Journal of Virological Methods 266 (2019) 103–113

Contents lists available at ScienceDirect

Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

The review of differential equation models of HBV infection dynamics

Miaolei Li, Jian Zu*

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, PR China

ARTICLE INFO

Keywords:
Hepatitis B virus
Mathematical model
Antiviral therapy
Pharmacokinetics
Drug resistance

ABSTRACT

Understanding the infection and pathogenesis mechanism of hepatitis B virus (HBV) is very important for the prevention and treatment of hepatitis B. Mathematical models contribute to illuminate the dynamic process of HBV replication in vivo. Therefore, in this paper we review the viral dynamics in HBV infection, which may help us further understand the dynamic mechanism of HBV infection and efficacy of antiviral treatment. Firstly, we introduce a family of deterministic models by considering different biological mechanisms, such as, antiviral therapy, CTL immune response, multi-types of infected hepatocytes, time delay and spatial diffusion. Particularly, we briefly describe the stochastic models of HBV infection. Secondly, we introduce the commonly used parameter estimation methods for HBV viral dynamic models and briefly discuss how to use these methods to estimate unknown parameters (such as drug efficacy) through two specific examples. We also discuss the idea and method of model identification and use a specific example to illustrate its application. Finally, we propose three new research programs, namely, considering HBV drug-resistant strain, coupling within-host and between-host dynamics in HBV infection and linking population dynamics with evolutionary dynamics of HBV diversity.

1. Introduction

Hepatitis B virus (HBV) infection remains a serious public health problem all over the world. When HBV invades the human body, it enters the hepatocyte by means of binding to the receptor. Then, HBV DNA is transported into the nucleus of hepatocyte and converted to cccDNA under the action of host enzymes. The cccDNA is not easily degradable, (Le Mire et al., 2005) which is an important reason why chronic hepatitis B is difficult to cure. Currently available antiviral agents are effective to block the infection by free virus particles, but are still far from ideal to eliminate HBV since the drugs do not directly target cccDNA (Hoofnagle et al., 2007). Besides, the long-term antiviral treatment has led to the problem of drug-resistant strains. That is why it is crucial to understand the infection and pathogenesis mechanism of HBV.

Mathematical models play an important role in understanding and predicting the dynamic process of HBV infection. In the 1990s, Nowak et al. proposed a basic model to analyze the effect of antiviral treatment on reducing viral loads (Nowak et al., 1996). Since then, many different mathematical models have been developed according to different biological mechanisms. Some models studied the interaction between antiviral treatment and HBV by considering drug efficacy or pharmacokinetics (Nowak et al., 1996; Murase and Sasaki, 2005; Eikenberry et al., 2009; Sypsa et al., 2005). Besides, some models investigated the effect of time delay or spatial diffusion, which took into account the time from infection to the release of free virus particles and free movement of virus particles in the liver respectively (Gourley et al., 2008; Huang et al., 2010; Guo and Cai, 2011; Wang et al., 2008). In particular, some models focused on understanding the effect of humoral immunity or CTL-mediated cellular immunity based on the body's defense or immune system (Yousfi et al., 2011; Fisicaro et al., 2009; Bertoletti and Ferrari, 2003; Perelson, 2002). Furthermore, some studies described the interaction of multi-types of infected hepatocytes with different copies of cccDNA (Ciupe et al., 2007; Li et al., 2014). Stochastic models which took into account the random factors were also investigated (Moneim et al., 2009; Xie et al., 2017; Luzyanina and Bocharov, 2014; Liu et al., 2017). These models have greatly enriched our understanding of HBV infection dynamics.

This paper aims to provide a quantitative reference for clinical treatment of hepatitis B virus. The structure of this paper is as follows. In Section 2, we review the viral dynamics of HBV infection by using differential equation models. In Section 3, we introduce some commonly used parameter estimation methods and the related theory of model selection. The application of these methods are also illustrated

Abbreviations: HBV, hepatitis B virus; IFN, interferon; PEG-IFN, peg-interferon; LAM, lamivudine; ADV, adefovir dipivoxil; TDF, tenofovir disoproxil fumarate; NA, nucleoside analogue

* Corresponding author.
E-mail address: jianzu@xjtu.edu.cn (J. Zu).

https://doi.org/10.1016/j.jviromet.2019.01.014
Received 6 August 2018; Received in revised form 30 December 2018; Accepted 24 January 2019
Available online 01 February 2019
0166-0934/

through several specific examples. In Section 4, we make a brief summary and propose three future research programs for HBV infection.

2. Viral dynamic models of HBV infection

In order to better understand the dynamic process of HBV infection, in recent years, a variety of dynamic models including deterministic and stochastic models for HBV infection have been established (Murase and Sasaki, 2005; Sypsa et al., 2005; Guo and Cai, 2011; Perelson, 2002; Ciupe et al., 2007; Bellecave et al., 2009). Specifically, deterministic models explore the effect of various biological mechanisms, including antiviral therapy, CTL-mediated immune response, multitypes of infected hepatocytes, time delay and spatial diffusion. Stochastic models describe the random fluctuation of infection rate and reveal the impact of noise on stochastic HBV kinetics. Next, we discuss the differential equation models of HBV infection by considering each of the above biological factors and illustrate the new research results in recent years.

2.1. HBV viral dynamic models with antiviral treatment

To account for the relationship between antiviral therapy and viral loads, Nowak et al. firstly proposed a basic model to explore the interaction between uninfected hepatocytes, x, infected hepatocytes, y, and free virus particles, v, in 1996 (Nowak et al., 1996). Based on this basic model, a large number of models have been proposed by considering antiviral treatment, noncytolytic loss and so on. The framework of these models can be described by

$$\begin{cases} \frac{dx}{dt} = f(x, y) - (1-\eta)bvx + \rho y, \\ \frac{dy}{dt} = (1-\eta)bvx - ay - \rho y, \\ \frac{dv}{dt} = (1-\varepsilon)ky - cv. \end{cases} \tag{1}$$

Here, b denotes the infection rate of uninfected hepatocytes, ε denotes the efficacy of antiviral treatment, η denotes the drug efficacy in blocking infection, a denotes the death rate of infected hepatocytes, k denotes the rate at which virus particles are produced from infected hepatocytes, ρ denotes the rate of noncytolytic loss, and c denotes the clearance rate of HBV. The function $f(x, y)$ describes the growth rate of uninfected hepatocytes, which is usually given by

$$f(x, y) = \begin{cases} f_1(x, y) = \lambda - dx, \\ f_2(x, y) = \lambda - dx + rx(1 - \frac{x+y}{K}), \\ f_3(x, y) = rx(1 - \frac{x+y}{K}). \end{cases}$$

Here, λ is the production rate of uninfected hepatocytes, d is the death rate of uninfected hepatocytes, r is the maximum proliferation rate of hepatocytes, K is the carrying capacity of liver. Nowak et al. investigated the case $f = f_1$ with $\varepsilon = 1$, $\eta = 1$, (Nowak et al., 1996) they showed that the model fitted to actual clinical data validly only for about 20 days, which implied that the drug efficacy was unlikely to reach 100%. They obtained that the half-life of HBV is about 1 day and the turnover rate of infected hepatocytes is about 13 days. Tsiang et al. assumed $\varepsilon < 1$ and studied the case $f = f_1$ with $\eta = 1$ (Tsiang et al., 1999), they obtained that the antiviral efficacy was 99.3% for a 30 mg daily dose of adefovir dipivoxil. Besides, the model explained the biphasic nature of viral decay and fitted to actual clinical data validly for about 80 days. Lewin et al. further investigated the case $f = f_1$ with $\eta = 0.5$, $\varepsilon < 1$ and $\rho \neq 0$ (Lewin et al., 2002). By considering a time delay from receiving treatment to the decline of HBV, they obtained that the mean delay was 1.6 days, the clearance rate of HBV may be underestimated without taking into account the time delay. In addition, they estimated that the antiviral efficacy of lamivudine was about 95%, and the combination therapy of lamivudine and famciclovir was more effective with an efficacy close to 99% (Lau et al., 2000). The study also found that some patients exhibited complex delay profiles except for biphasic or multiphasic delay of HBV (Lewin et al., 2002).

To explain the complex HBV DNA decay profiles that observed in treated patients, Dahari et al. studied the case $f = f_2$ with $\rho \neq 0$ and explored main reasons that lead to the decay pattern of HBV DNA (Dahari et al., 2009). The model could describe a large number of complex delay profiles (e.g. biphasic, flat partial response, triphasic and rebound). However, this model still failed to explain the stepwise decay pattern of HBV, which needs to be solved in the future.

