RESEARCH ARTICLE

Exploring an alternative explanation for the second phase of viral decay: Infection of short-lived cells in a drug-limited compartment during HAART

Steven Sanche1*, Thibault Mesplède2,3, Nancy L. Sheehan1,4, Jun Li1,5, Fahima Nekka1,5

1 Faculté de Pharmacie, Université de Montréal, Montreal, Quebec, Canada, 2 McGill University AIDS Centre, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada, 3 Department of Microbiology and Immunology, Faculty of Medicine, McGill University, Montreal, Quebec, Canada, 4 Chronic Viral Illness Service, McGill University Health Centre, Montreal, Quebec, Canada, 5 Centre de recherches mathématiques, Université de Montréal, Montreal, Quebec, Canada

* steven.sanche@umontreal.ca

Abstract

Most HIV-infected patients who initiate combination antiretroviral therapy experience a viral load decline in several phases. These phases are characterized by different rates of viral load decay that decrease when transitioning from one phase to the next. There is no consensus as to the origin of these phases. One hypothesis put forward is that short- and long-lived infected cells are responsible for the first and second phases of decay, respectively. However, significant differences in drug concentrations are observed in monocytes from various tissues, suggesting the first two phases of decay in viral loads could instead be attributed to short-lived cells being differently exposed to drugs. Compared to a well-exposed compartment, new cell infection can be expected in a compartment with limited drug exposure, thus leading to a slower viral load decay with potential virologic failure and drug resistance. In the current study, the latter hypothesis was investigated using a model of viral kinetics. Empirical datasets were involved in model elaboration and parameter estimation. In particular, susceptibility assay data was used for an in vitro to in vivo extrapolation based on the expected drug concentrations inside physiological compartments. Results from numerical experiments of the short-term evolution of viral loads can reproduce the first two phases of viral decay when allowing new short-lived cell infections in an unidentified drug-limited compartment. Model long-term predictions are however less consistent with clinical observations. For the hypothesis to hold, efavirenz, tenofovir and emtricitabine drug exposure in the drug-limited compartment would have to be very low compared to exposure in peripheral blood. This would lead to significant long-term viral growth and the frequent development of resistant strains, a prediction not supported by clinical observations. This suggests that the existence of a drug-limited anatomical compartment is unlikely, by itself, to explain the second phase of viral load decay.
Introduction

Viral loads in the plasma of patients initiating highly active antiretroviral therapy (HAART) generally decrease very rapidly during the first days of treatment before reaching a slower second phase of decay.[1, 2] In fact, up to four phases of decreasing viral load can be observed, each new phase being slower than the previous one.[3] These phases are the result of the complex interaction between host, drugs and virus. The existence of multiple phases of viral decay challenges our understanding of this interaction.[4]

In the following, we will demonstrate that there are multiple rational explanations for the first two phases of viral load decay. First, we will infer that a set of three assumptions is inconsistent with multiple phases of viral decay. Under the first assumption, viral loads during the first and second phases of viral decay mainly come from one infected cell population: CD4 cells having a half-life of virion production of about one day (short-lived). Under the second assumption, viral loads are proportional to the number of infected cells. This assumption is partially supported by results suggesting rapid virion clearance in lymphoid tissue and plasma (no accumulation of virions).[5, 6] Under the third assumption, HAART has the capacity to fully inhibit all new cell infections. If all of these assumptions were true, there would be only one phase of viral decay, as depicted by Fig 1A. Indeed, under assumption one and two, the viral load ($V$) would equate to a proportionality constant ($K$) times the number of short-lived infected CD4 cells ($C$) which would decay at a constant rate ($d$). Because of assumption three, $C$ cannot increase after treatment initiation. In other words, the viral load would be described by the following equation: $V(t) = K C(t_0) e^{-dt}$, where $t_0$ is the time of treatment initiation, and $t$ represents the time since $t_0$. This equation can only describe one phase of viral decay. This incoherence inevitably calls into question the above assumptions. We will now show how two of these assumptions have been revisited in order to give rise to two competing hypotheses, one of which is further studied herein.

In particular, the assumption of a single source of plasma virion (short-lived CD4 cells) has been largely disputed regarding its involvement in the slowdown of the viral load decay after the rapid first phase.[1, 7] Other sources of virions have been suggested to explain the occurrence of this second phase decay. One possible explanation put forward is the existence of infected cells that can survive for a longer period (long-lived) (Fig 1B).[1] Each of the four decay phases would correspond to a specific virion-producing cell population.[8] Recent work suggests the long-lived cells responsible for the second phase of decay could integrate viral DNA more slowly, explaining the delay before virion production.[9] These cells could potentially be resting CD4 cells or macrophages, which do exhibit slower integration when investigated in vitro.[10, 11] Other authors also mention dendritic cells or monocytes as a potential source of virions explaining the second phase of decay.[12–15] However, uncertainties prevail as to the kinetics of these infected cells and the total viral contribution of these cells in vivo. So far, kinetics estimates are solely based on data fitting using models that assume the involvement in viral decay of secondary sources of virions.[9] However, model assumptions can drastically change the values of these estimates since models solely involving cells with a uniform half-life can still fit the viral decay data quite well, as will be later seen.

