Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 246 (2007) 499–509
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Optimal drug treatment regimens for HIV depend on adherence

O. Krakovska*, L.M. Wahl

Department of Applied Mathematics, University of Western Ontario, London, Ont., Canada N6A 5B7

Received 1 September 2006; received in revised form 21 December 2006; accepted 21 December 2006
Available online 23 January 2007

Abstract

Drug therapies aimed at suppressing the human immunodeficiency virus (HIV) are highly effective, often reducing the viral load to below the limits of detection for years. Adherence to such antiviral regimens, however, is typically far from ideal. We have previously developed a model that predicts optimal treatment regimens by weighing drug toxicity against $CD4^+$ T-cell counts, including the probability that drug resistance will emerge. We use this model to investigate the influence of adherence on therapy benefit. For a drug with a given half-life, we compare the effects of varying the dose amount and dose interval for different rates of adherence, and compute the optimal dose regimen for adherence between 65% and 95%. Our results suggest that for optimal treatment benefit, drug regimens should be adjusted for poor adherence, usually by increasing the dose amount and leaving the dose interval fixed. We also find that the benefit of therapy can be surprisingly robust to poor adherence, as long as the dose interval and dose amount are chosen accordingly.

Keywords: HIV; Mathematical model; Theoretical immunology; Toxicity; Drug resistance; Adherence

1. Introduction

Highly active antiretroviral therapy (HAART) allows for the effective suppression of viral replication in human immunodeficiency virus (HIV)-infected individuals for years or even decades. Adherence to HAART regimens, however, is on average very low; estimates of mean adherence rates typically range from 45% to 85% (Arnsten et al., 2001; Howard et al., 2002; Weiser et al., 2003; Giordano et al., 2004; Hinkin et al., 2004), although average adherence has been reported to be as high as 93% (Oyugi et al., 2004).

Imperfect adherence is due to a number of medication- and patient-related factors; in particular, the severe and even life-threatening side effects of antiretroviral therapy are an increasingly important issue. For example, a substantial fraction of all hospital admissions for HIV-positive patients in the developed world are due to HAART-related toxicity, typically manifested as hepatitis (Martin-Carbonero et al., 2001). To reduce side effects, individuals maintaining HAART often reduce their own drug exposure through self-prescribed "drug holidays" (Kathleen, 2000; Siegel et al., 2000; Ammassari et al., 2001).

Imperfect adherence not only undermines the efficacy of therapy, but may also facilitate the emergence of drug resistance (Montaner et al., 1998; Harrigan et al., 2005; Walsh et al., 2002; Sethi et al., 2003; Bangsberg et al., 2003). Drug resistance is one of the main reasons for virologic failure during HAART (DeGruttola et al., 2000); viral suppression cannot be achieved in the face of multiply resistant viral strains.

The influence of adherence to HAART on treatment outcome has been investigated in a number of clinical studies. Paterson et al. (2000) reported a strong association between adherence rates and four measures of therapeutic success: virologic failure, number of days spent in the hospital, incidents of opportunistic infection and death. Similar results were described by Descamps et al. (2000), who found that lower adherence to indinavir therapy was associated with a higher risk of early virologic failure. Likewise, Roca et al. (2000) observed significant associations between adherence and baseline viral load. De Olalla

*Corresponding author. Tel.: +1 519 661 3649; fax: +1 519 661 3523.
E-mail address: olga.krakovska@alumni.uwo.ca (O. Krakovska).

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2006.12.038

et al. (2002) and Wood et al. (2003) found that the mortality rate among adherent patients is significantly lower than among non-adherent patients. Gross et al. (2001) found that adherence is important in achieving viral suppression with a newly initiated treatment regimen. In contrast, Maitland et al. (2005) found no association between virologic failure and adherence.

Although clinical studies such as these are essential in elucidating the complex effects of adherence in HAART, important insights can also be gained through the careful use of mathematical models. For example, Wahl and Nowak (2000) used the basic reproductive ratio associated with a particular treatment regimen to estimate the degree of adherence required to suppress viral replication and minimize the risk of drug resistance; both single and triple drug regimens were investigated. Phillips et al. (2001b) considered four different patterns of adherence in triple drug therapy and found patterns which lowered the probability that resistance would emerge during therapy. Huang et al. (2003) investigated the effect of imperfect adherence on viral rebound, and found that longer sequences of missed doses implied faster viral rebound. Smith (2006) used an impulsive differential equation model (Smith and Wahl, 2005) to estimate the number of doses that can be missed, as well as the number of subsequent doses necessary to suppress the growth of a 50-fold resistant strain, for 12 different drugs. In another group of studies (Wu et al., 2005; Verotta and Schaedeli, 2002; Ferguson et al., 2005) the adherence patterns of individual patients were recorded and used to estimate overall drug exposure. Wu et al. (2005) found that this overall drug exposure was a significant predictor of viral load. Verotta and Schaedeli (2002) fit clinical data describing viral dynamics to four basic models, and found that parameter estimates did not change significantly when adherence was explicitly included. Finally, Ferguson et al. (2005) demonstrated that adherence was a significant factor in predicting between-patient differences in viral load and T-cell count.

Our recent work has used mathematical modeling to investigate possible drug-sparing antiviral regimens—regimens which reduce toxicity while suppressing both viral replication and the risk of drug resistance (Krakovska and Wahl, 2007). Here we extend this work to consider, in detail, the effects of non-adherence. Our model incorporates accurate pharmacokinetics and quantitatively balances the benefits of therapy against both toxicity and the risk of drug resistance. We investigate changes in the optimal treatment regimen for patients with different levels of adherence, and estimate the effects of non-adherence on the overall benefit of therapy. These results sensitively depend on the underlying pharmacokinetics of the antiviral drug, particularly the drug half-life; we investigate a range of half-lives between 3 and 60 h. A surprising result of this investigation is that substantial treatment benefit can often be maintained, even when adherence is poor.

2. Methods

2.1. The model

Our model for in-host HIV dynamics (Krakovska and Wahl, 2007) consists of five populations: $CD4^+$ T-cells which are naïve (x), activated (r) or productively infected (y); and $CD8^+$ T-cells which are naïve (u) or activated (z). We include both naïve and activated cells in order to capture any possible boosting of the immune system during breaks in therapy. The concentration of cells in a small volume of plasma for each population is modeled as follows:

$$\dot{x} = \lambda_x - \beta(1-\eta)xy - d_x x - \alpha xy, \tag{1}$$

$$\dot{r} = \alpha b xy - \beta(1-\eta)ry - d_r r, \tag{2}$$

$$\dot{y} = \lambda_y + \beta(1-\eta)(r+x)y - d_y y - \rho zy, \tag{3}$$

$$\dot{u} = \lambda_u - d_u u - \xi ury, \tag{4}$$

$$\dot{z} = \xi g ury - d_z z. \tag{5}$$

The parameters, model assumptions and rationale for this system of equations have been described in detail elsewhere (Krakovska and Wahl, 2007). Briefly, $\lambda_x[\lambda_u]$ is the production rate of naïve $CD4^+[CD8^+]$ T-cells by thymus, $d_x[d_u]$ is the corresponding death rate, $d_r[d_z]$ is the death rate of activated $CD4^+[CD8^+]$ T-cells, λ_y is the input of productively infected cells from the latent source, d_y is the death rate of the productively infected cells and β is the infection rate. We make the conservative assumption that all infected cells are productively infected. In addition to these terms, our model includes T-cell activation and cloning. The mass-action term αxy represents the activation of naïve $CD4^+$ T-cells by antigen-presenting cells, which are assumed to be proportional to productively infected cells y. Activated $CD4^+$ T-cells can further clone themselves; a newly activated $CD4^+$ T-cell produces on average $b-1$ clones, such that the rate of increase in r through activation and cloning is αbxy. Activated $CD4^+$ are also required to activate naïve $CD8^+$ T-cells, with the overall activation rate ξury. Finally, activated $CD8^+$ T-cells can clone themselves at rate $g-1$, and kill productively infected cells at rate ρ. In all of the numerical work to follow, we use parameter values as described in Krakovska and Wahl (2007). A sensitivity analysis for these parameter values also appears in Krakovska and Wahl (2007).