When considering a logistic growth of hepatocytes, much more dynamic patterns would occur, such as stable periodic orbit. Hews et al. examined the case $f = f_3$ with $\rho = 0$ (Hews et al., 2010). By assuming a logistic growth of hepatocytes and a standard incidence function, the results showed that the new system had a stable periodic orbit and a stable steady state at the origin (all variables are 0). Furthermore, the disease free equilibrium is globally asymptotically stable when $R_0 < 1$, while the system can either converge to the endemic equilibrium or experience sustained oscillations if $R_0 > 1$, where

$$R_0 = \frac{bk}{ac}.$$

When a patient takes a certain dose of oral or injectable medicine, the drug concentration increases quickly within a certain time range and then decreases as the drug elimination rate increases. Hence, it is reasonable to assume that the drug concentration varies over time. In particular, (Sypsa et al., 2005) extended the basic model (1) by assuming that antiviral efficacy $\varepsilon(t)$ changes over time t due to the variation in drug concentration, and the relationship between drug concentration $C(t)$ of PEG-IFN once weekly and $\varepsilon(t)$ is given by

$$\varepsilon(t) = \frac{C(t)}{IC_{50} + C(t)},$$

where

$$C(t) = C_0 e^{-k(t-t_p)}.$$

Here t_p is the time corresponding to the decrease of drug concentration (Powers et al., 2003). IC_{50} denotes the drug concentration for half-maximal antiviral efficacy. They found that the average drug efficacy of PEG-IFN 100 or 200 μg once weekly combined with lamivudine 100 μg daily were 92.8% and 94.4% respectively, which were a little lower compared to 96.4% of lamivudine 100 μg daily. However, long term clinical data are needed to further validate the efficacy of combination therapy of lamivudine and PEG-IFN.

Borg et al. (2007) further explored the cumulative effect of drug concentration in vivo when patients received multiple weekly injections of PEG-IFN-α2b. The antiviral effectiveness for individual i is given by

$$\varepsilon_i(t) = \frac{C_i(t - t_p)^n}{IC_{50}^n + C_i(t - t_p)^n},$$

where

$$C_i(t) = \sum_{d:t_d<t} \frac{(F \cdot D_d)_i}{V_{di}} \frac{k_{a,i}}{k_{e,i} - k_{a,i}} (e^{-k^*_{a,i}(t-t_d)} - e^{-k^*_{e,i}(t-t_d)}).$$

Here F denotes the bioavailability, D_d denotes the dose of each injection d, $k_{a,i}$ denotes the absorption rate, $k_{e,i}$ denotes the elimination rate, t_d denotes the injection day and n denotes the Hill coefficient. They found that the drug concentration peaked 1 day after receiving PEG-IFN-α2b, then the drug concentration showed a prominent decline afterwards the first week and the maximal drug efficacy was 70% in patients who received multiple weekly injections of PEG-IFN-α2b during the first 4 weeks. Of course, more clinical therapeutic trials are needed to further confirm the therapeutic efficacy of PEG-IFN-α2b.

2.2. Viral dynamics of HBV infection with immune response

The occurrence and development of HBV infection are closely associated with the body's immune function. Especially, the cytotoxic T lymphocytes (CTLs) play a crucial role in generating appropriate immune killing and inducing hepatocyte apoptosis. Nowak et al. extended the basic model (1) by considering the CTL immune responses (Nowak and Bangham, 1996). Many other studies also developed a family of new models by considering the effect of humoral immunity, noncytolytic mechanism and so on. The general framework of these models is given by

$$\begin{cases} \frac{dx}{dt} = \lambda - dx - bvx + qyz, \\ \frac{dy}{dt} = bvx - ay - (p+q)yz, \\ \frac{dv}{dt} = ky - cv - nbvx, \\ \frac{dz}{dt} = g(y,z) - c_1 z, \end{cases} \tag{2}$$

where q denotes the noncytolytic mechanism, z denotes the concentration of CTLs, myz denotes the proliferation rate of CTLs, c_1 denotes the clearance rate of CTLs. The case $n = 1$ represents the loss of the free virus particle due to infection. The function $g(y, z)$ describes the production rate of CTLs, which is usually given by

$$g(y,z) = \begin{cases} g_1(y,z) = ly, \\ g_2(y,z) = myz, \\ g_3(y,z) = s + \frac{myz}{w+y}. \end{cases}$$

Here, s denotes the production rate of CTLs from the thymus. Wodarz et al. investigated the case $g = g_1$ with $n = 0$, $q \neq 0$ (Wodarz et al., 2002), they showed that the noncytolytic mechanism played an important role in reducing the replication of HBV. Pang et al. studied $g = g_3$ with $n = 0$, $q \neq 0$ (Pang et al., 2012), they found that the lack of noncytolytic mechanism can lead to a dramatic drop in susceptible hepatocytes, which implied that the noncytolytic mechanism might be responsible for repairing infected hepatocytes.

Murase et al. investigated the case $g = g_2$ with $n = 1$, which described the interaction between the humoral immune response and HBV (Murase and Sasaki, 2005). They showed that the loss of free virus particle did not affect the stability of interior equilibrium when the immune response was not considered. However, if taking into account the humoral immune response, the interior equilibrium can be unstable and a Hopf bifurcation may occur. Wang et al. further discussed the case of $g = g_2$, they obtained the global stability of disease-free equilibrium and the uniform persistence of endemic equilibrium (Wang et al., 2014). Their results further showed that the cytolytic and noncytolytic mechanisms played a significant role in the clearance of HBV.

2.3. Viral dynamics of HBV infection with multi-types of infected hepatocytes

With the development of molecular biology technology, many researchers explored the dynamic process of HBV infection at the cellular and molecular level (Nakabayashi and Sasaki, 2011; Elaiw, 2012; Murray et al., 2016; Qesmi et al., 2010). Ciupe et al. (2007) investigated a mathematical model by considering two types of infected hepatocytes y_1 (contains one copy of cccDNA) and y_2 (contains multiple copies of cccDNA) as well as a time delay τ from infection to the production of immune cells, which is given by

$$\begin{cases} \frac{dx}{dt} = r(x+y_1)(1 - \frac{x+y_1+y_2}{x_{\max}}) - bvx + \rho_1 y_1, \\ \frac{dy_1}{dt} = bvx - (\rho_1 + r_1)y_1 - \mu y_1 z + \rho_2 y_2, \\ \frac{dy_2}{dt} = ry_2(1 - \frac{x+y_1+y_2}{x_{\max}}) + r_1 y_1 - \rho_2 y_2 - \mu y_2 z, \\ \frac{dv}{dt} = k_1 y_1 + k_2 y_2 - c_1 v, \\ \frac{dz}{dt} = s + \alpha(y_1(t-\tau) + y_2(t-\tau))z(t-\tau) - c_2 z. \end{cases} \tag{3}$$

Here, ρ_1 represents the noncytolytic mechanism. r_1 and ρ_2 denote the conversion rate between two types of infected hepatocytes. They found that the infected hepatocytes which contained multiple cccDNA reached 81% on average and 99% of liver cells were infected at the peak of infection. This result was consistent with the high production rate of HBV in infected hepatocytes containing multiple cccDNA. They also confirmed the important role of CTL immune response in inhibiting production of free virus particles especially after viral loads peaked.

To explore the role of cccDNA in HBV infection, Li et al. further investigated a mathematical model by considering CTL immune response and humoral immunity based on model (3) (Li et al., 2014). They showed that the conversion rate r_1 of cccDNA from y_1 to y_2 was strongly associated with the status of HBV infection. They also found that untreated patients with hepatitis B had higher conversion rate r_1 and continuous treatment would effectively inhibit HBV level.

Besides, some researchers investigated the kinetics of HBV infection by considering multi-types of mRNA transcribed by cccDNA. Nakabayashi et al. studied a model with two types of mRNA: 3.5 kb RNA (mainly produce core particles) and 2.4 kb mRNA (mainly translate surface proteins) (Nakabayashi and Sasaki, 2011). They showed that the pattern of HBV replication was closely related to the ratio of production rate of 3.5 kb RNA to 2.4 kb mRNA. The viral load changed significantly and even showed explosive growth once the ratio exceeded a certain threshold. Furthermore, the creation of new HNF-1 binding site promoted the expression of 3.5 kb RNA and caused high viral loads, this result may provide a quantitative reference for better control of HBV infection.

2.4. Delay differential equation models of HBV infection

More realistically, it may take a time delay from HBV infection to the release of free virus particles. Considering the intracellular delay can more accurately reflect the dynamics of HBV infection. By incorporating the impact of a intracellular delay, Herz et al. developed a delay differential equation model to study the efficacy of antiviral treatment on HIV and HBV (Herz et al., 1996). They found that the half-life of free virus particles might be overestimated without considering the intracellular delay for HIV or HBV, the frequent sampling of plasma virus would more accurately reflect the clearance rate of HBV.

Gourley et al. investigated a mathematical model by considering the intracellular delay and a standard incidence function (Gourley et al., 2008). They showed that the model had a globally attractive endemic equilibrium. Eikenberry et al. (2009) further introduced a logistic growth of uninfected hepatocytes into the model in Gourley et al. (2008) and proposed the following model

$$\begin{cases} \frac{dx(t)}{dt} = rx(t)(1 - \frac{T(x)}{K}) - dx(t) - bv(t)\frac{x(t)}{T(x)}, \\ \frac{dp(t)}{dt} = -dp(t) + bv(t)\frac{x(t)}{T(x)} - be^{-d\tau}v(t-\tau)\frac{x(t-\tau)}{T(t-\tau)}, \\ \frac{dy(t)}{dt} = be^{-d\tau}v(t-\tau)\frac{x(t-\tau)}{T(t-\tau)} - ay(t), \\ \frac{dv(t)}{dt} = ky(t) - cv(t), \end{cases} \tag{4}$$

where

$$T(t) = x(t) + p(t) + y(t).$$

Here, $p(t)$ denotes the latent hepatocytes that have been infected but cannot produce free virus particles. They showed that the continuous periodic oscillation could occur in model (4), but not in the model with a constant growth of uninfected hepatocytes. Besides, they found that increasing the intracellular delay could lead to a decrease of basic reproduction number.