Recent work challenges the notion that HAART has the capacity to fully inhibit all new cell infections in all organs. Indeed, there can be a large difference in drug concentrations across tissues within the host.[16] For example, one study reports very poor penetration in lymph nodes for all tested antiretrovirals, and several poorly penetrated ileum and rectum tissues.[16] Testicular tissue is another example of a drug sanctuary for many antiretrovirals,[17] along with the central nervous system.[18] The extent of drug penetration in specific secondary lymphoid tissue is largely unknown for most antiretrovirals.[16] Motivated by this and by reported
evidence of ongoing viral replication for patients taking HAART,[19–21] we were led to question the role drug penetration may play in explaining the first phases of viral decay, as proposed by Murray et al. [2] The potential involvement of new cell infections in a physiological compartment with limited drug penetration provides an alternative perspective. Under this
hypothesis, cells infected before treatment initiation would produce the majority of virions lost
during the first phase of decay, while cells infected under treatment within a drug-limited
compartment would produce the virions lost during the second phase of decay. Long-lived
infected cells would have a negligible impact on viral load, infected cells being short-lived CD4
cells. A depiction of this hypothesis is provided (see Fig 1C).

Better characterizing the involvement of a drug-limited compartment is important, since a
compartment harbouring a large quantity of infected cells and with little drug exposure could
affect long-term treatment efficacy. Indeed, a large viral population implies a high chance for
resistant mutants to be already present before treatment initiation, as the virus naturally
mutates at a high frequency.[22] Low drug levels could then lead to suboptimal suppression of
mutants that resist to antiretroviral drugs. However, factors such as the decreased viral fitness
of resistant strains, compared to wild-type virus, could offset resistant viral growth.[23] Fur-
ther, concomitant use of multiple drugs could prevent the exponential growth of resistant
strains, as such strains may be sufficiently susceptible to one or more of the drugs.

We have developed a mathematical model accounting for these various factors. Our previ-
ous mechanistic model describes viral and short-lived infected CD4 cells dynamics in one
compartment, namely lymph nodes.[24] The model can explain the long-term risk of virologic
failure.[24] Herein, this model was modified to include two compartments, one with low and
another with high drug penetration. The compartment with low drug penetration, hereon
referred as the drug-limited compartment, partially allows the infection of CD4 cells by the
wild-type virus. The compartment is not associated to a specific tissue at this point, as its exis-
tence is hypothesized. The compartment with high drug penetration comprises tissues where
drug concentrations are sufficient to prevent new cell infection events from wild-type virus
(e.g. hypothetically the plasma and well-irrigated tissues).

Our approach was integrative and characteristic of the quantitative pharmacology of sys-
tems (QSP).[25] Compared to classical ‘top-down’ approaches that use empirical observations
to deduce a mathematical relationship between measured input and output, bottom-up ap-
proaches, including QSP, integrate all the information available on the involved processes to
build a mechanistic model that describes the hypothesized causal links between input and
output. This mechanistic model is then used to predict the output from input values. The pro-
duced predictions are compared to empirical observations in order to evaluate if the hypothe-
sized system can explain the phenomenon under investigation.

Our first objective is to evaluate the capacity of the new mechanistic model to predict viral
loads observed at the initiation of treatment, particularly the occurrence of two phases, the time
delay between phases and the rates of decay. For this purpose, the two-month viral load data of 6
patients taking nelfinavir, zidovudine and lamivudine were fitted.[1] Data was fitted using a least
square approach. As a second objective, we examined whether results from this model are consis-
tent with the observed risk of long-term virologic failure. Short- and long-term viral load data of
patients taking efavirenz, tenofovir disoproxil fumarate (DF) and emtricitabine combination ther-
apy were used.[26, 27] The former dataset was used to estimate model parameters using a least-
square approach, while the latter was used to compare with model predictions. Depending on
whether we assess the short- or long-term viral load evolution for treatment naïve patients, resis-
tance mutations are either excluded or included as part of our model, respectively.

Results

The QSP approach adopted here has previously led to a model that serves as a basis for this
study.[24] This previous model, which only looked at infected CD4+ lymphocytes in active
state as a source of virions, has been modified in order to simultaneously consider two
physiological compartments differently exposed to antiretroviral drugs. One compartment represents the blood compartment and well-irrigated tissues where it is assumed drug penetration is sufficient to prevent new cell infection by wild-type virus. The other compartment is drug-limited. In the following, the former compartment is identified with index 1, while the drug-limited compartment is identified with index 2. The model parameters requiring a posteriori estimation are: 1) the average fraction of total infection events not affected by the drugs in each compartment for the wild-type virus \( (f_u,1, f_u,2) \), and 2) the contribution of each compartment to the maximum plasma viral load or setpoint, in percent \((\varphi_1, \varphi_2)\). The values for parameters \( f_u,1 \) and \( f_u,2 \) are mathematically associated with drug concentrations in the respective compartments, with higher drug concentrations leading to smaller \( f_u \) values (see Methods and Eq 1 for detail).

As for parameters \( \varphi_1 \) and \( \varphi_2 \), an illustrative example would be if one compartment contributes 20% of the viral load when the patient reaches the peak viral load, then the value of \( \varphi \) associated with this compartment would be 20%. Since there are only two compartments, the value of \( \varphi \) for the other compartment would be 80%.

