

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Hepatitis C Viral Dynamics Using a Combination Therapy of Interferon, Ribavirin, and Telaprevir: Mathematical Modeling and Model Validation

Philip Aston, Katie Cranfield, Haley O'Farrell,
Alex Cassenote, Cassia J. Mendes-Correa,
Aluisio Segurado, Phuong Hoang,
George Lankford and Hien Tran

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.75761>

Abstract

Groundbreaking new drugs called direct acting antivirals have been introduced recently for the treatment of chronic Hepatitis C virus infection. We introduce a mathematical model for Hepatitis C dynamics treated with the direct acting antiviral drug, telaprevir, alongside traditional interferon and ribavirin treatments to understand how this combination therapy affects the viral load of patients exhibiting different types of response. We use sensitivity and identifiability techniques to determine which model parameters can be best estimated from viral load data. Parameter estimation with these best estimable parameters is then performed to give patient-specific fits of the model to partial virologic response, sustained virologic response and breakthrough patients.

Keywords: hepatitis C dynamics, inverse problem, subset selection, sensitivity analysis, identifiability analysis, automatic differentiation

1. Introduction

Over 200–300 million people worldwide are infected with a virus called Hepatitis C (HCV) that affects the liver, which was discovered in 1989 [1]. It is usually spread by blood-to-blood contact via intravenous drug use, poorly sterilized medical equipment and transfusions.

Scarring of the liver and ultimately cirrhosis are just a few of the more severe complications associated with HCV [2].

Six different genotypes of HCV exist due to the highly error prone RNA polymerase with the most common being genotype 1 that has the lowest levels of response to standard treatment [3, 4]. Genotype 1 patients have about a 50% chance for sustained virologic response (SVR), while non-genotype 1 patients have about an 80% chance for SVR [5]. The clinical data used for this study were provided by the University of Sao Paulo, School of Medicine in Sao Paulo, Brazil and consist of genotype 1 patients.

One of the first treatments for HCV was 6–12 months monotherapy with interferon glycoproteins. Interferon is naturally secreted from our bodies to fight off infection and monotherapy treatment with them is associated with around 10% SVR [6]. The addition of ribavirin (RBV), a drug believed to render some of the virus non-infectious, increased SVR to around 30% [6]. RBV monotherapy is not recommended because it does not give a significant benefit to SVR [7]. Until recently, the most common therapy was a combination of pegylated Interferon (IFN) and RBV for 24–48 weeks that yielded about a 45% SVR [5, 6]. One of the major differences between IFN and standard interferon glycoproteins is that the pegylation allows the drugs to stay in the body longer [8]. There have also been clinical trials with RBV monotherapy before and after IFN + RBV therapy as described in [9]. Recently, new drugs called direct-acting antiviral agents (DAAs) have raised the chance for SVR for HCV patients.

DAAs give an increase to about an 80% chance for SVR for genotype 1 [10]. According to the FDA, DAAs are drugs that interfere with specific steps in the HCV replication cycle by taking advantage of the biological makeup of HCV [11]. HCV is a single-stranded RNA molecule that is several nucleotides in length. During HCV's life cycle, it is translated into a polyprotein that is composed into structural and nonstructural proteins that aid in replication. During post-translational processing, DAAs called protease inhibitors block a key protease from the replication process and hinders further infection [10, 12]. Among the protease inhibitors available are boceprevir, telaprevir and simeprevir. Simeprevir is recommended over telaprevir and boceprevir because of both improved efficacy and less side effects, but telaprevir continues to be used because of its cost efficiency [13, 14].

Integration of mathematical modeling of viral dynamics with clinical data has led to further understanding of how different treatment strategies dictate viral load dynamics. One of the first mathematical models was given by Neumann et al. which attempted to describe HCV dynamics with interferon monotherapy [4]. Improvements were made to the Neumann's model to better describe different mechanisms in the liver during treatment including the regeneration of liver cells. Adjustments were also made to include the standard of care, IFN, and RBV. Some of these modifications can be found in [5, 15]. In particular, Snoeck et al. [5] had data after the end of the treatment phase so that the model can give a more accurate representation of its prediction of SVR. The introduction of DAAs has ushered in more mathematical models that include this type of therapy [16]. For example, mathematical models have been proposed using telaprevir monotherapy [17–20] and in combination

with IFN and RBV [18] that uses Bayesian feedback to estimate the parameters in the model. The challenges that come with modeling DAAs is that since they are relatively new, there is not as much data available [17]. It can be difficult to predict SVR because of a lack of data after the treatment phase ends due to how recent the drugs have been approved.

This chapter introduces a novel approach for the development of a mathematical model describing HCV dynamics given the triple-drug combination treatment of IFN, RBV, and the DAA telaprevir. In Section 2, we describe how we adapted a previously known HCV model to include telaprevir and the available clinical data. Section 3 describes the a priori analysis of sensitivity and identifiability and its incorporation into the parameter estimation problem. Section 4 gives the parameter estimation results using several patient specific clinical data including partial virologic response, sustained virologic response and breakthrough. Finally, concluding remarks are provided in Section 5.

2. Mathematical models of HCV dynamics

The original model for HCV dynamics in Neumann et al. [4] was frequently used to assess viral-load profiles after short-term treatment and is given by

$$\begin{aligned}\frac{dT}{dt} &= s - dT - (1 - \eta)\beta VT, \\ \frac{dI}{dt} &= (1 - \eta)\beta VT - \delta I, \\ \frac{dV}{dt} &= (1 - \varepsilon)pI - cV,\end{aligned}\tag{1}$$

where T and I denote the concentrations of healthy and infected hepatocytes, and V represents viral concentration in the liver fluid. One of the key contributions of the model was the understanding of the mechanism of IFN. It was unknown whether it acted through $\eta > 0$ (i.e., inhibiting the infection of healthy liver cells) or $\varepsilon > 0$ (i.e., reducing virion production in infected cells). In [4], it is determined that it is through ε which inhibits production of the virus. The drawback to (1) is that it cannot describe patients exhibiting breakthrough, relapse, and most importantly SVR. These responses are reasons that early viral response does not uniformly predict responses in the long term. Another important aspect is the handling of viral load measurements below the lower limit of quantification (LLOQ). Previous analysis omitted the data below LLOQ, but it can contain critical information regarding long-term treatment outcome. Snoeck et al. [5] present a mathematical model for the dynamics of HCV with the drug treatment combination of IFN and RBV that attempts to address both the long-term responses and the use of the LLOQ. The model described in [5] is given by the following system of nonlinear differential equations

$$\begin{aligned}
\frac{dT}{dt} &= s + rT \left(1 - \frac{T+I}{T_{max}}\right) - dT - \beta V_I T, \\
\frac{dI}{dt} &= \beta V_I T + rI \left(1 - \frac{T+I}{T_{max}}\right) - \delta I, \\
\frac{dV_I}{dt} &= (1 - \bar{\rho})(1 - \bar{\varepsilon})pI - cV_I, \\
\frac{dV_{NI}}{dt} &= \bar{\rho}(1 - \bar{\varepsilon})pI - cV_{NI},
\end{aligned} \tag{2}$$

where T (uninfected hepatocytes), I (infected hepatocytes), V_I (infectious virions) and V_{NI} (noninfectious virions) are natural states (international units IU/mL). This model was adapted from a standard model of viral infection [4]. The number of uninfected hepatocytes increases each day with reproduction rate s and regeneration rate r . That number decreases each day as those hepatocytes die naturally at a rate d or infected at a rate β . The maximum number of hepatocytes per mL is T_{max} . The number of infected hepatocytes increases when the healthy liver cells are infected and when the infected cells regenerate themselves. That number decreases when they die off naturally at a rate δ . Infected hepatocytes produce both infectious and noninfectious virions at a rate p . Virions are naturally cleared at a rate c . IFN inhibits virus production while RBV renders some of the virus noninfectious. The drug efficacies of IFN and RBV are represented by ε and ρ , respectively. The bounds for IFN and RBV are $0 < \varepsilon \leq 1$ and $0 < \rho \leq 1$ where the more effective the drug is, the closer the efficacy of the drug will be to 1. Snoeck uses data that extend beyond treatment for patients so the terms $\bar{\varepsilon}$ and $\bar{\rho}$ in (2) account for the exponential decays of the efficacies of the drugs after treatment has ceased. The exponential decay of the drug efficacies is given by

$$\bar{\varepsilon} = \varepsilon e^{-k(t-t_{end})_+}, \tag{3}$$

and

$$\bar{\rho} = \rho e^{-k(t-t_{end})_+}, \tag{4}$$

where k is the efficacy decay rate, t_{end} marks the end of treatment, and

$$(a)_+ = \begin{cases} a & \text{if } a \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The drug efficacies ε and ρ are related to the drug dosage levels by the following expressions

$$\varepsilon = \frac{\text{Dose}_{\text{PEG}}}{\text{ED}_{50\text{PEG}} + \text{Dose}_{\text{PEG}}}, \tag{5}$$

and

$$\rho = \frac{\text{Dose}_{\text{RBV}}}{\text{ED}_{50\text{RBV}} + \text{Dose}_{\text{RBV}}}, \tag{6}$$

where Dose_{PEG} is the weekly subcutaneous dose of IFN and $\text{ED}_{50\text{PEG}}$ is the estimated weekly dose that causes 50% inhibition of virion production. Dose_{RBV} represents the daily dose of

Parameter	Value
s	$6.17 \times 10^4 \frac{\text{hepatocyte}}{\text{mL} \cdot \text{day}}$
r	$.00562 \text{ day}^{-1}$
β	$8.7 \times 10^{-9} \frac{\text{mL}}{\text{virion} \cdot \text{day}}$
δ	$.139 \text{ day}^{-1}$
c	4.53 day^{-1}
T_{\max}	$1.85 \times 10^7 \frac{\text{hepatocytes}}{\text{mL}}$
d	$.003 \text{ day}^{-1}$
p	$25.1 \frac{\text{virions}}{\text{hepatocyte} \cdot \text{day}}$
ε	.896
ρ	.4-.6
k	$.0238 \text{ day}^{-1}$

Table 1. Typical values from [5].

RBV/kg body weight, and $ED_{50_{RBV}}$ represents the estimated daily dose in mg/kg that makes 50% of the virions noninfectious. Biologically, all state variables and parameters are non-negative. Typical values for model parameters used by Snoek et al. [5] are given in **Table 1**.

2.1. HCV model with DAA

Snoeck’s model is adapted to incorporate the DAA, telaprevir. Recall that a DAA targets specific parts of the genome of the virus to inhibit both replication and infection. The hindrance of replication of the virus in the infected hepatocytes results in the virus not being produced by those cells. This means that the DAA should be implemented as part of the infection term, βTV_I , for inhibiting infection and viral production terms, pV_I and pV_{NI} , for inhibiting replication of the virus in (2). However, after simulations and analysis, it is concluded in this study that the obstruction of the infection and replication of the virus by telaprevir can be described solely as an amplifier for mitigating the production of virions alongside IFN. With this assumption, the model in [5] is modified to include the triple drug combination of IFN, RBV and telaprevir as follows:

$$\begin{aligned}
 \dot{T} &= s + rT \left(1 - \frac{T+I}{T_{\max}} \right) - dT - \beta V_I T \\
 \dot{I} &= \beta V_I T + rI \left(1 - \frac{T+I}{T_{\max}} \right) - \delta I \\
 \dot{V}_I &= (1 - \bar{\rho})(1 - \bar{\varepsilon})(1 - \bar{\gamma})pI - cV_I \\
 \dot{V}_{NI} &= \bar{\rho}(1 - \bar{\varepsilon})(1 - \bar{\gamma})pI - cV_{NI},
 \end{aligned} \tag{7}$$

where $\bar{\gamma}$ represents the exponential decay of the telaprevir efficacy and is defined similarly as for $\bar{\varepsilon}$ and $\bar{\rho}$ (see (3) and (4)). In [21], existence and uniqueness of solutions to this updated

HCV dynamical model were established, and a steady-state stability analysis was also performed.

2.2. Treatment schedule

The data in this research uses the treatment schedule timeline as follows (also summarized in **Figure 1**).

1. The patient is treated with the triple drug combination of IFN + RBV + telaprevir for the first 12 weeks.
2. If at 12 weeks, viral load > 1000 IU/mL, then discontinue treatment. Otherwise, continue 12-week treatment of IFN + RBV.
3. If at 24 weeks, viral load $> \text{LLOQ}$ (12–15 IU/mL), then discontinue treatment. Otherwise, continue 12-week treatment of IFN + RBV.
4. If at 36 weeks, viral load $> \text{LLOQ}$, then discontinue treatment. Otherwise, continue 12-week treatment of IFN + RBV.
5. End of treatment at 48 weeks.

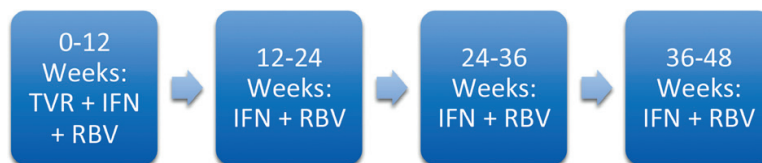


Figure 1. Treatment schedule for patients used for data received from patients treated at University of Sao Paulo, School of Medicine in Sao Paulo, Brazil.

3. Subset selection

The *forward problem* refers to using a model to predict the future behavior of a system given a set of parameters. The *inverse problem*, on the other hand, is the parameterization of a model from empirical data [22–24]. There have been extensive studies about parameter selection while solving the inverse problem for biological models and other applications that can be found in [3, 22, 25–27] and references therein. In this study, we use a simple algorithm to choose a subset of parameters to be estimated from clinical data based on both sensitivity and identifiability as follows:

1. Start with the full parameter set Q .
2. Remove parameters that are not locally sensitive to attain $Q_S \subset Q$.
3. Remove parameters that are not locally identifiable from Q_S to obtain sensitive and identifiable parameter set Q_{SI}

Since these are local analyses, this procedure is repeated over a large number of parameter sets and the parameters that appear most often in Q_{SI} are the parameters that are estimated. All other parameter values are fixed to values from the literature. A biological and structural explanation for some of the fixed parameters is given in the next section.

3.1. Fixed parameters

The assumptions for fixed parameters are the same as in [5]. Since the maximum number of hepatocytes in the liver is 2.50×10^{11} and HCV RNA is distributed in plasma and extracellular fluids with a volume of $\sim 1.35 \times 10^4$ ml, then $T_{max} = \frac{2.50 \times 10^{11}}{1.35 \times 10^4} = 1.85 \times 10^7$. d is obtained from hepatocyte turnover being every 300 days and $s = T_{max} \cdot d$ can be deduced in the absence of liver disease. p is always fixed because $p(1 - \varepsilon)$ appears in \dot{V} and \dot{V}_{NI} making p and ε , impossible to estimate uniquely. The rest of the parameters will be considered in the sensitivity analysis.

3.2. Sensitivity analysis

A sensitivity analysis is the process of understanding how the model output is affected by changes in the parameters. Sensitivity analyses are used in many branches of mathematics such as statistics, partial differential equations (PDEs), and control design [28, 29]. The parameters that give the most change in the output are said to be sensitive parameters. This is important in the *forward problem* because it allows an understanding of which parameters will give useful information. Once the parameters have been identified, a sensitivity analysis for the *inverse problem* is usually performed to determine the sensitive parameters. Parameters with minimal impact are fixed from the literature. There are two different types of sensitivity analysis: global and local. A global sensitivity analysis heavily depends on the structure of the model and quantifies how uncertainties in outputs can be apportioned to uncertainties in inputs. We refer the reader to [30] for a more comprehensive discussion. Our study uses a local sensitivity analysis that depends on the prescribed values of the parameters.

3.2.1. Sensitivity equations

The sensitivity analysis presented in this section uses a derivative-based approach. Consider the general form of an ODE model and a function z of its output

$$\begin{aligned} \frac{dy}{dt} &= f(t, y; q), \\ z &= g(t, y; q), \end{aligned} \tag{8}$$

whereby the vectors y and q contain the variables and parameters of the model, respectively. Since we are concerned with how our model output, z , is influenced by changes to our parameters, q , then we consider the partial derivative of z , $\frac{\partial z}{\partial q}$, with respect to q . One approach to computing this partial derivative is by solving the associated sensitivity equations. Differentiating both sides of the output Eq. (8) with respect to the parameter q yields

$$\begin{aligned}\frac{\partial z}{\partial q} &= \frac{\partial g}{\partial t} \frac{\partial t}{\partial q} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial q} + \frac{\partial g}{\partial q} \frac{\partial q}{\partial q} \\ &= \frac{\partial g}{\partial y} \frac{\partial y}{\partial q} + \frac{\partial g}{\partial q}\end{aligned}\quad (9)$$

since $\frac{\partial t}{\partial q} = 0$ and $\frac{\partial q}{\partial q} = 1$. The two components $\frac{\partial g}{\partial y}$ and $\frac{\partial g}{\partial q}$ can be directly calculated from g , but can be cumbersome to do by hand depending on the complexity of the function g . Thus, one can employ automatic differentiation to evaluate these derivatives. Since any mathematical function can be decomposed into elementary functions, automatic differentiation numerically implements the chain rule and basic arithmetic equations repeatedly to compute the total derivative of a function with accuracy to working machine precision [31]. This is achieved with table lookups and tabulating all the functional compositions [32, 33]. An automatic differentiation (AD) code developed by Martin Fink in MATLAB was employed [34]. Finally, to calculate $\frac{\partial y}{\partial q}$, it is noted that y is continuous in t and q . Since $\frac{\partial y}{\partial q}$ exists, by taking the partial derivative with respect to q of the state equations and reversing the order of differentiation [35], we obtain

$$\begin{aligned}\frac{\partial}{\partial q} \left(\frac{dy}{dt} \right) &= \frac{d}{dt} \left(\frac{\partial y}{\partial q} \right) = \frac{\partial f}{\partial t} \frac{\partial t}{\partial q} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial q} + \frac{\partial f}{\partial q} \frac{\partial q}{\partial q} \\ &= \frac{\partial f}{\partial y} \frac{\partial y}{\partial q} + \frac{\partial f}{\partial q}.\end{aligned}\quad (10)$$

Similar to $\frac{\partial g}{\partial y}$ and $\frac{\partial g}{\partial q}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial q}$ are calculated using automatic differentiation. From (10), the sensitivity equations are given by the following coupled system of differential equations

$$\begin{aligned}\frac{dy}{dt} &= f(t, y; q), \\ \frac{d}{dt} \left(\frac{\partial y}{\partial q} \right) &= \frac{\partial f}{\partial y} \frac{\partial y}{\partial q} + \frac{\partial f}{\partial q}.\end{aligned}\quad (11)$$

Solving the sensitivity equations yields $\frac{\partial y}{\partial q}$, which, in turn, gives $\frac{\partial z}{\partial q}$ from (9).

3.2.2. Model considerations and sensitivity results

The sensitivities of each parameter are ranked to obtain which parameters are most sensitive. Since there is a large range of parameter and viral load values, each parameter, q_j , is log scaled in association with the state variable, y , that is,

$$\frac{d \log_{10}(y)}{d \log_{10}(q_j)} = \frac{q_j}{y} \frac{dy}{dq_j}$$

is considered instead of $\frac{dy}{dq_j}$. This allows a comparison of the sensitivities of each parameter using similar magnitudes. The l_2 -norm is used to nondimensionalize the sensitivities over time so the following sensitivity coefficient is considered for each parameter

$$S_{ij} = \left\| \frac{\partial y_i}{\partial q_j} \right\|_2 = \left[\frac{1}{t_f - t_0} \int_{t_0}^{t_f} \left(\frac{\partial y_i}{\partial q_j} \left(\frac{q_j}{\max y_i} \right) \right)^2 dt \right]^{\frac{1}{2}}. \quad (12)$$

Eq. (12) is defined to be the relative ranking sensitivity of each variable y_i in y with respect to each individual parameter q_j .

Since the local sensitivity analysis depends on values in q , independent sets of parameters that have a log-normal distribution are created from the population-based model fit in Snoeck et al. [5]. That is, a sequence of independent parameter sets $\{q_k\}$ is generated from this distribution using the typical values from [5] as the mean. To determine pseudo-global sensitivities, a sensitivity coefficient, S_{ij}^k , is computed for each parameter in the k th parameter set. Then, if B parameter sets are to be analyzed, then an average for all the parameter sets is computed by

$$\bar{S}_{ij} = \frac{1}{B} \sum_{k=1}^B S_{ij}^k. \quad (13)$$

A cutoff is determined based on the ranking of the averages attained in (13). Those parameters above the cutoff are further examined in the identifiability analysis. This method is a version of what is referred to as Morris Screening in [30]. Similar to the work done here, the Morris algorithm [36] averages local derivative approximations to provide more global sensitivity measures. The difference being that the variance in the parameter sets is also considered. Here that variance would be given by

$$\sigma_{ij}^2 = \frac{1}{B-1} \sum_{k=1}^B \left(S_{ij}^k - \bar{S}_{ij} \right)^2. \quad (14)$$

As explained in [30], while the mean (13) quantifies the individual effect of the input on the output, the variance (14) estimates the combined effects of the input due to nonlinearities or interactions with other inputs. The reader is referred to [30, 36] and references therein for a more detailed analysis of Morris Screening. It is noted that only the marginal distributions are given in [5], so computations are ignorant of any covariances between parameters. The data that were used contain only the viral load observations. So the sensitivities of $V = V_I + V_{NI}$ are of interest. Therefore, (8) is considered where

$$y = [T \ I \ V_I \ V_{NI}]^T,$$

with output

$$z = V = V_I + V_{NI}.$$

Two different sets of time points are used during this analysis. The first and second set of time points come from the partial virologic response (PVR) case and Breakthrough case, respectively. This will provide a better illustration of sensitivities given that treatment decays in the Breakthrough case, but does not in PVR. The sensitivity rankings are given in **Figure 2** for over

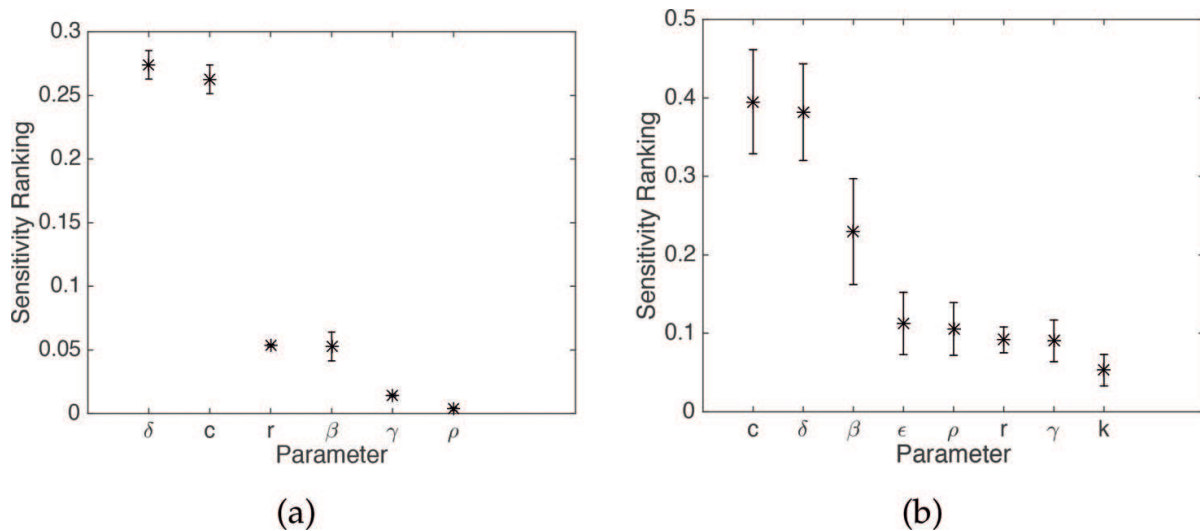


Figure 2. Sensitivity rankings using PVR (a) and Breakthrough time points (b).

2000 (a) and 400 (b) parameter sets, respectively. Error bars that are two standard deviations from the mean are included. The sensitive parameters for the PVR and Breakthrough time points are $Q_{PVR} = \{\delta, c, \beta, r, \gamma\}$ and $Q_{Brk} = \{\delta, c, \beta, r, \rho, \gamma, \epsilon\}$, respectively. These parameters are considered in the identifiability analysis. Note that γ is always considered in the identifiability analysis due to there not being a value from the literature to fix it to for this model. It is used to determine whether it affects the identifiability of other parameters.

3.3. Identifiability analysis

After deciding which parameters are sensitive, consideration is given to understanding which sensitive parameters can uniquely be identified from the data. In this study, we employed a sensitive-based approach for local identifiability analysis. To this end, we consider the parameters contained in q which minimize the cost function

$$J(q) = \frac{1}{N} \sum_{i=1}^N (V_d^i - V(t_i; q))^2,$$

with $V(t_i; q)$ denoting the model output and V_d^i denoting the corresponding data value at time point t_i for $i = 1, \dots, N$, where N is the number of data values. Similar to [37], let us assume that q^* is the minimum of this cost function. Then by using a Taylor series expansion around q^* , we obtain

$$V(t_i, q) = V(t_i; q^*) + \frac{dV(t_i; q^*)}{dq} (q - q^*) + \dots$$

If we only consider the first two elements of $V(t_i, q)$ under the assumption that $q \approx q^*$ and substitute this expression into the cost function we find that

$$\begin{aligned}
 J(q) &= \frac{1}{N} \sum_{i=1}^N \left(V_d^i - V(t_i; q^*) - \frac{dV(t_i; q^*)}{dq} (q - q^*) \right)^2, \\
 &= \frac{1}{N} \sum_{i=1}^N \left(\frac{dV(t_i; q^*)}{dq} (q - q^*) \right)^2,
 \end{aligned}
 \tag{15}$$

where we used the fact that q^* is the minimum of the cost function so that $V_d^i \approx V(t_i; q^*)$. Let

$$S = \frac{dV}{dq} = \begin{bmatrix} \frac{dV}{dq_1}(t_1) & \frac{dV}{dq_2}(t_1) & \dots & \frac{dV}{dq_l}(t_1) \\ \frac{dV}{dq_1}(t_2) & \frac{dV}{dq_2}(t_2) & \dots & \frac{dV}{dq_l}(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{dV}{dq_1}(t_N) & \frac{dV}{dq_2}(t_N) & \dots & \frac{dV}{dq_l}(t_N) \end{bmatrix},
 \tag{16}$$

be an $(N \times l)$ sensitivity matrix relating to the sensitivities $\frac{dV}{dq_j}(t_i)$ of the output with $i = 1, \dots, N$ and $j = 1, \dots, l$, where l denotes the number of parameters. The cost function of (15) is rewritten in terms of this sensitivity matrix

$$\begin{aligned}
 J(q) &= \frac{1}{N} (S(q - q^*))^T (S(q - q^*)), \\
 &= \frac{1}{N} (S\Delta q)^T (S\Delta q),
 \end{aligned}$$

where $\Delta q = q - q^*$. Rearranging $\Delta q = q - q^*$, we formulate the cost function in terms of $q^* + \Delta q$:

$$J(q^* + \Delta q) = \frac{1}{N} \Delta q^T S^T S \Delta q.
 \tag{17}$$

If we suppose that Δq is an eigenvector of $S^T S$ with $S^T S \Delta q = \lambda \Delta q$, then we have

$$\begin{aligned}
 J(q^* + \Delta q) &= \frac{1}{N} \Delta q^T (\lambda \Delta q), \\
 &= \frac{1}{N} \lambda \|\Delta q\|_2^2.
 \end{aligned}$$

We note that if Δq is an eigenvector with eigenvalue $\lambda = 0$, then the cost function to second-order approximation is $J(q^* + h\Delta q) = 0$. The least squares cost function does not change values when moving from q^* to $q^* + h\Delta q$, with h arbitrary. Thus, the parameters are locally unidentifiable at q^* . If $S^T S$ has very small eigenvalues, this can also be a problem for parameter identification. There have been studies about how the Fisher Information Matrix ($S^T S$) can be used for parameter identification [38, 39]. For example, in [38], they search all possible parameter combinations and choose them based on the rank of the

sensitivity matrix, S , and asymptotic standard error uncertainty. We use the following algorithm as described in [39] to determine which of the parameters in our model will be unidentifiable.

1. Create the matrix $S^T S$, compute its eigenvalues, and order them such that

$$|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_n|.$$

2. If $|\lambda_1|$ is less than some threshold ϵ (typically taken to be 10^{-4}), we say that there is a parameter that is unidentifiable.
3. The largest magnitude component of the eigenvector Δq_1 associated with the eigenvalue λ_1 corresponds to the least identifiable parameter. Remove the corresponding column from S and repeat step 1.

After performing this procedure, we now have a set of sensitive and locally identifiable parameters to estimate. The rest of the parameters are set to “typical values” found in the literature. The identifiability algorithm is applied to all the parameter sets of sensitive parameters, Q_{PVR} and Q_{Brk} , obtained in the previous section. It is observed from **Figure 3** that the parameters in $Q_{PVR} = \{\delta, c, \beta, \gamma\}$ are identifiable at least 50% of the time and the parameters in $Q_{Brk} = \{\delta, c, \beta, \gamma, \epsilon\}$ are identifiable at least 50% of the time. The parameters contained in Q_{PVR} and Q_{Brk} are those that will be estimated from the clinical data.

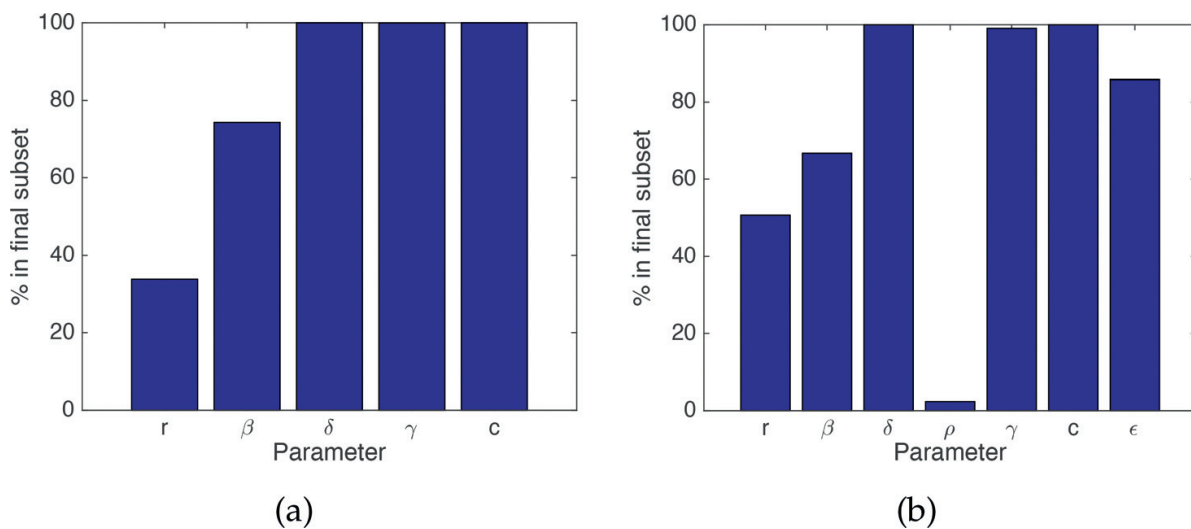


Figure 3. Final subset percentages using PVR (a) and Breakthrough time points (b).

4. Parameter estimation

The parameters in Q_{PVR} and Q_{Brk} are estimated using the weighted sum of squares of errors (WSSE) given by

$$J(q) = \sum_{i=1}^N w_i [\log(V_d^i) - \log(V(t_i; q))]^2, \tag{18}$$

where w_i is the weight for the error term $[\log(V_d^i) - \log(V(t_i; q))]$ at time t_i , V_d^i is data measurement of viral load at the i th time point and $V(t_i; q)$ is the model output with parameters q . We used both sampling and gradient based methods to minimize this function implemented in MATLAB. The model was fit to three data sets; namely, PVR, ETR (end-of-treatment response) and Breakthrough. PVR represents when the patient has an initial positive reaction to the therapy, but then the viral load rebounds during treatment and never goes below detection. ETR represents when the viral load drops below detection and does not rebound. Breakthrough represents when the patient’s viral load drops below detection, but rebounds. In our data, the LLOQ is 15 IU/ml. When the data drop below the LLOQ, least squared estimation does not suffice as a statistically rigorous methodology. Instead, we employ the expectation maximization (EM) [40] to compute maximum likelihood estimates of our patient specific parameters. For a detailed description of the EM algorithm, we refer the reader to [41]. The RBV dosage depends on the patient’s body weight and was sometimes modified during treatment due to different symptoms of the patients such as blood thinning. The patients experiencing PVR and Breakthrough had constant RBV dosage for the entire treatment while the patient exhibiting ETR had modified dosage. The RBV efficacy is fixed to $\rho = .1222$ from [22] for the PVR and Breakthrough patients. The efficacies for the ETR patient were modified based on time, t , in days since initial treatment and are presented in **Table 2**.

The parameters not in Q_{PVR} or Q_{Brk} are fixed to the values in **Table 3** from [5, 22]. As in [5], the infected steady state is used for the initial conditions for (7) because the patients considered had chronic infection. The values in **Table 4** are obtained after estimating the parameters in Q_{PVR} and Q_{Brk} . These estimates produce the model fits (graphs on the left) and residuals (graphs on the right) in **Figures 4–6**. It is noted that in **Figure 6**, the ETR patient’s viral load goes to zero, and the residuals for censored data are set to zero.

In practice, the mathematical model is never exact (model misspecification), and the data contain noise (human errors, instrument errors). Hence, confidence and prediction intervals are used to understand the extent of uncertainty involved in estimating our parameters. In

Parameter	$t \leq 27$	$27 < t \leq 83$	$t > 83$
ρ	.5127	.3185	.219

Table 2. Patient ETR’s RBV efficacies based on modified dosage.

Parameters	s	r	T_{\max}	d	p	ϵ
Values	6.17×10^4	.00562	1.85×10^7	.003	25.1	.6138

Table 3. Fixed parameter values from [5, 22].

Patient	PVR	ETR	Breakthrough
δ	$.1883 \pm .0462$.7211	.3293
c	2.717 ± 2.724	11.67	2.089
γ	$.9987 \pm .0015$.9999	.6575
β	$1.875 \times 10^{-5} \pm 1.688 \times 10^{-5}$	8.684×10^{-8}	2.259×10^{-6}
ϵ	.6138	.9829	.9875

Table 4. Values from parameter estimation for (7).

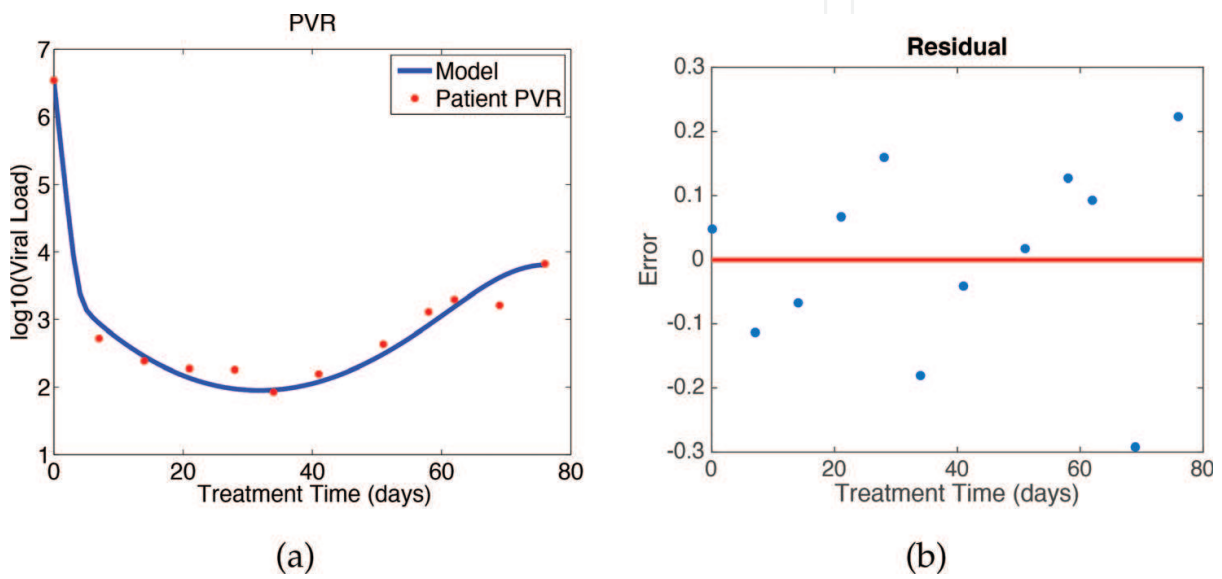


Figure 4. Viral load model fit (a) and residual plot (b) for PVR patient data.

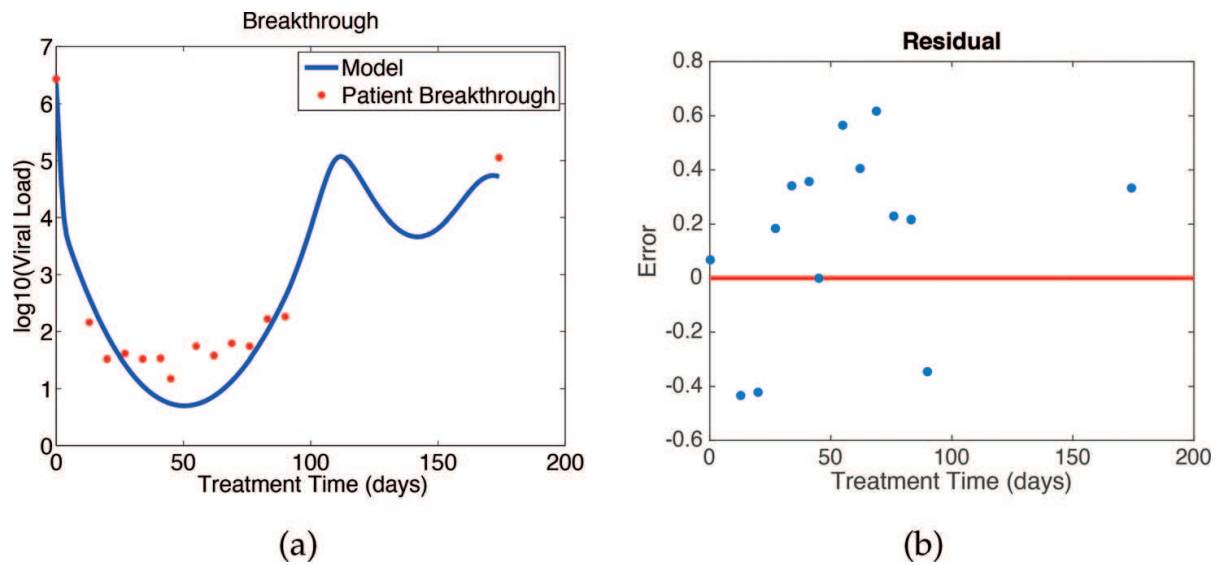


Figure 5. Viral load model fit (a) and residual plot (b) for Breakthrough patient data.

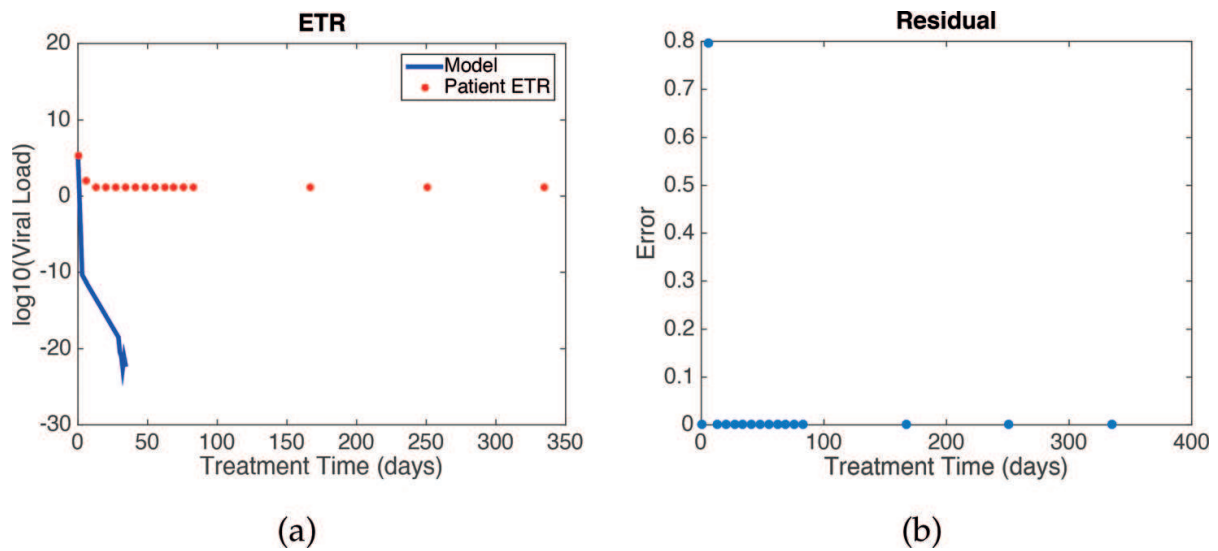


Figure 6. Viral load model fit (a) and residual plot (b) for ETR patient data.

calculating these intervals, standard errors are computed from the model predictions using the parameters that have been estimated. Moreover, 95% parameter and predictive confidence intervals and prediction intervals for the PVR parameters (attached as half-widths in **Table 4**) and predictions are calculated using the asymptotic theory outlined in [22, 27, 30, 41, 42]. The predictive confidence intervals and prediction intervals are shown in **Figure 7**.

5.1. Discussion

The higher values in c and δ in the ETR patient lead us to believe that the immune response along with the drugs has a stronger impact on the mutation and clearance of the virus. It is known that the immune response is strongly correlated with the clearance

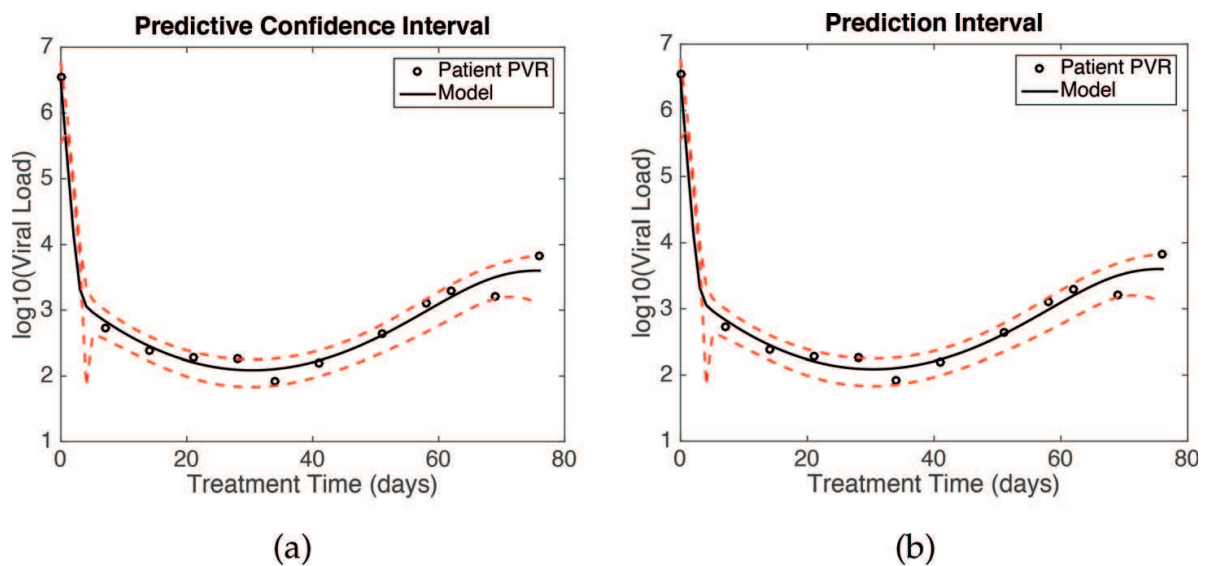


Figure 7. Predictive confidence intervals (a) and prediction intervals (b).

of the virus. Since the initial conditions of (7) are at the infected steady state, introduction of the drugs could be a mechanism to jump start the immune response. We note that even when the virus is not cleared, telaprevir still has a strong impact on viral load decay. This behavior corresponds with how powerful DAAs can be in reducing viral load even when it rebounds. The rebound could be because of mutations which are neglected in this model as stated earlier. There is a dip at around the 150th day in the Breakthrough response that is unquantifiable due to lack of information regarding the other three states or a dynamic immune response. However, this type of dip is observed in [5, 27] where data are available around this time. We conjecture that this dip is due to the immune response being stimulated by the spike in viral load and infection. The residuals in the PVR fit in **Figure 4** seem to be i.i.d. because the errors seem to be randomly distributed and are on both sides of the zero axis. This is unlike the Breakthrough fit in **Figure 5** which have most of the residuals above the zero axis. The predictive confidence intervals and prediction intervals look almost the same because the variance is very small, and the model fits the data very well. The reader is referred to [30] for further details on differences between the predictive confidence intervals and prediction intervals.

6. Conclusion

The missing data between weeks 12–24, 24–36 and 36–48 for the ETR and Breakthrough patients make parameter estimation challenging. The predictions would also be more robust if information concerning states T , I , and V_{NI} were available. These issues should be considered when making remarks about the estimations and confidence measures. DAAs were introduced in 2011, so there is not as much data available, but in the future, we hope for a larger quantity of data to make more precise estimations.

This chapter describes a model for patients with HCV who are treated with IFN, RBV, and telaprevir combination therapy. The development of this model was motivated by the desire for a model that can be validated and calibrated using sensitivity and identifiability techniques while simultaneously incorporating the new DAA, telaprevir. The model can be used to accurately describe patients exhibiting PVR, ETR, and Breakthrough.

Author details

Philip Aston¹, Katie Cranfield¹, Haley O'Farrell¹, Alex Cassenote², Cassia J. Mendes-Correa², Aluisio Segurado², Phuong Hoang³, George Lankford³ and Hien Tran^{3*}

*Address all correspondence to: tran@ncsu.edu

1 University of Surrey, Guildford, United Kingdom

2 University of Sao Paulo, Sao Paulo, Brazil

3 North Carolina State University, Raleigh, United States

References

- [1] Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis c. *Hepatology*. 2004;**39**:1147-1171
- [2] Rehermann B. Hepatitis c virus versus innate and adaptive immune responses: A tale of coevolution and coexistence. *The Journal of Clinical Investigation*. 2009;**119**:1745-1754
- [3] Baraldi R, Cross K, McChesney C, Poag L, Thorpe E, Flores K, Banks H. Mathematical modeling of HCV viral kinetics, Tech. Rep. Raleigh, NC: Center for Research in Scientific Computation; July 2013
- [4] Neumann A, Lam N, Dahari H. Hepatitis c viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science*. 1998;**282**:103-107
- [5] Snoeck E, Chanu P, Lavielle M, Jacqmin P, Jonsson E, Jorga K, Goggin T, Grippo J. A comprehensive hepatitis c viral kinetic model explaining cure. *Clinical Pharmacology & Therapeutics*. 2010;**87**:706-713
- [6] Kim AI, Saab S. Treatment of hepatitis c. *The American Journal of Medicine*. 2005;**118**:808-815
- [7] Brok J, Gluud LL, Gluud C. Ribavirin monotherapy for chronic hepatitis c. *The Cochrane Database of Systematic Reviews*. 2009;(4):CD005527
- [8] Veronese FM, Mero A. The impact of pegylation on biological therapies. *BioDrugs*. 2008;**22**:315-329
- [9] Quiles-Perez R, de Rueda PM, Maldonado AM-L, Martin-Alvarez A, Quer J, Salmeron J. Effects of ribavirin monotherapy on the viral population in patients with chronic hepatitis c genotype 1: Direct sequencing and pyrosequencing of the hcv regions. *Journal of Medical Virology*. 2014;**86**:1886-1897
- [10] Takayama K, Furusyo N, Ogawa E, Shimizu M, Hiramane S, Mitsumoto F, Ura K, Toyoda K, Murata M, Hayashi J. A case of successful treatment with telaprevir-based triple therapy for hepatitis c infection after treatment failure with vaniprevir-based triple therapy. *Journal of Infection and Chemotherapy*. 2014;**20**:577-581
- [11] Food and D. Administration, Guidance for Industry Chronic Hepatitis C Virus Infection: Developing Direct-Acting Antiviral Drugs For Treatment. Silver Spring, MD, USA: Center for Drug Evaluation and Research; 2013
- [12] Kiser JJ, Flexner C. Direct-acting antiviral agents for hepatitis c virus infection. *Annual Review of Pharmacology and Toxicology*. 2013;**53**:427-449
- [13] Bichoupan K, Martel-Laferrriere V, Sachs D, Ng M, Schonfeld EA, Pappas A, Crismale J, Stivala A, Khaitova V, Gardenier D, Linderman M, Perumalswami PV, Schiano TD, Odin JA, Liu L, Moskowitz AJ, Dieterich DT, Branch AD. Costs of telaprevir-based triple therapy for hepatitis c: \$189,000 per sustained virological response. *Hepatology*. 2014;**60**: 1187-1195

- [14] Hill A, Khoo S, Fortunak J, Simmons B, Ford N. Minimum costs for producing hepatitis c direct-acting antivirals for use in large-scale treatment access programs in developing countries. *Clinical Infectious Diseases Advance Access*. 2014
- [15] Dahari H, Ribeiro RM, Rice CM, Perelson AS. Mathematical modeling of subgenomic hepatitis c virus replication in huh-7 cells. *Journal of Virology*. 2007;**81**:750-760
- [16] Chatterjee A, Guedj J, Perelson AS. Mathematical modelling of HCV infection: What can it teach us in the era of direct-acting antiviral agents. *Antiviral Therapy*. 2012;**17**:1171-1182
- [17] Adiwijaya BS, Herrmann E, Hare B, Kieffer T, Lin C, Kwong AD, Garg V, Randle JCR, Sarrazin C, Zeuzem S, Caron PR. A multi-variant, viral dynamic model of genotype 1 HCV to assess the in vivo evolution of protease-inhibitor resistant variants. *PLOS Computational Biology*. 2010;**6**:e1000745
- [18] Adiwijaya BS, Kieffer TL, Henshaw J, Eisenhauer K, Kimko H, Alam JJ, Kauffman RS, Garg V. A viral dynamic model for treatment regimens with direct-acting antivirals for chronic hepatitis c infection. *PLOS Computational Biology*. 2012;**8**:e1002339
- [19] Guedj J, Perelson AS. Telaprevir-based therapy increases with drug effectiveness: Implications for treatment duration. *Hepatology*. 2011;**53**:1801-1808
- [20] Rong L, Ribeiro RM, Perelson AS. Modeling quasispecies and drug resistance in hepatitis c patients treated with a protease inhibitor. *Bulletin of Mathematical Biology*. 2012;**74**:1789-1817
- [21] Lankford G. Optimization, Modeling, and Control: Applications to Klystron Designing and Hepatitis C Virus Dynamics, Phd Thesis. Raleigh, North Carolina: North Carolina State University; 2016
- [22] Arthur JG, Tran H, Aston P. Feasibility of parameter estimation in hepatitis c viral dynamics models. *Journal of Inverse and Ill-Posed Problems*. 2016
- [23] Clermont G, Zenker S. The inverse problem in mathematical biology. *Mathematical Biosciences*. 2014;**260**:11-15
- [24] Zenker S, Rubin J, Clermont G. From inverse problems in mathematical physiology to quantitative differential diagnoses. *PLOS Computational Biology*. 2007;**3**:2072-2086
- [25] Banks H, Baraldi R, Cross K, Flores K, McChesney C, Poag L, Thorpe E. Uncertainty quantification in modeling HIV viral mechanics, Technical Report 16. Raleigh, NC, USA: Center for Research in Scientific Computation; December 2013
- [26] Banks H, Cintron-Arias A, Kappel F. Parameter selection methods in inverse problem formulation, Technical Report 03. Raleigh, NC, USA: Center for Research in Scientific Computation; November 2010
- [27] Banks H, Tran H. Mathematical and Experimental Modeling of Physical and Biological Processes. Boca Raton, FL, USA: Chapman and Hall/CRC; January 2009

- [28] Banks H, Bekele-Maxwell K, Bociu L, Noorman M, Tillman K. The complex-step method for sensitivity analysis of non-smooth problems arising in biology, Technical Report 11. Center for Research in Scientific Computation; October 2015
- [29] Wentworth MT, Smith RC, Banks H. Parameter selection and verification techniques based on global sensitivity analysis illustrated for an hiv model. *SIAM/ASA Journal on Uncertainty Quantification*. 2016;**4**:266-297
- [30] Smith RC. *Uncertainty Quantification: Theory, Implementation, and Applications*. Philadelphia, PA, USA: SIAM; 2014
- [31] Carmichael GR, Sandu A, Potra FA. Sensitivity analysis for atmospheric chemistry models via automatic differentiation. *Atmospheric Environment*. 1997;**31**:475-489
- [32] Griewank A. On Automatic Differentiation, *Mathematical Programming: Recent Developments and Applications 1989*. pp. 83-108
- [33] Neidinger RD. Introduction to automatic differentiation and matlab object-oriented programming. *SIAM Review*. 2010;**52**:545-563
- [34] Fink M. Automatic differentiation for Matlab: Version 1.0, June 2006. <http://www.mathworks.com/matlabcentral/fileexchange/15235-automatic-differentiation-for-matlab>
- [35] Trench WF. *Advanced Calculus*. New York, NY, USA: Harper & Row Publishers; 1978
- [36] Morris M. Factorial sampling plans for preliminary computational experiments. *Technometrics*. 1991:161-174
- [37] Miao H, Xia X, Perelson AS, Wu H. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Review*. Society for Industrial and Applied Mathematics. 2011;**53**:3-39
- [38] Cintron-Arias A, Banks H, Capaldi A, Lloyd AL. A sensitivity matrix based methodology for inverse problem formulation, Technical Report 09. Raleigh, NC, USA: Center for Research in Scientific Computation; April 2009
- [39] Quaiser T, Monnigmann M. Systematic identifiability testing for unambiguous mechanistic modeling - application to jak-stat, map kinase, and nf-kb signaling pathway models. *BMC Systems Biology*. 2009;**3**. Article number 50
- [40] Dempster A, Rubin NLD. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*. 1977;**39**:1-38
- [41] Attarian AR. Patient Specific Subset Selection, Estimation and Validation of an HIV-1 Model with Censored Observations under and Optimal Treatment Schedule, PhD thesis. North Carolina State University; 2012
- [42] Seber GAF, Wild CJ. *Nonlinear Regression*, Vol. 585 of Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley; 2003