More generally, some researchers incorporated immunity response or two time delays into the basic model of HBV infection (Wang et al., 2012; Meskaf et al., 2017; Pawelek et al., 2012). Wang et al. investigated a delay model by taking into account the CTL immune response (Wang et al., 2012). They showed that the globally dynamic properties of the model were entirely determined by two thresholds, defined as the viral invasion threshold and immune activation threshold, which are responsible for disease elimination, coexistence of free virus particles and CTL immune response, and immune exhaustion. In order to incorporate the time delay during the maturation process of HBV DNA-containing capsids, Manna et al. further studied a delay model by considering a second time delay (Manna and Chakrabarty, 2017). They showed that the second time delay did not affect the kinetics of HBV infection and there existed no periodic oscillation or bifurcation in the model. The further theoretical and experimental verification is needed to investigate the effect of time delay.

2.5. Viral dynamics of HBV infection with spatial diffusion

The HBV particles can move freely in the liver and change their locations over time. Considering the influence of spatial diffusion can more accurately describe the movement of virus particles. Wang et al. investigated a spatial diffusion model of HBV infection by assuming that virus particles can move freely in the liver (Wang and Wang, 2007). They obtained the following model

$$\begin{cases} \frac{\partial x}{\partial t} = \lambda - dx(s,t) - bx(s,t)v(s,t), \\ \frac{\partial y}{\partial t} = bx(s,t)v(s,t) - ay(s,t), \\ \frac{\partial v}{\partial t} = d_1 \Delta v + ky(s,t) - cv(s,t). \end{cases} \tag{5}$$

Here, $x(s,t)$, $y(s,t)$ and $v(s,t)$ denote the densities of uninfected hepatocytes, infected hepatocytes and free virus particles at position s and time t, respectively. d_1 denotes the diffusion coefficient and Δ denotes the Laplace operator. They showed that the increase of diffusion coefficient did not affect the existence and stability of traveling waves. Furthermore, the existence of traveling wave was also established, combining with the smallest wave speed, which can be used to accurately estimate the size of liver resection, thus ensuring the effectiveness of clinical surgery.

Xu et al. further studied a spatial diffusion model of HBV infection by considering a saturated infection rate (Xu and Ma, 2009), they also obtained that the diffusion coefficient did not affect the basic reproduction number although it may bring forward the time to reach endemic equilibrium. Similar conclusion was also found in other literature (Wang et al., 2011).

From Sections 2.1–2.5, we have introduced the deterministic models of HBV infection by considering different biological mechanisms, such as, antiviral therapy, CTL-mediated immune response, multitypes of infected hepatocytes, time delay and spatial diffusion. In the next section, we introduced a type of stochastic dynamic models of HBV infection.

2.6. Stochastic dynamic models of HBV infection

Deterministic models have been widely applied in characterizing the dynamic process of HBV infection and the population is typically assumed to be homogenous. Note that everyone has a different chance of contracting a disease from another infected person, those with poor immunity are particularly vulnerable to a disease. Hence, stochastic models that consider the heterogeneity of infectious process are more realistic to describe the dynamic process of HBV infection.

By assuming that the population is heterogeneous, (Moneim et al., 2009) derived a stochastic model of HBV infection with a random infection rate and used Monte Carlo methods to carry out numerical simulation. Xie et al. (2017) investigated a stochastic model by incorporating a random perturbation $\sigma\zeta(t)$ into the infection rate β and obtained the following model

$$\begin{cases} \frac{dx}{dt} = \left(rx\left(1 - \frac{x+y}{K}\right) - \beta xv\right)dt - \sigma xv dB(t), \\ \frac{dy}{dt} = (\beta xv - ay)dt + \sigma xv dB(t), \\ \frac{dv}{dt} = (ky - cv)dt, \end{cases} \tag{6}$$

where

$$dB(t) = \zeta(t)dt.$$

Here, $\zeta(t)$ denotes the Gaussian white noise, σ denotes the intensity of environmental forcing, $B(t)$ denotes the standard Weiner process, the remaining parameters have the same meanings as above. They showed that model (6) admits at least one periodic solution due to $\zeta(t)$ when assuming all parameters are positive-periodic. Particularly, they found that the random perturbation led to the formation of periodic solution. Luzyanina et al. (Luzyanina and Bocharov, 2014) further investigated the impact of noise intensity and noise type on stochastic HBV model with immune response. They found that the frequency of spontaneous recovery is below 10% when the intensity is very small, and is double when the fluctuation intensity is increased by 10 times. More generally, Liu et al. (Liu et al., 2017) examined a stochastic model by considering a logistic growth of hepatocytes. The sufficient conditions for disease extinction and the existence of unique stationary distribution were derived by means of stochastic Lyapunov function method. In the future, more realistic stochastic models need to be developed to reveal the impact of noise on HBV infection dynamics.

In order to get a better understanding of HBV infection dynamics, Table 1 summarized the differential equation models of HBV infection. In Table 1, we still use x to denote uninfected hepatocytes, y denotes infected hepatocytes (y_1, y_2 denotes the two types of infected hepatocytes), v denotes free virus particles, z denotes the concentration of the CTLs, p denotes the latent infected hepatocytes. The framework of these differential equation models roughly contains the following five types: xyv model, $xyvz$ model, xy_1y_2vz model, $xpyz$ model and xyv partial differential model.

3. Parameter estimation for HBV viral dynamic models

Accurately estimating the parameters of HBV viral dynamic models contributes to infer correct biological conclusions, thus guiding clinical treatment of hepatitis B. In this section, we discuss some commonly used methods of parameter estimation, such as, nonlinear least square (NLS) method and Bayesian estimation approach. Furthermore, we briefly illustrate their application through two specific examples. Finally, we introduce some basic idea in model identification.

3.1. Nonlinear least square method

The NLS method is the most basic way to estimate unknown parameters for a nonlinear model. In Table 2, we summarize partial estimation parameters by literature review, including the clearance rate of HBV and the clearance rate of infected hepatocytes. We can easily see that the most widely used parameter estimation method is the NLS method. (Sypsa et al., 2005; Tsiang et al., 1999; Lewin et al., 2002; Mihm et al., 2008).

Next we illustrate how to use NLS method to estimate unknown

Table 1
Summary of HBV viral dynamics by using differential equation models.

Factors considered	Framework	Main conclusion	References
The production rate of uninfected hepatocytes is constant with antiviral efficacy $\varepsilon = 1$.	***xyv* model**	Drug efficacy was unlikely to reach 100%, the half-life of HBV is about 1 day and the turnover rate of infected hepatocytes is about 13 days.	Nowak et al., 1996
The proliferation of infected hepatocytes and uninfected hepatocytes, as well as the noncytolytic mechanisms.	***xyv* model**	This model could describe a large number of complex decay profiles (e.g. biphasic, flat partial response, triphasic and rebound).	Dahari et al., 2009
Logistic growth of uninfected hepatocytes.	***xyv* model**	Exist a stable periodic orbit and a stable steady state (at the origin).	Hews et al., 2010
The intensity of CTL immune response is proportional to the number of infected hepatocytes, as well as the noncytolytic mechanism.	***xyvz* model**	The replication rate of HBV was reduced by noncytolytic mechanism.	Wodarz et al., 2002
The production rate of CTLs is a bilinear form of the concentration of infected hepatocytes and CTLs.	***xyvz* model**	Occurrence of Hopf bifurcation.	Murase and Sasaki, 2005
CTL cells from the thymus, cytolytic and noncytolytic mechanisms.	***xyvz* model**	Noncytolytic mechanism plays a crucial role in repairing infected hepatocytes.	Pang et al., 2012
A varied antiviral effectiveness.	***xyv* model**	The average drug efficacy of PEGylated interferon 100 or 200 μg weekly combined with lamivudine 100 μg daily are respectively 92.8% and 94.4%, which are lower compared to 96.4% of lamivudine 100 μg daily.	Sypsa et al., 2005
Two types of infected hepatocytes (contain one type of cccDNA or multiple types of cccDNAs).	***xy₁ y₂ vz* model**	81% ± 15% of the infected hepatocytes contain multiple cccDNAs on average and about 99% of liver cells are infected at the peak of infection.	Ciupe et al., 2007
Intracellular delay.	***xyv* model**	Frequent sampling of plasma virus would more accurately reflect the clearance rate of HBV.	Herz et al., 1996
The intracellular delay and latent infected hepatocytes, as well as a logistic growth of uninfected hepatocytes.	***xpyv* model**	Generate continuous periodic oscillation.	Eikenberry et al., 2009
Spatial diffusion.	***xyv* PDE model**	Establish the existence of traveling wave and accurately estimate the size of liver resection by using the smallest wave speed.	Wang and Wang, 2007
Incorporating a random perturbation into the infection rate.	***xyv* stochastic model**	The random perturbation lead to the formation of periodic solution.	Xie et al., 2017
The impact of noise intensity and noise type.	***xyv* stochastic model**	The frequency of spontaneous recovery is below 10% when the noise intensity is very small, and is double when the fluctuation intensity is increased by 10 times.	Luzyanina and Bocharov, 2014

Note: x, uninfected hepatocytes; y, infected hepatocytes; v, free virus particles; z, CTL-mediated immune response; p, latent infected hepatocytes.

parameters b, k and ε through a simple model

$$\begin{cases} \frac{dx}{dt} = \lambda - dx - (1-\eta)bvx, \\ \frac{dy}{dt} = (1-\eta)bvx - ay, \\ \frac{dv}{dt} = (1-\varepsilon)ky - cv. \end{cases} \quad (7)$$

The parameter meanings are the same as above. We assume $\lambda = 12$ cell $ml^{-1}\ day^{-1}$ based on past experience. The remaining parameters come from previous studies, namely, $d = 0.002$ day^{-1} from Dahari et al. (2009), $\eta = 0.5$ from Lewin et al. (2002), $a = 0.043$ and $c = 0.65$ from Tsiang et al. (1999). Let $\vec{\beta} = (b, k, \varepsilon)$, then $v(t, \vec{\beta})$ denotes the numerical solution $v(t)$ of the third equation in model (7) under the given parameter vector $\vec{\beta}$. The general steps of parameter estimation are as follows.