_Does a model with a drug-limited compartment hosting new short-lived CD4 cell infections have the capacity to predict viral load decay following treatment initiation?_

Viral load data were retrieved using WebPlotDigitizer software[29] for the 6 patients reported in Perelson et al.[1], displayed in Fig 2, blue dots. All patients were treatment naïve and initiated a therapy combining three antiretroviral drugs (nelfinavir, zidovudine and lamivudine). Patient demographics for this study can be found in Table 1 of the referenced article.[1]

The short-term viral load decay of the 6 patients can be recovered using our mechanistic model, while keeping the model structure as well as the _a priori_ determined parameter values. One of the two compartments is assumed to be sufficiently exposed to medication to prevent further CD4+ infection, i.e. \( f_u,1 = 0 \). This compartment represents physiological compartments where antiretroviral drug concentrations are sufficiently high to prevent any new cell infection (e.g. hypothetically the plasma and well-irrigated tissues). Results from data fitting procedures for the remaining parameters, using a least-square approach, are shown in Fig 2. With these parameter values, our model predicted the two phases of viral load decay for the 6 patients. Fractional values \( \varphi_2 \) ranging from 4 to 18% allowed the best adjustments. Also, when the drugs inhibit 41% to 80% of the infection in the drug-limited compartment (corresponding to \( 1-f_u,2 \)), the obtained predictions are the closest to the viral load observations.

To assess whether the above short-term findings still hold for longer-term virologic response, the model was used to simulate viral load dynamics on a period longer than two months. For this, we first determined a \( f_u,2 \) value using short-term viral load data for patients under efavirenz, tenofovir DF and emtricitabine (600, 300 and 200 mg daily) combination therapy.[26] Secondly, values around the \( f_u,2 \) estimate were translated in terms of drug concentrations in the drug-limited compartment. Finally, these concentrations served to simulate the viral load evolution over a period of approximately one year. This time, resistant strains were allowed to emerge. Virologic failure at 48 weeks was the simulated outcome. To compare the simulation results with clinical data, the virologic failure threshold was set to 400 copies per mL of plasma.[27]

_What values of \( f_u,2 \) and \( \varphi_2 \) allow reproducing the first two phases of decay in a typical patient under efavirenz, tenofovir DF and emtricitabine?_

Karris et al. reports a mixed-effect biexponential regression model for the viral loads of 25 patients initiating the combination treatment containing efavirenz, tenofovir DF and emtricitabine.[26] Patient demographics for this dataset can be found in Table 2 of the referenced
To determine $f_{u,2}$ values, we focused on the typical patient's viral loads, in the sense that the patient is characterized by median values of the regression model parameters. In Fig 3, we superimposed the best fitting decay curve predicted by our model on the regression curve representing the viral load decay over a two-month period for the typical patient. The values of $f_{u,2}$ and $\phi_2$ used in our model were obtained in the same way as done for Fig 2. The values for these parameters allowing the adjustment in Fig 3 are 40% and 8%, respectively. A value of 40% for $f_{u,2}$ corresponds to an inhibition of 60% of the infection activity in the drug-limited compartment.

What drug concentrations lead to an inhibition of 60% of the infection in the drug-limited compartment?

The parameter value $f_{u,2} = 40\%$ was translated into concentration values of efavirenz, tenofovir and emtricitabine in the drug-limited compartment. In fact, the association between $f_{u,2}$ and
concentrations were mathematically described for many drugs using *in vitro* experiments.[30] The *in vivo* relationships can be deduced from these associations by assuming that intracellular concentrations dictate drug efficacy, independently of the medium in which the cells reside.[24] The concentration inside the mononuclear cells of a compartment can be derived from extracellular plasma concentrations using two parameters, \( k_p \) and \( k_l \) (see Eq 1 from Material and Methods). At steady-state, the median extracellular plasma concentrations \( C_p \) were extracted from population pharmacokinetics studies.[31–33] Parameter \( k_p \), reflecting the effect of plasma protein binding on drug efficacy, is independent of compartmentalization and has been determined empirically elsewhere.[34] Parameter \( k_l \), representing the ratio of drug concentrations inside peripheral blood monocyte to the equivalent in the compartment, is the sole unknown value to determine \( f_{u,2} \). When a single drug is used, the relationship between \( f_{u,2} \) and \( k_l \) is one-to-one, and the latter can be uniquely determined. When multiple drugs are used concomitantly, the procedure is more complex as the values \( k_l \) for each drug need to be determined (\( k_{EFV} \); \( k_{TFV} \) and \( k_{FTC} \) for efavirenz, tenofovir and emtricitabine, respectively). Further, different combinations of drug concentrations can lead to the same value of \( f_{u,2} \), after accounting for interaction in drug effects.[35] This translates into multiple combinations of \( k_{EFV} \), \( k_{TFV} \) and \( k_{FTC} \) leading to an inhibition of 60 ± 5% of the infection activity.[24]

The space of parameter values is displayed in Fig 4. In general, a higher \( k_l \) value indicates less drug penetration in the hypothesized compartment compared to peripheral blood, while a value of \( k_l < 1 \) indicates a higher drug concentration in the compartment. The \( k_{EFV} \) value ranges between 88 and 2593, \( k_{TFV} \) between 0.6 and 130 and \( k_{FTC} \) between 32 and 4000. Inguinal lymph nodes are known to be less exposed to efavirenz, tenofovir and emtricitabine.[16] For comparison purposes, we here give the reported values of \( k_{EFV} \), \( k_{TFV} \) and \( k_{FTC} \) in this tissue, which are