The drug effectiveness, η, is the degree to which the infection rate β is reduced by the current drug concentration, with $0 \leqslant \eta \leqslant 1$. We use Michaelis–Menten dynamics to model changes in η with changing drug concentrations, $\eta(t) = c(t)/(c(t) + IC_{50})$ (Wahl and Nowak, 2000), where $c(t)$ is the intracellular drug concentration, and IC_{50} is the concentration necessary to inhibit viral replication by 50%. Although *in vivo* IC_{50} values are very difficult to measure, for the purposes of this model we are able to use drug

concentrations normalized by IC_{50}, $C(t) = c(t)/IC_{50}$, such that $\eta(t) = C(t)/(C(t)+1)$. This approach has the additional advantage of removing one free parameter from the model.

We assume that the normalized drug concentration, $C(t)$, can be modeled using an impulsive differential equation as follows (Smith and Wahl, 2004, 2005): doses are taken at fixed times $0, \tau, 2\tau, \ldots$, where τ is a constant dose interval. Between doses, the drug concentration decays exponentially with decay rate w, determined by the drug half-life $t_{1/2}$ as $w = \ln 2/t_{1/2}$:

$$\frac{dC}{dt} = -wC, \quad t \neq 0, \tau, 2\tau, \ldots \tag{6}$$

with $C(0) = 0$. At the dosing times, the drug concentration increases instantaneously by the single dose amount, C^m, if a dose is taken:

$$C(t_k^+) = \begin{cases} C(t_k^-) + C^m, & t = 0, \tau, 2\tau, \ldots \quad \text{if dose is taken,} \\ C(t_k^-), & t = 0, \tau, 2\tau, \ldots \quad \text{if dose is missed.} \end{cases} \tag{7}$$

This impulsive system implicitly assumes that the time-to-peak is negligible compared to the intracellular decay rate. This formulation also assumes that drug concentrations add linearly with increasing dose, which is clearly an oversimplification of the underlying pharmacodynamics. Thus the total drug concentration accrues over several doses (see Fig. 1), reaching a periodic orbit if adherence is perfect, as elucidated in detail in Smith and Wahl (2004). To clarify notation, C^m is the instantaneous increase in drug concentration with each dose, while C_{max} is the peak concentration reached during the periodic orbit, for perfect adherence. We will also use the term "single dose efficacy", denoted by η^m, which we define as the instantaneous increase in drug efficacy with each dose: $\eta^m = C^m/(C^m+1)$. For convenience, we also refer to η^m as simply the "dose amount". Analogously, η_{max} is the peak efficacy reached during the periodic orbit (peak values in Fig. 1).

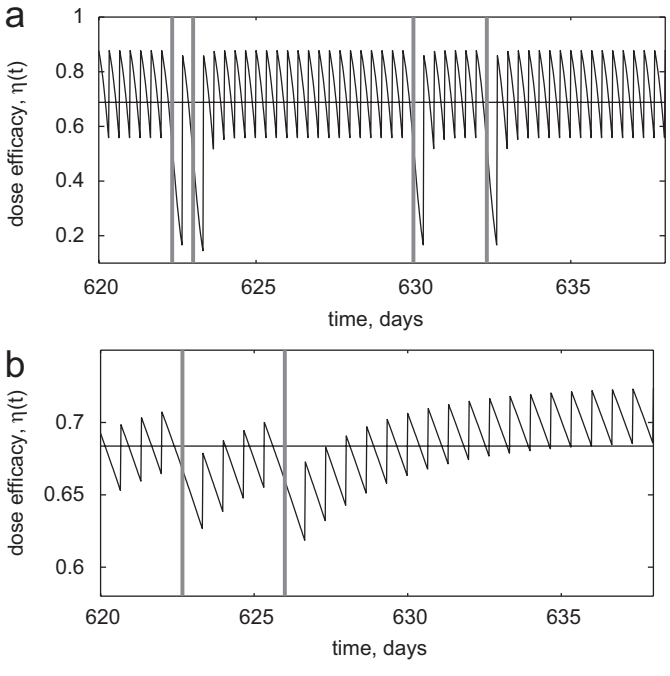

Fig. 1. Changes in the drug efficacy $\eta(t)$ versus time, determined by Michaelis–Menten dynamics. Drug concentration is determined as a numerical solution to the impulsive differential equations (6) and (7), with 90% adherence. The top panel gives the corresponding value of $\eta(t)$ for a drug with a 3 h half-life; for the bottom panel $t_{1/2} = 60$ h. Note that depending on the drug half-life, several consecutive doses may be required to return to the periodic orbit after a dose is missed (vertical lines).

2.2. Treatment benefit

To compare treatment regimens, we have previously developed a metric that incorporates the three main goals of HIV therapy (Krakovska and Wahl, 2007): (i) maximize the population of $CD4^+$ T-cells; (ii) minimize toxicity; (iii) minimize the risk of drug resistance. In this approach, we assess the benefit of a given treatment regimen, by integrating the population of naïve $CD4^+$ T-cells, x, over the course of therapy. To assess drug toxicity, we assume that side effects increase linearly with drug concentration, $C(t)$, and normalize for different drugs by dividing by the average drug concentration when each drug is taken with perfect adherence. Thus, if C^a gives the mean drug concentration when the impulsive equations (6) and (7) reach the periodic orbit, we gauge overall toxicity by integrating $C(t)/C^a$ over the treatment period $[t_0, t_f]$:

$$\text{Toxicity} \propto \int_{t_0}^{t_f} \frac{C(t)}{C^a}\, dt. \tag{8}$$

To assess the risk of drug resistance, we estimate a probability distribution for the time at which resistance emerges, t_r. Before the emergence of drug resistance, the costs of therapy, in terms of side effects, are subtracted from the treatment benefit, in terms of the naïve $CD4^+$ T-cell count. Therefore, treatment benefit during the time interval before resistance emerges, $[t_0, t_r]$, is given by $\int_{t_0}^{t_r} x - q(C(t)/C^a)\, dt$. After resistance emerges, we assume a worst-case scenario, that is, we assume that the emergent strain is fully resistant to current therapy, the patient is taken off therapy and the $CD4^+$ T-cell count falls to 200 cells/μl, consistent with a diagnosis of AIDS. Therefore the treatment benefit on $[t_r, t_f]$ is $\int_{t_r}^{t_f} 200\, dt$, since $C(t)$ is zero in this interval.

To estimate the time at which resistance emerges, we use the simplest possible model, assuming that the rate at which resistance emerges per unit time, or the "failure rate", F, is constant. This is clearly a simplifying assumption, since drug concentrations in particular may fluctuate on an hourly time scale. However, for the optimal treatment regimens we compute, this rate is uniformly small (see Appendix B) and the expected fluctuations in F are small as well. We thus find that t_r is exponentially distributed, with probability density Fe^{-Ft_r}. The mean failure rate F can be computed, using clinical data, as

described in Appendix B. We note that F depends on the drug concentration, $C(t)$, over the entire course of therapy, as well as on the populations of susceptible and infected $CD4^+$ T-cells at all times.

The overall benefit of therapy, T, over a time period $[0, t_f]$ is then given by

$$T = \int_{t_0}^{t_f} p(t_r)\left[\int_{t_0}^{t_r}\left(x - q\frac{C}{C^a}\right)\mathrm{d}t + \int_{t_r}^{t_f} 200\,\mathrm{d}t\right]\mathrm{d}t_r + (1 - P(t_f))\int_{t_0}^{t_f}\left(x - q\frac{C}{C^a}\right)\mathrm{d}t, \quad (9)$$

where e^{-Ft_f} gives the probability that resistance emerges after t_f. Thus the first integral estimates the benefit of therapy if resistance emerges before t_f and therapy is discontinued, while the second term estimates the benefit if resistance is not predicted to emerge during the treatment interval. The parameter q is a scalar which gives the weight between $CD4^+$ T-cell counts, x, and normalized side effects C/C^a; following Krakovska and Wahl (2007), we use the methods described in Appendix A to estimate this weight.

For the purposes of this investigation, we consider a hypothetical antiviral pharmaceutical, with a fixed half-life. Our goal is to predict optimal dose amounts and dose intervals for this drug, and to examine how these optima might change when imperfect adherence is considered. To compute the benefit of therapy in Eq. (9), however, we need the average drug concentration, C^a, when the drug is taken as typically prescribed with perfect adherence. The typical dose interval for a drug of a given half-life can be estimated easily by comparing drug intervals and half-lives for currently available antivirals (see Table 1). To estimate the typical dose amount, we make the somewhat arbitrary assumption that drugs will typically be prescribed such that the minimum (trough) drug concentration, when taken with perfect adherence, corresponds to a drug efficacy of 90%. This estimate falls in the range estimated by Montaner et al. (2001). Thus

$$\frac{C_{max}\mathrm{e}^{-w\tau}}{C_{max}\mathrm{e}^{-w\tau} + 1} = 0.9. \quad (10)$$

Integrating over a single dose interval, we can then compute the average drug concentration, C^a, for a typical dosing regimen:

$$C^a = \frac{9\mathrm{e}^{w\tau}(1 - \mathrm{e}^{-w\tau})}{w}. \quad (11)$$

We consider a treatment interval of two years, and assume that the system is at the infected equilibrium at the start of therapy. This allows us to capture an upper bound on the probability that resistance emerges, thus giving conservative estimates of the benefit of therapy.

We consider adherence between 65% and 95%. In modeling imperfect adherence, we assume that each dose is taken independently with a probability equal to the adherence level, and a dose is never taken when it is not prescribed.

We consider drugs with four different half-lives for each adherence level: a drug with a 3 h half-life, standardly prescribed every 8 h; a drug with a 6 h half-life, prescribed every 12 h; and drugs with 20 and 60 h half-lives, both prescribed every 24 h. These half-lives and dosing intervals span the range of currently available protease and reverse transcriptase inhibitors, as shown in Table 1. We then numerically compute the target function value corresponding to other dose intervals and dose amounts, for a given level of adherence. The dose amount is varied such that the single dose efficacy, η^m, changes from 0.25 to 1, and the dose interval is varied from 8 to 100 h in increments of 4 h. Note, however, that not all of these dose intervals are clinically feasible; for example, a dosing interval of 20 h would not be realistically sustainable. We thus limit dosing intervals to $\tau = 8, 12, 24, 48, 72$ or 96 h when discussing "practically feasible" schedules.

The treatment benefit, T, is a random variable which depends on the particular sequence of missed doses. We thus estimate T by numerically evaluating Eq. (9) over 250 simulation runs for each dose interval, dose amount and adherence level.

Table 1
Half-lives and typical dose intervals for antiviral drugs, including protease inhibitors (PI), nucleoside analog reverse transcription inhibitors (NRTI) and non-nucleoside reverse transcription inhibitors (NNRTI)

Drug	Class	$t_{1/2}$ (h)	Dose interval (h)	Dosage (mg)
Indinavir	PI	1.5[a]	8	800
Abacavir	NRTI	18[b]	12	300
Stavudine	NRTI	3.5[b]	12	40
Zidovudine	NRTI	7[b]	24	200
Ritonavir	PI	4[a]	12	600
Nelfinavir	PI	4.25[a]	12	1200
Delavirdine	NNRTI	5.8[a]	12	400
Amprenavir	PI	8.85[a]	12	1200
Lamivudine	NRTI	12.5[b]	24	300
Didanosine	NRTI	16[b]	24	400
Nevirapine	NNRTI	16.5[a]	24	200
Efavirenz	NNRTI	64[a]	24	600

[a]Serum half-life.
[b]Intracellular half-life.

3. Results

Fig. 1 provides an illustrative example of changes in the drug effectiveness, $\eta(t)$, versus time. When a single dose is missed (vertical lines), the drug continues to decay with its usual half-life; we illustrate $\eta(t)$ for drugs with a short (3 h) and long (60 h) half-life, and adherence of 90%. Depending on the half-life, recovery of the drug concentration to the periodic orbit may require several doses to be taken in succession after missed doses. The horizontal line on each panel illustrates the mean drug efficacy achieved with this level of adherence.

Fig. 2 illustrates, for 90% adherence, how the overall treatment benefit, T, changes with dose interval (τ) and single dose efficacy, η^m. We tested dose intervals ranging from 8 to 100 h, in increments of 4 h. Panels show results for drug half-lives of 3, 6, 20 and 60 h. Stars show the best attainable treatment benefit, T_{max}; circles show the best practical treatment benefit, achievable with a clinically feasible dose interval. At 90% adherence, we predict that drugs with half-lives of 3, 6 or 20 h should be taken every 8 h, the minimum dose interval we tested; stars and circles overlap at an 8 h dose interval in these three cases. For short half-life drugs, the range of single dose efficacy at which treatment benefit outweighs the cost of toxicity is narrow, while the peak on the surface is considerably broader for longer half-lives. For longer half-lives, a range of "optimal" regimens may be nearly equivalent, with higher doses recommended for longer dose intervals. The sharp drop in treatment benefit for high dose amounts and short intervals is due to drug toxicity; more drug is taken than is necessary to control the virus in this range.

We also computed the mean basic reproductive ratio, R_0, for each combination of single dose efficacy and dose interval by estimating the average drug efficacy achieved for each combination and a particular level of adherence. Results for 90% adherence are shown in Fig. 3. We find that the optimal drug regimen reduces R_0 to just below unity. Individual lines in this figure show how the treatment benefit changes for fixed dose amount when the dose interval is varied. When R_0 is allowed to exceed one, the treatment benefit drops sharply, whereas the effect of reducing R_0 below its optimal value is less sharp; the decrease in T on the left is due to increased drug toxicity.

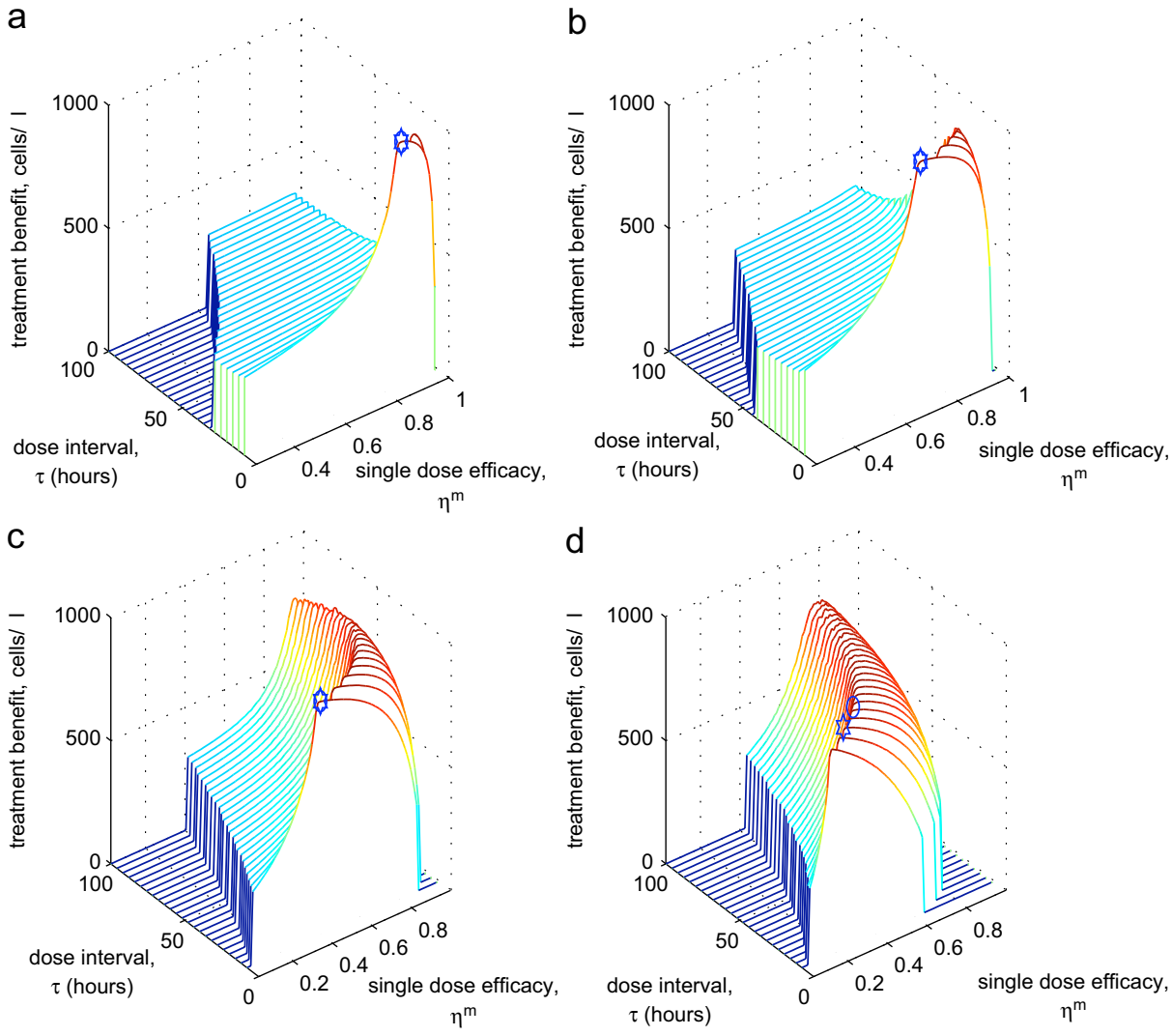

Fig. 2. Treatment benefit for various dosing regimens, determined for 90% adherence and for drug half-lives $t_{1/2} = 3, 6, 20$ and 60 h (panels a,b, c and d, respectively). A numerical solution of Eq. (9), subject to the cell dynamics described by Eqs. (1)–(5), was found as described in the text for dose intervals, τ, between 8 and 100 h, and single dose efficacy, η^m, between 0.025 and 1. Stars show the peak on the surface, that is, the best attainable treatment benefit; circles show the best practical treatment benefit, achievable with a clinically feasible dose interval. Note that for short half-life drugs, the range of single dose efficacy for which treatment benefit outweighs the cost of toxicity is narrow, while the peak on the surface is considerably broader for longer half-lives.

We varied adherence between 65% and 95%, and regenerated surfaces such as those shown in Fig. 2 for 90% adherence. Fig. 4 shows how the optimal dose regimen changes with adherence. For a drug half-life of 3 h, the optimal dose interval is always at the minimum, 8 h. However, the recommended dose amount decreases with increasing adherence (left panel, squares). Similarly, the benefit of therapy increases with adherence (stars). Results for half-lives of 6 and 20 h were qualitatively identical to the results illustrated here.

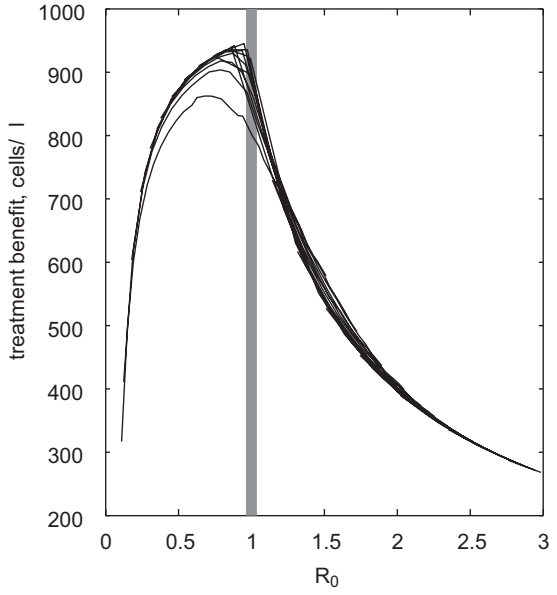

Fig. 3. Treatment benefit versus basic reproductive ratio, R_0. Treatment benefit was computed as described in the legend to Fig. 2. For each combination of single dose efficacy and dose interval, the average drug effectiveness, $\bar{\eta}$, was estimated over 250 simulation runs and a treatment interval of 2 years. The basic reproductive ratio was then calculated as $R_0 = \beta\lambda_x(1-\bar{\eta})/d_x d_y$. Individual lines in this figure show how the treatment benefit changes for fixed dose amount when the dose interval is varied.

For a 60 h half-life (right panel), we find that the optimal regimen is extremely sensitive to adherence. The figure shows the recommended dose interval and single dose amount for adherence between 0.65 and 0.95. Circles show drug regimens that are practically achievable, while stars show the numerically optimal solutions. We find that as adherence increases, smaller dose amounts, with the same dose interval, are beneficial. As adherence increases further, the optimal regimen "jumps" to a longer dose interval, and larger dose amount. We further note that the recommended dose intervals range from 8 to 24 h, and that the recommended dose amount changes by nearly a factor of 2 for adherence between 75% and 95%. These changes in dose amount and dose interval counter-balance each other, however, such that the corresponding peak efficacy, η_{max}, changes only modestly, and is in the range of 0.71–0.79 for all the optima shown here.

The final goal of our investigation was to evaluate the performance of "optimal" dose regimens, determined for a specific level of adherence, at other adherence levels. Results are shown in Fig. 5 for short and long drug half-lives. We find that the benefit of therapy is reduced if a regimen which works best at one adherence level is taken with *either* higher or lower adherence. This reduction, however, is not symmetrical: our model predicts that a dose regimen which is optimal at high adherence performs very poorly at lower adherence (steep slope on the left of each curve); in contrast, a regimen designed for low adherence

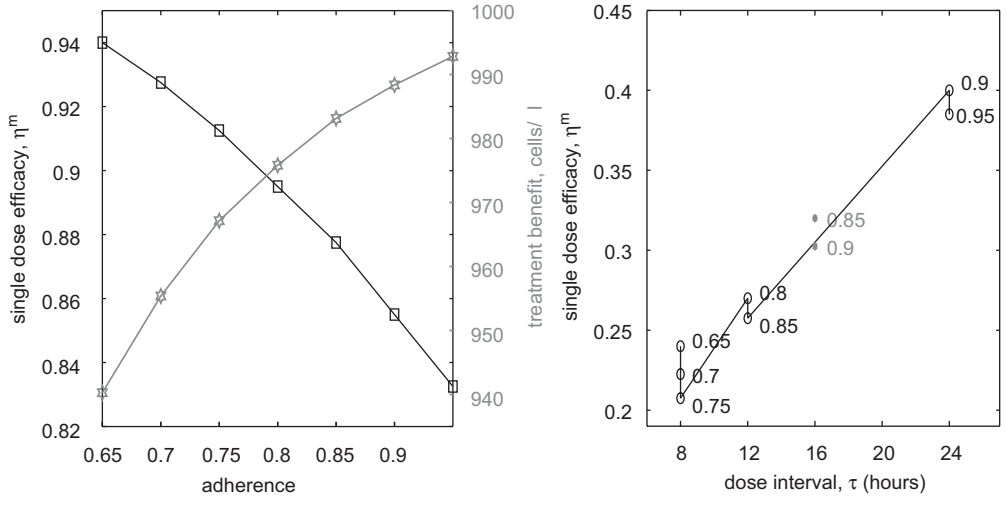

Fig. 4. Changes in the optimal dose regimen with adherence. The optimal dose regimen was determined as the combination of dose interval, τ, and single dose efficacy, η^m, that maximizes treatment benefit, as illustrated by the stars and circles in Fig. 2. The panel on the left shows results for $t_{1/2} = 3$ h. As adherence increases, the single dose efficacy η^m decreases, and treatment benefit increases. Although not illustrated in the figure, the dose interval remained the same, $\tau = 8$ h. Results for 6 and 20 h half-lives were qualitatively identical to this case (not shown). The panel on the right shows how the "optimal" single dose efficacy η^m and dose interval are adjusted for adherence when $t_{1/2} = 60$ h. Adherence is labeled to the right of each point; the solid line shows the path of change as adherence increases. Circles give drug regimens that are practically achievable, while stars show numerically best solutions.

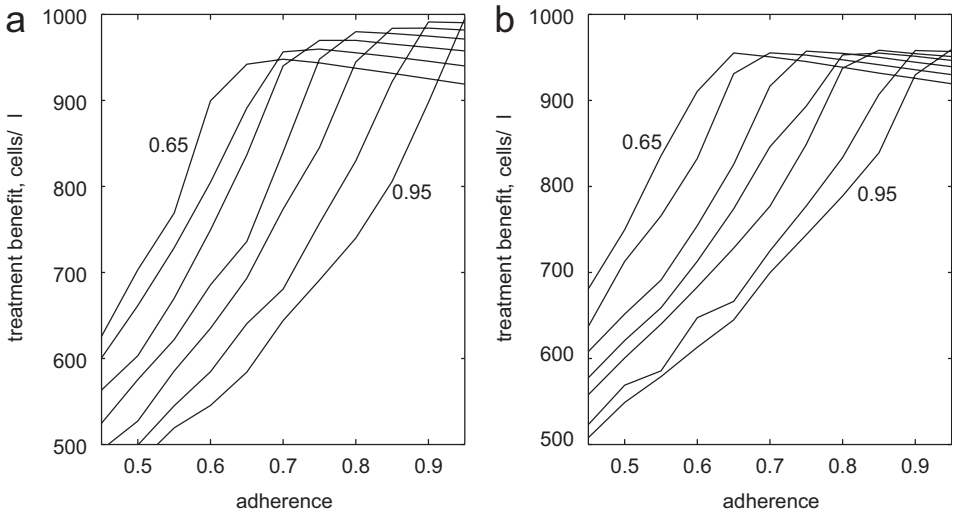

Fig. 5. Performance of "optimal" dosing regimens, determined for a specific level of adherence, at other adherence levels. The optimal dose amount and dose interval for each adherence rate were first determined as described in the text and the legend to Fig. 2. The treatment benefit for each of these optimal regimens was then computed at different adherence rates by computing the mean value of T for 250 independent simulation runs at each rate of adherence between 0.45 and 0.95 in increments of 0.05. The panel on the left gives results for $t_{1/2} = 3$ h; the panel on the right gives results for $t_{1/2} = 60$ h. Note that the maximum benefit attainable is approximately the same for all adherence levels, particularly when the half-life is long.

still performs relatively well at higher adherence levels. Note that when adherence is high, the maximum treatment benefit attainable is slightly higher for short half-life drugs (panel (a)); this is possibly due to a reduction in side effects because of the faster decay of the drug from the system.

Despite this subtle difference, the peak values for each curve in Fig. 5 have approximately similar heights, especially for drugs with long half-lives (Fig. 5(b)). This unexpected result predicts that poor adherence does not necessarily compromise the treatment benefit that is attainable, when doses are missed independently and randomly, as long as the dose amount and dose interval are tailored to the adherence level. This counter-intuitive result will be taken up again in the Discussion.

4. Discussion

In general, our results predict that the optimal dosing regimen is sensitive to adherence. Our results suggest that drug regimens could be adjusted for poor adherence, typically by increasing the dose amount, while the dose interval remains fixed. For drugs with very long half-lives, we predict that both the recommended dose amount and dose interval should be adjusted. In the example illustrated in Fig. 4, the recommended dose amount is cut in half, while the dose interval decreases from 24 to 8 h if adherence decreases from 95% to 75%. We note, however, that decreasing the dose interval might itself further compromise adherence, a factor which we have not yet considered. Nonetheless, the most interesting prediction of this model is that similarly effective therapy can often be achieved, even for low adherence, if the dosing regimen is adjusted appropriately. In contrast, we find that poor adherence severely compromises the benefit of therapy if the treatment regimen is not adjusted.

In particular, our model predicts, not surprisingly, that therapies designed for 90% adherence do not perform well at 65% adherence. More interestingly, we also predict that if therapies which work best at 65% adherence are taken with 90% adherence, the benefit of therapy is also diminished: since the virus is well controlled by the prescribed regimen at 65% adherence, 90% adherence likely increases toxicity without improving viral suppression. These quantitative findings are important since dose amounts and timings are often validated under the conditions of controlled clinical trials, while adherence rates in a random clinical population may vary more widely (Bangsberg et al., 2000). Ideally, patient adherence should be assessed before prescribing therapy, and monitored periodically, in order to prescribe a regimen that is most effective for the individual.

The model we describe focuses on estimates of optimal dose regimens for a single antiviral with a known half-life. This approach might also be applicable to a combination of antivirals taken together (in a single pill), as long as the drug half-lives were similar. An avenue for future work is the extension of this model to combination therapies during which drugs are taken independently, and adherence by the same individual to different drugs might vary. Future extensions to this approach should also include more realistic patterns of non-adherence (random drug holidays, imperfect timing of successive doses), more accurate intracellular pharmacokinetics and better estimates of the rates at which drug resistance emerges, since these are known to depend on the type of drug used (Bangsberg et al., 2004).

The benefit of therapy, as estimated in our model, explicitly includes the negative effects of drug toxicity. While the success of HAART in prolonging survival and delaying the progression to AIDS cannot be overstated, persons living with HIV/AIDS are now typically expected to participate in drug therapy for years or even decades. It is critical that well-tolerated therapeutic regimens, with realistic expectations for adherence, are developed over the next several years. The theoretical predictions we describe here represent a small step in that direction.

Acknowledgments

This work is supported by the Natural Sciences and Engineering Research Council of Canada, the Ontario Ministry of Science, Technology and Industry and by the SHARCNET parallel computing facility. We are indebted to Vasily Vakorin and Jane Heffernan for many helpful discussions, and to two anonymous referees for insightful comments.

Appendix A. Estimating the weight of drug toxicity, q

One of the more difficult aspects of this work is to balance, quantitatively, the costs of side effects against the benefits of improved helper T-cell counts. As in Krakovska and Wahl (2007), we do this by considering clinical practice. In particular, a standard guideline is to begin triple-drug therapy when $CD4^+$ T-cell levels fall to 350 cells/μl or less.

Consider two possible treatment strategies: one is to initiate treatment at the moment of infection, t_0; the other is to delay treatment until t_{350}, the time at which $CD4^+$ T-cell counts reach 350 cells/μl. The clinical practice outlined above implicitly assumes that the drug toxicity avoided by this delay "is worth" the loss of $CD4^+$ T-cells experienced during this interval. Thus $\int_{t_0}^{t_{350}}(x_{\max}-q)\,dt = \int_{t_0}^{t_{350}} x(t)\,dt$, where $x(t)$ gives the time course of naïve $CD4^+$ T-cell counts in the absence of therapy. Although the time course of the loss of x will clearly differ from patient to patient, assuming this loss is roughly linear, we find $q = (x_{\max}-350)/2$.

Appendix B. Estimating the probability that resistance emerges

To estimate the rate at which resistance emerges, the failure rate F, we compute the probability that new drug-resistant mutations arise during the therapeutic interval $[t_0, t_f]$, and multiply by the probability that each new mutation forms a lineage of infected cells which will ultimately survive in the population of infected cells. The latter probability is given by the fixation probability, π, as described below. The former probability is given by the overall number of newly infected cells, multiplied by the mutation rate μ: $\mu W \int_{t_0}^{t_f} \beta(1-\eta)(x+r)y\,dt$. Here W is a conversion coefficient which is necessary to compute the number of drug-resistant mutations which arise in the whole body (particularly in lymph tissue), given the rate at which such new mutations occur in 1 μl of blood.

The fixation probability π is governed by the basic reproductive ratio of the drug-resistant viral strain, R_m. Assuming complete drug resistance, we have

$$R_m = \frac{\beta(1-\sigma)(x+r)}{d_y+\rho z} \leqslant \frac{\beta(x+r)}{d_y}. \tag{12}$$

For a linear birth–death process (Karlin and Taylor, 1975), we have $\pi = 1 - 1/R_m$.

Therefore, the rate at which resistance emerges, F, can be expressed as

$$F = \frac{W\mu\beta}{t_f}\int_{t_0}^{t_f}(1-\eta(t))(x(t)+r(t))y(t) \times\left(1-\frac{d_y}{\beta(x(t)+r(t))}\right)dt. \tag{13}$$

Clearly W is unknown; however, we can find a reasonable numerical estimate of the product $W\mu\beta$ in the following way. A distribution of adherence rates has been observed for patients on continuous drug regimens (Eldred and Cheever, 1998). Likewise, overall immunological failure rates, due to drug resistance, have been estimated in various studies (Grabar et al., 2000; Phillips et al., 2001a; Martinez-Picado et al., 2003; Mocroft et al., 2004; Lohse et al., 2005; Phillips et al., 2005). We assume a distribution of adherence rates, and for each half-life and standard dose interval, solve for the value of $W\mu\beta$ which gives the appropriate overall failure rate. The failure rate we assume is that resistance emerges in 12.5% of patients in six years (Mocroft et al., 2004). The adherence rate distribution is based on Eldred and Cheever (1998): 30% of patients have very low adherence (0.25), 20% have low adherence (0.65) and 50% have high adherence (0.9). This yields values of $W\mu\beta$ in the range of $4.2-8.1\times10^{-8}$. We can then use the appropriate value of $W\mu\beta$, for a particular dose amount, dose interval and adherence rate, to compute the predicted failure rate, F.

The reciprocal of the failure rate, $T_F = 1/F$, gives the expected time until resistance emerges. We illustrate the behavior of this variable in Fig. 6. On the left (Fig. 6(a)), we illustrate that treatment benefit is severely compromised when the expected time until resistance emerges is not sufficiently long. This allows our optimization routine to reject treatment regimens which allow drug resistance to develop. However, we note that for the *optimal* drug regimens predicted by our model, T_F is very long, and thus has a negligible effect on the benefit of treatment. In particular, for the optimal regimens found in this study, T_F was always longer than 1000 years.

T_F itself depends directly on the mutation rate, but also depends in a more complicated way on drug efficacy. We illustrate the latter effect by fixing $\eta(t)$ to a constant level, η^c, and computing T_F from Eq. (13). Fig. 6(b) illustrates that if the mean drug efficacy is high, the expected time

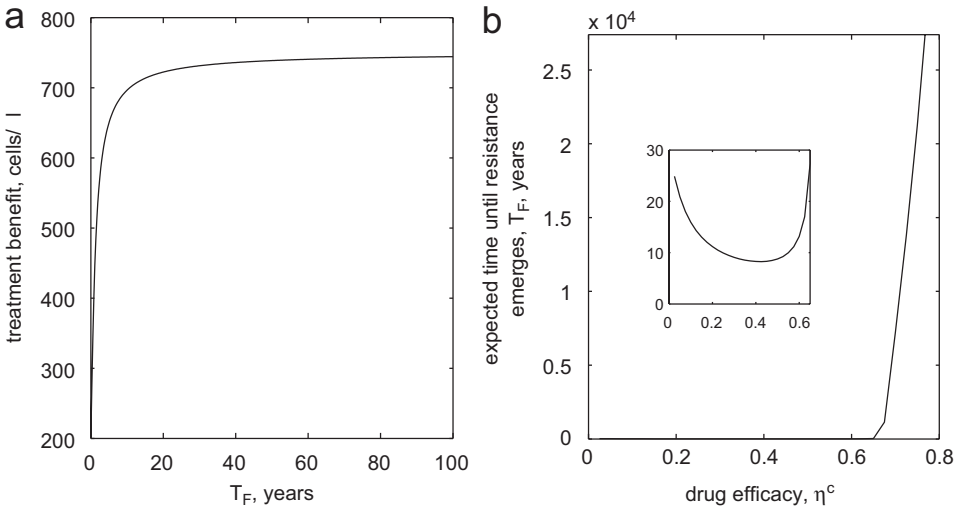

Fig. 6. Behavior of the expected time until resistance emerges. In panel (a), treatment benefit was determined in Eq. (9) for continuous therapy (perfect adherence), for different values of T_F, the expected time until resistance emerges. Note that T_F enters Eq. (9) as the failure rate $F = 1/T_F$. We find that treatment benefit is insensitive to T_F once this time is sufficiently long; however, regimens with short values of T_F will not perform well. Panel (b) illustrates the dependence of T_F on mean treatment efficacy, η^c. Here T_F was computed by setting $\eta(t)$ to a constant value, η^c in Eqs. (1)–(5), computing the cell dynamics, and using these numerical results and η^c in Eq. (13). Either large or small values of η^c result in very long times before resistance will emerge (for small values of η^c, refer to the inset).

until resistance emerges is long. In the inset, we see that T_F is also very long if η^c is small, that is, if the infected cells are not exposed to the drug. In contrast, intermediate drug levels promote the emergence of the drug resistance. Once again, although the optimal drug regimens we report have very long values of T_F, this functional behavior is important in eliminating regimens which may facilitate the emergence of drug resistance.

As an aside, we predict that the product $W\mu\beta$ is in the range of $4.2-8.1 \times 10^{-8}$. We also note that the variability observed clinically in HIV viral load has been shown to be best matched by an individual-based simulation in which the infection is driven by 300–500 independent infectious units (Heffernan and Wahl, 2006). Since our model assumes 50 independent infected cells per μl, this suggests that the conversion factor W should be in the range of 6–10. Using $\beta = .0028$, this predicts that μ is in the range of $1.5-4.7 \times 10^{-6}$. We note that μ gives the rate at which mutations which confer drug resistance are predicted to occur, based on the clinical data cited above. The values predicted here are smaller than the point mutation rate, but larger than the 2-point mutation rate for double mutations. This is perhaps not surprising since most of the patients in the studies cited above, from which we estimate $W\mu\beta$, were prescribed combination therapy, but resistance presumably arose during substantial periods of lower concentration for one or more drugs, due to low adherence.

References

Ammassari, A., Murri, R., Pezzotti, P., Trotta, M.P., Ravasio, L., De Longis, P., Lo Caputo, S., Narciso, P., Pauluzzi, S., Carosi, G., Nappa, S., Piano, P., Izzo, C.M., Lichtner, M., Rezza, G., Monforte, A., 2001. Self-reported symptoms and medication side effects influence adherence to highly active antiretroviral therapy in persons with HIV infection. J. Acquir. Immune Defic. Syndr. 5 (28), 445–449.

Arnsten, J.H., Demas, P.A., Farzadegan, H., Grant, R.W., Gourevitch, M.N., Chang, C.J., Buono, D., Eckholdt, H., Howard, A.A., Schoenbaum, E.E., 2001. Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. Clin. Infect. Dis. 338 (8), 1417–1423.

Bangsberg, D.R., Hecht, F.M., Charlebois, E.D., Zolopa, A.R., Holodniy, M., Sheiner, L., Bamberger, J.D., Chesney, M.A., Moss, A., 2000. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. AIDS 14 (4), 357–366.

Bangsberg, D.R., Charlebois, E.D., Grant, R.M., Holodniy, M., Deeks, S.G., Perry, S., Conroy, K.N., Clark, R., Guzman, D., Zolopa, A., Moss, A., 2003. High levels of adherence do not prevent accumulation of HIV drug resistance mutations. AIDS 17 (13), 1925–1932.

Bangsberg, D.R., Moss, A.R., Deeks, S.G., 2004. Paradoxes of adherence and drug resistance to HIV antiretroviral therapy. J. Antimicrob. Chemother. 53 (5), 696–699.

de Olalla, G.P., Knobel, H., Carmona, A., Guelar, A., Lopez-Colomes, J.L., Cayla, J.A., 2002. Impact of adherence and highly active antiretroviral therapy on survival in HIV-infected patients. J. Acquir. Immune Defic. Syndr. 30 (1), 105–110.

DeGruttola, V., Dix, L., D'Aquila, R., Holder, D., Phillips, A., Ait-Khaled, M., Baxter, J., Clevenbergh, P., Hammer, S., Harrigan, R., Katzenstein, D., Lanier, R., Miller, M., Para, M., Yerly, S., Zolopa, A., Murray, J., Patick, Aand Miller, V., Castillo, S., Pedneault, L., Mellors, J., 2000. The relation between baseline HIV drug resistance and response to antiretroviral therapy: re-analysis of retrospective and prospective studies using a standardized data analysis plan. Antivir. Ther. 5 (1), 41–48.

Descamps, D., Flandre, P., Calvez, V., Peytavin, G., Meiffredy, V., Collin, G., Delaugerre, C., Robert-Delmas, S., Bazin, B., Aboulker, J.P., Pialoux, G., Raffi, F., Brun-Ve'zinet, Franc, O., and for the Trile'ge (Agence Nationale de Recherches sur le SIDA 072) Study Team, 2000. Mechanisms of virologic failure in previously untreated HIV-infected

patients from a trial of induction-maintenance therapy. J. Am. Med. Assoc. 283, 205–211.
Eldred, L., Cheever, L., 1998. Update on adherence to HIV therapy. Hopkins HIV Rep. 10 (1), 10–11.
Ferguson, N.M., Donnelly, C.A., Hooper, J., Ghani, A.C., Fraser, C., Bartley, L.M., Rode, R.A., Vernazza, P., Lapins, D., Mayer, S.L., Anderson, R.M., 2005. Adherence to antiretroviral therapy and its impact on clinical outcome in HIV-infected patients. J. R. Soc. Interface 2 (4), 349–363.
Giordano, T.P., Guzman, D., Clark, R., Charlebois, E.D., Bangsberg, D.R., 2004. Measuring adherence to antiretroviral therapy in a diverse population using a visual analogue scale. HIV Clinical Trials 5 (2), 74–79.
Grabar, S., Pradier, C., Le Corfec, E., Lancar, R., Allavena, C., Bentata, M., Berlureau, P., Dupont, C., Fabbro-Peray, P., Poizot-Martin, I., Costagliola, D., 2000. Factors associated with clinical and virological failure in patients receiving a triple therapy including a protease inhibitor. AIDS 14 (2), 141–149.
Gross, R., Bilker, W.B., Friedman, H.M., Strom, B.L., 2001. Effect of adherence to newly initiated antiretroviral therapy on plasma viral load. AIDS 15 (16), 2109–2117.
Harrigan, P.R., Hogg, R.S., Dong, W.W., Yip, B., Wynhoven, B., Woodward, J., Brumme, C.J., Brumme, Z.L., Mo, T., Alexander, C.S., Montaner, J.S., 2005. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. J. Infect. Dis. 191 (3), 339–347.
Heffernan, J.M., Wahl, L.M., 2006. Natural variation in HIV infection: Monte Carlo estimates that include CD8 effector cells. J. Theor. Biol. 243, 191–204.
Hinkin, C.H., Hardy, D.J., Mason, K.I., Castellon, S.A., Durvasula, R.S., Lam, M.N., Stefaniak, M., 2004. Medication adherence in HIV-infected adults: effect of patient age, cognitive status, and substance abuse. AIDS 18 (Suppl. 1), S19–S25.
Howard, A.A., Arnsten, J.H., Lo, Y., Vlahov, D., Rich, J.D., Schuman, P., Stone, V.E., Smith, D.K.f., Schoenbaum, E.E.f.t.H.S.G., 2002. A prospective study of adherence and viral load in a large multi-center cohort of HIV-infected women. AIDS 16 (16), 2175–2182.
Huang, Y., Rosenkranz, S.L., Wu, H., 2003. Modeling HIV dynamics and antiviral response with consideration of time-varying drug exposures, adherence and phenotypic sensitivity. Math. Biosci. 184 (2), 165–186.
Karlin, S., Taylor, H.M., 1975. A First Course in Stochastic Processes. Academic Press, New York.
Kathleen, J.R., 2000. Barriers to and facilitators of HIV-positive patients' adherence to antiretroviral treatment regimens. AIDS Patient Care STDs 14 (3), 155–168.
Krakovska, O., Wahl, L.M., 2007. Costs versus benefits: best possible and best practical treatment regimens for HIV. J. Math. Biol., in press, doi:10.1007/s00285-006-0059-1.
Lohse, N., Obel, N., Kronborg, G., Laursen, A., Pedersen, C., Larsen, C.S., Kvinesdal, B., Sorensen, H.T., Gerstoft, J., 2005. Declining risk of triple-class antiretroviral drug failure in Danish HIV-infected individuals. AIDS 19 (8), 815–822.
Maitland, D., Moyle, G., Hand, J., Mandalia, S., Boffito, M., M, N., Gazzard, B., 2005. Early virologic failure in HIV-1 infected subjects on didanosine/tenofovir/efavirenz: 12-week results from a randomized trial. AIDS 19 (11), 1183–1188.
Martin-Carbonero, L., Soriano, V., Valencia, E., Garcia-Samaniego, J., Lopez, M., Gonzalez-Lahoz, J., 2001. Increasing impact of chronic viral hepatitis on hospital admissions and mortality among HIV-infected patients. AIDS Res. Hum. Retroviruses 17 (16), 1467–1471.
Martinez-Picado, J., Negredo, E., Ruiz, L., Shintani, A., Fumaz, C.R., Zala, C., Domingo, P., Vilaro, J., Llibre, J.M., Viciana, P., Hertogs, K., Boucher, C., D'Aquila, R.T., Clotet, B., SWATCH Study Team, 2003. Alternation of antiretroviral drug regimens for HIV infection. A randomized, controlled trial. Ann. Intern. Med. 139 (2), 81–89.
Mocroft, A., Ledergerber, B., Viard, J.P., Staszewski, S., Murphy, M., Chiesi, A., Horban, A., Hansen, A.B., Phillips, A.N., Lundgren, J.D., EuroSIDA Study Group, 2004. Time to virological failure of 3 classes of antiretrovirals after initiation of highly active antiretroviral therapy: results from the EuroSIDA Study Group. J. Infect. Dis. 190 (11), 1947–1956.
Montaner, J., Hill, A., Acosta, E., 2001. Practical implications for the interpretation of minimum plasma concentration/inhibitory concentration ratios. Lancet 357 (9266), 1438–1440.
Montaner, J.S., Reiss, P., Cooper, D., Vella, S., Harris, M., Conway, B., Wainberg, M.A., Smith, D., Robinson, P., Hall, D., Myers, M., Lange, J.M., 1998. A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients: the INCAS trial. J. Am. Med. Assoc. 279 (12), 930–937.
Oyugi, J.H., Byakika-Tusiime, J.B., Charlebois, E.D., Kityo, C., Mugerwa, R., Mugyenyi, P., Bangsberg, D.R., 2004. Multiple validated measures of adherence indicate high levels of adherence to generic HIV antiretroviral therapy in a resource-limited setting. JAIDS 36 (5), 1100–1102.
Paterson, D.L., Swindells, S., Mohr, J., Brester, M., Vergis, E.N., Squier, C., Wagener, M.M., Singh, N., 2000. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann. Intern. Med. 133 (1), 21–30.
Phillips, A.N., Staszewski, S., Weber, R., Kirk, O., Francioli, P., Miller, V., Vernazza, P., Lundgren, J.D., Ledergerber, B., Swiss HIV Cohort Study, Frankfurt HIV Clinic Cohort, and EuroSIDA Study Group, 2001a. HIV viral load response to antiretroviral therapy according to the baseline CD4 cell count and viral load. J. Am. Med. Assoc. 286(20), 2560–2567.
Phillips, A.N., Youle, M., Johnson, M., Loveday, C., 2001b. Use of a stochastic model to develop understanding of the impact of different patterns of antiretroviral drug use on resistance development. AIDS 15 (17), 2211–2220.
Phillips, A.N., Dunn, D., Sabin, C., Pozniak, A., Matthias, R., Geretti, A.M., Clarke, J., Churchill, D., Williams, I., Hill, T., Green, H., Porter, K., Scullard, G., Johnson, M., Easterbrook, P., Gilson, R., Fisher, M., Loveday, C., Gazzard, B., Pillay, D., UK Collaborative Group on HIV Drug Resistance, and UK CHIC Study Group, 2005. Long term probability of detection of HIV-1 drug resistance after starting antiretroviral therapy in routine clinical practice. AIDS 19(5), 487–494.
Roca, B., Gomez, C.J., Arnedo, A., 2000. Adherence, side effects and efficacy of stavudine plus lamivudine plus nelfinavir in treatment-experienced HIV-infected patients. J. Infect. 41 (1), 50–54.
Sethi, A.K., Celentano, D.D., Gange, S.J., Moore, R.D., Gallant, J.E., 2003. Association between adherence to antiretroviral therapy and human immunodeficiency virus drug resistance. Clin. Infect. Dis. 37 (8), 1112–1118.
Siegel, K., Schrimshaw, E.W., Raveis, V.H., 2000. Accounts for non-adherence to antiviral combination therapies among older HIV-infected adults. Psychol. Health Med. 5 (1), 29–42.
Smith, R.J., 2006. Adherence to antiretroviral HIV drugs: how many doses can you miss before resistance emerges? Proc. R. Soc. London B Biol. Sci. 273 (1586), 617–624.
Smith, R.J., Wahl, L.M., 2004. Distinct effects of protease and reverse transcriptase inhibition in an immunological model of HIV-1 infection with impulsive drug effects. Bull. Math. Biol. 66 (5), 1259–1283.
Smith, R.J., Wahl, L.M., 2005. Drug resistance in an immunological model of HIV-1 infection with impulsive drug effects. Bull. Math. Biol. 67 (4), 783–813.
Verotta, D., Schaedeli, F., 2002. Non-linear dynamics models characterizing long-term virological data from AIDS clinical trials. Math. Biosci. 176 (2), 163–183.
Wahl, L.M., Nowak, M.A., 2000. Adherence and resistance: predictions for therapy outcome. Proc. Biol. Sci. 267 (1445), 835–843.
Walsh, J.C., Pozniak, A.L., Nelson, M.R., Mandalia, S., Gazzard, B.G., 2002. Virologic rebound on HAART in the context of low treatment adherence is associated with a low prevalence of antiretroviral drug resistance. J. Acquir. Immune Defic. Syndr. 30 (3), 278–287.

Weiser, S., Wolfe, W., Bangsberg, D., Thior, I., Gilbert, P., Makhema, J., Kebaabetswe, P., Dickenson, D., Mompati, K., Essex, M., Marlink, R., 2003. Barriers to antiretroviral adherence for patients living with HIV infection and AIDS in Botswana. JAIDS 34 (3), 281–288.

Wood, E., Hogg, R.S., Yip, B., Harrigan, P.R., O'Shaughnessy, M.V., Montaner, J.S., 2003. Effect of medication adherence on survival of HIV-infected adults who start highly active antiretroviral therapy when the $CD4^+$ cell count is 0.200 to $0.350 \times 10(9)$ cells/L. Ann. Intern. Med. 139 (10), 810–816.

Wu, H., Huang, Y., Acosta, E.P., Rosenkranz, S.L., Kuritzkes, D.R., Eron, J.J., Perelson, A.S., Gerber, J.G., 2005. Modeling long-term HIV dynamics and antiretroviral response: effects of drug potency, pharmacokinetics, adherence, and drug resistance. J. Acquir. Immune Defic. Syndr. 39 (3), 272–283.