3.1.1. Step 1. Data collection

Collect the HBV DNA load for n days in a patient receiving antiviral therapy, which is given by

$v_0, v_1, v_2, \ldots, v_{n-1}$.

The initial values of three variables in model (7) are set to x_0, y_0 and v_0, respectively.

3.1.2. Step 2. NLS fitting

Assuming that $v(t_i, \vec{\beta})$ is the numerical solution corresponding to

Table 2
Summary of key parameter values in HBV viral dynamic models.

References	Treatment	Estimation methods or software	η	ε	c	a	The half-life of HBV (hour)	The turnover rate of infected hepatocytes (day)
6	PEG-IFN 100 mg/ 200 mg	NLS	assumed 0.5	0.826	0.24~6.847	0.071	2.4~69.2	2.7~75
	PEG-IFN 100 mg + LAM			0.928		0.120		
	PEG-IFN 200mg + LAM			0.944		0.146		
	LAM			0.964		0.119		
15	—	MCMC	—	—	0.67	—	1	990
22	ADV 30mg	NLS	assumed 1	0.993 ± 0.008	0.65 ± 0.15	0.043 ± 0.015	26.4 ± 7.2	18 ± 7
23	LAM	NLS	assumed 0.5	0.951	—	—	28.5	2.4~120
25	IFN-α + NAs	Berkeley-Madonna	0.5	—	3.5	0.25	4.752	2.77
			0.5	0.991	0.7	0.22	23.76	3.15
			0.2	0.9988	0.18	0.25	92.4	2.772
			0.5	0.997	1	0.06	16.632	11.55
32	PEG-IFN-α2b	NLS	—	0.7738 ± 0.2879	1.3	0.07 ± 0.12	12.79	9.9 ± 0.058

Abbreviation: LAM = Lamivudine; ADV = Adefovir Dipivoxil; IFN = Interferon; PEG-IFN = Peg-Interferon; NA = Nucleoside Analogue; MCMC = Markov Chain Monte Carlo; NLS = Nonlinear Least Square.

the t_i day and parameter vector $\vec{\beta}$, where $v(t_i, \vec{\beta})$ can be solved by using Runge-Kutta algorithm. In order to obtain the best fitting, we need to find the optimal $\hat{\vec{\beta}}$ to minimize the residual sum of squares $S(\vec{\beta})$, that is,

$$\hat{\vec{\beta}} = \arg\min_{\vec{\beta}} S(\vec{\beta}) = \arg\min_{\vec{\beta}} \sum_{i=0}^{n-1} (v(t_i,\vec{\beta}) - v_i)^2$$

3.1.3. Step 3. Solve the NLS problem numerically

Given the initial parameter values $\vec{\beta}^0 = (b_0, k_0, \varepsilon_0)$ and set the upper and lower bound for each parameter b, k and ε based on previous studies. Using Matlab function 'fmincon' or 'lsqnonlin', we can obtain the optimal parameter vector $\hat{\vec{\beta}} = (\hat{b}, \hat{k}, \hat{\varepsilon})$.

3.1.4. Step 4. Sensitivity analysis

In order to identify whether the parameter estimation is robust or not, we perform the sensitivity analysis on the initial input parameters $\vec{\beta}^0 = (b_0, k_0, \varepsilon_0)$ by varying their values in a reasonable range.

By selecting the observed data from the patient 64 in Sypsa et al. (2005) using the above method we obtain that parameters $\hat{b} = 1.7880 \times 10^{-7}$, $\hat{k} = 0.3196$ and $\hat{\varepsilon} = 0.7233$, and the corresponding 95% confidence intervals are (0, 0.7104), (0, 1.03) and (0.0129, 1.0), respectively. Here, if the lower bound of confidence interval is negative, then we set it to 0. If the upper bound of confidence interval of $\hat{\varepsilon}$ is larger than 1, then we set it to 1.0. The relevant Matlab procedure is given in Appendix A. Fig. 1 shows the comparison between the numerical solution of model (7) and the observed data. The results show that the simulated viral loads fit well with the clinical data. Therefore, for this patient we obtain that the antiviral efficacy ε is about 72.33% (95% CI, 1.29%–100%).

However, by performing sensitivity analysis we find that the estimated parameters are strongly dependent on the selection of initial input parameter values and the confidence intervals are too wide, which are the disadvantages of NLS method and in need of improvement.

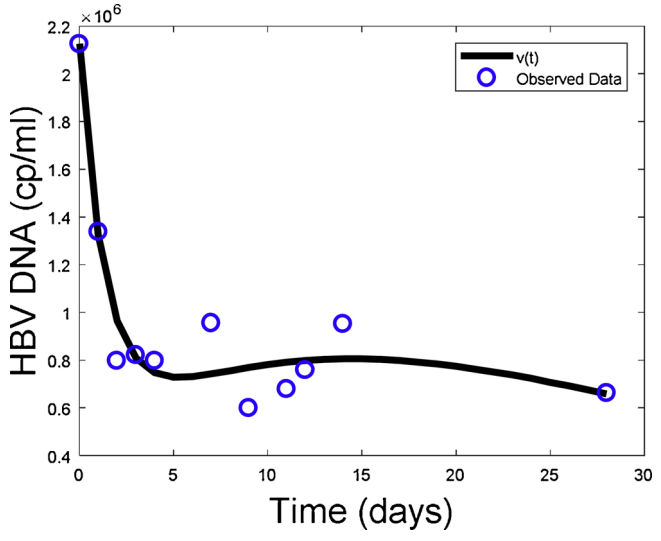

Fig. 1. Fitting the observed viral loads by using NLS method in model (7). Here we chose the viral loads which were measured at days 0, 1, 2, 3, 4, 7, 9, 12, 14 and 28. Assuming that $\lambda = 12$ cell $ml^{-1}\, day^{-1}$, the remaining parameter values are: $d = 0.002\ day^{-1}$ from Dahari et al. (2009), $\eta = 0.5$ from Lewin et al. (2002), $a = 0.043$ and $c = 0.65$ from Tsiang et al. (1999). The initial values for uninfected and infected hepatocytes are assumed to be 1×10^7 and 2.57×10^6, respectively.

3.2. Bayesian estimation approach

Currently, the Markov Chain Monte Carlo (MCMC) method has become more and more popular, which is widely used in actual research and data fitting. There are two main MCMC methods typically used: Gibbs sampling and Metropolis-Hastings ((M-H)) algorithm. In particular, the (M-H) algorithm has become more generalized MCMC method than Gibbs sampling. For a given sample set Y that consists of n random variables, how to use the observed data to estimate unknown parameters $\vec{\beta} = (b, k, \varepsilon, y_0)$ in model (7) is a problem of Bayesian estimation, where y_0 is the initial value of infected hepatocytes. Here, we still take model (7) as an example to briefly introduce the basic steps of MCMC (M − H) algorithm.

3.2.1. Step 1. Collect sample observations

Here, we collect the observed data of HBV DNA load $v_0, v_1, v_2, \ldots, v_{n-1}$ for n days in a patient receiving antiviral therapy. The meaning of $v(t_i, \vec{\beta})$ is the same as above.

3.2.2. Step 2. Determine the conditional density function

Given $Y = (v_0, v_1, \ldots, v_{n-1})$, the conditional density function of parameter vector $\vec{\beta}$ is given by

$$P(\vec{\beta}|Y) = \frac{P(Y|\vec{\beta})P(\vec{\beta})}{P(Y)}.$$

Since $P(Y)$ is a constant relative to parameter vector $\vec{\beta}$, the conditional density function is usually rewritten as follows

$$P(\vec{\beta}|Y) \propto P(Y|\vec{\beta})P(\vec{\beta}).$$

3.2.3. Step 3. Construct the likelihood function

We assume that the viral load of day t_i follows a Poisson distribution with parameter $v(t_i, \vec{\beta})$. Then the likelihood function is given by

$$L(Y|\vec{\beta}) = \prod_{i=0}^{n-1} P(v_i, \vec{\beta}) = \frac{\prod_{i=0}^{n-1} [v(t_i, \vec{\beta})^{v_i}] e^{-\sum_{i=0}^{n-1} v(t_i,\vec{\beta})}}{(v_0 v_1 \ldots v_{n-1})!}.$$

3.2.4. Step 4. Establish the probability of joint posterior distribution

By selecting the non-informative prior distribution $P(\vec{\beta}) \propto$ constant, then the probability of joint posterior distribution is given by

$$\begin{aligned} P(\vec{\beta}|Y) &\propto L(Y|\vec{\beta})P(\vec{\beta}) \\ &\propto \prod_{i=0}^{n-1} P(v_i|\vec{\beta}) \\ &\propto \prod_{i=0}^{n-1} \frac{v(t_i,\vec{\beta})^{v_i}}{v_i!} e^{-v(t_i,\vec{\beta})}. \end{aligned}$$

3.2.5. Step 5. Select the initial value of parameters

Here, we select the initial parameter vector $\vec{\beta}^0 = (b_0, k_0, \varepsilon_0, y_0)$ based on previous studies. We use the random walk method to generate candidate parameter $\vec{\beta}' = (b', k', \varepsilon', y'_0)$ and assume that the proposal distribution $q(\vec{\beta}', \vec{\beta})$ satisfies

$$q(\vec{\beta}', \vec{\beta}) = q(\vec{\beta}, \vec{\beta}')$$

3.2.6. Step 6. Determine the acceptance probability

Since the proposal distribution is symmetric, the acceptance probability can be simplified as

$$\alpha(\vec{\beta}', \vec{\beta}) = \min\{1, \frac{q(\vec{\beta}, \vec{\beta}')P(b', k', \varepsilon', y'_0 | Y)}{q(\vec{\beta}', \vec{\beta})P(b, k, \varepsilon, y_0 | Y)}\}$$
$$= \min\{1, \frac{\frac{v(t_i,\vec{\beta})^{v_i}}{v_i!}e^{-v(t_i,\vec{\beta}')}}{\frac{v(t_i,\vec{\beta}')^{v_i}}{v_i!}e^{-v(t_i,\vec{\beta})}}\}$$
$$= \min\{1, (\frac{v(t_i,\vec{\beta})^{v_i}}{v(t_i,\vec{\beta}')^{v_i}})^{v_i} e^{v(t_i,\vec{\beta})-v(t_i,\vec{\beta}')}\}$$

Select the random number ζ from the Uniform distribution $U(0, 1)$. When $\alpha(\vec{\beta}', \vec{\beta}) \geq \zeta$, accept the candidate parameter $\vec{\beta}'$ and denoted by $\vec{\beta}^{(t+1)} = \vec{\beta}'$, otherwise reject $\vec{\beta}'$. Select a burn-in period of m times and a cycle of N times, calculate the mean value $(\hat{b}, \hat{k}, \hat{\varepsilon}, \hat{y}_0)$ of the last $N - m$ times about four parameters and take them as the estimated values of $\vec{\beta}$.

3.2.7. Step 7. Sensitivity analysis

In order to identify whether the parameter estimation is robust or not, we perform sensitivity analysis on the initial input parameters by varying their values in a reasonable range.

Here, we assume that $N = 500000$ and $m = 300000$. Given the same observed data and parameters as in Fig. 1, by using the above method, we obtain the estimated values $\hat{b}$, $\hat{k}$, $\hat{\varepsilon}$ and $\hat{y}_0$ as well as their 95% confidence intervals (CI), namely, $\hat{b} = 1.7985 \times 10^{-7}$ (95% CI, 1.787×10^{-7}-1.810×10^{-7}), $\hat{k} = 0.3377$ (95% CI, 0.3367-0.3388), $\hat{\varepsilon} = 0.7379$ (95% CI, 0.7374-0.7384) and $\hat{y}_0 = 2.5647 \times 10^6$ (95% CI, 2.5629×10^6-2.5664×10^6). Fig. 2 shows that the simulated curve using MCMC (M−H) algorithm fits well with the clinical data. Therefore, by using the MCMC (M−H) method we obtain the antiviral efficacy ε for patient 64 is about 73.79% (95% CI, 73.74%–73.84%). The relevant Matlab procedure is given in Appendix B. By performing the sensitivity analysis, we can see that the parameter values estimated by the MCMC (M−H) algorithm are relatively stable and well fitted. Besides, the confidence intervals we obtained are relatively narrow, which mean that the parameter values estimated by the MCMC (M−H) algorithm are more credible. Furthermore, the procedure of MCMC (M−H) algorithm is relatively easy, which opens up a new way for the application of Bayesian method.

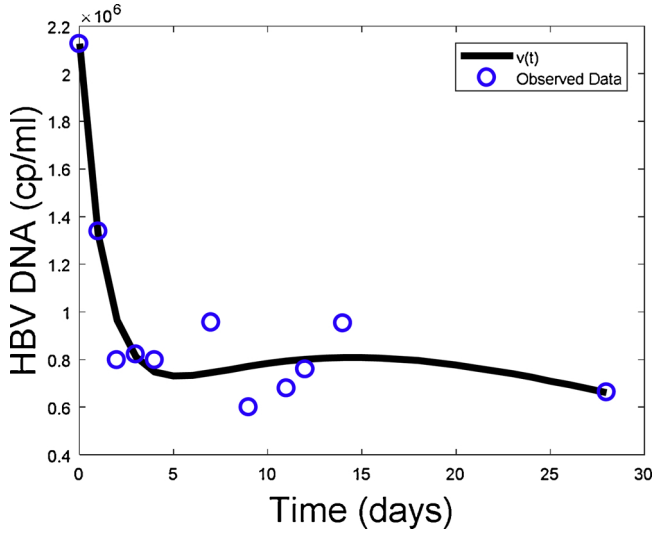

Fig. 2. Fitting the observed viral loads by using MCMC (M−H) algorithm in model (7). The selected observed data of HBV DNA load, other fixed parameters and initial values for uninfected and infected hepatocytes are the same as in Fig. 1.

3.3. Model identification method

Mathematical models play a significant role in predicting the dynamic process of HBV infection. Complex mathematical models may fit well with sample observations, but are prone to overfitting. Hence, it is essential to balance the complexity and accuracy of the model. The parameter selection criterion considering the analytical expression of error estimation is usually adopted. These criteria mainly include Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are given by

$AIC = 2K - 2\ln L(\vec{\theta})$, $BIC = -2\ln L(\vec{\theta}) + K\ln n$.

Here, K denotes the number of unknown parameters and reflects the simplicity of model, L represents the likelihood function and reflects the accuracy of model, $\vec{\theta}$ is the unknown parameter vector of candidate model and n denotes the number of sample observations. The best model is the one with the smallest AIC and BIC values.

Moreover, in order to evaluate the accuracy of predictive values, the mean absolute percentage error (MAPE) and the root mean square percentage error (RMSPE) are usually used to assess the predictive performance of model (Wang et al., 2009; Stekler, 1991). Their definitions are given by

$$MAPE = \frac{1}{n}\sum_{i=0}^{n-1}\left|\frac{v(t_i, \vec{\theta}) - v_i}{v(t_i, \vec{\theta})}\right| \times 100\%,$$

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}\left|\frac{v(t_i, \vec{\theta}) - v_i}{v(t_i, \vec{\theta})}\right|^2} \times 100\%$$

The meanings of $v(t_i, \vec{\theta})$ and v_i are the same as above.

Next, we discuss how to use the criteria AIC and BIC combined with MAPE and RMSPE to evaluate the predictive performance of candidate models through the following model

$$\begin{cases} \frac{dx}{dt} = \lambda - dx - (1-\eta)bvx + \rho y, \\ \frac{dy}{dt} = (1-\eta)bvx - ay - \rho y, \\ \frac{dv}{dt} = (1-\varepsilon)ky - cv, \end{cases} \quad (8)$$

where ρ denotes the non-cytolytic mechanism. The general steps are given below.

3.3.1. Step 1. Collect the observed data

Here, we select the observed viral load $v_0, v_1, v_2, \ldots, v_{n-1}$ for n days in a patient receiving antiviral therapy as a set of sample observations.

3.3.2. Step 2. Select parameters to be evaluated

Here, we choose b, k, ε, y_0 as unknown parameters and ρ, η as candidate parameters that need to be evaluated. In this case, a total of four candidate models need to be evaluated (see Table 3). We denote the parameter vectors as $\vec{\theta}_1 = (b, k, \varepsilon, y_0)$, $\vec{\theta}_2 = (\rho, b, k, \varepsilon, y_0)$, $\vec{\theta}_3 = (\eta, b, k, \varepsilon, y_0)$ and $\vec{\theta}_4 = (\rho, \eta, b, k, \varepsilon, y_0)$, which corresponds to these four candidate models.

Table 3
Candidate models and the corresponding AIC, BIC values.

Candidate models	Assumption	AIC	BIC
1	$\rho = 0, \eta = 0$	1.7556×10^5	1.7557×10^5
2	$\rho \neq 0, \eta = 0$	1.7556×10^5	1.7556×10^5
3	$\rho = 0, \eta \neq 0$	1.7553×10^5	1.7553×10^5
4	$\rho \neq 0, \eta \neq 0$	1.7555×10^5	1.7555×10^5

Note: The observed data of HBV DNA load, other fixed parameters and the initial values for uninfected and infected hepatocytes are the same as in Fig. 1.

3.3.3. Step 3. Construct the likelihood function

Assuming that the viral load of day t_i follows a Poisson distribution with parameter $v(t_i, \vec{\theta}_j)$, $j = 1, 2, 3, 4$, then we obtain the likelihood function corresponding to each parameter vector $\vec{\theta}_j j = 1, 2, 3, 4$,

$$L(v_0, v_1, \ldots, v_{n-1}|\vec{\theta}_j) = \prod_{i=0}^{n-1} P(v_i|\vec{\theta}_j) = \frac{\prod_{i=0}^{n-1}[v(t_i, \vec{\theta}_j)^{v_i}]e^{-\sum_{i=0}^{n-1} v(t_i,\vec{\theta}_j)}}{(v_0 v_1 \ldots v_{n-1})!}.$$

3.3.4. Step 4. Calculate AIC and BIC values

Let $v(t_i, \vec{\theta}_j)$ denote the numerical solution corresponding to the parameter vector $\vec{\theta} j$, K_j denotes the total number of parameters corresponding to parameter vector $\vec{\theta} j$. Then, we obtain the values of AIC and BIC for each parameter vector $\vec{\theta} j$, where

$$AIC_j = 2K_j - 2\sum_{i=0}^{n-1}[\ln v(t_i, \vec{\theta}_j)^{v_i} - v(t_i, \vec{\theta}_j)] + 2\ln(v_0 v_1 \ldots v_{n-1})!,$$

$$BIC_j = -2\sum_{i=0}^{n-1}[\ln v(t_i, \vec{\theta}_j)^{v_i} - v(t_i, \vec{\theta}_j)] + 2\ln(v_0 v_1 \ldots v_{n-1})! + K_j \ln n$$

Given the same observed data and parameters as in Fig. 1, we obtain the corresponding values of AIC and BIC for candidate model 1–4 (see Table 3). We can easily see that the candidate model 3 is the optimal model, because both the AIC and BIC values are the smallest. For candidate model 3, by using MCMC (M−H) algorithm, the estimated parameter values and corresponding 95% CI are listed in Table 4.

3.3.5. Step 5. Calculate the values of MAPE and RMSPE

We recalculate the model (8) with the estimated parameters in Table 4 and obtain the values of MAPE and RMSPE, which are 10.42% and 14.16%, respectively. Based on the criteria of MAPE and RMSPE (10%˜20% means a good prediction), we conclude that the estimated parameters by (M−H) algorithm are reasonable.

The criteria AIC and BIC are widely used in model selection. Besides, some researchers further improve the guideline of model selection. Burnham et al. proposed a modified AIC criterion which could be applied to any sample size (Burnham and Anderson, 2004). The extended Bayesian Information Criterion (EBIC) makes model to focus on simplicity by adding new penalties to the BIC criterion, which further enriches the criteria of model selection.

4. Summary and discussion

In this paper, we review the differential equation models of HBV infection dynamics, which help us understand the dynamic mechanism of HBV infection and the efficacy of receiving antiviral treatment. Firstly, we introduce a family of deterministic models by considering different biological mechanisms, such as, antiviral therapy, CTL immune response, multi-types of infected hepatocytes, time delay and spatial diffusion. Then, we briefly introduce some stochastic models of HBV infection. In addition, we discuss some classic methods of parameter estimation such as nonlinear least square method and Bayesian estimation approach. Their applications are also illustrated through two specific models. Particularly, we find that the parameter values estimated by the MCMC (M−H) algorithm are stable and well fitted. The confidence intervals we obtained are also relatively narrow, which mean that the parameter values estimated by the MCMC (M−H) algorithm are more credible. Finally, in order to balance the accuracy and simplicity of mathematical models, we introduce the basic idea of model identification through a specific example. Specifically, we explore the classic criteria AIC and BIC as well as their applications. By combining these criteria with clinical observed data, we can accurately estimate the key clinical parameter, such as drug efficacy.

Although various models have been developed to characterize the dynamic process of HBV infection, there are still some remaining problems that cannot be explained. As a result of long-term antiviral therapy, the mechanism about the rapid emergence of drug resistance has not yet been fully revealed. The corresponding study of drug resistance with multiple strains in HBV is rarely reflected. Here, we propose a HBV model with multiple strains based on the model in Rong et al. (2010). By dividing HBV particles into drug-sensitive strain v_s and drug-resistant strain v_r, the model is given by

$$\begin{cases} \frac{dx}{dt} = \lambda - dx + rx(1 - \frac{x+y_s+y_r}{K}) - (1-\eta_s)bv_s x - (1-\eta_r)bv_r x + \rho y_s \\ \qquad + \rho y_r, \\ \frac{dy_s}{dt} = (1-\eta_s)bv_s x - ay_s - \rho y_s, \\ \frac{dy_r}{dt} = (1-\eta_r)bv_r x - ay_r - \rho y_r, \\ \frac{dv_s}{dt} = (1-\varepsilon_s)k_s y_s - cv_s, \\ \frac{dv_r}{dt} = (1-\varepsilon_r)k_r y_r - cv_r. \end{cases} \tag{9}$$

Here, y_s and y_r represent the infected hepatocytes which are infected by v_s and v_r, respectively. The remaining parameters have the same meanings as above. Through exploring the coexistence condition for v_s and v_r, this model may provide a theoretical reference for understanding the mechanism of drug resistance.

In addition, whether the within-host dynamics of HBV infection will affect the macroscopic propagation of between-host has not yet been studied in detail, which is also a future research direction. Here, we propose a coupled HBV infection model based on the model in Feng et al. (2012) which is given by

Table 4
Estimated parameter values in candidate model 3 by using the MCMC (M−H) algorithm.

Parameter	Meaning	Estimated value	95% confidence interval
η	Drug efficacy of inhibition infection	0.5117	(0.4980, 0.5254)
b	Infection rate	1.9300×10^{-7}	$(1.872 \times 10^{-7}, 1.988 \times 10^{-7})$
k	Production rate of viral particles	0.3391	(0.3377,0.3405)
ε	Antiviral efficacy	0.7390	(0.7382,0.7398)
y_0	Initial value of infected hepatocytes	2.5630×10^6	$(2.5618 \times 10^6, 2.5642 \times 10^6)$

Note: The observed data of HBV DNA load, other fixed parameters and the initial values for uninfected and infected hepatocytes are the same as in Fig. 1. The corresponding values of MAPE and RMSPE are 10.42% and 14.16%, respectively.

$$\begin{cases} \frac{dS}{dt} = (1-w)(b_1(N-C) + b_1(1-q_1)C) - \beta(v)\frac{SC}{N} - \theta S - m_1 S, \\ \frac{dC}{dt} = b_1 q_1 C + \eta_1 \beta(v)\frac{SC}{N} - (m_1 + \alpha - \gamma)C, \\ \frac{dR}{dt} = w(b_1(N-C) + b_1(1-q_1)C) + \gamma C + (1-\eta_1)\beta(v)\frac{SC}{N} + \theta S - m_1 R, \\ \frac{dx}{dt} = \lambda - dx + rx(1 - \frac{x+y}{K}) - (1-\eta)bvx + qyz, \\ \frac{dy}{dt} = (1-\eta)bvx - ay - (p+q)yz, \\ \frac{dv}{dt} = (1-\varepsilon)ky - cv, \\ \frac{dz}{dt} = myz - c_1 z. \end{cases} \quad (10)$$

where the transmission coefficient $\beta(v)$ is a monotonically increasing saturation function of viral load v. Here, S, R and C represent the number of susceptible individuals, recovered individuals and chronic HBV carriers at time t respectively, N is the total population, b_1 is the birth rate, m_1 and α are respectively the natural mortality and disease-related mortality, ω is the neonatal vaccine coverage rate, γ is the natural recovery rate, q_1 is the vertical transmission rate from mother to infant, θ is the immune protective rate of adults, η_1 is the conversion rate from acute to chronic infected individuals. This coupled model may generate new dynamic properties or threshold conditions, which may reveal more complex biological phenomena of hepatitis B.

Interestingly, the virulence of HBV may adaptively evolve, thus affecting the macroscopic propagation of hepatitis B. Higher virulence in vivo may result in a higher transmission rate and mortality. By assuming that the transmission rate $\beta(\alpha)$ is a monotonically increasing function of virulence, α, we propose another HBV infection model by linking population dynamics with evolutionary dynamics of HBV virulence, which is given by

$$\begin{cases} \frac{dS}{dt} = (1-w)(b_1(N-C) + b_1(1-q_1)C) - \beta(\alpha)\frac{SC}{N} - \theta S - m_1 S, \\ \frac{dC}{dt} = b_1 q_1 C + \eta_1 \beta(\alpha)\frac{SC}{N} - (m_1 + \alpha - \gamma)C, \\ \frac{dR}{dt} = w(b_1(N-C) + b_1(1-q_1)C) + \gamma C + (1-\eta_1)\beta(\alpha)\frac{SC}{N} + \theta S \\ \qquad - m_1 R, \\ \frac{d\alpha}{dt} = G\frac{\partial f(\hat{\alpha},\alpha)}{\partial \hat{\alpha}}\Big|_{\hat{\alpha}=\alpha}, \end{cases} \quad (11)$$

where the invasion fitness of mutant HBV infections, $f(\hat{\alpha}, \alpha)$, is given by

$$f(\hat{\alpha}, \alpha) = b_1 q_1 + \eta_1 \beta(\hat{\alpha})\frac{S}{N} - \hat{\alpha} - m_1 + \gamma$$

Here, G denotes the speed of evolutionary adaptation. This model can reveal the impact of evolution of HBV virulence on the epidemic trend of hepatitis B. Besides, this coupled model also describe the impact of macroscopic propagation dynamics on the evolution of HBV virulence, which may provide a novel insight for the prevention and treatment of hepatitis B.

Funding

This work was supported by the National Natural Science Foundation of China [grant numbers 11571272, 11631012]; the National Science and Technology Major Project of China [grant number 2012ZX10002001]; the Natural Science Foundation of Shaanxi Province [grant number 2015JQ1011]; and the China Postdoctoral Science Foundation [grant number 2014M560755].

Authors contributions

Miaolei Li participated in the design and drafted the manuscript. Jian Zu revised the draft. All authors read and approved the final version of the manuscript.

Conflicts of interest

We declare that we have no competing interests.

Acknowledgements

We are very grateful to referees and editors for their careful reading and valuable comments. We would also like to thank Profs. Yicang Zhou and Yanni Xiao for their valuable discussion.

Appendix A. The Matlab code for parameter estimation of model (7) by using NLS method

```
function NLS
clear all
clc
Y = [2.125*10^6 1.3375*10^6 7.975*10^5 8.2*10^5 7.975*10^5
9.55*10^5 5.9821*10^5 6.7857*10^5 7.5893*10^5 9.5179*10^5
6.625*10^5]; % The observed data
Ic = [1*10^7 2.57*10^6 2.125*10^6]; % Initial conditions
lb = [0 0 0]; % Constraint condition, the lower limits of parameters
ub = [1 10 10^-6]; % Constraint condition, the upper limits of
parameters
par1guess = [0.73 0.23 4.43*10^-7]; % The initial values of para-
meters
options = optimset('Algorithm','trust-region-
reflective','Display','final','MaxIter',20000,'MaxFunEvals',20000);
% Optimization options
[par1,resnorm,residual,exitflag,output,lambda,jacobian]
=lsqnonlin(@LSminzu,par1guess,lb,ub,options);
Estimated_Paras = par1% Estimated parameters
sigma_par1 = std(par1);
par1_ci = [par1-1.96*sigma_par1; par1 + 1.96*sigma_par1]'
par1(1)
par1(2)
par1(3)

% Comparison of simulated curve in model (7) with observed data

[T2,W2] = ode45(@xyvmodel,[0:1:28],Ic,[],par1);
figure(1)
plot(T2,W2(:,3),'k-','LineWidth',4);
hold on
plot([0 1 2 3 4 7 9 11 12 14 28],Y,'bo','LineWidth',2,'MarkerSize',9);
xlabel('Time (days)');
ylabel('HBV DNA (cp/ml)');
legend('v(t)','Observed Data');

% Define the objective function of NLS

function diff = LSminzu(par1,Y)
[T,W] = ode45(@xyvmodel,[0:1:28],[1*10^7 2.57*10^6 2.125*10^
6],[],par1);
M = W(:,3);
MM = M([0 1 2 3 4 7 9 11 12 14 28]/1 + 1);
Y = [2.125*10^6 1.3375*10^6 7.975*10^5 8.2*10^5 7.975*10^5
9.55*10^5 5.9821*10^5 6.7857*10^5 7.5893*10^5 9.5179*10^5
6.625*10^5];
diff = MM'-Y;

% Definition of the model (7)

dw = zeros(3,1);
function dw = xyvmodel(t,w,par1)
```

```
epsilon = par1(1);k = par1(2);b = par1(3);
dw = [12-0.002*w(1)-0.5*b*w(1)*w(3);0.5*b*w(1)*w(3)-0.043*w(2);
k*(1-epsilon)*w(2)-0.65*w(3)];
```

Appendix B. The Matlab code for parameter estimation of model (7) by using MCMC (M − H) algorithm

```
function MCMC_MH
clear all
clc
Y = [2.125*10^6 1.3375*10^6 7.975*10^5 8.2*10^5 7.975*10^5 9.55*10^5 5.9821*10^5 6.7857*10^5 7.5893*10^5 9.5179*10^5 6.625*10^5];
b = 2.5*10^-7;epsilon = 0.738;k = 1;yo = 2.57*10^6; % The initial values of parameters
Olik = Loglik(b,epsilon,k,yo,Y); % Calculate the Logarithm likelihood function
w = [0.00000001,0.00,010.01,100]; % The maximum range of random walks
N = 500000; m = 300000; % MCMC cycle index and burn-in periods
MHPar = zeros(N + 1,4);
MHPar(1,:) = [b,epsilon,k,yo]; % MCMC algorithm and Random walks to update parameters
for i = 2:N + 1
NPar = [b,epsilon,k,yo] + w.*(2*rand(1,4)-1); % generate candidate parameters randomly
Nb = NPar(1);Nepsilon = NPar(2);Nk = NPar(3);Nyo = NPar(4);
if(Nb > 0&&Nepsilon > 0&&Nk > 0&&Nyo > 0) % Accept or reject
Nlik = Loglik(Nb,Nepsilon,Nk,Nyo,Y);
alpha = min(1,exp(Nlik-Olik));
if rand < alpha
b = Nb;epsilon = Nepsilon;k = Nk;yo = Nyo;
Olik = Nlik;
end
end
MHPar(i,:) = [b,epsilon,k,yo];
end
```

% Calculate the estimated values of parameters

```
b = mean(MHPar(m:end,1))
sigma_b = std(MHPar(m:end,1));
b_ci = [b-1.96*sigma_b, b + 1.96*sigma_b]
epsilon = mean(MHPar(m:end,2))
sigma_epsilon = std(MHPar(m:end,2));
epsilon_ci = [epsilon-1.96*sigma_epsilon, epsilon + 1.96*sigma_epsilon]
k = mean(MHPar(m:end,3))
sigma_k = std(MHPar(m:end,3));
k_ci = [k-1.96*sigma_k, k + 1.96*sigma_k]
yo = mean(MHPar(m:end,4))
sigma_yo = std(MHPar(m:end,4));
yo_ci = [yo-1.96*sigma_yo, yo + 1.96*sigma_yo]
figure(1)
```

% Output the random sequence of parameters

```
subplot(2,4,1);plot(MHPar(:,1));
ylabel('b','FontSize',8,'rotation',0);
subplot(2,4,2);plot(MHPar(:,2));
ylabel('epsilon','FontSize',8,'rotation',0);
subplot(2,4,3);plot(MHPar(:,3));
ylabel('k','FontSize',8,'rotation',0);
subplot(2,4,4);plot(MHPar(:,4));
ylabel('y0′,'FontSize',8,'rotation',0);
```

% Output the histogram of random sequence of parameters

```
subplot(2,4,5);histogram(MHPar(m:end,1),30);
ylabel('b','FontSize',8,'rotation',0);
subplot(2,4,6);histogram(MHPar(m:end,2),30);
ylabel('epsilon','FontSize',8,'rotation',0);
subplot(2,4,7);histogram(MHPar(m:end,3),30);
ylabel('k','FontSize',8,'rotation',0);
subplot(2,4,8);histogram(MHPar(m:end,4),30);
ylabel('y0′,'FontSize',8,'rotation',0);
% Calculate the Logarithm likelihood function
function [NLik] = Loglik(b,epsilon,k,yo,Y)
n = realpop(b,epsilon,k,yo);
NLik = sum(log(n).*Y-n-gammaln(Y + 1));
end
```

% Definition of the model (7)

```
function [n] = realpop(b,epsilon,k,yo)
W2 = zeros(291);W2(1) = yo;
F = @(t,w)[12.0-0.002*w(1)-0.5*(b)*w(1)*w(3);0.5*(b)*w(1)*w(3)
−0.043*w(2);k*(1-epsilon)*w(2)-0.65*w(3)];
[T,W] = ode45(F,[0:1:28],[1*10^7 W2(1) 2.125*10^6]);
MM = W([0 1 2 3 4 7 9 11 12 14 28] + 1,3);
for o = 1:11
n(o) = MM(o);
end
end
```

% Comparison of simulated curve in model (7) with the observed data

```
Y = [2.125*10^6 1.3375*10^6 7.975*10^5 8.2*10^5 7.975*10^5 9.55*10^5 5.9821*10^5 6.7857*10^5 7.5893*10^5 9.5179*10^5 6.625*10^5];
F = @(t,w)[12.0-0.002*w(1)-0.5*(b)*w(1)*w(3);0.5*(b)*w(1)*w(3)
−0.043*w(2);k*(1-epsilon)*w(2)-0.65*w(3)];
[T,W] = ode45(F,[0:1:28],[1*10^7 yo 2.125*10^6]);
figure(2)
plot(T,W(:,3),'k-','LineWidth',4);
hold on
plot([0 1 2 3 4 7 9 11 12 14 28],Y,'bo','LineWidth',2,'MarkerSize',9);
xlabel('Time (days)');
ylabel('HBV DNA (cp/ml)');
legend('v(t)','Observed Data');
end
```

References

Bellecave, P., Gouttenoire, J., Gajer, M., et al., 2009. Hepatitis B and C virus coinfection: a novel model system reveals the absence of direct viral interference. Hepatology 50, 46–55.

Bertoletti, A., Ferrari, C., 2003. Kinetics of the immune response during HBV and HCV infection. Hepatology 38, 4–13.

Borg, M.J., Hansen, B.E., Herrmann, E., et al., 2007. Modelling of early viral kinetics and pegylated interferon-alpha2b pharmacokinetics in patients with HBeAg-positive chronic hepatitis B. Antivir. Ther. 12, 1285–1294.

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: understanding aic and bic in model selection. Soc. Method Res. 33, 261–304.

Ciupe, S.M., Ribeiro, R.M., Nelson, P.W., et al., 2007. Modeling the mechanisms of acute hepatitis B virus infection. J. Theor. Biol. 247, 23–35.

Dahari, H., Shudo, E., Ribeiro, R.M., et al., 2009. Modeling complex decay profiles of hepatitis B virus during antiviral therapy. Hepatology 49, 32–38.

Eikenberry, S., Hews, S., Nagy, J.D., et al., 2009. The dynamics of a delay model of HBV infection with logistic hepatocyte growth. Math. Biosci. Eng. 6, 283–299.

Elaiw, A.M., 2012. Global properties of a class of virus infection models with multitarget cells. Nonlinear Dynam. 69, 423–435.

Feng, Z., Velasco-Hernandez, J., Tapia-Santos, B., et al., 2012. A model for coupling within-host and between-host dynamics in an infectious disease. Nonlinear Dynam.

68, 401–411.
Fisicaro, P., Valdatta, C., Boni, C., et al., 2009. Early kinetics of innate and adaptive immune responses during hepatitis B virus infection. Gut 58, 974–982.
Gourley, S.A., Yang, K., Nagy, J.D., 2008. Dynamics of a delay differential model of hepatitis B virus. J. Biol. Dynam. 2, 140–153.
Guo, B.Z., Cai, L.M., 2011. A note for the global stability of a delay differential equation of hepatitis B virus infection. Biosci. Eng. 8, 689–694.
Herz, A.V., Bonhoeffer, S., Anderson, R.M., et al., 1996. Viral dynamics in vivo: limitations on estimates of intracellular delay and virus decay. Proc. Natl. Acad. Sci. U. S A. 93, 7247–7251.
Hews, S., Eikenberry, S., Nagy, J.D., et al., 2010. Rich dynamics of a hepatitis B viral infection model with logistic hepatocyte growth. J. Math. Biol. 60, 573–590.
Hoofnagle, J.H., Doo, E., Liang, T.J., et al., 2007. Management of hepatitis B: summary of a clinical research workshop. Hepatology 45, 1056–1075.
Huang, G., Takeuchi, Y., Wanbiao, M., 2010. Lyapunov functionals for delay differential equations model of viral infections. Siam. J. Appl. Math. 70, 2693–2708.
Lau, G.K., Tsiang, M., Hou, J., et al., 2000. Combination therapy with lamivudine and famciclovir for chronic hepatitis b-infected chinese patients: a viral dynamics study. Hepatology 32, 394–399.
Le Mire, M.F., Miller, D.S., Foster, W.K., et al., 2005. Covalently closed circular DNA is the predominant form of duck hepatitis B virus DNA that persists following transient infection. J. Virol. 79, 12242–12252.
Lewin, S.R., Ribeiro, R.M., Walters, T., et al., 2002. Analysis of hepatitis B viral load decline under potent therapy: complex decay profiles observed. Hepatology 34, 1012–1020.
Li, Q., Lu, F., Deng, G., et al., 2014. Modeling the effects of covalently closed circular dna and dendritic cells in chronic HBV infection. J. Theor. Biol. 357, 1–9.
Liu, Q., Jiang, D., Shi, N.Z., et al., 2017. Dynamical behavior of a stochastic HBV infection model with logistic hepatocyte growth. Acta Math. Sci. 37, 927–940.
Luzyanina, T., Bocharov, G., 2014. Stochastic modeling of the impact of random forcing on persistent hepatitis B virus infection. Math. Comput. Simulat. 96, 54–65.
Manna, K., Chakrabarty, S.P., 2017. Global stability of one and two discrete delay models for chronic hepatitis B infection with HBV DNA-containing capsids. Comput. Appl. Math. 36, 525–536.
Meskaf, A., Allali, K., Tabit, Y., 2017. Optimal control of a delayed hepatitis B viral infection model with cytotoxic T-lymphocyte and antibody responses. Int. J. Dynam. Control 5, 893–902.
Mihm, U., Chan, H.L., Zeuzem, S., et al., 2008. Virodynamic predictors of response to pegylated interferon and lamivudine combination treatment of hepatitis B e antigen-positive chronic hepatitis B. Antivir. Ther. 13, 1029–1037.
Moneim, I.A., Al-Ahmed, M., Mosa, G.A., 2009. Stochastic and Monte carlo simulation for the spread of thehepatitis B. Aust. J. Basic Appl. Sci. 3, 1607–1615.
Murase, A., Sasaki, T.T., 2005. Stability analysis of pathogen-immune interaction dynamics. J. Math. Biol. 51, 247–267.
Murray, J.M., Stancevic, O., Ltgehetmann, M., et al., 2016. Variability in long-term hepatitis B virus dynamics under antiviral therapy. J. Theor. Biol. 391, 74–80.
Nakabayashi, J., Sasaki, A., 2011. A mathematical model of the intracellular replication and within host evolution of hepatitis type B virus: understanding the long time course of chronic hepatitis. J. Theor. Biol. 269, 318–329.
Nowak, M.A., Bangham, C.R., 1996. Population dynamics of immune responses to persistent viruses. Science 272, 74–79.
Nowak, M.A., Bonhoeffer, S., Hill, A.M., et al., 1996. Viral dynamics in hepatitis B virus infection. Proc. Natl. Acad. Sci. U. S. A. 93, 4398–4402.
Pang, J., Cui, J.A., Hui, J., 2012. The importance of immune responses in a model of hepatitis B virus. Nonlinear Dynam. 67, 723–734.
Pawelek, K.A., Liu, S., Pahlevani, F., et al., 2012. A model of HIV-1 infection with two time delays: mathematical analysis and comparison with patient data. Math. Biosci. 235, 98–109.
Perelson, A.S., 2002. Modelling viral and immune system dynamics. Nat. Rev. Immunol. 2, 28–36.
Powers, K.A., Dixit, N.M., Ribeiro, R.M., et al., 2003. Modeling viral and drug kinetics: hepatitis C virus treatment with pegylated interferon alfa-2b. Semin. Liver Dis. 23, 13–18.
Qesmi, R., Wu, J., Wu, J., et al., 2010. Influence of backward bifurcation in a model of hepatitis B and C viruses. Math. Biosci. 224, 118–125.
Rong, L., Dahari, H., Ribeiro, R.M., et al., 2010. Rapid emergence of protease inhibitor resistance in hepatitis C virus. Sci. Transl. Med. 2 30ra32-30ra32.
Stekler, H.O., 1991. Macroeconomic forecast evaluation techniques. Int. J. Forecast. 7, 375–384.
Sypsa, V.A., Mimidis, K., Tassopoulos, N.C., et al., 2005. A viral kinetic study using pegylated interferon alfa-2b and/or lamivudine in patients with chronic hepatitis B/HBeAg negative. Hepatology 42, 77–85.
Tsiang, M., Rooney, J.F., Toole, J.J., et al., 1999. Biphasic clearance kinetics of hepatitis B virus from patients during adefovir dipivoxil therapy. Hepatology 29, 1863–1869.
Wang, K.F., Wang, W.D., 2007. Propagation of HBV with spatial dependence. Math. Biosci. 210, 78–95.
Wang, K.F., Wang, W.D., Song, S.P., 2008. Dynamics of an HBV model with diffusion and delay. J. Theor. Biol. 253, 36–44.
Wang, W.C., Chau, K.W., Cheng, C.T., et al., 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J. Hydrol. 374, 294–306.
Wang, S., Feng, X., He, Y., 2011. Global asymptotical properties for a diffused hbv infection model with CTL immune response and nonlinear incidence. Acta Math. Sci. 31, 1959–1967.
Wang, X., Elaiw, A., Song, X., 2012. Global properties of a delayed HIV infection model with CTL immune response. Appl. Math. Comput. 218, 9405–9414.
Wang, K.F., Jin, Y., Fan, A.J., 2014. The effect of immune responses in viral infections: a mathematical model view. Discrete Contin. Dyn. Syst. Ser. B. 19, 3379–3396.
Wodarz, D., Christensen, J.P., Thomsen, A.R., 2002. The importance of lytic and nonlytic immune responses in viral infections. Trends Immunol. 23194–23200.
Xie, F., Shan, M., Lian, X., et al., 2017. Periodic solution of a stochastic HBV infection model with logistic hepatocyte growth. Appl. Math. Comput. 293, 630–641.
Xu, R., Ma, Z., 2009. An HBV model with diffusion and time delay. J. Theor. Biol. 257, 499–509.
Yousfi, N., Hattaf, K., Tridane, A., 2011. Modeling the adaptive immune response in HBV infection. J. Math. Biol. 63, 933–957.