![Fig 3. Regression line (black) for the typical patient undertaking a treatment combining efavirenz, tenofovir DF and emtricitabine, model fit (red), and associated \( f_{u,2} \) and \( \phi_2 \) values. \( f_{u,2} \) is the average fraction of total infection events not affected by the drugs and \( \phi_2 \) is the fraction of maximum plasma viral load. Both parameters refer to the drug-limited compartment. Regression curve based on data from Karris et al.[26].](https://doi.org/10.1371/journal.pone.0198090.g003)
16.7, 2.9 and 5, respectively, implying a greater drug exposure in lymph nodes as compared to the hypothesized drug-limited compartment.[16]

What is the expected risk of virologic failure over a long-term period if the less exposed compartment allowing the infection of new short-lived CD4 cells is responsible for the second phase of viral load decay?

The values $k_i^{EFV}$, $k_i^{TFV}$ and $k_i^{FTC}$ associated with the drug-limited compartment were used to compute drug efficacy through viral load curves over 48 weeks of therapy with efavirenz, tenofovir DF and emtricitabine (600, 300 and 200mg daily). For this in silico experiment, we considered the possibility of the emergence of resistant strains through single nucleotide mutations responsible for the K103N, Y181C, G190S, M184V or K65R substitutions, as used in our previous study.[24] In the simulations, these strains could be generated de novo or be selected. We used reported empirical curves of strain-specific drug efficacy and viral fitness.[28, 30] Virtual patients expressed inter-individual variability in pharmacokinetics and immune response, modeled using patient population data.[31–33, 36, 37] This inter-individual variability was added so groups of virtual patients are representative of real patient populations, as in Sanche et al.[24] Due to the uncertainty in drug penetration levels in the second compartment (the $k_i$ values), one simulation was carried out for each of 200 sampled trios of parameter values $k_i^{EFV}$, $k_i^{TDF}$ and $k_i^{FTC}$, selected using simple random sampling (points shown in red in Fig 4). For each trio of parameter values, we calculated the proportion of individuals expected to experience virologic failure over 100 simulated virtual patients.

To give the reader a better sense of the performed simulations, the predicted viral load of one of the simulated patients, as a function of time since treatment initiation, is shown in Fig 5. The transition from phase one to phase two can be observed approximately a week after treatment initiation (Fig 5, upper left). The long-term viral load originating from each compartment is also presented (Fig 5, upper right). In this example, the K103N mutation was
selected in the drug-limited compartment (Fig 5, bottom left). There was no selection for other strains (Fig 5, bottom right).

We used a virologic failure threshold of 400 copies/mL to match the definition used in a previously published clinical study.[27] Patient demographics for this study can be found in Fig 1 of the referenced article.[27] A box plot for the predicted risk of virologic failure was compared to clinical data.[27] As illustrated in Fig 6, we obtained a range of values for the risk of virologic failure going from 22% to 67% at the end of the trials. Most virtual cases of virologic failure were with resistance (>20% of the viral load attributed to resistant viral strains), with K103N and K65R being the most common mutations conferring resistance (data not shown). The smallest value over all 95% confidence intervals for the predicted risk of virologic failure was 17.4%. In comparison, the risk of virologic failure (>400 copies/mL) based on clinical observation is 2% of patients (4/210 patients), with a 95% confidence interval of [0.5;3.8]%.[27] The p-value for the absence of difference of the proportion of failure between the empirical sample and the model-based sample associated with the smallest predicted risk is smaller.
than 1E-08. Those who left the clinical study for reasons unrelated to their viral load observations were discarded before computing the empirical risk of failure (pregnancy (n = 4), adverse events (n = 9), lost to follow-up (n = 12), withdrew consent (n = 5), other (n = 4)),[27] explaining the difference with the value reported in the original paper (16% without a response).[27]

We then questioned whether it was possible to reproduce the empirical risk of failure using $\phi_2$ and $f_{u,2}$ values around their estimates from the main analysis (8% and 40%, respectively). For the purpose of this sensitivity analysis, additional simulations were performed by first allowing $\phi_2$ to take values of 6%, then 10%. For each new $\phi_2$ value, we sampled 50 points in the space of parameter values for $k_{\text{EFV}}$, $k_{\text{TFV}}$, and $k_{\text{FTC}}$ associated with an unchanged $f_{u,2}$ and, for each point, computed the risk of failure over simulations of N = 100 virtual patients. Predicted probabilities of failure fell within the range obtained in the main analysis (22% to 67%). Using a similar approach, we computed the risk of failure assuming $f_{u,2}$ values of 30%, then 50%, with an unchanged $\phi_2$. The smallest risk of failure over these simulations was 17% (95%CI [10.2;25.8]%). We further extended our analysis to $f_{u,2}$ values of 20% and 60%. For these parameter values, the smallest predicted risk of failure was 9% (95%CI [4.2;16.4]%) and was achieved for a $f_{u,2}$ value of 20%. Under this $f_{u,2}$ value, the slope of second phase decay for the typical individual would be more than 2.5 times greater than its empirical estimate.

**Discussion**

The viral load of most HIV-infected patients initiating HAART decays in distinct phases.[4] During each phase, the viral load decays at an approximately constant rate. However, this rate decreases substantially from one phase to the next. During the final phase, the decay is so slow that low-level viremia (<50 copies/mL) persists over many years of therapy.[38] To date, the causes behind these phases, which characterize the viral decay profiles, still draw attention and fuel a debate. One hypothesis advances that the second phase of viral decay would be caused by infected cells that could survive longer than short-lived CD4 cells.[1] Evidence suggests these cells could be CD4 cells exhibiting a slower integration of viral DNA.[9] Another hypothesis stipulates that new short-lived infected cells from a drug-limited compartment could cause the phenomenon.[2] We investigated the latter hypothesis.

Contrary to the role long-lived cells could play in this phenomenon, much less attention has been paid to the involvement of drug penetration in tissues. Evidence for the existence of tissues where new cell infections occur during therapy,[19–21] and suggesting high variability in drug penetration in terms of specific lymphoid tissue,[16] adds elements of controversy to the ongoing debate. In the current work, we focus on the first two phases of viral decay without the specific differentiation between the sub-phases 1a and 1b that have been recently revealed.[9, 39] With the objective to shed additional light on the existence of multi-phasic viral load
decay, we here compared predictions from a mathematical model to short-term viral load data as well as to long-term clinical outcomes.

We overlaid our model predictions on short-term viral load data of 6 patients. *De visu*, the agreement between model predictions and patient observations is substantial (Fig 2), and compares with the results obtained through a model assuming short- and long-lived virion-producing cells.[1] This suggests both models are plausible mathematical explanations for the short-term viral load decay after treatment initiation. In fact, the results indicate that in order to reproduce the short-term viral loads, the infection activity in the less drug-exposed compartment has to contribute from 4% to 18% of the total plasma viral load before treatment initiation. Further, the drugs in this compartment would need to prevent 41% to 80% of all new cell infections for the model to accurately fit the data. This suggests the associated tissues need to harbour a substantial amount of the total infection activity before treatment initiation (infection hot spot). Moreover, the level of drug exposure would likely be much smaller in this compartment than what is prohibitive in infection assays.[30]

The existence of a drug-limited compartment allowing new short-lived CD4 infections could affect long-term drug resistance. To study this, we based ourselves on a model able to reproduce the observed risk of virologic as a function of the adherence level of a population of patients initiating treatment with efavirenz, tenofovir DF and emtricitabine.[24] This model accurately predicted the risk of virologic failure observed in patients after long-term therapy (1 to 1.5 years) under various regimen and drug adherence.[24] Herein, this model was modified to consider two compartments differentially exposed to drugs. The principles governing viral kinetics in the compartments were kept: both models used an *in vivo* extrapolation of drug susceptibility assays along with empirical data on *in vivo* antiviral activity.[24] Further, the extent of the immune response or whether resistant virus is present when treatment starts, which may also impact viral growth, are equivalently considered in both models.[24]

We used the modified model to simulate the viral activity within a population of patients initiating treatment with efavirenz, tenofovir DF and emtricitabine. Here, we first deduced that even if 100% infections are prevented in well-irrigated tissues, 60% of wild-type virus infections have to be inhibited in a second compartment to explain the first two phases of decays observed in patients (Fig 3, $f_{u,2} = 40\%$). This value is consistent with what was obtained in Fig 2. Second, we determined the drug concentrations within this compartment corresponding to such level of inhibition. Finally, we used these concentrations to simulate long-term viral suppression in small groups of patients. To palliate the uncertainty in the combination of drug concentrations leading to 60% inhibition of cell infection events, we undertook 200 simulations, one per selected trio of drug penetration parameters $k_{EFV}^d$, $k_{TFV}^d$ and $k_{FTC}^d$. We used a simple random sample, using the log scale of $k_l$ values, to obtain the sampled values in red in Fig 4. The chosen points are well spread across the domain of allowable values, indicating a good representation of the entire space. Results in Fig 6 suggest a minimum of 22% of patients would experience virologic failure at the end of 48 weeks of treatment under the evaluated hypothesis (lower bound of 95% CI of 17.4%). The predicted risk is not consistent with the 2% risk of virologic failure reported in a clinical study we used to validate our model (95% confidence interval of empirical virologic is [0.5%;3.8%]).[21] Sensitivity analyses also showed this inconsistency remains once uncertainty in parameter values has been accounted for. Even if the median slope of second phase viral decay was 2.5 times its empirical estimate ($f_{u,2}>20\%$), it was still impossible to achieve overlap between predicted and empirical confidence intervals of the risk of failure.

The mechanistic nature of our model enables evaluating the likelihood of the hypothesis of new short-lived CD4 cell infections in a drug-limited compartment explaining the second
phase of viral decay.[25] The model consists in the physiological, pharmacological and viral elements interacting to dictate the evolution of the viral population within its host. While there is an increasing body of evidence suggesting the existence of drug-limited tissues impacting the viral load dynamics, such as lymph nodes,[16] the results of the current study suggest persistent short-lived CD4 infections in a drug-limited compartment cannot–by itself–explain the second phase of viral decay. Here, all parameter values, with the exception of the ones that are linked to the hypothesis in question (φ₂ and f_u₂), were determined a priori from reported experiments. Plasma drug concentrations were simulated using models derived from population pharmacokinetic studies.[31–33] The relationship between drug concentrations and drug effect on each separated strain was based on an in vitro to in vivo extrapolation of drug efficacy.[16, 28, 30, 34, 35] The within-host growth rate of viral loads was also based on reported distributions.[36, 37] Finally, the probability of emergence of specific mutations as well as the fitness of each strain were well described in the literature.[28] However, it was not possible to find parameter values consistent with the hypothesis and allowing coherent predictions of long-term virologic responses.

The model we formulated can suggest other explanations for the second phase of viral load decay. Because mass action law dictates the dynamics of the infection in the model,[40] how quickly the number of infected cells grows or declines over time is dependent on the availability of target cells. A lower density of target cells in a compartment means that virions have a higher chance to be cleared before they can infect a cell, which leads to a decline in the infection rate while the decay rate of infected cells remains the same. When the density of target cells is too low, the model predicts a constant decline of virion production over time with a rate dependent on density. In fact, we reproduced two phases of viral decay with our model when new cell infections are completely prevented by the drug in one compartment, while in another compartment unexposed to drugs, virion production slowly declines due to the loss of target cells (data not shown). Because of the absence of drug exposure in the second compartment, no long-term selection of resistant mutants occurs.

In reality, a reduction in the number of target cells could be initiated by immune contraction.[41] It is hypothesized that this phenomenon, which implies a rapid decay of effector T cells, can be triggered by a reduction in antigen availability.[41] We hypothesize that one explanation for the first 2 phases of decay could be that the sharp decrease in the density of virions in one compartment that is highly exposed to drugs triggers a contraction phase. This contraction phase, associated with systemic cytokine signaling,[41] could promote the death or inactivation of CD4 cells in another compartment unexposed to drugs.

In the current model, the effect of the immune response is represented via a single parameter that summarizes the growth rate of viral infection in absence of drug action (reproduction number).[24] This is a consequence of a lack of quantitative knowledge about changes in patients’ immunity from disease onset. In particular, little is known about the evolution of the density of activated CD4 cells in infected tissues. Although the fraction of cells expressing CD38+HLA-DR+ among CD4 cells is significantly reduced in the blood of patients after treatment initiation,[42] more precise estimates of local changes in various tissues could lead to a better understanding of the interaction between immunity and the virus.

In summary, the specific cause for the phases of viral decay is still uncertain. On the one hand, the hypothesis advancing the existence of infected cells exhibiting a decay dynamic that is different from short-lived CD4 cells has literature support.[1, 9] On the other hand, the hypothesis stipulating that such cells have a negligible impact on virion production and completely attributing the slower decay of virion production to new short-lived cell infections in a drug limited compartment finds long term inconsistencies. Alternative models that would measure the contribution of both poor drug exposure in tissues and the immune response
remain to be studied. A better understanding of the dynamics of activated immune cells before and after treatment is initiated could shed light on the matter.

**Materials and methods**

**The model and simulations**

An overview of the model is given below. It includes the description of the algorithm and main parameters, as well as the discussion of the source of parameter values. The model is adapted from Sanche et al. [24], the main difference being the consideration of two compartments with different exposures to the drugs.

The model is implemented using an algorithm which iteratively computes the number of events involving active CD4 cells during small time intervals: i) the number of virion-producing cells that die out, ii) the number of new cell infections, and iii) the fraction of these infections that involve a newly mutated virus. Step i) is computed using a constant cell death rate \((d_v)\). Step ii) is computed from reproduction numbers \(R_{ij}\), where \(i\) and \(j\) indicate the strain of the virus and the compartment, respectively. Details on reproduction numbers are given below. The fraction of de novo mutation is based on probabilities of SNP mutations estimated from empirical data and reported elsewhere. [43] Total infection activity is translated in plasma virions, by assuming one virion-producing cell supplies one plasma virion, a relationship that is consistent with independent data. [5, 44]

The model assumes compartments are isolated from each other in terms of viral infection, i.e. a virus produced by one cell can only infect cells within the same compartment. This has been confirmed with experimental data: the infection of new cells is considered an essentially local phenomenon and the genetic makeup of the viral populations suggests a high degree of compartmentalization. [21, 45] The modeled viral dynamics is very similar in each compartment, the main difference being the level of drug exposure influencing the reproduction numbers.

Reproduction numbers \(R_{ij}\) are expressed as \(R_0(1-s_i)f_{iu,j}\), where \(R_0\) is the mean number of CD4 cells becoming infected by viruses produced by a single infected cell when susceptible cells are abundant and when no drug is present, \(s_i\) is a fitness cost for strain \(i\), and \(f_{iu,j}\) is the fraction of CD4 infection events unaffected by the drugs for strain \(i\) in compartment \(j\). In this study, \(f_{iu,j}\) was either estimated for wild-type virus from viral load data (index \(i\) is omitted in this case, since only wild-type virus was considered), or computed from drug concentrations. In the latter case, plasma drug concentrations are first simulated from reported pharmacokinetic models. Since the concentrations vary over time \(t\), so do the \(f_{iu,j}\). In the particular case where only one drug is used, \(f_{iu,j}\) takes a relatively simple form (Eq 1):

\[
f_{iu,j}(t) = \frac{1}{1 + \left( \frac{C_p(t)}{\rho_i \cdot IC_{50}^{j} k_i} \right)^{m(1+\sigma_i)}}
\]

where \(C_p(t)\) is the plasma drug concentration at time \(t\), \(k_p\) is the coefficient adjusting for plasma protein binding, \(k_i\) is the coefficient adjusting for the degree of drug penetration in the compartment \(j\), respectively, \(IC_{50}^{j}\) is the concentration inhibiting 50% of infection events by the wild-type strain *in vitro* in a medium devoid of plasma proteins, \(m\) is the Hill coefficient, and finally \(\rho_i\) and \(\sigma_i\) are two factors adjusting the values of \(IC_{50}^{j}\) and \(m\), respectively, for the resistant viral strain \(i\). For the wild-type strain, \(\rho_i = 1\) and \(\sigma_i = 0\). It should be noted that to estimate the impact of concomitant drug use, \(f_{iu,j}(t)\) no longer takes a simple analytical form (see Jilek et al. and Sanche et al.). [24, 35]
Parameters $\Phi_i$ are used in all simulations to limit the number of cells each compartment may contain. The overall number of cells is itself dictated by parameter $\lambda$, which is an entry rate of uninfected activated CD4 cells.

Many parameter values are patient-specific. The death rate $d_i$ is randomly selected for each patient from a distribution of values based on empirical measures of the first phase decay.[46] The distribution has a median value corresponding to a half-life of about 0.7 day. The same applies for $\lambda$ (directly linked to the distribution of viral set points)[28] and $R_0$ (based on the growth rate of viral loads during rebounds).[36, 37] Inter-individual variability for pharmacokinetic parameters is based on the reported values of population pharmacokinetics models.[31, 33, 34, 47] All other model parameters have a priori assigned values reported in Sanche et al. [24], with the exception of $\Phi_j$ and $f_{u,j}$.

Viral load data fitting

Using the developed mechanistic model, data fitting refers to finding model parameters $\Phi_2$ and $f_{u,2}$ for which viral load predictions are closest to empirical observations. The process was carried out in two steps to reduce numerical calculations. In the first step, we generated combinations of $\Phi_2$ and $f_{u,2}$ taking values between 5% and 95% in 5% increments (361 combinations).

For each combination, the viral loads of 20 virtual patients were simulated. Values of patient-specific parameters $d_u$, $\lambda$ and $R_0$ were randomly selected from their corresponding patient population distributions. Overall, the viral loads of a total of 7,220 virtual patients were simulated. For each set of real viral load data, we computed the sum of the squared difference between predictions and observations, using a logarithm scale for viral loads, in order to identify the virtual patient with the closest viral load predictions. This numerical experiment was repeated for values of $\Phi_2$ and $f_{u,2}$ around those retained for the identified virtual patient, noted $\Phi'_2$ and $f'_{u,2}$. More precisely, all discrete percentage values of $\Phi_2$ in $\phi'_2 \pm 4\%$ and $f_{u,2}$ in $f'_{u,2} \pm 4\%$ were investigated (81 combinations, 1620 virtual patients). We report the virtual patient’s viral load being the closest to the observations, using the same criteria as above.

Identification of $k^{EFV}_i$, $k^{TFV}_i$ and $k^{FTC}_i$ in the less drug-exposed compartment

Multiple trios of values $k^{EFV}_i$, $k^{TFV}_i$ and $k^{FTC}_i$ can lead to the same level of drug efficacy under combination treatment. We identified the space of parameter values for $k^{EFV}_i$, $k^{TFV}_i$ and $k^{FTC}_i$ for which $f_{u,2}$ was within 5% of a target value $\tilde{f}_{u,2}$. This was done in an iterative fashion, by exploring a wider space of parameter values $[k^{EFV}_{\min_i}; k^{EFV}_{\max_i}] \times [k^{TFV}_{\min_i}; k^{TFV}_{\max_i}] \times [k^{FTC}_{\min_i}; k^{FTC}_{\max_i}]$.

We used Eq 1 to identify the minimum and maximum values of $k_i$ for each drug. The rationale for this procedure was the following. First, we found $k_{i,\min}$ such that if $k_i < k_{i,\min}$ for each of the three considered drugs, then the value of $f_{u,2}$ for the combined drug effect will necessarily be beneath $\tilde{f}_{u,2} - 5\%$. In other words, the use of a single drug would be more than sufficient to inhibit $1-f_{u,2}$ infection events in the less exposed compartment. Further, we found $k_{i,\max}$ for which if $k_i > k_{i,\max}$ implies almost no effect from the drug (too small drug concentration in the tissue to have an effect). More precisely, we computed the ratios $C_p(t)/k_i$ that correspond to $f_{u,2} = f_{\tilde{u},2} - 10\%$ and $f_{u,2} = 99\%$, where the values for $m$, $k_p$ and $IC_{50}$ are reported elsewhere.[24, 30, 34] We then derived $k_{i,\min}$ and $k_{i,\max}$ using average steady-state plasma concentrations $C_p(t)$ (2.54mg/L for efavirenz, 0.183mg/L for tenofovir and 0.55mg/L for emtricitabine).[31–33] Finally, we discretized the range of values for $k_i$ for each of the three drugs using 200 points between $k_{i,\min}$ and $k_{i,\max}$ leading to 8,000,000, or 200$^3$ determined trio values. For each of these trios, we could compute the expected value of $f_{u,2}$ for the drug combination, by using a procedure further accounting for the degree of independence of drug
effects.\cite{24, 35} We retained the subset of $\left( k_l^{EFV}, k_l^{TFV}, k_l^{FTC} \right)$ for which the expected value $f_{u,2} \in [\hat{f}_{u,2} - 5\%; \hat{f}_{u,2} + 5\%]$.  

**Statistical analysis**

Parameter estimation was performed to fit our model to short term empirical data. As explained in the section entitled Data fitting for viral load data, this procedure was based on the least square difference between model predictions and empirical viral load observations in log$_{10}$ scale. The goodness of fit of the model was evaluated visually. We computed the 95\% confidence intervals for the risk of long-term virologic using the exact method, both for the empirical and the smallest predicted risks.\cite{48} We also computed the p-value associated to a null hypothesis of equal risk of failure in both the empirical and the model-based samples and assuming the difference of proportion is normally distributed (central limit theorem).\cite{48}

**Acknowledgments**

This research is supported by the FRQS Scholarship held by SS. Support was also provided by NSERC - Industrial Chair in Pharmacometrics—Novartis, Pfizer and Inventiv Health Clinical as well as FRQNT (FN). The authors acknowledge the help of Nesrine Fazez for her contribution to the figures.

**Author Contributions**

Conceptualization: Steven Sanche.

Data curation: Steven Sanche.

Formal analysis: Steven Sanche, Jun Li, Fahima Nekka.

Investigation: Steven Sanche, Thibault Mesplède, Nancy L. Sheehan.

Methodology: Steven Sanche, Thibault Mesplède, Nancy L. Sheehan, Jun Li, Fahima Nekka.

Project administration: Steven Sanche.

Validation: Steven Sanche, Thibault Mesplède, Nancy L. Sheehan.

Visualization: Steven Sanche.

Writing – original draft: Steven Sanche.

Writing – review & editing: Steven Sanche, Thibault Mesplède, Nancy L. Sheehan, Jun Li, Fahima Nekka.

**References**

1. Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, et al. Decay characteristics of HIV-1-infected compartments during combination therapy. Nature. 1997; 387(6629):188–91. Epub 1997/05/08. https://doi.org/10.1038/387188a0 PMID: 9144290.

2. Murray JM, Emery S, Kelleher AD, Law M, Chen J, Hazuda DJ, et al. Antiretroviral therapy with the integrase inhibitor raltegravir alters decay kinetics of HIV, significantly reducing the second phase. AIDS. 2007; 21(17):2315–21. Epub 2007/12/20. https://doi.org/10.1097/QAD.0b013e3282f12377 PMID: 18090280.

3. Maldarelli F. HIV-infected cells are frequently clonally expanded after prolonged antiretroviral therapy: implications for HIV persistence. J Virus Erad. 2015; 1(4):237–44. Epub 2015/01/01. PMID: 27482422; PubMed Central PMCID: PMCPMC484654.

4. Blankson JN, Persaud D, Siliciano RF. The challenge of viral reservoirs in HIV-1 infection. Annu Rev Med. 2002; 53:557–93. Epub 2002/01/31. https://doi.org/10.1146/annurev.med.53.082901.104024 PMID: 11818490.
22. Abram ME, Ferris AL, Das K, Quinones O, Shao W, Tuske S, et al. Mutations in HIV-1 reverse transcriptase affect the errors made in a single cycle of viral replication. J Virol. 2014; 88(13):7589–601. Epub 2014/04/25. https://doi.org/10.1128/JVI.00302-14 PMID: 24760888; PubMed Central PMCID: PMCPMC4054409.

23. Hu Z, Kuritzkes DR. Altered viral fitness and drug susceptibility in HIV-1 carrying mutations that confer resistance to nonnucleoside reverse transcriptase and integrase strand transfer inhibitors. J Virol. 2014; 88(16):9268–76. Epub 2014/06/06. https://doi.org/10.1128/JVI.00699-14 PMID: 24899199; PubMed Central PMCID: PMCPMC4136249.

24. Sanchez S, Sheehan N, Mesplede T, Wainberg MA, Li J, Nekka F. A Mathematical Model to Predict HIV Virological Failure and Elucidate the Role of Lymph Node Drug Penetration. CPT Pharmacometrics Syst Pharmacol. 2017; 6(7):469–76. Epub 2017/05/31. https://doi.org/10.1002/psp4.12200 PMID: 28556627.

25. Craig M, Gonzalez-Sales M, Li J, Nekka F. Approaching Pharmacometrics as a Paleontologist Would: Recovering the Links Between Drugs and the Body Through Reconstruction. CPT Pharmacometrics Syst Pharmacol. 2016; 5(3):158–60. Epub 2016/04/14. https://doi.org/10.1002/psp4.12069 PMID: 27069779; PubMed Central PMCID: PMCPMC4809624.

26. Karris MY, Jain S, Day TR, Perez-Santiago J, