

Article

A New Model for the Dynamics of Hepatitis C Infection: Derivation, Analysis and Implications

Philip J. Aston

Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, UK; P.Aston@surrey.ac.uk

Received: 25 January 2018; Accepted: 10 April 2018; Published: 13 April 2018



Abstract: We review various existing models of hepatitis C virus (HCV) infection and show that there are inconsistencies between the models and known behaviour of the infection. A new model for HCV infection is proposed, based on various dynamical processes that occur during the infection that are described in the literature. This new model is analysed, and three steady state branches of solutions are found when there is no stem cell generation of hepatocytes. Unusually, the branch of infected solutions that connects the uninfected branch and the pure infection branch can be found analytically and always includes a limit point, subject to a few conditions on the parameters. When the action of stem cells is included, the bifurcation between the pure infection and infected branches unfolds, leaving a single branch of infected solutions. It is shown that this model can generate various viral load profiles that have been described in the literature, which is confirmed by fitting the model to four viral load datasets. Suggestions for possible changes in treatment are made based on the model.

Keywords: HCV infection; mathematical model; steady state solutions; bifurcations

1. Introduction

Viral diseases are major causes of human morbidity and mortality worldwide. Hepatitis C virus (HCV) infection is one of the major contributors in this regard. The WHO recently published a Global Hepatitis Report [1], which provides global estimates regarding various aspects of HCV infection. They report that in 2015:

- Globally, an estimated 71 million people were living with chronic HCV infection;
- An estimated 1.75 million new HCV infections occurred worldwide, while 399,000 people died from end-stage HCV infection and 843,000 were cured;
- 20% of HCV-infected persons (14 million) have been diagnosed, and of these, 7.4% (1.1 million) had started treatment;
- HCV infection affects all regions with the highest reported prevalence in the Eastern Mediterranean and European Regions.

The report also noted that incidence of HCV infection in the USA doubled between 2010 and 2014. Thus, HCV infection is still an issue of major importance for global public health.

In May 2016, the World Health Assembly adopted the Global Health Sector Strategy, which committed to a 65% reduction in mortality and a 90% reduction in incidence by 2030 [1]. Clearly, given the above statistics, there is much work to do in order to achieve these targets. A good understanding of the disease mechanism is vital in designing new drugs for treatment of the infection. This understanding is somewhat hampered by the fact that data can only be easily collected for the viral load, and so, other quantities, such as the concentration of healthy and infected hepatocytes, cannot be determined from patient data.

Mathematical modelling allows the inclusion of all the relevant variables and can help to give insight into disease progression and the effect of treatment. Many different mathematical models have been proposed for HCV infection (as well as for many other viral infections), and we add our own new model to this collection. One of the early HCV models by Neumann et al. [2] was adapted from similar models of HBV and HIV infections. Other models have since been proposed, many of which are variations on this original model. However, these models are not consistent with some of the known facts about HCV infection. Thus, we review some of the information about HCV infection that is available in the biological literature and propose a new model that is consistent with this information.

Having derived a new model, we then analyse the steady state solutions and, unusually for a nonlinear model, we are able to give an analytic solution for the infected steady state solutions in a special case. There are many different viral load profiles described in the literature, and we describe how each of these can be achieved with our model as well as fitting the model to four different viral load profiles. We also make some suggestions for changes in treatment of HCV infection, based on our model predictions.

2. A Review of Existing Models of HCV Infection

A number of different mathematical models has been proposed for modelling the dynamics of HCV infection. We first review some of the existing models, before proposing one of our own.

2.1. The Neumann Model

Neumann et al. [2] adapted models of HIV and HBV infections to HCV and derived the equations:

$$\dot{T} = s - dT - (1 - \eta)\beta VT \quad (1)$$

$$\dot{I} = (1 - \eta)\beta VT - \delta I \quad (2)$$

$$\dot{V} = (1 - \epsilon)pI - cV \quad (3)$$

where T is the concentration of healthy hepatocytes, I is the concentration of infected hepatocytes, V is the concentration of virions and the dot represents the derivative with respect to time. The treatment parameters η and ϵ correspond to a reduction in the rate of production of infected cells and virions, respectively.

These equations have an uninfected steady state:

$$T_u = \frac{s}{d}, \quad I_u = V_u = 0 \quad (4)$$

which exists for all η and ϵ , together with an infected steady state:

$$\begin{aligned} T_i &= \frac{\delta c}{\beta p(1 - \eta)(1 - \epsilon)} \\ I_i &= \frac{s}{\delta} - \frac{cd}{\beta p(1 - \eta)(1 - \epsilon)} \\ V_i &= \frac{sp(1 - \epsilon)}{\delta c} - \frac{d}{\beta(1 - \eta)} \end{aligned}$$

The two branches intersect at a bifurcation point at:

$$(1 - \eta)(1 - \epsilon) = \frac{\delta cd}{\beta sp} \quad (5)$$

2.2. The First Dahari Model

Dahari et al. [3] analysed the Neumann model and showed that the solutions could exhibit biphasic decline of the viral load under treatment, with an initial rapid decline followed by a slower

decline, but that it could not show triphasic behaviour of the viral load, where a plateau exists between the rapid and slow declines. They therefore developed an extended model by adapting the Neumann model, including extra terms for the replication of the healthy and infected hepatocytes. Their model equations are given by:

$$\dot{T} = s - dT + rT \left(1 - \frac{T+I}{T_{\max}}\right) - (1-\eta)\beta VT \quad (6)$$

$$\dot{I} = (1-\eta)\beta VT + rI \left(1 - \frac{T+I}{T_{\max}}\right) - \delta I \quad (7)$$

$$\dot{V} = (1-\epsilon)pI - cV \quad (8)$$

These equations also have an uninfected steady state with $I_u = V_u = 0$ and T_u found as the positive solution of the quadratic equation:

$$rT^2 - T_{\max}(r-d)T - sT_{\max} = 0 \quad (9)$$

We note that this quadratic equation has one positive and one negative solution, and clearly, the positive solution is the one of interest. Substituting $T = T_{\max} + \delta T$ into (9) gives the quadratic equation:

$$r\delta T^2 + T_{\max}(r+d)\delta T + T_{\max}(T_{\max}d - s) = 0 \quad (10)$$

The number of positive or negative solutions of this equation depends on the sign of $T_{\max}d - s$. If this term is positive, then both the solutions are negative, and so, the two solutions of (9) will both lie below T_{\max} . However, if this term is negative, then (10) will have one positive and one negative solution. Clearly, the negative solution will correspond to the negative solution of (9), and so, the positive solution must correspond to $T > T_{\max}$. Now, T_{\max} is supposed to be the maximum value for the healthy hepatocytes, and so, we maintain this upper bound provided that the condition:

$$T_{\max} > \frac{s}{d}$$

is satisfied.

Equations (6)–(8) also have an infected steady state given by:

$$I_i = \frac{c(T_{\max}(s + (r-d)T_i) - rT_i^2)}{T_i(rc + (1-\eta)(1-\epsilon)p\beta T_{\max})}$$

$$V_i = \frac{\epsilon p(T_{\max}(s + (r-d)T_i) - rT_i^2)}{T_i(rc + (1-\eta)(1-\epsilon)p\beta T_{\max})}$$

where T_i is the positive solution of the quadratic equation:

$$(1-\eta)^2(1-\epsilon)^2\beta^2 p^2 T_{\max} T^2 + c((1-\eta)(1-\epsilon)\beta p T_{\max}(r-\delta) - cr(\delta-d))T - rsc^2 = 0 \quad (11)$$

which also has one positive and one negative solution. Clearly, the positive solution is the only solution of interest. The two branches again intersect at a bifurcation point, which occurs at:

$$(1-\eta)(1-\epsilon) = \frac{c(rT_i - T_{\max}(r-\delta))}{p\beta T_{\max} T_i}$$

Note, however, that T_i also depends on η and ϵ .

2.3. The Second Dahari Model

If ribavirin is included as part of the treatment, then it causes some of the virions to be non-infectious. Dahari et al. [4] modelled this by separating the virions into two groups, namely infectious (V_I) and non-infectious (V_{NI}). The equations that they derived are given by:

$$\dot{T} = s - dT + r_T T \left(1 - \frac{T+I}{T_{\max}}\right) - (1-\eta)\beta V_I T \quad (12)$$

$$\dot{I} = (1-\eta)\beta V_I T + r_I I \left(1 - \frac{T+I}{T_{\max}}\right) - \delta I \quad (13)$$

$$\dot{V}_I = (1-\rho(t))(1-\epsilon)pI - cV_I \quad (14)$$

$$\dot{V}_{NI} = \rho(t)(1-\epsilon)pI - cV_{NI} \quad (15)$$

where $\rho(t)$ is the fraction of virions that are rendered non-infectious due to the effect of ribavirin. Snoeck et al. [5] simplified these equations by taking $r_T = r_I$ and by making ρ a constant rather than a time-dependent function. We now review the solutions of these simplified equations.

The first three Equations (12)–(14) involve only the three variables T , I and V_I , and so, these equations decouple from the fourth. These equations are essentially the same as Equations (6)–(8), except that V is replaced by V_I and p is replaced by $(1-\rho)p$. Thus, the uninfected steady state is the same as for the first Dahari equations, together with $V_{NI} = 0$. The infected steady state is readily obtained from the first Dahari steady state by replacing p with $(1-\rho)p$ in the above equations together with:

$$V_{NI,i} = \frac{\rho(1-\epsilon)p}{c} I_i$$

Similarly, the bifurcation point at which the two branches intersect occurs at:

$$(1-\eta)(1-\epsilon)(1-\rho) = \frac{c(rT_i - T_{\max}(r-\delta))}{p\beta T_{\max} T_i}$$

where T_i is the positive solution of (11).

2.4. Other Models

There are also a number of variants of the above models in the literature. Herrman et al. [6] used Equations (2) and (3) from the Neumann model, but used a different equation for the healthy hepatocytes, given by:

$$\dot{T} = \gamma(T(0) + I(0) - T - I)$$

We note that this equation does not include the term $-\beta VT$ associated with the virus infecting the healthy cells. Furthermore, the rate of production of the healthy cells depends on the difference between the maximum value $T(0) + I(0)$ and the total hepatocyte population $T + I$. After an initial time period, they change δ in (2) to $M\delta$ for some $M > 1$ to represent an inflated loss of infected cells after an initial delay.

Song and Neumann [7] also use Equations (2) and (3) from the Neumann model and supplement this with the equation for T given by:

$$\dot{T} = s - dT + aT(1 - T/T_{\max}) - \beta VT$$

In this case, the proliferation is modelled by a logistic function, but the reason for this choice is not given. They also adapt this model by including a saturation term for the generation of infected cells where they replace the βVT term by $\beta VT/(1 + \alpha V)$.

An earlier model by Dahari et al. [8] is very similar to the first Dahari model described above, but with a time-dependent component for the death rate of infected cells and a time-dependent

treatment parameter $\epsilon(t)$, as well as an extra equation for the dynamics of ALT (alanine aminotransferase). There is also a term $q(t)I$, which is added to the T equation and subtracted from the I equation, which represents non-cytolytic clearance of the virus from infected cells. These extra terms were dropped in the later models by Dahari et al. This early model was analysed in more detail by Reluga et al. [9], but with q constant, which is then described as a spontaneous cure term.

3. A New Model of HCV Infection

Clearly, there are many mathematical models of HCV infection, which are generally similar in nature, but with a variety of different terms occurring in the equations, which represent different biological processes. We now add to this collection of models by developing our own model. We consider the available literature concerning the various dynamical processes that occur during HCV infection and treatment and include those that we consider to be the most important in the model.

3.1. Cell Regeneration

The liver is unique among the organs of the body in that it can regenerate itself if part of the liver mass is removed [10]. We therefore begin by considering the dynamics of healthy hepatocytes in an uninfected liver. All of the models described in Section 2 include the terms:

$$\dot{T} = s - dT \quad (16)$$

The standard interpretation of this equation is that new hepatocytes are formed at a constant rate s and that a certain fraction d of them die. However, hepatocyte turnover is low in a normal liver with more than 99% in the quiescent phase of the cell cycle. But if part of the liver is removed in a partial hepatectomy, then the liver mass is replaced within seven days by replication of the mature hepatocytes [11]. Thus, if the concentration T is measured relative to the original volume of the liver, then we model this cell division process by:

$$\dot{T} = r(T_{\max} - T)$$

so that the concentration is reduced when part of the liver is removed and the division process stops completely when T reaches T_{\max} . Including the term for cell death gives the equation:

$$\dot{T} = r(T_{\max} - T) - dT = rT_{\max} - (r + d)T \quad (17)$$

where d must be small to ensure slow turnover in a healthy liver. This equation has steady state $T_s = rT_{\max}/(r + d)$. We note that this equation is essentially the same as (16) with $s = rT_{\max}$ and d replaced by $r + d$. Thus, this equation seems reasonable, but the interpretation of the parameters requires care.

It has been found that the liver regenerates itself after surgery with approximately exponential convergence to a steady state [12], as can be seen in Figure 1 (although the regeneration in the first seven days seems to occur more rapidly than predicted by the fitted curve). The solution of Equation (17) also consists of exponential convergence to the steady state. However, care must be exercised in linking these as the plots in Figure 1 are for liver volume, whereas T is a concentration of hepatocytes. However, T/T_s is the ratio of the reduced number of hepatocytes to the number of hepatocytes in the liver before surgery, which is the same as the ratio of liver volume after surgery to volume before surgery. Dividing (17) by T_s gives:

$$\frac{\dot{T}}{T_s} = \frac{rT_{\max}}{T_s} - (r + d)\frac{T}{T_s} = (r + d) \left(1 - \frac{T}{T_s}\right)$$

The solution of this equation is:

$$\frac{T}{T_s} = 1 - Ae^{-(r+d)t}$$

where the constant A is determined from the initial condition. This solution can be multiplied by 100 to give a percentage figure, which does not affect the exponential term. Of course, the data shown in Figure 1 indicate that the liver does not completely regain its previous volume, and so, we cannot match this solution precisely to the data. However, we assume that the exponential rate of convergence is the same, which implies that the coefficient of the exponential of the percentage of original liver volume is $r + d$.

Fitting a function of the form $y = a - be^{-ct}$ to the data shown in Figure 1 gives the curves shown. The corresponding parameter values are $a = 79.8889, b = 23.8424, c = 0.0153121$ for the female data and $a = 86.0996, b = 29.1096, c = 0.0111078$ for the male data. Thus, we have $r + d = 1.53 \times 10^{-2} \text{ day}^{-1}$ for the female data and the slightly lower value of $r + d = 1.11 \times 10^{-2} \text{ day}^{-1}$ for the male data.

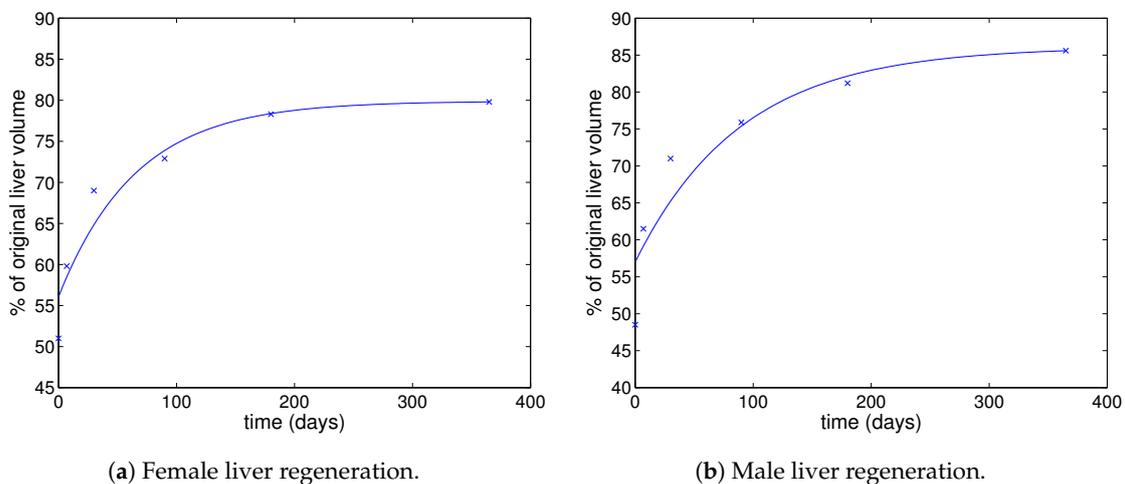


Figure 1. Data for female and male liver regeneration from [12] fitted with a curve of the form $y = a - be^{-ct}$, which is the form of the solution of (17).

For an HCV-infected liver, it is assumed in Dahari et al. [8] that the regeneration of the liver occurs through “blind homeostasis” in, which the infected cells regenerate in the same way as the healthy cells. We make this assumption here, but will revisit this issue later. Thus, if we define $H = T + I$, to be the total concentration of hepatocytes, then the equation for H should be the same as our previous Equation (17), but with T replaced by H , which is therefore given by:

$$\dot{H} = rT_{\max} - (r + d)H \tag{18}$$

We note that considering only the regeneration terms in the first Dahari model [8] and the variant of the second Dahari model by Snoeck et al. [5], the regeneration terms in the equations for T and I are:

$$\begin{aligned} \dot{T} &= rT \left(1 - \frac{T + I}{T_{\max}} \right) \\ \dot{I} &= rI \left(1 - \frac{T + I}{T_{\max}} \right) \end{aligned}$$

Adding these equations gives:

$$\dot{H} = rH \left(1 - \frac{H}{T_{\max}} \right)$$

which is a logistic equation for the growth of hepatocytes. The dynamics of the logistic equation ends with exponential convergence to a steady state, but has slower growth at an earlier stage, particularly for initial conditions that are not close to the stable steady state. This does not agree with the above results that show exponential convergence to the steady state, but with more rapid initial growth.

Thus, we allocate the growth of H , given by $r(T_{\max} - H)$, to the two pools T and I in relative proportion to the size of the pools, noting that the size of the pools is not constant over time. We therefore have the equations:

$$\begin{aligned}\dot{T} &= r \frac{T}{T+I} (T_{\max} - (T+I)) \\ \dot{I} &= r \frac{I}{T+I} (T_{\max} - (T+I))\end{aligned}$$

Adding these equations gives the regeneration term in (18) as we required.

It is generally assumed that the death rates of the healthy and infected cells are different, as the immune system tries to get rid of the infected hepatocytes, resulting in a higher death rate. Thus, adding in the death terms gives the equations:

$$\begin{aligned}\dot{T} &= r \frac{T}{T+I} (T_{\max} - (T+I)) - dT \\ &= \frac{rT_{\max}T}{T+I} - (r+d)T \\ \dot{I} &= r \frac{I}{T+I} (T_{\max} - (T+I)) - \delta I \\ &= \frac{rT_{\max}I}{T+I} - (r+\delta)I\end{aligned}$$

Setting $\delta = d$ and adding these equations then gives (18). However, since the death rate for infected cells is higher than for healthy cells, we assume that $\delta > d$.

We made the assumption above that the healthy and infected cells regenerate in the same way, which implies that the total number of hepatocytes H is not changed by progression of the infection (except for the increased death rate of the infected cells). However, it is known that “the vast majority of liver diseases are characterized by various levels of damage, loss, and impaired regeneration of mature hepatocytes” [13]. Hepatocyte loss due to disease has been classed as mild (<30%), moderate (30–50%) or severe (>50%) [14], and so, it is clear that infection does result in quite significant loss of hepatocytes. As noted above, the infection also impairs the regeneration of mature hepatocytes, and so, this should be included in the model. We do this by assuming that the regeneration rates of healthy and infected hepatocytes are different (as is done in DebRoy et al. [15]), and so, our equations now become:

$$\begin{aligned}\dot{T} &= r_T \frac{T}{T+I} (T_{\max} - (T+I)) - dT \\ &= \frac{r_T T_{\max} T}{T+I} - (r_T + d)T \\ \dot{I} &= r_I \frac{I}{T+I} (T_{\max} - (T+I)) - \delta I \\ &= \frac{r_I T_{\max} I}{T+I} - (r_I + \delta)I\end{aligned}$$

where the two constants r_T and r_I determine the rate of proliferation of the healthy and infected hepatocytes respectively. DebRoy et al. [15] assume that $r_I > r_T$ so that the proliferation of infected cells is greater than that for uninfected cells. It has been noted that chronic HCV infection results in increased expression of proliferation markers [11], which suggests an increase in the proliferation rate due to infection. However, it has also been found that progression to the S phase of the cell

cycle is blocked, so that there is overall a reduction in the number of cells that complete the cell cycle. This is related to the virus in the cells since “cell cycle progression (is) blocked by individual viral proteins” [11]. We therefore assume that $r_I < r_T$ so that proliferation of infected cells is reduced. This is also the assumption made by Dahari et al. [4]. A combination of healthy cells becoming infected, a lower rate of regeneration for infected cells and a higher rate of cell death for infected cells is likely to result in a large decrease in the total number of hepatocytes as the infection progresses, as has been observed.

3.2. Stem Cells

Thus far, we have considered hepatocyte production only as a result of cell division. However, liver stem cells can also generate hepatocytes, and so, there are two mechanisms for hepatocyte production [10,13,16]. In a healthy liver, which has undergone a partial hepatectomy (where part of the liver has been removed), the healthy hepatocytes self-replicate to restore liver mass, and the contribution of liver stem cells to regeneration “seems to be minimal if any” [10]. However, if there is liver injury, which affects this mechanism of regeneration, such as due to chronic viral hepatitis, then the stem cells become more active as an alternative mechanism of hepatocyte production [10,13,16]. Moreover, “a 50% loss of hepatocytes, together with a significant decrease in proliferation of the remaining mature hepatocytes, is required for an extensive activation of hepatic progenitors” [13]. Thus, it seems that the activation of stem cells is proportional to the degree of infection, with no effect in a healthy liver and an extensive effect in a severely-infected liver. We model this by adding a term sI to the equation for T , giving the model equations, which include these two mechanisms of hepatocyte production together with cell death as:

$$\begin{aligned}\dot{T} &= sI + \frac{r_T T_{\max} T}{T + I} - (r_T + d)T \\ \dot{I} &= \frac{r_I T_{\max} I}{T + I} - (r_I + \delta)I\end{aligned}$$

3.3. Infection

Infected hepatocytes are formed when virions enter healthy hepatocytes. The principle of mass action implies that the infection rate is proportional to the product of the concentrations of virions and healthy cells. Adding the infection terms into our earlier equations now gives:

$$\begin{aligned}\dot{T} &= sI + \frac{r_T T_{\max} T}{T + I} - (r_T + d)T - \beta VT \\ \dot{I} &= \frac{r_I T_{\max} I}{T + I} - (r_I + \delta)I + \beta VT \\ \dot{V} &= -cV - \beta VT\end{aligned}$$

where we have also included an elimination term in the V equation. We note that all of the models discussed above omit the infection term in the V equation. In Dahari et al. [4], they point out that if T is assumed to be constant, then $-cV - \beta VT = -(c + \beta T)V = -\tilde{c}V$, and so, the infection term can be considered to be included in the elimination term $-cV$. However, as the models all include a differential equation for T , it is clearly not constant. Moreover, in the case of severe infection, the concentration of healthy hepatocytes may be very low, and so, the assumption that T is approximately constant is not valid. Thus, the infection term should also be included in the V equation. We note that models for HIV infection are very similar to these models, and in this case, the infection term is also included in all three Equations [17].

3.4. The Revised Model

Finally, assuming that infected cells produce virions at a constant rate p , we add an extra term into the V equation. Treatment with interferon alpha is assumed to both reduce the rate of new infection (η) and to reduce the rate of production of virions (ϵ). Adding in these treatment parameters, we obtain our revised model equations:

$$\dot{T} = sI + \frac{r_T T_{\max} T}{T + I} - (r_T + d)T - (1 - \eta)\beta VT \quad (19)$$

$$\dot{I} = \frac{r_I T_{\max} I}{T + I} - (r_I + \delta)I + (1 - \eta)\beta VT \quad (20)$$

$$\dot{V} = (1 - \epsilon)pI - cV - (1 - \eta)\beta VT \quad (21)$$

If the effect of ribavirin is included, which causes some of the virions to be non-infectious, then the model becomes:

$$\dot{T} = sI + \frac{r_T T_{\max} T}{T + I} - (r_T + d)T - (1 - \eta)\beta V_I T \quad (22)$$

$$\dot{I} = \frac{r_I T_{\max} I}{T + I} - (r_I + \delta)I + (1 - \eta)\beta V_I T \quad (23)$$

$$\dot{V}_I = (1 - \rho)(1 - \epsilon)pI - cV_I - (1 - \eta)\beta V_I T \quad (24)$$

$$\dot{V}_{NI} = \rho(1 - \epsilon)pI - cV_{NI} \quad (25)$$

Again, the last equation decouples from the other three, and so, the dynamics of the system is determined by Equations (22)–(24), which are essentially the same as Equations (19)–(21) with V_I replaced by V . Thus, we only consider Equations (19)–(21) in detail.

3.5. Spontaneous Clearance

It has been reported that “about 15–30% of asymptomatic patients and more than 50% of symptomatic patients with acute hepatitis C spontaneously clear the virus during the early phase of infection” [8]. A more recent estimate is that “15–45% of [HCV] infected persons spontaneously clear the virus within 6 months of infection without any treatment” [18]. For infected patients who are cured by treatment, it is also likely that the treatment does not completely eliminate all infected hepatocytes and virions, but that the body is able to clear a small remnant once treatment has stopped. Both of these situations can occur in the equations by ensuring that the uninfected steady state is stable when there is no treatment ($\epsilon = \eta = 0$) for initial conditions sufficiently close to the steady state. However, the Neumann model and the two Dahari models described above do not allow for this possibility as the model equations with the treatment parameters ϵ and η set to zero all have the uninfected steady state as unstable.

The model described in Dahari et al. [8] includes terms (qI) for non-cytolytic clearance of the virus from infected cells. These terms are described in Reluga et al. [9] as being associated with spontaneous cure. In [9], a quasi-steady state approximation is used to reduce the three equations to two. The eigenvalues of the Jacobian evaluated at the uninfected steady state are $\lambda_1 < 0$ and $\lambda_2(q) = \lambda_2(0) - q$, where $\lambda_2(0)$ is the difference of positive terms and so could be positive or negative. Therefore, clearly, the addition of the parameter q ensures that a larger region of the parameter space corresponds to spontaneous cure since it moves the eigenvalue λ_2 to the left on the real line. However, this term was dropped in later models used by this group [3,4].

Spontaneous clearance and clearance of infection after cessation of treatment can both be realised in a model by ensuring that the uninfected steady state is stable with no treatment, but with another nearby steady state also existing, which has one unstable and two stable eigenvalues. The two-dimensional stable manifold would then act as a surface that separates out trajectories that converge to the uninfected steady state or move away from this region to a stable infected steady

state. Chronic infection would then occur if either the initial conditions at the start of infection lie above this two-dimensional stable manifold, or if, for particular parameter values, the uninfected steady state had become unstable, in which case a complete cure with treatment would also be impossible.

All the models have an uninfected steady state with $I = V = 0$ and $T > 0$. One way in which the scenario just described could occur is if the bifurcation from the uninfected steady state to an infected steady state occurs close to $\epsilon = 0$. If the bifurcation point occurs for $\epsilon < 0$, at an unphysical parameter value, with the bifurcating branch occurring to the right of the bifurcation point, then there would be two steady state solutions at $\epsilon = 0$ with the uninfected steady state being stable and the infected steady state being unstable (see Figure 2a). This situation allows for spontaneous cure or chronic infection, depending on the initial conditions, and for complete cure for a patient once the treatment has reduced the infected hepatocytes and virus levels to a sufficiently low level. However, only a small change in parameter values could move this bifurcation point so that it occurs for a positive value of ϵ (see Figure 2c), which makes the uninfected steady state without treatment unstable. In this case, spontaneous clearance of the infection is not possible, and there will always be relapse of the infection on cessation of treatment, as is observed in some cases [5]. Clearly, only a small change in parameters is required to switch between these cases, which could be due to the variability between patients and would also explain the variability in outcomes for different patients. Various different scenarios will be considered in more detail in Section 8.

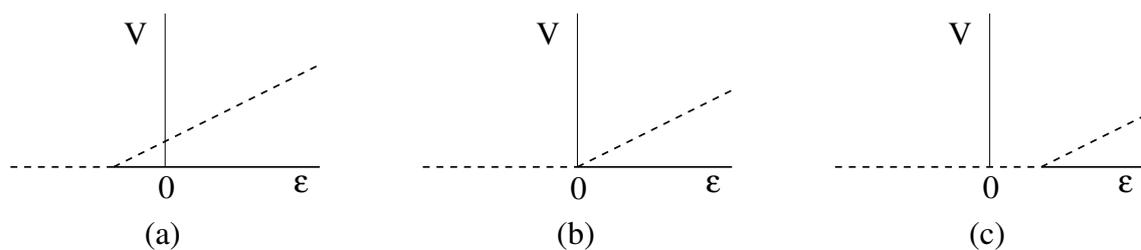


Figure 2. The bifurcation from the uninfected steady state. The bifurcation point occurs for (a) $\epsilon < 0$, (b) $\epsilon = 0$, (c) $\epsilon > 0$. Stable steady states are indicated by solid lines and unstable steady states by dashed lines.

3.6. Non-Dimensionalisation

Before considering the properties of our new model in detail, we non-dimensionalize it in order to reduce the number of parameters. We non-dimensionalize Equations (19)–(21) by rescaling the variables by the uninfected steady state value of T , which is $T_u = r_T T_{\max} / (r_T + d)$ (note that $T_u < T_{\max}$). We therefore define the new non-dimensional variables:

$$x = \frac{(r_T + d)T}{r_T T_{\max}}, \quad y = \frac{(r_T + d)I}{r_T T_{\max}}, \quad z = \frac{(r_T + d)V}{r_T T_{\max}} \tag{26}$$

together with a new non-dimensional time variable:

$$\tau = (r_T + d)t$$

Equations (19)–(21) involve two treatment parameters ϵ and η . Interferon therapy is assumed to partially block viral production (ϵ) and to reduce the rate of production of infected cells (η) [19], which implies that both of the treatment parameters are non-zero. For a bifurcation analysis, it is more convenient to have a single bifurcation parameter, and so, we also make the assumption that:

$$\eta = \alpha \epsilon \tag{27}$$

for some $\alpha > 0$. It has been shown that the major effect of interferon is to block production or release of virions from an infected cell [2], and this implies that:

$$\alpha < 1 \quad (28)$$

The equations in terms of all these new variables are then given by:

$$x' = Sy + \frac{x}{x+y} - x - (1 - \alpha\epsilon)Bxz \quad (29)$$

$$y' = \frac{1}{1+R} \left(\frac{y}{x+y} - y \right) - Dy + (1 - \alpha\epsilon)Bxz \quad (30)$$

$$z' = (1 - \epsilon)Py - Cz - (1 - \alpha\epsilon)Bxz \quad (31)$$

where the prime denotes derivatives with respect to τ , and the non-dimensional parameters are given by:

$$S = \frac{s}{r_T + d'}, \quad B = \frac{\beta r_T T_{\max}}{(r_T + d')^2}, \quad R = \frac{r_T}{r_I} - 1, \quad D = \frac{r_T \delta - r_I d}{r_T (r_T + d')}, \quad P = \frac{p}{r_T + d'}, \quad C = \frac{c}{r_T + d'} \quad (32)$$

We note that our previous assumptions that $r_I < r_T$ and $d < \delta$ imply that $D > 0$. We also explain our choice of the parameter R . The coefficient multiplying the first term in (30) is r_I/r_T , and this satisfies $0 < r_I/r_T < 1$. We generally define parameters so that they are positive, but in this case, we have two conditions to satisfy. However, we note that these two conditions are equivalent to the single condition $r_T/r_I > 1$, since it is implicitly assumed that the ratio is finite. Thus, we define $R = r_T/r_I - 1$ and note that $R \rightarrow \infty$ as $r_I/r_T \rightarrow 0$ and $R = 0$ when $r_I/r_T = 1$. Thus, the single condition $R > 0$ ensures that both conditions on the ratio r_I/r_T are satisfied. We note therefore that all of these non-dimensional parameters must be positive.

We also require that $\epsilon \in [0, 1)$, which ensures that the treatment factor $1 - \epsilon$ is positive, which then also implies that $1 - \alpha\epsilon$ is positive using (28).

4. Validation of the New HCV Model

In order to validate our new HCV model, we need to check some fundamental mathematical and biological properties of the model. From a mathematical perspective, we need to ensure that the equations are well-posed (a solution exists, is unique and depends continuously on the initial conditions). From a biological perspective, we must show that the solutions are non-negative for all time and are bounded.

All of our variables must be non-negative as they are related to physical quantities, and so, we first show that the non-dimensional Equations (29)–(31) have an invariant region within the octant $x, y, z \geq 0$. Using polar coordinates defined by $x = r \cos \theta$, $y = r \sin \theta$, we see that:

$$\frac{x}{x+y} = \frac{\cos \theta}{\cos \theta + \sin \theta}, \quad \frac{y}{x+y} = \frac{\sin \theta}{\cos \theta + \sin \theta}$$

and so, the quotients that occur in Equations (29) and (30) are well defined in the limit $r \rightarrow 0$ along rays with fixed θ for $0 \leq \theta \leq \pi/2$. However, this also implies that the vector field is not uniquely defined along the z axis, and so, we exclude a neighbourhood of this axis. In a modelling context, the z axis ($x = y = 0$) corresponds to there being no healthy or infected hepatocytes, which is unrealistic.

Theorem 1. We define the octant $\mathcal{O} = \{(x, y, z) \in \mathbf{R}^3 : x, y, z \geq 0\}$ and a cylinder around the z axis $\mathcal{C} = \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 < r_0^2, z \geq 0\}$. The region $\mathcal{R} = \mathcal{O} \setminus \mathcal{C}$ is invariant under the flow of Equations (29)–(31) for sufficiently small $r_0 > 0$. Moreover, the line $y = z = 0, x \geq r_0$ is an invariant line,

corresponding to the dynamics of a healthy liver, on which x converges exponentially to the steady state $x = 1$ for all $x(0) \geq r_0$.

Proof. We first consider the line given by $y = z = 0, x \geq r_0$. On this line, we have $y' = z' = 0$, and so, it is invariant for all time; and the flow is given by:

$$x' = 1 - x$$

The solution of this equation consists of exponential convergence to the steady state $x = 1$ for all $x(0) \geq r_0$.

The region \mathcal{R} is bounded by the planes $\mathcal{P}_1 (x = 0, y \geq r_0, z \geq 0)$, $\mathcal{P}_2 (x \geq r_0, y = 0, z \geq 0)$ and $\mathcal{P}_3 (x^2 + y^2 \geq r_0^2, z = 0)$ together with the surface $\mathcal{S} = \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 = r_0^2\} \cap \mathcal{O}$. We will show that the vector field given by Equations (29)–(31) on these surfaces is directed inside the region \mathcal{R} except for the invariant line.

Taking the inner product of the vector field with the inward pointing normal on the plane \mathcal{P}_1 , we see from (29) that:

$$x'|_{\mathcal{P}_1} = Sy$$

and so, $x'|_{\mathcal{P}_1} > 0$ on this plane since $y \geq r_0 > 0$. Thus, the vector field is directed inside \mathcal{R} on this plane.

Similarly, from (30), we have:

$$y'|_{\mathcal{P}_2} = (1 - \alpha\epsilon)Bxz$$

and so, $y'|_{\mathcal{P}_2} > 0$ for $z > 0$ (since $1 - \alpha\epsilon > 0$ as $\alpha, \epsilon < 1$) and $y'|_{\mathcal{P}_2} = 0$ for $z = 0$, which is the invariant line. Thus, the vector field is directed inside \mathcal{R} on this plane, except for the invariant line.

From (31), we also have:

$$z'|_{\mathcal{P}_3} = (1 - \epsilon)Py$$

and so, $z'|_{\mathcal{P}_3}$ is positive everywhere except on the invariant line, since $1 - \epsilon > 0$ (as $\epsilon < 1$).

Finally, we consider the surface \mathcal{S} . The normal to this surface points in the r direction when $x = r \cos \theta, y = r \sin \theta$. Differentiating the equation $x^2 + y^2 = r^2$ with respect to τ and using (29), (30) gives:

$$r' = \frac{1 + R \cos^2 \theta}{(1 + R)(\cos \theta + \sin \theta)} + O(r)$$

Therefore, $r'|_{r=r_0} > 0$, since θ is in the first quadrant, and this implies that $r'|_{r=r_0} > 0$ for sufficiently small $r_0, 0 \leq \theta \leq \pi/2, z \geq 0$. Thus, the vector field points inside \mathcal{R} also on this surface.

Thus, we conclude that any trajectory with $(x(0), y(0), z(0)) \in \mathcal{R} = \mathcal{O} \setminus \mathcal{C}$ must satisfy $(x(t), y(t), z(t)) \in \mathcal{R}$ for all $t > 0$, and so, \mathcal{R} is invariant. \square

We now show that our model equations are well-posed on the region \mathcal{R} .

Theorem 2. *The model Equations (29)–(31) are well-posed on the region \mathcal{R} defined in Theorem 1.*

Proof. The well-posed condition requires the local existence and uniqueness of solutions of the model equations. For the system of differential equations given by $u' = f(u, t), u(t_0) = u_0$ with $u \in \mathbf{R}^n$, if the function f is continuous and satisfies a Lipschitz condition in u on a region $|t - t_0| \leq \alpha, \|u - u_0\| \leq \beta$, then it is well-posed [20], and there is then a unique solution on an interval $|t - t_0| \leq \delta$ for some $\delta \leq \alpha$.

For our model equations given by (29)–(31), the continuity and Lipschitz conditions are satisfied for any such bounded region in the positive octant if the terms $x/(x + y)$ in (29) and $y/(x + y)$ in (30) are excluded. Once the cylinder \mathcal{C} in the positive octant is excluded, then these nonlinear terms also satisfy a Lipschitz condition on any bounded region since $x + y$ is bounded below by r_0 . Therefore, our model equations are well-posed. \square

It remains to show that the solutions are bounded for all time, which we can do provided that a condition on the parameters holds, and this will also imply the global existence of solutions.

Theorem 3. *If:*

$$\min(S, (1 - \epsilon)P)(1 + R) - (1 + D(1 + R)) < 0 \quad (33)$$

then there is a bounded invariant region contained in the region \mathcal{R} (defined in Theorem 1). In this case, the solution of Equations (29)–(31) exists and is bounded for all $t \geq 0$ for any finite initial conditions.

We note that this theorem gives a sufficient condition for bounded solutions. If this condition is not satisfied, it may still be the case that the solutions are bounded. We note that the condition (33), when divided by $(1 + R)$, requires that the coefficient of the linear decay term in y in (30) must exceed one of the coefficients of the linear growth terms in y in Equation (29) or (31).

Proof. We show that the region \mathcal{R} defined in Theorem 1 (the positive octant with a cylinder around the z axis removed) together with the additional restriction that:

$$wx + y + (1 - w)z \leq k$$

is a bounded invariant region for $k > 0$ sufficiently large and for all $w \in (0, 1)$ provided that a condition on the parameters is satisfied. We note that this region is not bounded for $w = 0$ and $w = 1$, which is why we require that $w \in (0, 1)$. This region is shown in Figure 3.

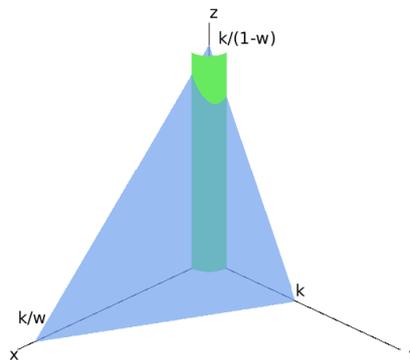


Figure 3. The invariant region in the positive octant bounded by the cylinder \mathcal{C} (see Theorem 1) and the plane (34).

Since we have already shown that the region \mathcal{R} is invariant, it remains to show that the vector field is directed inside the bounded region on the plane:

$$wx + y + (1 - w)z = k \quad (34)$$

when the condition (33) holds.

Using the model Equations (29)–(31), we see that:

$$\begin{aligned} wx' + y' + (1 - w)z' &= -wx + \left(wS - \frac{1}{1 + R} - D + (1 - w)(1 - \epsilon)P \right) y - (1 - w)Cz \\ &\quad + \frac{wx}{x + y} + \left(\frac{1}{1 + R} \right) \left(\frac{y}{x + y} \right) \end{aligned}$$

and substituting for z using (34) gives:

$$\begin{aligned} wx' + y' + (1-w)z' &= (C-1)wx + \left(wS - \frac{1}{1+R} - D + (1-w)(1-\epsilon)P + C \right) y - Ck \\ &\quad + \frac{wx}{x+y} + \left(\frac{1}{1+R} \right) \left(\frac{y}{x+y} \right) \end{aligned} \quad (35)$$

We first omit the nonlinear terms and so consider the function:

$$f(x, y) = (C-1)wx + \left(wS - \frac{1}{1+R} - D + (1-w)(1-\epsilon)P + C \right) y - Ck$$

which is a linear function of x and y . We first show that this function is negative at the three corner points of the triangle where the plane (34) intersects the positive octant \mathcal{O} , ignoring the excluded cylinder \mathcal{C} at this stage. The function evaluated at these corner points is given by:

$$\begin{aligned} f(0, 0) &= -Ck \\ f(0, k) &= \left(wS - \frac{1}{1+R} - D + (1-w)(1-\epsilon)P \right) k \\ f(k/w, 0) &= -k \end{aligned}$$

If we assume that:

$$wS - \frac{1}{1+R} - D + (1-w)(1-\epsilon)P < 0 \quad (36)$$

then $f(x, y)$ is negative at all three of the corner points for all $k > 0$. Since it is a linear function of x and y , this implies that it must also be negative on the boundary and the interior of the triangular region generated by these three points.

We have so far ignored the two nonlinear terms in (35), which are both positive. We note that:

$$0 \leq \frac{x}{x+y} \leq 1 \quad \text{for all } (x, y) \in \mathcal{R}$$

and so:

$$0 \leq \frac{wx}{x+y} \leq w \quad \text{for all } (x, y) \in \mathcal{R}$$

Similarly,

$$0 \leq \left(\frac{1}{1+R} \right) \left(\frac{y}{x+y} \right) \leq \left(\frac{1}{1+R} \right) \quad \text{for all } (x, y) \in \mathcal{R}$$

Thus, these two nonlinear terms are uniformly bounded above. Now, the corner points of the linear function $f(x, y)$ scale with k , and so, $f(x, y)$ will scale with k for all x and y . Hence, when adding the bounded nonlinear terms to $f(x, y)$, the resulting function will be negative over the whole of the triangular region for sufficiently large k . Hence, the vector field points inside the invariant region over the whole of the plane (34) intersected with the positive octant provided that (36) holds and k is chosen sufficiently large, and the bounded region (which excludes the cylinder \mathcal{C}) is then invariant. For any given initial conditions, k must also be chosen sufficiently large so that the initial conditions are contained within the invariant region. The solution must therefore be bounded for all time.

We next consider the condition (36) in more detail. Suppose that $(1-\epsilon)P < S$. Then, when $w \in (0, 1)$, we have that:

$$wS + (1-w)(1-\epsilon)P \in ((1-\epsilon)P, S)$$

Thus, if $(1-\epsilon)P - 1/(1+R) - D < 0$ ($w = 0$), then by the intermediate value theorem, there exists a small $w > 0$ such that (36) holds also. A similar argument holds if $(1-\epsilon)P > S$ with the condition $S - 1/(1+R) - D < 0$. Thus, the optimum value of $wS + (1-w)(1-\epsilon)P$ to choose to be as small as possible is S if $S < (1-\epsilon)P$ or $(1-\epsilon)P$ if $S > (1-\epsilon)P$, which is equivalent to

$\min(S, (1 - \epsilon)P)$, and this gives the condition (33) in the statement of the theorem after multiplying through by $(1 + R)$.

It is a standard result in ode theory that if the function is continuous and satisfies a Lipschitz condition for all $t \geq 0$ and on a given domain, then solutions either exist for all time or blow up in finite time [20]. On the bounded domain described above, Equations (29)–(31) satisfy these conditions, using the same argument as in the proof of Theorem 2, and since, we have shown that the solutions are bounded, if (33) holds, then this implies that the solution exists for all $t \geq 0$. \square

5. Analysis of the New HCV Model

We now consider the steady states and their bifurcations of the new HCV model given by Equations (19)–(21), together with the dynamical properties of the model.

5.1. Assumptions

In our analysis, we make a number of assumptions on the parameters. Rather than having them scattered throughout the text, we now state our main assumptions together for later reference. The key assumptions that we make on the parameters are as follows:

$$\alpha \in (0, 1) \quad (37)$$

$$\epsilon \in [0, 1) \quad (38)$$

$$(B + C)(B - CR) > BCP(1 + R) \quad (39)$$

$$D = \frac{BP}{B + C} \quad (40)$$

The justification for the first two of these conditions was given in Section 3.6, and we assume that these always hold. When we require Assumptions (39) and (40), this will be stated.

Note that if (39) holds, then:

$$B > CR \quad (41)$$

must also hold. Moreover, if (39) holds, then:

$$B + C > CP \left(\frac{B(1 + R)}{B - CR} \right)$$

We note that $B(1 + R) > B > B - CR$, and so, $B(1 + R)/(B - CR) > 1$. It then follows that:

$$B > C(P - 1) \quad (42)$$

5.2. Steady States ($S = 0$)

We first consider the steady states of our revised model (29)–(31) in the special case when $S = 0$. In Section 5.4, we will show how these states are perturbed when $S > 0$.

When $S = 0$, there is a branch of uninfected steady states given by:

$$x_u = 1, \quad y_u = z_u = 0 \quad (43)$$

for all $\epsilon \in [0, 1)$.

Due to the absence of the S term, there is also a state of pure infection given by:

$$x_p = 0, \quad y_p = \frac{1}{1 + D(1 + R)}, \quad z_p(\epsilon) = \frac{(1 - \epsilon)P}{C(1 + D(1 + R))} \quad (44)$$

since, if the only mechanism for cell regeneration is by division and there are no healthy hepatocytes ($x = 0$), then that state will persist as there are no healthy cells to divide. We note that x_p and y_p are constant, but that z_p depends on ϵ .

There is also a branch of infected steady states given by:

$$\begin{aligned}x_i(\epsilon) &= \frac{C(1 + D(1 + R))z_i(\epsilon) - (1 - \epsilon)P}{f(z_i(\epsilon), \epsilon)} \\y_i(\epsilon) &= \frac{z_i(\epsilon)[BCR(1 - \alpha\epsilon)z_i(\epsilon) - (B(1 - \alpha\epsilon) + C)]}{f(z_i(\epsilon), \epsilon)}\end{aligned}$$

where:

$$f(z_i(\epsilon), \epsilon) = B(1 - \alpha\epsilon)[(1 - \epsilon)PR - 1 - D(1 + R)]z_i(\epsilon) - (1 - \epsilon)P \quad (45)$$

and $z_i(\epsilon)$ is a solution of the quadratic equation:

$$\begin{aligned}B^2CR(1 - \alpha\epsilon)^2z^2 + B(1 - \alpha\epsilon)(C(R + D(1 + R)) - B(1 - \alpha\epsilon))z \\+ (1 + R)[B(1 - \alpha\epsilon)(D - (1 - \epsilon)P) + CD] = 0\end{aligned} \quad (46)$$

These solutions will only be valid provided that $f(z_i(\epsilon), \epsilon) \neq 0$ for all $\epsilon \in [0, 1)$. We note that $f(z, \epsilon)$ is quadratic in ϵ and that for all $z > 0$:

$$\begin{aligned}f(z, 1) &= -B(1 - \alpha)[1 + D(1 + R)]z < 0 \\f\left(z, \frac{1}{\alpha}\right) &= \frac{(1 - \alpha)P}{\alpha} > 0\end{aligned}$$

since $\alpha < 1$ (see (37)). Thus, there exists a function $\tilde{\epsilon}(z)$ such that $f(z, \tilde{\epsilon}(z)) = 0$ with $1 < \tilde{\epsilon}(z) < 1/\alpha$ for all $z > 0$, which is clearly outside our range of interest.

The coefficient of ϵ^2 in $f(z, \epsilon)$ is $\alpha BPRz$, which is positive for all $z > 0$, and this implies that the second value of ϵ that gives $f(z, \epsilon) = 0$ must occur for $\epsilon < 1$. In order to avoid the infected steady state branch blowing up in our range of interest, we require that this second root occur for $\epsilon < 0$, and this will be the case provided that $f(z, 0) < 0$ for all $z > 0$. Now:

$$f(z, 0) = B(PR - 1 - D(1 + R))z - P$$

The sign of the z coefficient in this expression is not clear. However, we assume that (40) holds, and so, substituting for D gives:

$$f(z, 0) = -\frac{B(P(B - CR) + B + C)}{B + C}z - P$$

which must be negative for all $z > 0$ using (41). Therefore, assuming that (40) and (41) hold, then $f(z, \epsilon) \neq 0$ for all $\epsilon \in [0, 1)$, and so, the infected branch of solutions does not blow up in our region of interest.

5.3. Bifurcations ($S = 0$)

We now consider bifurcations that occur in our equations with $S = 0$. In particular, there are bifurcations occurring when the branch of infected steady states intersects the uninfected branch and when it intersects the pure infection branch.

5.3.1. Bifurcation on the Uninfected Branch

Lemma 1. *With $S = 0$, there are two bifurcation points that occur on the uninfected branch of steady state solutions (43) that each give rise to a bifurcating branch of infected steady state solutions. One bifurcation occurs for $\epsilon > 1/\alpha > 1$, and so is out of our range of interest. The other bifurcation point occurs at $\epsilon = \epsilon_0 < 1$ where:*

$$\epsilon_0 \begin{cases} < 0 & \text{if } BP < D(B + C) \\ = 0 & \text{if } BP = D(B + C) \\ > 0 & \text{if } BP > D(B + C) \end{cases}$$

The uninfected branch is stable for $\epsilon_0 < \epsilon < 1$ and unstable for $\epsilon < \epsilon_0$.

Proof. The Jacobian matrix $J(x, y, z)$ of Equations (29)–(31) (with $S = 0$) evaluated at the uninfected steady state is given by:

$$J(1, 0, 0) = \begin{bmatrix} -1 & -1 & -(1 - \alpha\epsilon)B \\ 0 & -D & (1 - \alpha\epsilon)B \\ 0 & (1 - \epsilon)P & -C - (1 - \alpha\epsilon)B \end{bmatrix} \quad (47)$$

Clearly, one eigenvalue of this matrix is -1 . We note that the remaining two eigenvalues must be real since, if they were complex, then y would oscillate around the steady state value $y = 0$ and similarly for z . However, by our invariance result (Theorem 1), this is not possible.

A bifurcation point on the branch of uninfected steady states occurs when $\det J(1, 0, 0) = 0$, which gives:

$$\alpha BP\epsilon^2 + (\alpha D - (1 + \alpha)P)B\epsilon + BP - D(B + C) = 0 \quad (48)$$

If $\epsilon = 1 + \mu$, then (48) expressed in terms of μ is given by:

$$\alpha BP\mu^2 + (\alpha D - (1 - \alpha)P)B\mu - D(B(1 - \alpha) + C) = 0 \quad (49)$$

We note that the constant term is negative (using (37)), and the μ^2 coefficient is positive; so the quadratic equation (49) has one positive and one negative solution. This implies that the two solutions of (48) must lie on opposite sides of $\epsilon = 1$, and so, one bifurcation point occurs for $\epsilon = \epsilon_1 > 1$ and the other for $\epsilon = \epsilon_0 < 1$.

A similar analysis with $\epsilon = 1/\alpha + \mu$ also shows that the two solutions in μ must have opposite signs, and this implies that $\epsilon_1 > 1/\alpha > 1$.

Since ϵ_1 is always positive, the sign of the other solution ϵ_0 must match the sign of the constant coefficient of (48) since the coefficient of ϵ^2 is positive, as stated.

To determine the stability of the uninfected branch, we need to know the signs of the three eigenvalues of $J(1, 0, 0)$. We have already noted that one eigenvalue is -1 , and the remaining two eigenvalues are found from the lower 2×2 block in the Jacobian matrix, which we denote by J_1 . We note that:

$$\text{tr}(J_1) = -C - D - (1 - \alpha\epsilon)B < 0$$

since $\alpha\epsilon < 1$ (using (37), (38)) and:

$$\det(J_1) = -\alpha BP\epsilon^2 - (\alpha D - (1 + \alpha)P)B\epsilon - BP + D(B + C)$$

The determinant is negative for $\epsilon < \epsilon_0$ since the coefficient of ϵ^2 is negative. A negative determinant implies that the eigenvalues have the opposite sign, and so, there is one positive and one negative eigenvalue, which implies that the uninfected branch is unstable for $\epsilon < \epsilon_0$.

At the bifurcation point $\epsilon = \epsilon_0$, there is one zero eigenvalue. We note that a double zero eigenvalue is not possible since the other bifurcation point occurs at $\epsilon = \epsilon_1 > 1/\alpha$. Thus, the determinant changes

sign at $\epsilon = \epsilon_0$ and so is positive for $\epsilon_0 < \epsilon < 1$. A positive determinant and negative trace implies that both of the eigenvalues are negative, and so, the uninfected branch is stable in this range. \square

Following our discussion regarding spontaneous clearance in Section 3.5, we note that the stability of the uninfected branch for our model agrees with that shown in Figure 2. We initially assume that the bifurcation occurs precisely at $\epsilon = 0$ (Figure 2b), as this is the point between the two cases discussed. By Lemma 1, this occurs when $BP = D(B + C)$, and solving this for D gives the relation (40).

We now consider the bifurcating branch of infected solutions that arises from the bifurcation point at $\epsilon = 0$.

Lemma 2. *When (40) holds, there is a transcritical bifurcation at $\epsilon = 0$ on the uninfected branch of solutions. The bifurcating branch of infected solutions is given by:*

$$\begin{aligned}x &= 1 - c(BP + B + C)\epsilon + O(\epsilon^2) \\y &= c(B + C)\epsilon + O(\epsilon^2) \\z &= cP\epsilon + O(\epsilon^2)\end{aligned}$$

where:

$$c = \frac{(1 + R)(B + (1 + \alpha)C)}{(B + C)(B - CR) - BCP(1 + R)} \quad (50)$$

This branch is unstable for $\epsilon > 0$ and stable for $\epsilon < 0$.

Proof. To find the bifurcation equation, we solve (29)+(1 + R)(30) and (31) for x and y in terms of z and ϵ , which gives:

$$x = 1 - \left(\frac{BP + B + C}{P}\right)z + zO(z, \epsilon) \quad (51)$$

$$y = \left(\frac{B + C}{P}\right)z + zO(z, \epsilon) \quad (52)$$

Substituting these into (29) gives the bifurcation equation:

$$c_1z\epsilon + c_2z^2 + zO((z, \epsilon)^2) = 0 \quad (53)$$

where:

$$\begin{aligned}c_1 &= -BP^2(1 + R)(B + (1 + \alpha)C) \\c_2 &= BP[(B + C)(B - CR) - BCP(1 + R)]\end{aligned}$$

The two quadratic terms are required for the normal form of a transcritical bifurcation [21]. The trivial solution $z = 0$ of (53) corresponds to the uninfected branch. The non-trivial solution gives a low order solution of the bifurcating infected branch, which is given by:

$$z = -\frac{c_1}{c_2}\epsilon + O(\epsilon^2) = cP\epsilon + O(\epsilon^2)$$

Substituting this back into (51) and (52) gives the stated expansions for x and y in terms of ϵ .

With a transcritical bifurcation, the stability of the bifurcating branches is the opposite of the trivial solution, and so, since the uninfected branch is stable for $\epsilon > 0$ (see Lemma 1), the bifurcating branch must be unstable for $\epsilon > 0$ and stable for $\epsilon < 0$. \square

Clearly, Lemma 2 can be generalised to the case of a bifurcation point occurring at $\epsilon = \epsilon_0$ when (40) does not hold.

We note that only one half of the bifurcating branch of infected solutions lies within our invariant region \mathcal{R} , and this is the half for which $c\epsilon > 0$, since then, $y, z \geq 0$. In this case, we also have $x \leq 1$, as we would expect. Referring back to Figure 2, we would like the valid half of the bifurcating branch to occur for $\epsilon > 0$, and this occurs provided that $c > 0$. Clearly the numerator of c is positive, and Assumption (39) ensures that the denominator is also positive so that $c > 0$. We then have precisely the situation sketched in Figure 2b.

Now, if condition (40) is perturbed slightly, then the bifurcation point will occur for ϵ either positive or negative. If it occurs for $\epsilon < 0$, then the uninfected steady state with no treatment ($\epsilon = 0$) is stable, but with a nearby unstable steady state (Figure 2a). Conversely, if the bifurcation point occurs with $\epsilon > 0$, then the uninfected steady state with no treatment is unstable (Figure 2c).

5.3.2. Bifurcation on the Pure Infection Branch

In addition to the bifurcation involving the uninfected and infected branches of solutions, there is also a second bifurcation where the infected branch intersects the pure infection branch.

Lemma 3. *With $S = 0$, there are two bifurcation points that occur on the pure infection branch of steady state solutions (44) that each give rise to a bifurcating branch of infected steady state solutions. One bifurcation occurs for $\epsilon > 1/\alpha > 1$ and so is out of our range of interest. The other bifurcation point occurs at $\epsilon = \epsilon_2 < 1$ where:*

$$\epsilon_2 \begin{cases} < 0 & \text{if } BP < CD(1+R)(1+D(1+R)) \\ = 0 & \text{if } BP = CD(1+R)(1+D(1+R)) \\ > 0 & \text{if } BP > CD(1+R)(1+D(1+R)) \end{cases}$$

The pure infection branch is stable for $\epsilon < \epsilon_2$ and unstable for $\epsilon_2 < \epsilon < 1$.

Proof. The bifurcation points can be found by considering the Jacobian matrix derived from Equations (29)–(31) evaluated at the pure infection steady state. This matrix is lower triangular with diagonal entries given by:

$$\begin{aligned} J(x_p, y_p, z_p(\epsilon))_{1,1} &= \frac{1}{y_p} - (1 - \alpha\epsilon)Bz_p(\epsilon) - 1 \\ J(x_p, y_p, z_p(\epsilon))_{2,2} &= -\frac{1}{(1+R)y_p} \\ J(x_p, y_p, z_p(\epsilon))_{3,3} &= -C \end{aligned} \quad (54)$$

Since the matrix is lower triangular, these diagonal entries are the eigenvalues. Clearly, the second and third eigenvalues are negative. The bifurcation points therefore occur when $J(x_p, y_p, z_p(\epsilon))_{1,1} = 0$. Substituting for y_p and $z_p(\epsilon)$ from (44) gives the quadratic equation in ϵ :

$$p(\epsilon) = -BP\alpha\epsilon^2 + BP(1+\alpha)\epsilon + CD(1+R)(1+D(1+R)) - BP = 0 \quad (55)$$

If $\epsilon = 1 + \mu$, then (55) expressed in terms of μ is given by:

$$p(1+\mu) = -BP\alpha\mu^2 + BP(1-\alpha)\mu + CD(1+R)(1+D(1+R)) = 0$$

Since the constant term is positive and the μ^2 coefficient is negative, this quadratic equation has solutions of opposite sign, which implies that the two solutions of (55) lie on opposite sides of $\epsilon = 1$. Clearly, the bifurcation point $\epsilon = \epsilon_3 > 1$ is out of our range of interest, and so, we now consider the other bifurcation point at $\epsilon = \epsilon_2 < 1$.

A similar analysis with $\epsilon = 1/\alpha + \mu$ also shows that the two solutions in μ must have opposite signs, and this implies that $\epsilon_3 > 1/\alpha > 1$.

Since ϵ_3 is always positive, the sign of ϵ_2 will be the opposite of the sign of the constant coefficient in (55), since the ϵ^2 coefficient is negative.

The stability of the pure infection steady state is determined by the sign of $J(x_p, y_p, z_p(\epsilon))_{1,1}$ since the other two eigenvalues of $J(x_p, y_p, z_p(\epsilon))$ are negative. Combining the two terms in (54) gives a rational function with a positive denominator. The sign is therefore determined by the numerator, which is the expression on the left-hand side of (55). Since $\epsilon_2 < 1 < \epsilon_3$ and the coefficient of ϵ^2 is negative, then clearly, $J(x_p, y_p, z_p(\epsilon))_{1,1} < 0$ for $\epsilon < \epsilon_2$, and so, the pure infection steady state is stable in this range. There is no possibility of a double root occurring as the two roots lie on opposite sides of $\epsilon = 1$, and so, this quadratic must change sign at the bifurcation point. Therefore, the pure infection solution is unstable for $\epsilon_2 < \epsilon < 1$ as claimed. \square

It is natural to assume that the pure infection state is stable when there is no treatment ($\epsilon = 0$), and from Lemma 3, we see that this occurs provided that:

$$BP - CD(1 + R)(1 + D(1 + R)) > 0 \quad (56)$$

which implies that $\epsilon_2 > 0$. However, we will show in Section 5.4.4 that this is in fact not required for our later analysis of the model. We also note that with the assumption (40), the inequality (56) becomes:

$$(B + C)(B - CR) - BCP(1 + R)^2 > 0$$

Clearly, this is very similar to the inequality in (39), but the fact that the last term is squared implies that the sign of this expression cannot be determined using (39) and so could be positive or negative.

We now consider the bifurcating branch of infected solutions that arises from the bifurcation point at $\epsilon = \epsilon_2$.

Lemma 4. *The bifurcating branch of infected solutions arising from the transcritical bifurcation on the pure infection branch at $\epsilon = \epsilon_2$ is given by:*

$$\begin{aligned} x &= -\frac{c_3}{c_4}(\epsilon - \epsilon_2) + O((\epsilon - \epsilon_2)^2) \\ y &= y_p + c_5(\epsilon - \epsilon_2) + O((\epsilon - \epsilon_2)^2) \\ z &= z_p(\epsilon_2) + \left(c_6 - \frac{P}{C}y_p\right)(\epsilon - \epsilon_2) + O((\epsilon - \epsilon_2)^2) \\ &= z_p(\epsilon) + c_6(\epsilon - \epsilon_2) + O((\epsilon - \epsilon_2)^2) \end{aligned} \quad (57)$$

where:

$$\begin{aligned} c_3 &= B \left((1 - \alpha\epsilon_2) \frac{Py_p}{C} + \alpha z_p(\epsilon_2) \right) \\ c_4 &= (1 - \alpha\epsilon_2) B z_p(\epsilon_2) \left((1 - \alpha\epsilon_2) B \left(\frac{1}{C} - (1 + R) z_p(\epsilon_2) \right) - \frac{R}{y_p} \right) \\ c_5 &= -((1 - \alpha\epsilon_2) B (1 + R) y_p z_p(\epsilon_2) - 1) \frac{c_3}{c_4} \\ c_6 &= z_p(\epsilon_2) \left((1 - \alpha\epsilon_2) B \left(\frac{1}{C} - (1 + R) z_p(\epsilon_2) \right) + \frac{1}{y_p} \right) \frac{c_3}{c_4} \end{aligned}$$

This branch is stable for $\epsilon > \epsilon_2$ and unstable for $\epsilon < \epsilon_2$.

Proof. Since the pure infection steady state has $x = 0$, we will derive the bifurcation equation in terms of x and $\epsilon - \epsilon_2$. To do this, we solve Equations (30) and (31) for y and z in terms of x and $\epsilon - \epsilon_2$, which gives:

$$y = y_p + ((1 - \alpha\epsilon_2)B(1 + R)y_p z_p(\epsilon_2) - 1)x + xO(x, (\epsilon - \epsilon_2)) \quad (58)$$

$$z = z_p(\epsilon_2) + z_p(\epsilon_2) \left((1 - \alpha\epsilon_2)B \left((1 + R)z_p(\epsilon_2) - \frac{1}{C} \right) - \frac{1}{y_p} \right) x - \frac{P}{C}y_p(\epsilon - \epsilon_2) + O((x, (\epsilon - \epsilon_2))^2) \quad (59)$$

Substituting these into (29) gives the bifurcation equation:

$$c_3x(\epsilon - \epsilon_2) + c_4x^2 + xO((x, (\epsilon - \epsilon_2))^2) \quad (60)$$

where:

$$c_3 = B \left((1 - \alpha\epsilon_2) \frac{Py_p}{C} + \alpha z_p(\epsilon_2) \right)$$

$$c_4 = (1 - \alpha\epsilon_2)Bz_p(\epsilon_2) \left((1 - \alpha\epsilon_2)B \left(\frac{1}{C} - (1 + R)z_p(\epsilon_2) \right) - \frac{R}{y_p} \right)$$

The two quadratic terms are required for the normal form of a transcritical bifurcation [21]. The trivial solution $x = 0$ of (60) corresponds to the pure infection branch. The non-trivial solution gives a low order solution of the bifurcating infected branch, which is given by (57). Substituting for x back into (58) and (59) gives the stated expansions for y and z in terms of $(\epsilon - \epsilon_2)$. We also note from the pure infection steady states (44) that:

$$z_p(\epsilon_2) - \frac{P}{C}y_p(\epsilon - \epsilon_2) = z_p(\epsilon)$$

and this is used to derive the stated second form of z .

The stability of the bifurcating branches at a transcritical bifurcation is the opposite of the trivial solution, and so, since the pure infection branch is stable for $\epsilon < \epsilon_2$ (see Lemma 3), the bifurcating branch must be unstable for $\epsilon < \epsilon_2$ and stable for $\epsilon > \epsilon_2$. \square

We note that $c_3 > 0$ using (37) and the fact that $\epsilon_2 < 1$ (see Lemma 3). However, c_4 could be positive or negative depending on the values of the parameters in the model. Since x cannot be negative, we see from (57) that:

- if $c_4 < 0$, then $x(\epsilon)$ has a positive slope, and valid solutions exist locally only for $\epsilon \geq \epsilon_2$;
- if $c_4 > 0$, then $x(\epsilon)$ has a negative slope, and valid solutions exist locally only for $\epsilon \leq \epsilon_2$.

Furthermore, if $c_4 > 0$, then:

$$(1 - \alpha\epsilon_2)B \left(\frac{1}{C} - (1 + R)z_p(\epsilon_2) \right) > \frac{R}{y_p}$$

and this implies that:

$$c_6 > \frac{z_p(\epsilon_2)(1 + R)c_3}{c_4y_p}$$

and so, $c_6 > 0$ also. However, if $c_4 < 0$, then c_6 could be positive or negative.

The solutions of Equations (29)–(31) near the bifurcation point are sketched in Figure 4 for the three possible cases associated with different signs of c_4 and c_6 .

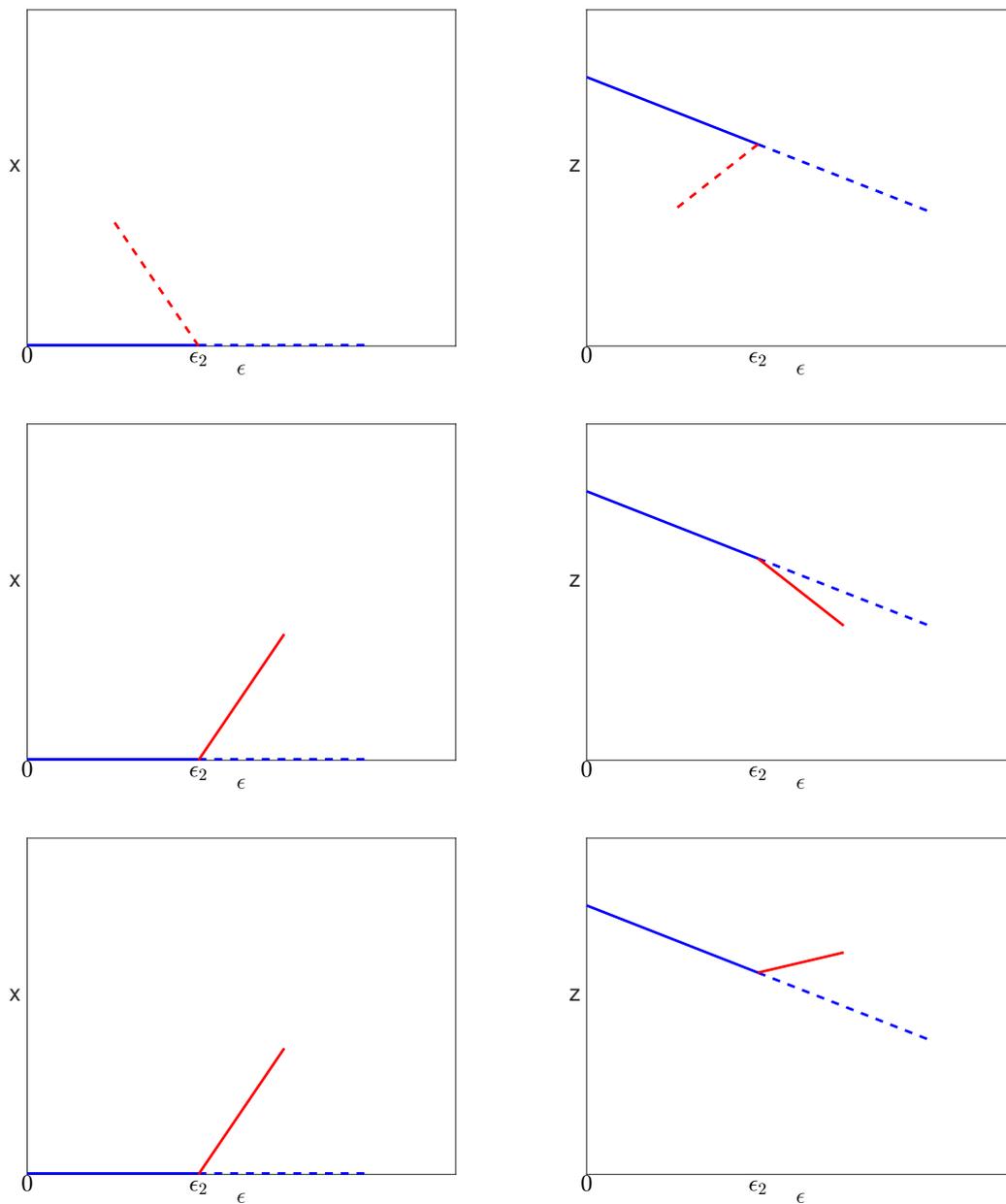


Figure 4. Solutions near the bifurcation involving the infected (red) and pure infection (blue) branches. Solid lines indicate stable solutions, while dashed lines represent unstable solutions. **Top:** $c_4 > 0$. **Middle:** $c_4 < 0, c_6 < 0$. **Bottom:** $c_4 < 0, c_6 > 0$.

5.3.3. The Infected Steady State Branch ($S = 0$)

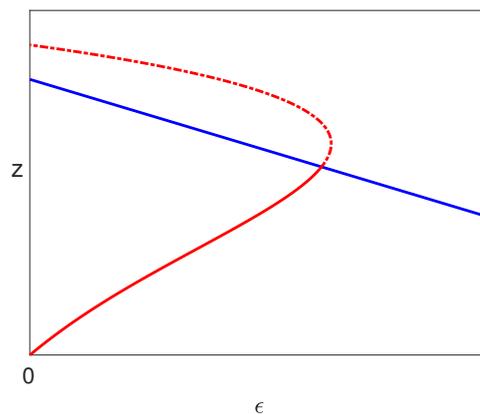
We have found that the infected steady state branch intersects both the uninfected and the pure infection branches of steady states at transcritical bifurcation points. We next consider what happens to the infected branch in between these two bifurcations.

Theorem 4. We assume that (39) and (40) hold and that:

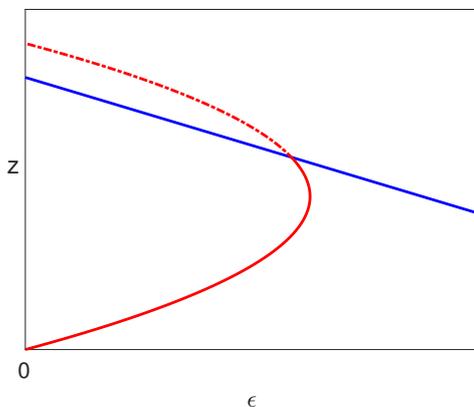
$$R > \frac{\sqrt{2}-1}{2} \tag{61}$$

Then, the bifurcation point at $\epsilon = 0$ on the uninfected branch of solutions and the bifurcation point at $\epsilon = \epsilon_2$ on the pure infection branch of solutions are connected by a continuous branch of infected steady state solutions. There is a single limit point on this branch of infected solutions, which occurs between the two bifurcations when $c_4 < 0$ or after the bifurcation on the pure infection branch when $c_4 > 0$. The limit point occurs at $\epsilon = \epsilon_4$ where $\epsilon_2 < \epsilon_4 < 1/\alpha$. The three possible cases, associated with different signs of the coefficients c_4 and c_6 , are shown in Figure 5.

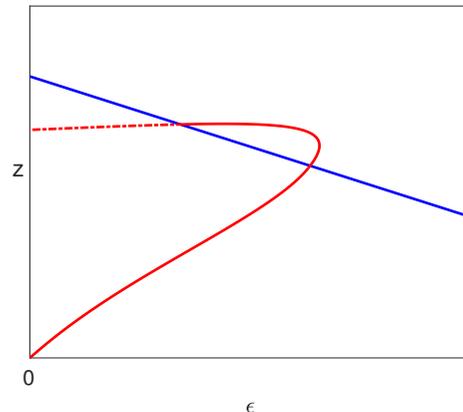
Remark 1. We conjecture that Theorem 4 holds for all $R > 0$, so that condition (61) is unnecessary. However, we have been unable to prove this. Some evidence to support this conjecture is included in the proof. Condition (61) is not overly restrictive though, as it implies that $r_I/r_T < 2/(\sqrt{2} + 1) = 0.8284$.



(a) The infected branch for $c_4 > 0$.



(b) The infected branch for $c_4 < 0, c_6 < 0$.



(c) The infected branch for $c_4 < 0, c_6 > 0$.

Figure 5. The pure infection (blue) and infected (red) steady state branches. Note that the dashed-dotted lines indicate invalid solutions.

Proof. Substituting for D using (40) into (46), we obtain an equation involving z and ϵ to be solved for the infected branch of solutions given by:

$$g(z, \epsilon) = BCR(B + C)(1 - \alpha\epsilon)^2 z^2 + (BCP(1 + R) + (B + C)(CR - B(1 - \alpha\epsilon)))(1 - \alpha\epsilon)z + \epsilon P(1 + R)((B + C)(1 - \alpha\epsilon) + \alpha C) = 0 \tag{62}$$

To find limit points on the infected steady state branch [21], we must solve the two equations:

$$g(z, \epsilon) = g_z(z, \epsilon) = 0$$

and check the non-degeneracy conditions:

$$g_{zz}(z, \epsilon) \neq 0, \quad g_\epsilon(z, \epsilon) \neq 0 \quad (63)$$

Since $g(z, \epsilon)$ is quadratic in z , then clearly, $g_z(z, \epsilon)$ is linear in z , and the z coefficient is strictly positive. Thus, the second equation can be solved for z and substituted back into the first equation, which gives the quadratic equation in ϵ :

$$A_1 \epsilon^2 + B_1 \epsilon + C_1 = 0 \quad (64)$$

where:

$$\begin{aligned} A_1 &= -\alpha B(B+C)^2(4CPR(1+R) + \alpha B) \\ B_1 &= CP(1+R)(2R(B+(1+\alpha)C) - \alpha B) + \alpha(B+C)(B-CR) \\ C_1 &= -((B+C)(B-CR) - BCP(1+R))^2 \end{aligned}$$

The discriminant of the quadratic is given by:

$$\Delta = 16BC^2PR(1+R)(B+C)^2\Delta_1\Delta_2 \quad (65)$$

where:

$$\Delta_1 = BP(1+R) + \alpha(B-CR) \quad (66)$$

$$\Delta_2 = (B+C)(\alpha B + R(B+C)) - \alpha BCP(1+R) \quad (67)$$

We have assumed that (39) holds, and this implies (41) also; this gives $\Delta_1 > 0$. We also note that Δ_2 can be expressed as:

$$\Delta_2 = ((B+C)^2 - \alpha BCP)R + \alpha B(B+C - CP)$$

Assumption (39) implies that (42) holds, which implies that the second term is positive. Moreover,

$$\begin{aligned} (B+C)^2 - \alpha BCP &> (B+C)CP - \alpha BCP \quad \text{using (42)} \\ &= CP(B(1-\alpha) + C) \\ &> 0 \end{aligned}$$

using (37), and this implies that $\Delta_2 > 0$ also. Therefore, $\Delta > 0$, and so, the quadratic Equation (64) will have two distinct solutions. We therefore expect to have two limit points on the infected solution branch(es). Now, $g_{zz}(z, \epsilon) = 2BCR(B+C)(1-\alpha\epsilon)^2 > 0$ using (37), (38), and so, the first non-degeneracy condition in (63) is satisfied. However, it is not easy to verify the second non-degeneracy condition. We now use a different approach to get further information regarding the infected solution branches, which will also confirm the existence of two limit points.

To get a more complete picture of the solutions, we substitute:

$$z = z_0 + \frac{z_1 + \delta z}{\tilde{\epsilon}_1 + \delta \epsilon}, \quad \epsilon = \tilde{\epsilon}_0 + \tilde{\epsilon}_1 + \delta \epsilon \quad (68)$$

into (62). The parameters $z_0, z_1, \tilde{\epsilon}_0$ and $\tilde{\epsilon}_1$ can be solved for in terms of the parameters so that our equation reduces to:

$$\frac{h_1 \delta \epsilon^2}{h_0} - \frac{h_2 \delta z^2}{h_0} = 1 \tag{69}$$

where:

$$\begin{aligned} z_0 &= \frac{1}{2CR}, & z_1 &= \frac{BP(1+R) + R(B+C)}{2\alpha BR(B+C)}, \\ \tilde{\epsilon}_0 &= \frac{1}{\alpha}, & \tilde{\epsilon}_1 &= -\frac{C[\alpha R(B+C) + P(1+R)(2R(B+C(1-\alpha)) + \alpha B)]}{\alpha(B+C)(\alpha B + 4CPR(1+R))}, \\ h_0 &= \frac{P(1+R)\Delta_1\Delta_2}{4BRh_1}, & h_1 &= \frac{\alpha(B+C)(\alpha B + 4CPR(1+R))}{4CR}, & h_2 &= \alpha^2 BCR(B+C) \end{aligned}$$

Clearly, $h_1, h_2 > 0$, and we showed above that $\Delta_1, \Delta_2 > 0$; so, $h_0 > 0$ also. Thus, (69) is the equation of a hyperbola, and the two solution branches are given in parametric form by:

$$\delta z(\beta) = \sqrt{\frac{h_0}{h_2}} \sinh \beta, \quad \delta \epsilon(\beta) = \pm \sqrt{\frac{h_0}{h_1}} \cosh \beta$$

These solutions exist for all δz and for $|\delta \epsilon| \geq \sqrt{h_0/h_1}$. This gives rise to parametric solutions of (62) given by:

$$z(\beta) = z_0 + \frac{z_1 + \delta z(\beta)}{\tilde{\epsilon}_1 + \delta \epsilon(\beta)}, \quad \epsilon(\beta) = \tilde{\epsilon}_0 + \tilde{\epsilon}_1 + \delta \epsilon(\beta) \tag{70}$$

Limit points on these branches occur when:

$$\frac{d\epsilon}{dz} = \frac{d\epsilon/d\beta}{dz/d\beta} = 0$$

Now, $\beta = 0$ is the unique solution of the equation $d\epsilon/d\beta = 0$ and:

$$\left. \frac{dz}{d\beta} \right|_{\beta=0} = \frac{\sqrt{h_0/h_2}}{\tilde{\epsilon}_1 \pm \sqrt{h_0/h_1}}$$

For this derivative to be finite, we clearly require $\tilde{\epsilon}_1 \pm \sqrt{h_0/h_1} \neq 0$. Now, $\tilde{\epsilon}_1 < 0$, and so, $\tilde{\epsilon}_1 - \sqrt{h_0/h_1} < 0$; however, it is possible that $\tilde{\epsilon}_1 + \sqrt{h_0/h_1} = 0$. However, we will show later that the right-hand (+) branch is outside our range of interest, and so, it does not matter whether this quantity is zero or non-zero. Thus, on the left-hand (-) branch, we have:

$$\left. \frac{dz}{d\beta} \right|_{\beta=0} = \frac{\sqrt{h_0/h_2}}{\tilde{\epsilon}_1 - \sqrt{h_0/h_1}} \neq 0$$

and so, $d\epsilon/dz = 0$ when $\beta = 0$. This point will be a quadratic limit point provided that the non-degeneracy condition $d^2\epsilon/dz^2|_{\beta=0} \neq 0$ is satisfied. It is a matter of calculation to show that for the left-hand branch:

$$\left. \frac{d^2\epsilon}{dz^2} \right|_{\beta=0} = \left. \frac{d^2\epsilon/d\beta^2}{(dz/d\beta)^2} \right|_{\beta=0} = -\sqrt{\frac{h_2}{h_0 h_1}} \left(\tilde{\epsilon}_1 - \sqrt{h_0/h_1} \right)^2 \neq 0$$

again using the fact that $\tilde{\epsilon}_1 - \sqrt{h_0/h_1} < 0$, and so, we do indeed have a quadratic limit point when $\beta = 0$. Thus, the limit point on the left-hand branch of solutions occurs at:

$$z(0) = z_0 + \frac{z_1}{\tilde{\epsilon}_1 - \sqrt{h_0/h_1}}, \quad \epsilon(0) = \tilde{\epsilon}_0 + \tilde{\epsilon}_1 - \sqrt{h_0/h_1}$$

and two solutions exist for each $\epsilon < \epsilon(0)$, which is also confirmed by the negative second derivative.

We note that there are no solutions of (69) when $\delta\epsilon = 0$, and this corresponds to $\epsilon = \tilde{\epsilon}_0 + \tilde{\epsilon}_1$. We now show that $\epsilon_2 < \tilde{\epsilon}_0 + \tilde{\epsilon}_1 < 1/\alpha$. The right-hand inequality is clearly satisfied since $\tilde{\epsilon}_0 = 1/\alpha$ and $\tilde{\epsilon}_1 < 0$. To verify the left-hand inequality, we consider the quadratic equation (55). We have already shown in the proof of Lemma 3 that the two solutions ϵ_2 and ϵ_3 of this equation satisfy $\epsilon_2 < 1 < \epsilon_3$. Since the coefficient of ϵ^2 is negative, this implies that the quadratic function $p(\epsilon)$ is positive if and only if $\epsilon_2 < \epsilon < \epsilon_3$. If $p(\tilde{\epsilon}_0 + \tilde{\epsilon}_1) > 0$, then this implies that $\epsilon_2 < \tilde{\epsilon}_0 + \tilde{\epsilon}_1$ as required. It can be shown that:

$$p(\tilde{\epsilon}_0 + \tilde{\epsilon}_1) = \frac{BPC(a_3P^3 + a_2P^2 + a_1P + a_0)}{\alpha(B+C)^2(\alpha B + 4CPR(1+R))^2} \quad (71)$$

where $a_0, a_1, a_3 > 0$, using (37) and (41). The remaining coefficient, a_2 , is given by:

$$a_2 = C(1+R)^2(b_2B^2 + b_1B + b_0)$$

where $b_0, b_1 > 0$ and:

$$b_2 = (8\alpha^2 - 8\alpha + 4)R^2 + 4\alpha^2R - \alpha^2$$

Now, b_2 is negative for $\alpha > 0$ and sufficiently small R . Substituting $R = (\sqrt{2} - 1)/2 + \tilde{R}$ into b_2 gives:

$$b_2 = (8\alpha^2 - 8\alpha + 4)\tilde{R}^2 + \frac{4(2\sqrt{2} - 1)}{7}(7\alpha^2 + (2\sqrt{2} - 6)\alpha + 3 - \sqrt{2})\tilde{R} + (3 - 2\sqrt{2})(1 - \alpha)^2$$

The first and second order coefficients of \tilde{R} are positive, and the constant coefficient is non-negative for all $\alpha \in [0, 1]$. Thus, $b_2 > 0$ for $\tilde{R} > 0$ and for all $\alpha \in [0, 1]$, or equivalently, for all $R > (\sqrt{2} - 1)/2$. In this case, $a_2 > 0$ also, which then implies that $p(\tilde{\epsilon}_0 + \tilde{\epsilon}_1) > 0$, as required.

Thus, we have proved that $p(\tilde{\epsilon}_0 + \tilde{\epsilon}_1) > 0$ using Condition (61). However, even if $b_2 < 0$, there are still many other positive terms in $p(\tilde{\epsilon}_0 + \tilde{\epsilon}_1)$, and so, it may be the case that Condition (61) is not necessary. We note that the minimum of b_2 in the region $R \geq 0, \alpha \in [0, 1]$ occurs at $R = 0, \alpha = 1$. Substituting $R = 0, \alpha = 1$ into the cubic polynomial in P in (71) gives:

$$(a_3P^3 + a_2P^2 + a_1P + a_0)|_{R=0, \alpha=1} = B^2(1+P)(B+C-CP)$$

Using assumption (42), this is positive. Moreover, expanding the cubic polynomial in a Taylor series about the point $(R, \alpha) = (0, 1)$, we find that the first order terms are positive, and so, for sufficiently small $R > 0$ and $\alpha < 1$, the cubic polynomial in P will be positive. However, this does not guarantee that it is positive for all $R > 0$ and $\alpha \in [0, 1]$, although we conjecture that this is in fact the case.

Now, the two branches of solutions (70) occur on either side of the gap in ϵ , and so, one branch exists for $\epsilon < \tilde{\epsilon}_0 + \tilde{\epsilon}_1$ and the other for $\epsilon > \tilde{\epsilon}_0 + \tilde{\epsilon}_1$. The left-hand branch is therefore the only branch of infected solutions for $\epsilon < \tilde{\epsilon}_0 + \tilde{\epsilon}_1$, and so, the two bifurcating branches of infected solutions arising from the two bifurcations described above must be part of this single branch.

Thus, there is a single branch of infected solutions that connects the two bifurcation points, as claimed. There must also be a limit point on this branch, which occurs for $\epsilon > \epsilon_2$.

The only steady state solutions with $x = 0$ are the uninfected or pure infection solutions. Thus, all other solutions, in particular the infected solutions, have $x \neq 0$. In the case of $c_4 > 0$ (Figure 4 (top)), the only way that the valid infected branches arising from the two bifurcations can connect on a curve with a single limit point is for the limit point to occur on the invalid solutions after the bifurcation on the pure infection branch, as shown in Figure 5a. When $c_4 < 0$ (Figure 4 (middle, bottom)), the valid infected solution emanates from the pure infection branch to the right of the bifurcation point, and the only way that this can connect to the bifurcation on the uninfected branch at $\epsilon = 0$ is for there to be a limit point in between the two bifurcations.

We note that the left branch of (69) exists for all δz , but does not exist for all z . As $\delta\epsilon \rightarrow -\infty$ ($\beta \rightarrow \pm\infty$), the left branch of the hyperbola asymptotes to the straight lines $\delta z = \pm\sqrt{h_1/h_2} \delta\epsilon$. Substituting these into $z(\beta)$ given by (70), we see that z converges to the constant values:

$$z = z_0 \pm \sqrt{h_1/h_2} = \frac{1}{2CR} \left(1 \pm \sqrt{1 + 4CPR(1 + R)/(\alpha B)} \right)$$

as $\epsilon \rightarrow -\infty$. Clearly one of these asymptotes is negative and the other positive, as we would expect. \square

Finally, we recall the assumption we made in Theorem 4 that D is defined by (40). If this is not the case, then the same method described above can be used to derive Equation (69), and $z_0, \bar{\epsilon}_0, h_1$ and h_2 are unchanged. However, the other coefficients now involve the parameter D . In this case, it is not possible to determine the sign of the coefficient h_0 in (69). As long as h_0 remains positive, the same picture as described above will hold qualitatively. However, if h_0 changes sign and becomes negative, then the structure of the solutions of (69) changes, so that there are no limit points occurring. Therefore, if (40) does not hold, then there is an extra condition $h_0 > 0$ that is required to give the same solution structure.

5.4. Steady States ($S > 0$)

We have studied in detail the solutions of our model in the special case when $S = 0$, which is when there is no production of hepatocytes from stem cells. When S is small and positive, this will result in a small perturbation of these solutions, and we now consider this case.

5.4.1. Bifurcation on the Uninfected Branch

When $S > 0$, the uninfected steady state (43) and the bifurcation point that is found by solving (48) between the uninfected and infected steady state branches both remain the same. Therefore, the condition (40) again ensures that this bifurcation point occurs at $\epsilon = 0$.

5.4.2. Bifurcation on the Pure Infection Branch

When $S > 0$, the state of pure infection no longer exists since there is a second mechanism for generating healthy hepatocytes due to stem cell production. Thus, for small values of S , the bifurcation involving the infected and pure infection branches will unfold. There are two possible ways that a transcritical bifurcation can unfold, depending on the sign of the perturbation term.

When $S > 0$, the bifurcation equation (60) in x simply gains an extra term Sy where y is given by (58) and so is given by:

$$h(x, \delta\epsilon) = S(y_p + O(x, x^2, x\delta\epsilon)) + c_3x\delta\epsilon + c_4x^2 + xO(x, \delta\epsilon) \tag{72}$$

where $\delta\epsilon = \epsilon - \epsilon_2$. The constant term is the only term needed to determine the unfolding of a transcritical bifurcation [21]. Limit points for this equation are found by solving $h(x, \delta\epsilon) = h_x(x, \delta\epsilon) = 0$ and are then given by:

$$x^2 = \frac{Sy_p}{c_4} + O(S^2)$$

Since $S > 0$, we see that for sufficiently small S :

- if $c_4 < 0$, then no limit points exist;
- if $c_4 > 0$, then there are two limit points at $x = \pm\sqrt{Sy_p/c_4}$.

We recall that $c_3 > 0$, and so, the sign of c_4 also determines the slope of the bifurcating branch (see Figure 4). The unfolding of the bifurcation in these two cases is shown in Figure 6. We note that when $S > 0$, in both cases, one of the branches locally has $x < 0$, and one has $x > 0$; so, there is only one valid branch of solutions.

When $\epsilon_2 > 0$, the unfolding of the solutions in Figure 5 is shown in Figure 7.

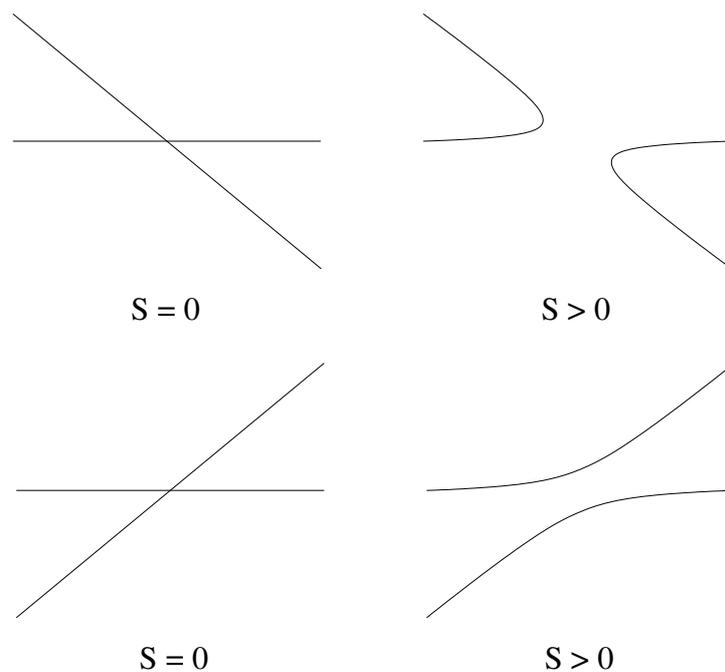
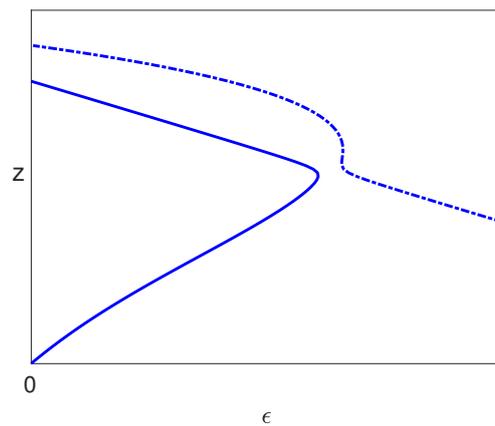


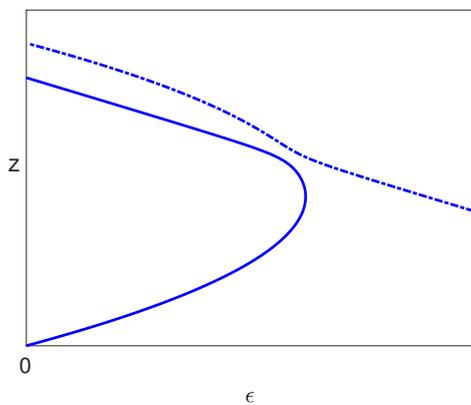
Figure 6. Unfolding of the transcritical bifurcation given by (72) in the (ϵ, x) plane for $c_4 > 0$ (top) and $c_4 < 0$ (bottom).

5.4.3. Infected Branch of Solutions

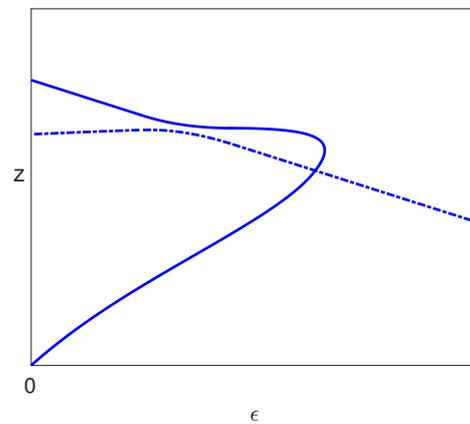
For the infected solution branch away from the bifurcation points, the implicit function theorem implies that a small perturbation in the equations due to $S > 0$ results in an $O(S)$ perturbation in the infected solutions, and so, there is only a small perturbation in this branch of solutions. Thus, we see that in all cases, the form of the infected steady states is qualitatively similar since the branch of infected states emanates from the trivial solution at the bifurcation point at $\epsilon = 0$ and increases with ϵ , goes round a limit point and then decreases, as shown in Figure 7. Moreover, this is the only valid branch of infected steady state solutions.



(a) Steady state solutions for $c_4 > 0$.



(b) Steady state solutions for $c_4 < 0, c_6 < 0$.



(c) Steady state solutions for $c_4 < 0, c_6 > 0$.

Figure 7. Steady state solutions where the bifurcation between the infected and pure infection branches has been unfolded as $S > 0$, assuming that $\epsilon_2 > 0$. The dashed-dotted lines are invalid solutions as $x < 0$.

5.4.4. The Case of $\epsilon_2 < 0$

Finally, we consider one more situation, namely when condition (56) does not hold. Thus, we now assume that:

$$BP - CD(1 + R)(1 + D(1 + R)) < 0 \tag{73}$$

This implies that $\epsilon_2 < 0$ by Lemma 3. and so, the bifurcation point on the pure infection branch with $S = 0$ occurs for negative ϵ . The infected branch of solutions still has a limit point in this case, but there is no connection between these two branches in our range of interest given by $\epsilon \in [0, 1)$. Moreover, the pure infection steady state with no treatment ($\epsilon = 0$) is unstable, since Condition (56) was required to ensure stability. In this case, the unfolding of the bifurcation when $S > 0$ occurs in the same way, but outside our range of interest, and the pure infection branch becomes invalid. The solutions in this case are shown in Figure 8. We note that the valid solution branch is similar in this case as in the previous cases considered, with a single limit point occurring.

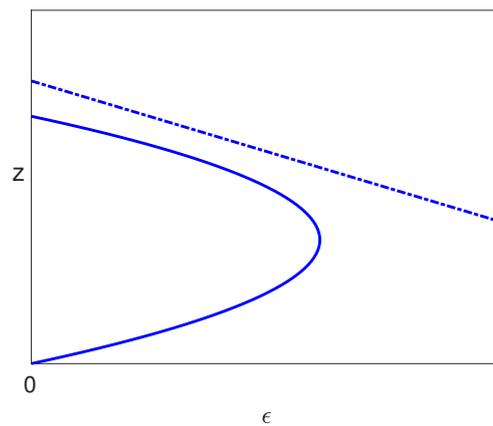


Figure 8. The steady state solutions when (73) holds and $S > 0$. The dashed-dotted line consists of invalid solutions as $x < 0$.

6. Stability

All the cases considered above, with Assumptions (39) and (40), give rise, when $S > 0$, to an uninfected steady state branch $x = 1, y = z = 0$ together with a branch of infected steady state solutions that bifurcates from $\epsilon = 0$ and has a single limit point. We have also seen that the uninfected branch of solutions is stable for $\epsilon \in (0, 1)$, and this implies that the bifurcating branch of infected solutions initially has one unstable eigenvalue together with two stable eigenvalues. The only way that the infected branch can become stable again is if the unstable eigenvalue passes back through zero, which would correspond to a limit point. We have also ascertained that there is precisely one limit point on the branch of infected solutions (for sufficiently small S). Since the solutions past the limit point are stable when $S = 0$, they will still be stable for small $S > 0$, and so, the one unstable eigenvalue must pass through zero at the limit point to give a stable branch of solutions. Thus, the bifurcation diagram for $S > 0$ (and sufficiently small) is as shown in Figure 9a.

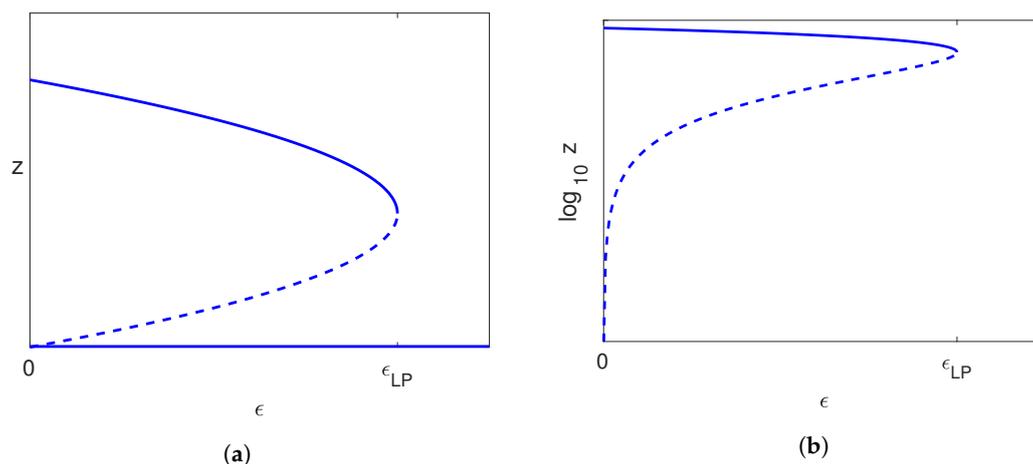


Figure 9. The bifurcation diagram, where solid lines indicate stable solutions and dashed lines indicate unstable solutions. Note that the vertical scale is either (a) z or (b) $\log_{10} z$.

The other possibility is that the two stable eigenvalues along the unstable section of the infected branch could collide and become complex and then cross the imaginary axis in a Hopf bifurcation. They would then have to cross back again in a reverse Hopf bifurcation before the limit point in order for the solution to stabilise after going round the limit point. However, to determine whether or not

such Hopf bifurcations occur analytically from the model is a very challenging problem. It is also of little interest, since any bifurcating periodic orbits would be unstable.

7. Comparison with the Neumann/Dahari Models

We now compare the predictions from our new model of HCV infection with those of the Neumann and Dahari models that we reviewed in Section 2. Our model is similar to previous models in that it involves the three variables T , I and V . However, the steady state solutions, and therefore the dynamics, of the new model are quite different from previous models in several ways. These differences result in different predictions for the dynamics of the infection during treatment and suggest possible different treatment regimes.

A typical bifurcation diagram for the Neumann/Dahari models is shown in Figure 10a, which we compare with the bifurcation diagram for our model, as shown in Figure 9a. We also note that experimental data are plotted with $\log_{10} V$ on the vertical axis rather than V , and so, we also show the bifurcation diagrams plotted for $\log_{10} z$ in Figures 9b and 10b. We also note from (26) that $\log_{10} V = \log_{10} z + k$ where $k = \log_{10}(r_T T_{\max}/(r_T + d))$, and so, the non-dimensionalisation only results in a shift on the vertical axis.

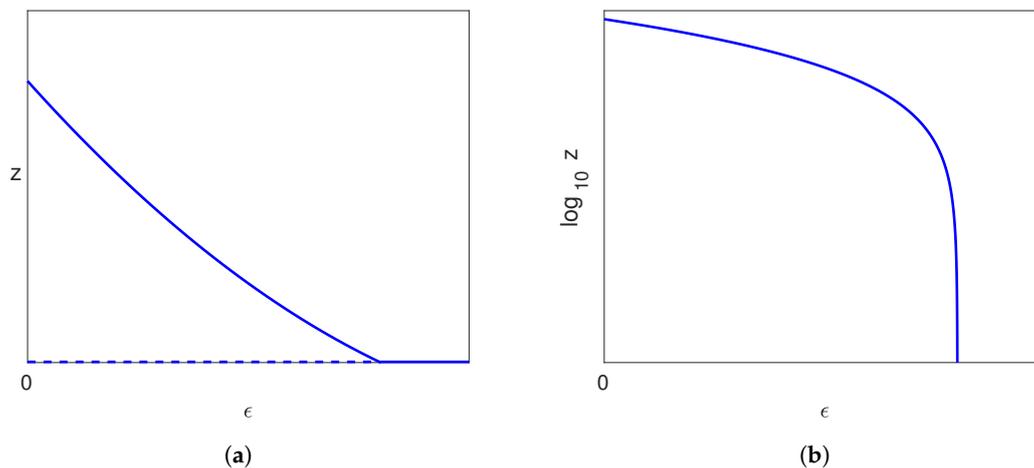


Figure 10. A typical bifurcation diagram for the Neumann/Dahari models, where solid lines indicate stable solutions and dashed lines indicate unstable solutions. Note that the vertical scale is either (a) z or (b) $\log_{10} z$.

We now make a number of comparisons between these models.

- For the Neumann/Dahari models, treatment will only be effective once the treatment factor ϵ exceeds a critical value determined by the bifurcation point, regardless of the viral load when treatment commences. For our model, if the viral load is close to the infected steady state before treatment starts, then similarly, the treatment factor ϵ must exceed the critical value ϵ_{LP} determined by the limit point for the treatment to be effective. However, if the infection is caught and treated in the early stages, while the viral load is still relatively low, then our model predicts that a lower drug dose, with a corresponding smaller value of the treatment parameter ϵ , will be effective.
- As mentioned previously, once treatment is stopped, the prediction of the Neumann/Dahari models is that the infection will take hold again unless the infected hepatocytes and virus have been completely eliminated during treatment. The prediction from our model, if the bifurcation point on the uninfected solution branch occurs at a negative value of ϵ ($\epsilon_0 < 0$), is that the body will be able to eliminate a small amount of infected hepatocytes and virus cells without further

treatment once their levels have been reduced sufficiently. On the other hand, if the bifurcation point on the uninfected branch occurs at a positive value of ϵ ($\epsilon_0 > 0$), then our model predicts in this case that the infection will take hold again on cessation of treatment unless the infected hepatocytes and virus cells have been completely eliminated during treatment. However, our model also predicts that continuing with a low level of drug treatment, corresponding to a small value of ϵ , will stop the infection recurring in this case.

- The Neumann/Dahari models suggest that treatment will only be effective if the treatment parameter ϵ is greater than the critical value during the whole period of treatment, which is the way that patients are generally treated in practice. Our model suggests that the drug dose could be reduced as treatment progresses and that this will still be effective, provided that it is not reduced too far too quickly. If this is indeed the case, it could save some of the costs of treatment, and a lower drug dosage may also mean a reduction in side effects, which would benefit the patient.

8. Description of Observed Viral Load Profiles

The usual approach with a new model is to fit it to data in order to show that there are values of the model parameters that give a good fit to these data. This is a useful approach, but does not give any insight into the mechanisms involved in the different cases. Thus, we now consider many of the observed behaviours of the viral load under treatment that are reported in the literature and show how our model can explain these observations. This also helps to explain the possible mechanism associated with the observations in some cases. In the next section, we then fit our model to four datasets to show that this can also be done in practice.

A number of different viral load profiles were reported in [5], and we first consider these. However, we also consider various other observations in the literature. In this section, we define $\epsilon^* \in (0, 1)$ to be the value of the treatment parameter in the model, which may of course vary for different patients.

8.1. Sustained Virologic Response

Sustained virologic response is where the viral load rapidly decreases and is undetectable at the completion of treatment at 24 weeks [5]. For our model, this is easily realised by having $\epsilon_{LP} < \epsilon^*$ so that the trajectory in our model converges to the only available uninfected steady state, which is stable provided that $\epsilon_0 < \epsilon^*$. To see this rapid one phase decline, it is likely that treatment commenced not too long after infection, since a delay in treatment is often associated with biphasic or triphasic decline of the viral load (see Section 8.6). Once treatment is stopped, the viral load remains undetectable, and this will be the case provided that $\epsilon_0 < 0$, or equivalently $BP < D(B + C)$ (see Lemma 1 and the discussion in Section 3.5).

8.2. Relapse

Relapse is similar to sustained virological response (SVR) during treatment, in that there is a rapid decline in viral load. However, once treatment stops, the patient relapses as the viral load increases back to pre-treatment levels. This relapse occurs if $\epsilon_0 > 0$ as discussed in Section 3.5. In this case, continuing with a lower drug dose may be sufficient to keep the viral load under control.

8.3. Partial Virologic Response

There is a partial virologic response (PVR) if an initial decrease in viral load is followed by an increase during treatment. This could be explained by our model if $\epsilon^* < \epsilon_{LP}$ and with the viral load quite high before the start of treatment. In this case, the trajectory of the model would converge to the infected steady state at ϵ^* , and it is quite possible that it could initially overshoot this steady state and then return back to it.

8.4. Breakthrough

Breakthrough is similar to PVR, except that at some point during treatment, the viral load is undetectable before increasing again during treatment. Thus, the mechanism would be similar to that described for PVR, but with the infected steady state occurring at a lower viral load, so that there is a larger initial drop in the viral load, below the level of detection, before increasing again back up to the infected steady state.

8.5. Null Response

Some patients show no significant reduction in viral load under treatment. This could occur when $\epsilon^* < \epsilon_{LP}$ and where the infected steady state at ϵ^* is similar to that at $\epsilon = 0$. This could occur when $c_4 < 0$ and $c_6 > 0$, as shown in Figure 7c.

8.6. Biphasic and Triphasic Decline

A common observation of the viral load during treatment is that there is biphasic decline [3], although triphasic decline, with a flat region between two declining phases, has also been observed [3,4]. We now show that both of these patterns of decline in the viral load under treatment can occur in our model.

We assume that S , the parameter associated with the generation of hepatocytes due to stem cells, is small (and positive). When there is no treatment ($\epsilon = 0$), the infected steady state is then given by:

$$\begin{aligned}x &= \frac{C}{BP - CD(1 + R)(1 + D(1 + R))}S + O(S^2) \\y &= y_p + O(S) \\z &= z_p(0) + O(S)\end{aligned}$$

where y_p and $z_p(0)$ are the pure infection steady states given by (44). We assume that (56) holds so that this solution is both valid (as $x > 0$) and stable.

If we assume that a patient has been infected for a long time before treatment, then this implies that at the start of treatment, $x(0)$, $y(0)$ and $z(0)$ will be close to these steady state values. We therefore now assume that $x = O(S)$, and so, we set $x = S\tilde{x}$. Substituting for x in (29)–(31) and taking only the leading order terms in S gives the reduced equations:

$$\tilde{x}' = y + \frac{\tilde{x}}{y} - \tilde{x} - (1 - \alpha\epsilon)B\tilde{x}z \quad (74)$$

$$y' = \frac{1}{1 + R}(1 - y) - Dy \quad (75)$$

$$z' = (1 - \epsilon)Py - Cz \quad (76)$$

We note that equation (75) is a linear equation in y with stable steady state $y = y_p$. Since we are assuming that $y(0)$ is close to this steady state, dynamically, y will continue to converge towards the steady state and is not influenced by either x or z .

Equation (76) has the stable steady state $z = z_p(\epsilon)$. Once treatment starts with $\epsilon = \epsilon^*$, this steady state drops from $z_p(0)$ to $z_p(\epsilon^*) = (1 - \epsilon^*)z_p(0)$. Since y is approximately constant, this implies at the start of treatment that z will decay exponentially towards the new steady state value. If we assume that $y = y_p$, then the solution of (76) is:

$$z = z_p(\epsilon^*) + (z(0) - z_p(\epsilon^*))e^{-Ct} \quad (77)$$

This is the observed first phase of rapid decline in the viral load. We note that Equation (74) can be written as:

$$\tilde{x}' = y + \frac{\tilde{x}}{y} - \tilde{x} - B\tilde{x}z + (\alpha B\tilde{x}z)\epsilon$$

Thus, the effect of increasing ϵ at the start of treatment is to increase \tilde{x}' , and this will result in \tilde{x} increasing also.

Once z has dropped to close to the steady state $z_p(\epsilon^*)$, and assuming that y is also close to y_p , then $y' = O(S)$ and $z' = O(S)$, and so, y and z evolve on a slow time scale. Thus, they will remain approximately constant for a time of $O(1/S)$. This is the flat middle phase in the triphasic decline.

Eventually, the $O(S)$ terms in the y and z equations will result in y and z being displaced from their leading order steady states, and the third phase decline towards the uninfected steady state will start. This rate of decline is determined by the eigenvalue of the Jacobian matrix evaluated at the uninfected steady state given in (47) that is closest to zero. Clearly, one eigenvalue is $\lambda_1 = -1$, and this is unlikely to be the closest to zero. The other two eigenvalues are found from the lower 2×2 matrix. The characteristic equation, which has to be solved for these eigenvalues, is given by:

$$p(\lambda) = \lambda^2 + ((1 - \alpha\epsilon)B + C + D)\lambda - BP(1 - \epsilon)(1 - \alpha\epsilon) + D((1 - \alpha\epsilon)B + C) = 0 \quad (78)$$

Assuming that $\epsilon^* > \epsilon_0$, then the uninfected steady state is stable, and so, the two solutions of this characteristic equation are both negative. Suppose that these solutions are $-\lambda_1$ and $-\lambda_0$ with $-\lambda_1 < -\lambda_0 < 0$. In this case, the rate of decline to the uninfected steady state is determined by the eigenvalue closest to zero, which is $-\lambda_0$. The third phase decline is generally observed to be slower than the first phase, and this will be the case if $C > \lambda_0$. Now, $p(\lambda) < 0$ if and only if $-\lambda_1 < \lambda < -\lambda_0$. Thus, if $p(-C) < 0$, then this implies that $-C < -\lambda_0$ as required. Now:

$$p(-C) = -(1 - \alpha\epsilon^*)B(C + (1 - \epsilon^*)P - D)$$

Clearly, this is negative if $C + (1 - \epsilon)P > D$, and so, this is the condition that ensures that the third phase decline is slower than the first phase.

Thus, we have shown that our new model can exhibit triphasic decline of the viral load, as is seen in the data. This analysis also suggests that this pattern of decline is associated with late commencement of treatment so that the initial conditions for the model are close to the infected steady state. This is also the conclusion reached by Dahari et al. [3].

Biphasic decline is similar to triphasic decline, but where the middle flat phase is very short. The first order term in S in Equations (75) and (76) is $\pm S(1 - \alpha\epsilon^*)B\tilde{x}z$, and so, anything that increases the magnitude of this term will reduce the length of the middle phase. This includes a larger value of S or B or a larger value of $z_p(\epsilon^*)$. We note that an increase in B will also result in a more rapid increase in x , which will also help to shorten this phase. Thus, our model can also exhibit biphasic decline of the viral load.

8.7. Initial Increase in Viral Load

Hsu et al. [22] reported that in some patients, there is an initial increase in viral load when treatment is started and that this initial increase is associated with a higher likelihood of achieving SVR. They used different models to investigate this behaviour and concluded that a modification of the Neumann model gave the best fit provided that $c = \delta$ and $\eta = 1$, which implies that the treatment always provides a complete block on de novo infection. Guedj et al. [23] questioned these unrealistic assumptions and the analysis in [22] and suggested that the initial increase in viral load was due to the infected steady state not having been reached so that viral loads were increasing before the start of therapy. They also suggested that the correlation between the initial increases and SVR was an indication of the effectiveness of the therapy. Rong and Perelson [24] also considered this effect in their multiscale model for direct acting antiviral agents. Their model always showed a small initial increase

in viral load, even when starting from steady state. Of course, this result cannot be directly compared with the experimental results of Hsu et al. [22] since these results were obtained for patients treated with interferon (IFN) and ribavirin (RBV) only.

To understand this effect from our model, we express (31) as:

$$z' = Py - Cz - Bxz - \epsilon(Py - \alpha Bxz) \quad (79)$$

Without treatment ($\epsilon = 0$), we expect z to be increasing towards the infected steady state, and so $z' > 0$, which implies that:

$$Py - Cz - Bxz > 0 \quad (80)$$

We note that:

$$Py - \alpha Bxz = (Py - Cz - Bxz) + (Cz + (1 - \alpha)Bxz) > 0$$

assuming that (80) holds and using (37).

If we assume that the infected steady state (x_i, y_i, z_i) has been reached before the start of treatment, then at $t = 0$, we have:

$$z'(0) = -\epsilon^*(Py_i - \alpha Bx_i z_i) = -\epsilon(Cz_i + (1 - \alpha)Bxz)$$

which is negative, and so, we have an immediate decline in viral load. However, if the steady state has not been reached before treatment commences, then with $(x, y, z) = (x(0), y(0), z(0))$ at the start of treatment, we have:

$$z'(0) = Py(0) - Cz(0) - Bx(0)z(0) - \epsilon^*(Py(0) - \alpha Bx(0)z(0))$$

which is the difference of two positive terms, and so could be negative or positive. Larger values of ϵ^* are more likely to give $z'(0) < 0$, and so, the effectiveness of treatment does not explain the correlation between the initial increase and SVR, as suggested in [23]. We propose an alternative explanation for this correlation.

Soon after infection, the viral load will be increasing rapidly while y and z are both quite small, and so, the effect of the treatment will be small also, so that $z'(0) > 0$, thus giving an initial increase in viral load. However, if the start of treatment is delayed and the system is getting towards the infected steady state, then $z'(0)$ will be quite small before treatment commences. However, near the steady state, $Py(0)$ will be relatively large, so that the net effect gives $z'(0) < 0$ once treatment starts. Thus, according to this model, an initial increase in viral load is associated with early initiation of treatment, which is likely to correlate with a higher rate of SVR, as reported in [22].

8.8. Direct Acting Antiviral Agents

Recently, new direct acting antiviral (DAA) agents have become available, which are more effective than treatment with IFN and RBV alone. If DAAs are used as a monotherapy, then it is found that drug-resistant virus cells quickly form, rendering the treatment ineffective; so, they are used in combination with IFN and RBV, and this combination is found to be highly effective [25]. More complex models for the action of DAAs have been proposed [26]. It has been found that treatment that includes DAAs has notable differences compared to treatment with only IFN and RBV, which include (i) a more rapid and longer first phase decline and (ii) a more rapid second phase decline [25].

We note that the action of a DAA is essentially to further block viral replication, although by different mechanisms than the older drugs [27]. The simplest way that this can be modelled is to increase the parameter ϵ in (31). If it is also assumed that the DAAs do not have any additional effect in reducing the rate of production of infected cells (parameter η in Equations (19)–(21)), then an increase in ϵ must be accompanied by a corresponding reduction in the parameter α given in (27), in order to

keep $\eta = 1 - \alpha\epsilon > 0$ constant. We claim that this simple change is sufficient to explain both of the above observed effects, which we now justify.

- (i) It has been noted that the first phase decline when treating with DAAs is both longer and faster than when using IFN and RBV. Since the rate of the first phase decline is essentially given by the parameter C (or c for the original equations), it has been suggested that both ϵ and c should be increased in the models for treatment with DAAs [25]. However, we claim that an increase in ϵ alone is sufficient to produce a longer and faster first phase. To see this, we consider the decline in viral load during the first phase that is given in (77). For this solution, we find that the initial rate of decline is:

$$\begin{aligned} \left. \frac{d(\log_{10} z)}{d\tau} \right|_{\tau=0} &= - \left(1 - \frac{z_p(\epsilon^*)}{z_p(0)} \right) C \log_{10} e \\ &= -\epsilon^* C \log_{10} e \end{aligned}$$

since $z_p(\epsilon^*) = (1 - \epsilon^*)z_p(0)$, and so, the initial slope increases as the treatment factor ϵ^* increases, with maximum slope only being achieved when $\epsilon^* = 1$ (which corresponds to $z_p(\epsilon^*) = 0$). Clearly, the steady state $z_p(\epsilon^*)$ is also reduced as ϵ^* is increased. These two effects result in an increase in the length of the decline in the viral load together with a more rapid decline. This is illustrated in Figure 11.

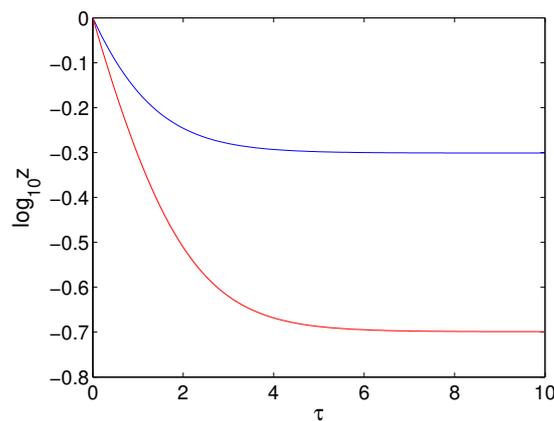


Figure 11. The decline in viral load given by (77) with $z_p(0) = 1$, $C = 1$ and $z_p(\epsilon^*) = 0.5$ (blue) or $z_p(\epsilon^*) = 0.2$ (red).

- (ii) The second observation made for DAAs is that the second phase is also faster than that for treatment with IFN and RBV. We have seen in Section 8.6 that for our model, the rate of decay in the second phase is proportional to $e^{-\lambda_0 t}$ where $-\lambda_0 < 0$ is the solution of the characteristic equation (78) that is closest to zero. With our assumption that $\eta = 1 - \alpha\epsilon$ is constant, (78) becomes:

$$p(\lambda(\epsilon)) = \lambda(\epsilon)^2 + (\eta B + C + D)\lambda(\epsilon) - \eta B P(1 - \epsilon) + D(\eta B + C) = 0 \tag{81}$$

Differentiating this equation with respect to ϵ and solving for $\lambda'(\epsilon)$ gives:

$$\lambda'(\epsilon) = - \frac{\eta B P}{(\eta B + C + D + 2\lambda(\epsilon))}$$

and evaluating at $\epsilon = \epsilon^*$ gives:

$$\lambda'(\epsilon^*) = - \frac{\eta B P}{(\eta B + C + D - 2\lambda_0)} \tag{82}$$

for $\lambda(\epsilon^*) = -\lambda_0$. The sign of the denominator is not clear. However, substituting $\lambda = -(\eta B + C + D)/2$ into $p(\lambda)$ given by (81) gives:

$$p\left(-\frac{1}{2}(\eta B + C + D)\right) = -\left((1 - \epsilon)\eta BP + \frac{1}{4}(\eta B + C - D)^2\right) < 0$$

using (38) and the assumption that $\eta > 0$. The quadratic coefficient of $\lambda(\epsilon)$ in (81) is positive, and so, $p(\lambda(\epsilon))$ is only negative between the two roots $-\lambda_1$ and $-\lambda_0$, which implies that:

$$-\lambda_1 < -\frac{1}{2}(\eta B + C + D) < -\lambda_0$$

It follows from this that the denominator of (82) is positive, and so, $\lambda'(\epsilon^*) < 0$. Hence, if ϵ is increased from ϵ^* , then the eigenvalue $-\lambda_0$ will decrease to first order, thereby increasing the rate of the second phase decline.

Thus, the observed effects of direct acting antiviral agents, in association with IFN and RBV, can be included in our model by simply keeping $1 - \alpha\epsilon$ constant and reducing ϵ .

9. Data Fitting

We claimed in Section 8 that our new model is capable of generating many of the observed profiles of viral load under treatment. We now fit our model to some data in order to demonstrate that this is indeed the case.

For this data fitting, we use the model equations in dimensional form given by (19)–(21). We first note that it is not possible to fit all of the parameters in the model. In particular, the parameters η and β always occur in the combination $(1 - \eta)\beta$, and so, during treatment, we can only hope to determine this single quantity from the data. Similarly, ϵ and p occur in the combination $(1 - \epsilon)p$, and so again, we can only determine this single quantity during treatment.

We also observe that the five parameters r_T , r_I , T_{\max} , d and δ occur only in the four groups $r_T T_{\max}$, $r_I T_{\max}$, $r_T + d$, $r_I + \delta$, and so, it will not be possible to identify all five of these parameters from the data. Instead of using these four parameter groups, we use $r_T T_{\max}$, $r_T + d$, R and D , where R and D are non-dimensional parameters defined in (32). We write Equations (19)–(21) in terms of these parameters as:

$$\dot{T} = sI + \frac{r_T T_{\max} T}{T + I} - (r_T + d)T - \beta^* VT \quad (83)$$

$$\dot{I} = \frac{1}{1 + R} \left(\frac{r_T T_{\max} I}{T + I} - (r_T + d)I \right) - D(r_T + d)I + \beta^* VT \quad (84)$$

$$\dot{V} = p^* I - cV - \beta^* VT \quad (85)$$

where:

$$\beta^* = (1 - \eta)\beta, \quad p^* = (1 - \epsilon)p$$

Values of the parameter groups $r_T T_{\max}$ and $r_T + d$ will be found by fitting the model to data. We note that by including the non-dimensional parameter R , and requiring it to be positive, ensures that the conditions $0 < r_I/r_T < 1$ are satisfied. Moreover, requiring $D > 0$ is consistent with the conditions $r_I < r_T$ and $d < \delta$, but does not guarantee that they hold. We also note that it is not possible to determine whether Condition (39) holds since the parameter β cannot be determined. The sign of ϵ_0 , the value of ϵ at which the bifurcation on the uninfected branch occurs (see Lemma 1), cannot be determined either since β and p cannot be determined.

The initial value $V(0)$ will be taken from the data and the other initial values, $T(0)$ and $I(0)$, will be regarded as unknown parameters. Thus, there is a total of 10 parameters and initial values to be found by fitting the model Equations (83)–(85) to the data, all of which are required to be positive.

These parameters are found using the method of least squares in which the sum of the squares of the differences between the log of the viral load data points and the log of the viral load predicted by the model at the given time points is minimised. This task was performed using MATLAB with the differential equations described in terms of the log of each of the variables.

We considered various datasets that show partial virologic response (PVR), breakthrough, null response and triphasic behaviour. The first three datasets are taken from [5], while the last one is taken from [6]. We fitted our model to the data during treatment only. The data together with the fitted curve $V(t)$ and the predicted curves $T(t)$ and $I(t)$ from the model for each case are shown in Figure 12. The parameter values used are given in Table 1. These should not be regarded as definitive parameter values since very similar fits to the viral load data can be found using quite different sets of parameter values.

We did not consider data that shows sustained virologic response (SVR) since these datasets typically contain only two data points before the viral load goes below the lower limit of quantification (see [5]), and there will be many parameter combinations that will fit the two data points. The model can of course show this behaviour for appropriate parameter values.

For the breakthrough data, we note that the three lowest data points are recorded as 50 IU/mL, which is the lower limit of quantification, and so, the actual values of the viral load will be lower than this. However, we do not know the correct values of the viral load, and so, for the data fitting, we kept the outer two of these values, but ignored the middle value.

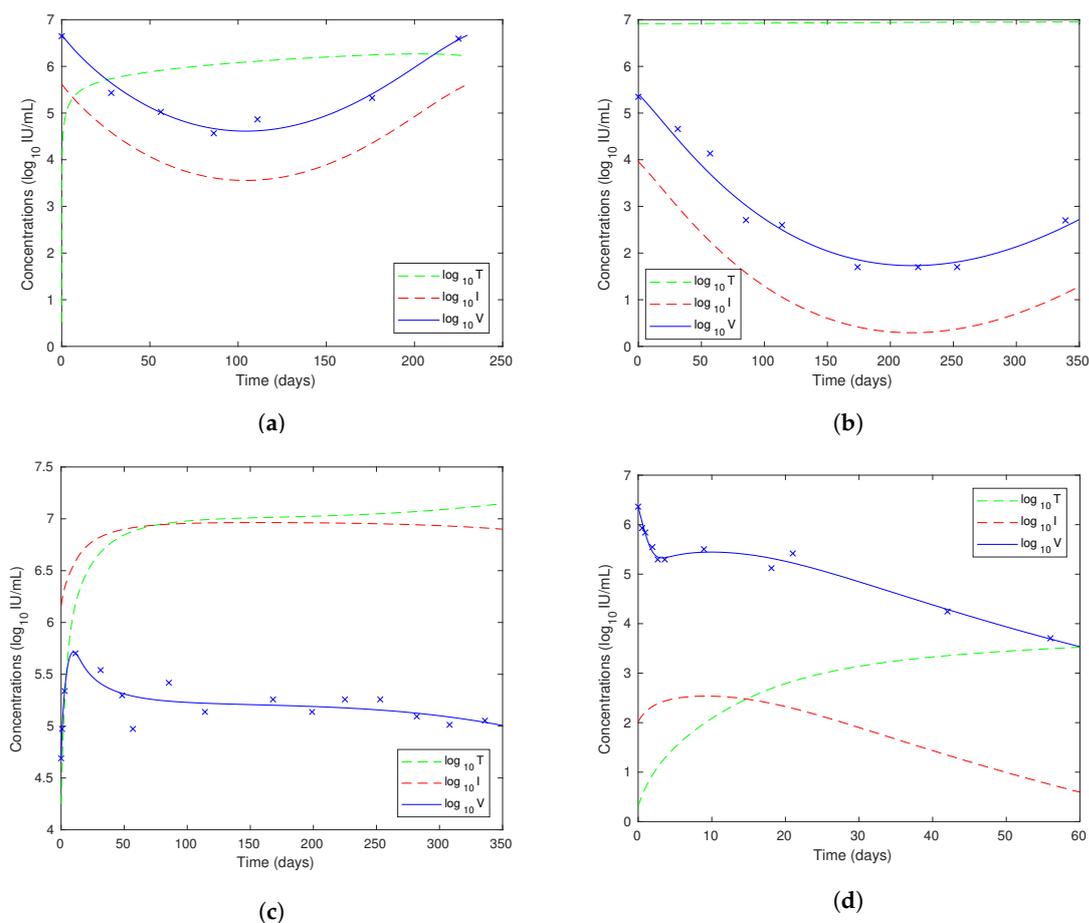


Figure 12. Plots of the viral load V (blue), healthy hepatocytes T (green) and infected hepatocytes I (red) against time fitted to the viral load datasets for (a) partial virologic response (PVR), (b) breakthrough, (c) null response, (d) triphasic.

We now make a number of observations regarding these results.

- In the PVR, breakthrough and triphasic cases, we have $I(t) < V(t)$ for all t , but in the null response case, $I(t)$ is significantly higher than $V(t)$.
- The initial viral load is highest in the PVR case, and we predicted in Section 8.3 that PVR would be associated with a high initial viral load.
- In the null response case, it is interesting to observe that the fitted viral load V and infected hepatocyte concentration I both reduce towards zero, but very slowly. The start of the decline in these variables can be observed from around 300 days in Figure 12c. At 1000 days after the start of treatment, the predicted values are $\log_{10} I(1000) = 5.1202$ and $\log_{10} V(1000) = 2.8515$.
- In Section 8.6, we stated that triphasic decline might be expected when the patient has been infected for a long time before treatment, which means that the viral load will be high while the healthy hepatocyte concentration will be very low. This is precisely the situation observed in Figure 12d.
- In Section 8.7, we showed that an initial increase in viral load at the start of treatment is possible, and we see this in the null response case.
- We saw in Section 3.1 that the regeneration rate for a healthy liver is $1.15 \times 10^{-2} \text{ day}^{-1}$ for females and $1.11 \times 10^{-2} \text{ day}^{-1}$ for males. This is effectively the parameter $r_T + d$ in our new model. The value of this parameter when fitted to data is a little lower than this for the PVR and breakthrough cases and is slightly higher in the null response and triphasic cases.
- The condition (33) for the solution to be bounded for all $t \geq 0$ in terms of the parameters we are using here is given by:

$$\min(s, p^*)(1 + R) - (r_T + d)(1 + D(1 + R)) < 0$$

The term on the LHS for each of the fits to the data is given by: PVR: -0.1790 ; breakthrough: -2003 ; null response: -0.1753 ; triphasic: -0.2500 . All of these values are negative, which ensures that each of the solutions exists and is bounded for all time by Theorem 3.

Table 1. Parameter values for fitting the model to the data.

	PVR	Breakthrough	Null Response	Triphasic
$s \text{ (day}^{-1}\text{)}$	1.1178×10^{-1}	1.5104×10^{-4}	4.6260×10^{-3}	3.1259×10^{-3}
$r_T T_{\max} \text{ (IU/ml/day)}$	1.0645×10^4	2.8556×10^4	1.2920×10^6	1.1149×10^2
$r_T + d \text{ (day}^{-1}\text{)}$	1.9927×10^{-3}	2.9890×10^{-3}	3.8518×10^{-2}	1.7882×10^{-2}
R	3.0078×10^1	1.1686×10^3	2.6011	2.0350×10^{-1}
D	5.8954×10^1	5.7302×10^2	1.1064	1.0962×10^1
$\beta^* \text{ (ml/IU/day)}$	8.3376×10^{-9}	7.1149×10^{-9}	1.9493×10^{-7}	3.3281×10^{-8}
$p^* \text{ (day}^{-1}\text{)}$	2.0396×10^2	9.4025×10^1	3.4868×10^{-2}	1.1646×10^3
$c \text{ (day}^{-1}\text{)}$	1.7908×10^1	3.3659	2.7784×10^{-4}	1.4294
$T(0) \text{ (IU/ml)}$	3.3246	8.2935×10^6	1.7755×10^4	1.9948
$I(0) \text{ (IU/ml)}$	4.1752×10^5	9.5880×10^3	1.4523×10^6	1.0355×10^2

10. Conclusions

We have proposed a new mathematical model of HCV infection, which involves the same three variables (concentrations of healthy and infected hepatocytes and of virions) as the earlier models of Neumann and Dahari, but which has a significantly different structure to the steady state solutions. The typical bifurcation diagram for our model is shown in Figure 9 and consists of an uninfected steady state branch with a bifurcating branch of infected steady state solutions on which there is a single limit point. Allowing the bifurcation point to occur at positive or negative values of the treatment parameter means that the model can include spontaneous clearance, as well as relapse at the end of

treatment. In the case $S = 0$ (no generation of hepatocytes from stem cells), we showed that there is a pure infection branch of solutions, and an infected steady state branch was found analytically that connects the uninfected and pure infection branches. When $S > 0$, the bifurcation between the pure infected and infected branches unfolds, generating a single valid branch of infected solutions on which there is a limit point (see Figure 7). We have been able to describe these solutions of the model using only the assumption (39) and the requirement that the quantity $BP - D(B + C)$ is close to zero.

We have shown in Section 8 that our model is able to show the many profiles of viral load reported in the literature, and the model has been fitted to four datasets in Section 9. Moreover, based on the bifurcation diagram shown in Figure 9 for our model, we also made some predictions regarding treatment (see Section 7), which we summarise as follows:

- If the infection is caught and treated in the early stages, then our model predicts that a lower drug dose may be effective in eliminating the infection.
- If the viral load relapses on cessation of treatment, then continuing with a low level of drug treatment may keep the viral load low.
- The infected branch from the bifurcation on the uninfected branch to the limit point has a positive slope, and this suggests that the drug dose could be reduced as treatment progresses, which could save some of the costs of treatment and give a reduction in side effects for the patient.

The Neumann HCV model was obtained by modifying earlier models for HBV and HIV infections. It would now be interesting to see whether this process could be reversed by adapting this new model for HCV infection for other viral infections.

Another avenue of interest would be to fit this model to viral load data while a patient is on treatment in order to make patient-specific recommendations from the model regarding the future treatment plan. An important step in this process would be to determine the feasibility of estimating each of the parameters in the model, as has been done for other HCV models [28].

Acknowledgments: I am grateful to many colleagues for helpful conversations regarding mathematical modelling of HCV infection, including Piet van der Graaf, Mark Ransley, Edwin Rwemigabo, Hien Tran, Tom Banks, Joseph Arthur, Aluisio Seguardo, Cássia Mendes Corrêa, João Renato Rebello Pinho and Bernadette Moore.

Conflicts of Interest: The author declares no conflict of interest.

References

1. World Health Organization. *Global Hepatitis Report 2017*; World Health Organization: Geneva, Switzerland, 2017. [[CrossRef](#)]
2. Neumann, A.U.; Lam, N.P.; Dahari, H.; Gretch, D.R.; Wiley, T.E.; Layden, T.J.; Perelson, A.S. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon- α therapy. *Science* **1998**, *282*, 103–107. [[CrossRef](#)]
3. Dahari, H.; Lo, A.; Ribeiro, R.M.; Perelson, A.S. Modelling hepatitis C virus dynamics: Liver regeneration and critical drug efficacy. *J. Theor. Biol.* **2007**, *247*, 371–381. [[CrossRef](#)]
4. Dahari, H.; Ribeiro, R.M.; Perelson, A.S. Triphasic decline of hepatitis C virus RNA during antiviral therapy. *Hepatology* **2007**, *46*, 16–21. [[CrossRef](#)]
5. Snoeck, E.; Chanu, P.; Lavielle, M.; Jacqmin, P.; Jonsson, E.N.; Jorga, K.; Goggin, T.; Grippo, J.; Jumbe, N.L.; Frey, N. A comprehensive hepatitis C viral kinetic model explaining cure. *Clin. Pharm. Therapeut.* **2010**, *87*, 706–713. [[CrossRef](#)]
6. Herrmann, E.; Lee, J.H.; Marinos, G.; Modi, M.; Zeuzem, S. Effect of ribavirin on hepatitis C viral kinetics in patients treated with pegylated interferon. *Hepatology* **2003**, *37*, 1351–1358. [[CrossRef](#)]
7. Song, X.; Neumann, A.U. Global stability and periodic solution of the viral dynamics. *J. Math. Anal. Appl.* **2007**, *329*, 281–297. [[CrossRef](#)]
8. Dahari, H.; Major, M.; Zhang, X.; Mihalik, K.; Rice, C.M.; Perelson, A.S.; Feinstone, S.M.; Neumann, A.U. Mathematical modeling of primary hepatitis C infection: Noncytolytic clearance and early blockage of virion production. *Gastroenterology* **2005**, *128*, 1056–1066. [[CrossRef](#)]

9. Reluga, T.C.; Dahari, H.; Perelson, A.S. Analysis of hepatitis C virus infection models with hepatocyte homeostasis. *SIAM J. Appl. Math.* **2009**, *69*, 999–1023. [CrossRef]
10. Tanaka, M.; Itoh, T.; Tanimizu, N.; Miyajima, A. Liver stem/progenitor cells: Their characteristics and regulatory mechanisms. *J. Biochem.* **2011**, *149*, 231–239. [CrossRef]
11. Marshall, A.; Rushbrook, S.; Davies, S.E.; Morris, L.S.; Scott, I.S.; Vowler, S.L.; Coleman, N.; Alexander, G. Relation between hepatocyte G1 arrest, impaired hepatic regeneration, and fibrosis in chronic hepatitis C virus infection. *Gastroenterology* **2005**, *128*, 33–42. [CrossRef]
12. Pomfret, E.A.; Pomposelli, J.J.; Gordon, F.D.; Erbay, N.; Price, L.L.; Lewis, W.D.; Jenkins, R.L. Liver regeneration and surgical outcome in donors of right lobe liver grafts. *Transplantation* **2003**, *76*, 5–10.
13. Darwiche, H.; Petersen, B.E. Biology of the adult hepatic progenitor cell: “Ghosts in the machine”. *Prog. Mol. Biol. Transl. Sci.* **2010**, *97*, 229–249. [CrossRef]
14. Katoonizadeh, A.; Nevens, F.; Verslype, C.; Pirenne, J.; Roskams, T. Liver regeneration in acute severe liver impairment: A clinicopathological correlation study. *Liver Int.* **2006**, *26*, 1225–1233. [CrossRef]
15. DebRoy, S.; Bolkerzand, B.M.; Martcheva, M. Bistability and long-term cure in a within-host model of hepatitis C. *J. Biol. Sys.* **2011**, *19*, 533–550. [CrossRef]
16. Turner, R.; Lozoya, O.; Wang, Y.; Cardinale, V.; Gaudio, E.; Alpini, G.; Mendel, G.; Wauthier, E.; Barbier, C.; Alvaro, D.; et al. Human hepatic stem cell and maturational liver lineage biology. *Hepatology* **2011**, *53*, 1035–1045. [CrossRef]
17. Adams, B.M.; Banks, H.T.; Davidian, M.; Kwon, H.D.; Tran, H.T.; Wynne, S.N.; Rosenberg, E.S. HIV dynamics: Modeling, data analysis, and optimal treatment protocols. *J. Comp. Appl. Math.* **2005**, *184*, 10–49.
18. Hepatitis C Fact Sheet. World Health Organization, 2017. Available online: <http://www.who.int/mediacentre/factsheets/fs164/en/> (accessed on 11 April 2018). [CrossRef]
19. Rong, L.; Perelson, A.S. Treatment of hepatitis C virus infection with interferon and small molecule direct antivirals: Viral kinetics and modeling. *Crit. Rev. Immunol.* **2010**, *30*, 131–148.
20. Grimshaw, R. *Nonlinear Ordinary Differential Equations*; Blackwell: Oxford, UK, 1990.
21. Wiggins, S. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2nd ed.; Texts in Applied Mathematics; Springer: New York, NY, USA, 2003, Volume 2. [CrossRef]
22. Hsu, C.S.; Hsu, S.J.; Chen, H.C.; Tseng, T.C.; Liu, C.H.; Niu, W.F.; Jeng, J.; Liu, C.J.; Lai, M.Y.; Chen, P.J.; et al. Associate of IL28B gene variations with mathematical modeling of viral kinetics in chronic hepatitis C patients with IFN plus ribavirin therapy. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3719–3724. [CrossRef]
23. Guedj, J.; Dahari, H.; Perelson, A.S. Understanding the nature of early HCV RNA blips and the use of mathematical modeling of viral kinetics during IFN-based therapy. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E302. [CrossRef]
24. Rong, L.; Perelson, A.S. Mathematical analysis of multiscale models for hepatitis C virus dynamics under therapy with direct-acting antiviral agents. *Math. Biosci.* **2013**, *245*, 22–30. [CrossRef]
25. Dahari, H.; Guedj, J.; Perelson, A.S.; Layden, T.J. Hepatitis C viral kinetics in the era of direct acting antiviral agents and IL28B. *Curr. Hepat. Rep.* **2011**, *10*, 214–227. [CrossRef]
26. Rong, L.; Guedj, J.; Dahari, H.; Coffield, D.J., Jr.; Levi, M.; Smith, P.; Perelson, A.S. Analysis of hepatitis C virus decline during treatment with the protease inhibitor danoprevir using a multiscale model. *PLOS Comp. Biol.* **2013**, *9*, e1002959.
27. TenCate, V.; Sainz, B., Jr.; Cotler, S.J.; Uprichard, S.L. Potential treatment options and future research to increase hepatitis C virus treatment response rate. *Hepat. Med.* **2010**, *2*, 125–145. [CrossRef]
28. Arthur, J.G.; Tran, H.T.; Aston, P.J. Feasibility of parameter estimation in hepatitis C viral dynamics models. *J. Inverse Ill-Posed Probl.* **2017**, *25*, 69–80.

