training in Industrially Focused Mathematical Modelling (EP/L015803/1) in collaboration with Biosensors Beyond Borders Ltd. (RS-P).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jtbi.2023.111491>.

References

- Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S., 2006. Kinetics of influenza a virus infection in humans. *J. Virol.* 80 (15), 7590–7599.
- Bai, F., Huff, K.E.S., Allen, L.J.S., 2018. The effect of delay in viral production in within-host models during early infection. *J. Biol. Dyn.* 13 (sup1), 47–73.
- Banerjee, S., Guedj, J., Ribeiro, R.M., Moses, M., Perelson, A.S., 2016. Estimating biologically relevant parameters under uncertainty for experimental within-host murine West Nile virus infection. *J. R. Soc. Interface* 13 (117), 20160130.
- Beauchemin, C.A., Handel, A., 2011. A review of mathematical models of influenza a infections within a host or cell culture: lessons learned and challenges ahead. *BMC Public Health* 11 (Suppl 1), S7.
- Bellman, R., Åström, K., 1970. On structural identifiability. *Math. Biosci.* 7 (3–4), 329–339.
- Bender, B.S., Croghan, T., Zhang, L., Small, P.A., 1992. Transgenic mice lacking class I major histocompatibility complex-restricted T cells have delayed viral clearance and increased mortality after influenza virus challenge. *J. Exp. Med.* 175 (4), 1143–1145.
- Best, K., Perelson, A.S., 2018. Mathematical modeling of within-host Zika virus dynamics. *Immunol. Rev.* 285 (1), 81–96.
- Blomberg, B.B., Frasca, D., 2011. Quantity, not quality, of antibody response decreased in the elderly. *J. Clin. Invest.* 121 (8), 2981–2983.
- Burnham, K.P., Anderson, D.R., 2004. *Model Selection and Multimodel Inference*. Springer, New York.
- Canini, L., Carrat, F., 2010. Population modeling of influenza A/H1N1 virus kinetics and symptom dynamics. *J. Virol.* 85 (6), 2764–2770.
- Canini, L., Conway, J.M., Perelson, A.S., Carrat, F., 2014. Impact of different oseltamivir regimens on treating influenza a virus infection and resistance emergence: Insights from a modelling study. In: Pond, S.L.K. (Ed.), *PLoS Comput. Biol.* 10 (4), e1003568.
- Canini, L., Woolhouse, M.E.J., Maines, T.R., Carrat, F., 2016. Heterogeneous shedding of influenza by human subjects and its implications for epidemiology and control. *Sci. Rep.* 6 (1).
- Cao, P., McCaw, J., 2017. The mechanisms for within-host influenza virus control affect model-based assessment and prediction of antiviral treatment. *Viruses* 9 (8), 197.
- Carrat, F., Vergu, E., Ferguson, N.M., Lemaître, M., Cauchemez, S., Leach, S., Valleron, A.-J., 2008. Time lines of infection and disease in human influenza: A review of volunteer challenge studies. *Am. J. Epidemiol.* 167 (7), 775–785.
- Centers for Disease Control and Prevention (CDC), 2020. Types of influenza viruses. <https://www.cdc.gov/flu/about/viruses/types.htm>, Accessed on 01/11/21.
- Cheng, Y.-H., You, S.-H., Lin, Y.-J., Chen, S.-C., Chen, W.-Y., Chou, W.-C., et al., 2017. Mathematical modeling of postinfection with influenza a virus and *Streptococcus pneumoniae*, with implications for pneumonia and COPD-risk assessment. *Int. J. Chronic Obstr. Pulm. Dis.* 12, 1973–1988.
- Chis, O.-T., Banga, J.R., Balsa-Canto, E., 2011. Structural identifiability of systems biology models: A critical comparison of methods. *PLoS One* 6 (11), e27755.
- Daniels, B.C., Nemenman, I., 2015. Automated adaptive inference of phenomenological dynamical models. *Nature Commun.* 6 (1).
- de Jong, M.D., Simmons, C.P., Thanh, T.T., Hien, V.M., Smith, G.J.D., Chau, T.N.B., et al., 2006. Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nat. Med.* 12 (10), 1203–1207.
- Dietzen, D.J., 2018. Amino acids, peptides, and proteins. In: *Principles and Applications of Molecular Diagnostics*. Elsevier, pp. 345–380.
- Dobrovolny, H.M., Baron, M.J., Gieschke, R., Davies, B.E., Jumbe, N.L., Beauchemin, C.A.A., 2010. Exploring cell tropism as a possible contributor to influenza infection severity. *PLoS One* 5 (11), e13811.
- Dobrovolny, H.M., Reddy, M.B., Kamal, M.A., Rayner, C.R., Beauchemin, C.A.A., 2013. Assessing mathematical models of influenza infections using features of the immune response. *PLoS One* 8 (2), e57088.
- Dou, D., Revol, R., Östbye, H., Wang, H., Daniels, R., 2018. Influenza a virus cell entry, replication, virion assembly and movement. *Front. Immunol.* 9.
- Fadai, N.T., Sachak-Patwa, R., Byrne, H.M., Maini, P.K., Bafadhel, M., Nicolau, D.V., 2021. Infection, inflammation and intervention: mechanistic modelling of epithelial cells in COVID-19. *J. R. Soc. Interface* 18 (175).
- García-Sastre, A., 2001. Inhibition of interferon-mediated antiviral responses by influenza A viruses and other negative-strand RNA viruses. *Virology* 279 (2), 375–384.
- González-Parra, G., Dobrovolny, H.M., 2015. Assessing uncertainty in A2 respiratory syncytial virus viral dynamics. *Comput. Math. Methods Med.* 2015, 1–9.
- Grant, E.J., Chen, L., Quinones-Parra, S., Pang, K., Kedzierska, K., Chen, W., 2014. T-Cell immunity to influenza A viruses. *Crit. Rev. Immunol.* 34 (1), 15–39.
- Hadjichrysanthou, C., Cauët, E., Lawrence, E., Vegvari, C., de Wolf, F., Anderson, R.M., 2016. Understanding the within-host dynamics of influenza a virus: from theory to clinical implications. *J. R. Soc. Interface* 13 (119), 20160289.
- Handel, A., Liao, L.E., Beauchemin, C.A., 2018. Progress and trends in mathematical modelling of influenza a virus infections. *Curr. Opin. Syst. Biol.* 12, 30–36.
- Handel, A., Longini, I.M., Antia, R., 2007. Neuraminidase inhibitor resistance in influenza: Assessing the danger of its generation and spread. *PLoS Comput. Biol.* 3 (12), e240.
- Handel, A., Longini, I.M., Antia, R., 2009. Towards a quantitative understanding of the within-host dynamics of influenza a infections. *J. R. Soc. Interface* 7 (42), 35–47.
- Haq, K., McElhaney, J.E., 2014. Immunosenescence: influenza vaccination and the elderly. *Curr. Opin. Immunol.* 29, 38–42.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57 (1), 97–109.
- Hayden, F.G., Fritz, R., Lobo, M.C., Alvord, W., Strober, W., Straus, S.E., 1998. Local and systemic cytokine responses during experimental human influenza a virus infection. Relation to symptom formation and host defense. *J. Clin. Invest.* 101 (3), 643–649.
- Hernandez-Vargas, E.A., Wilk, E., Canini, L., Toapanta, F.R., Binder, S.C., Uvarovskii, A., et al., 2014. Effects of aging on influenza virus infection dynamics. *J. Virol.* 88 (8), 4123–4131.
- Honce, R., Schultz-Cherry, S., 2019. Impact of obesity on influenza a virus pathogenesis, immune response, and evolution. *Front. Immunol.* 10.
- Hui, S., Chu, L., Peiris, J., Chan, K., Chu, D., Tsui, W., 2006. Immune response to influenza vaccination in community-dwelling Chinese elderly persons. *Vaccine* 24 (25), 5371–5380.
- Iwasaki, T., Nozima, T., 1977. Defense mechanisms against primary influenza virus infection in mice: I. The roles of interferon and neutralizing antibodies and thymus dependence of interferon and antibody production. *J. Immunol.* 118 (1), 256–263.
- Iwasaki, A., Pillai, P.S., 2014. Innate immunity to influenza virus infection. *Nat. Rev. Immunol.* 14 (5), 315–328.
- Janeway, Jr., C.A., Travers, P., Walport, M., Shlomchik, M.J., 2001. *Principles of innate and adaptive immunity*. In: *Immunobiology: The Immune System in Health and Disease*, fifth ed. Garland Science.
- Kamal, M.A., Gieschke, R., Lemenuel-Diot, A., Beauchemin, C.A.A., Smith, P.F., Rayner, C.R., 2015. A drug-disease model describing the effect of oseltamivir neuraminidase inhibition on influenza virus progression. *Antimicrob. Agents Chemother.* 59 (9), 5388–5395.
- Kilbourne, E.D., 2006. Influenza pandemics of the 20th century. *Emerg. Infect. Diseases* 12 (1), 9–14.
- Lambert, N.D., Ovsyannikova, I.G., Pankratz, V.S., Jacobson, R.M., Poland, G.A., 2012. Understanding the immune response to seasonal influenza vaccination in older adults: a systems biology approach. *Expert Rev. Vaccin.* 11 (8), 985–994.
- Lee, H.Y., Topham, D.J., Park, S.Y., Hollenbaugh, J., Treanor, J., Mosmann, T.R., et al., 2009. Simulation and prediction of the adaptive immune response to influenza a virus infection. *J. Virol.* 83 (14), 7151–7165.
- Li, Y., Handel, A., 2014. Modeling inoculum dose dependent patterns of acute virus infections. *J. Theoret. Biol.* 347, 63–73.
- Liu, Q., Hong Zhou, Y., Qiu Yang, Z., 2015. The cytokine storm of severe influenza and development of immunomodulatory therapy. *Cell. Mol. Immunol.* 13 (1), 3–10.
- Lukens, S., DePasse, J., Rosenfeld, R., Ghedin, E., Mochan, E., Brown, S.T., et al., 2014. A large-scale immuno-epidemiological simulation of influenza a epidemics. *BMC Public Health* 14 (1).
- Maines, T.R., Szretter, K.J., Perrone, L., Belser, J.A., Bright, R.A., Zeng, H., et al., 2008. Pathogenesis of emerging avian influenza viruses in mammals and the host innate immune response. *Immunol. Rev.* 225 (1), 68–84.
- Manchanda, H., Seidel, N., Krumbholz, A., Sauerbrei, A., Schmidtke, M., Guthke, R., 2014. Within-host influenza dynamics: A small-scale mathematical modeling approach. *Biosystems* 118, 51–59.
- Mancuso, P., 2013. Obesity and respiratory infections: Does excess adiposity weigh down host defense? *Pulm. Pharmacol. Ther.* 26 (4), 412–419.
- Meeker, W.Q., Escobar, L.A., 1995. Teaching about Approximate Confidence Regions based on maximum likelihood estimation. *Amer. Statist.* 49 (1), 48.
- Miao, H., Xia, X., Perelson, A.S., Wu, H., 2011. On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM Rev.* 53 (1), 3–39.
- Montaseri, G., Boianelli, A., Hernandez-Vargas, E.A., Meyer-Hermann, M., 2018. PK/PD-based adaptive tailoring of oseltamivir doses to treat within-host influenza viral infections. *Prog. Biophys. Mol. Biol.* 139, 31–42.
- Moore, J., Ahmed, H., Jia, J., Akondy, R., Ahmed, R., Antia, R., 2017. What controls the acute viral infection following yellow fever vaccination? *Bull. Math. Biol.* 80 (1), 46–63.
- Nabel, G.J., Fauci, A.S., 2010. Induction of unnatural immunity: prospects for a broadly protective universal influenza vaccine. *Nat. Med.* 16 (12), 1389–1391.
- Nguyen, V.K., Klawonn, F., Mikolajczyk, R., Hernandez-Vargas, E.A., 2016. Analysis of practical identifiability of a viral infection model. In: Vera, J. (Ed.), *PLoS One* 11 (12), e0167568.
- Nicholson, K.G., Wood, J.M., Zambon, M., 2003. Influenza. *Lancet* 362 (9397), 1733–1745.
- Outlaw, M.C., Dimmock, N.J., 1991. Insights into neutralization of animal viruses gained from study of influenza virus. *Epidemiol. Infect.* 106 (2), 205–220.
- Parkin, J., Cohen, B., 2001. An overview of the immune system. *Lancet* 357 (9270), 1777–1789.

- Peiris, J., Hui, K.P., Yen, H.-L., 2010. Host response to influenza virus: protection versus immunopathology. *Curr. Opin. Immunol.* 22 (4), 475–481.
- Pillai, P.S., Molony, R.D., Martinod, K., Dong, H., Pang, I.K., Tal, M.C., et al., 2016. Mx1 reveals innate pathways to antiviral resistance and lethal influenza disease. *Science* 352 (6284), 463–466.
- Pinky, L., Dobrovolny, H.M., 2016. Coinfections of the respiratory tract: Viral competition for resources. *PLoS One* 11 (5), e0155589.
- Pinky, L., Gonzalez-Parra, G., Dobrovolny, H.M., 2019. Effect of stochasticity on coinfection dynamics of respiratory viruses. *BMC Bioinform.* 20 (1).
- Raue, A., Karlsson, J., Saccomani, M.P., Jirstrand, M., Timmer, J., 2014. Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics* 30 (10), 1440–1448.
- Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25 (15), 1923–1929.
- Rawlins, E.L., Hogan, B.L.M., 2008. Ciliated epithelial cell lifespan in the mouse trachea and lung. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* 295 (1), L231–L234.
- Sachak-Patwa, R., Byrne, H.M., Thompson, R.N., 2021. Accounting for cross-immunity can improve forecast accuracy during influenza epidemics. *Epidemics* 34, 100432.
- Saenz, R.A., Quinlivan, M., Elton, D., MacRae, S., Blunden, A.S., Mumford, J.A., et al., 2010. Dynamics of influenza virus infection and pathology. *J. Virol.* 84 (8), 3974–3983.
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Statist.* 461–464.
- Siekmann, I., Sneyd, J., Crampin, E.J., 2012. MCMC can detect nonidentifiable models. *Biophys. J.* 103 (11), 2275–2286.
- Smith, A.M., 2018a. Host-pathogen kinetics during influenza infection and coinfection: insights from predictive modeling. *Immunol. Rev.* 285 (1), 97–112.
- Smith, A.M., 2018b. Validated models of immune response to virus infection. *Curr. Opin. Syst. Biol.* 12, 46–52.
- Smith, A.M., Adler, F.R., Ribeiro, R.M., Gutenkunst, R.N., McAuley, J.L., McCullers, J.A., Perelson, A.S., 2013. Kinetics of coinfection with influenza A virus and *Streptococcus pneumoniae*. *PLoS Pathog.* 9 (3), e1003238.
- Sompayrac, L.M., 2019. How the Immune System Works. John Wiley & Sons.
- Takeuchi, O., Akira, S., 2009. Innate immunity to virus infection. *Immunol. Rev.* 227 (1), 75–86.
- van der Vries, E., Stittelaar, K.J., van Amerongen, G., Kroeze, E.J.B.V., de Waal, L., Fraaij, P.L.A., et al., 2013. Prolonged influenza virus shedding and emergence of antiviral resistance in immunocompromised patients and ferrets. *PLoS Pathog.* 9 (5), e1003343.
- Wang, X., Sontag, D., Wang, F., 2014. Unsupervised learning of disease progression models. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '14. ACM Press.
- World Health Organization, 2020. Influenza (Seasonal). [https://www.who.int/en/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal)), Accessed on 04/11/20.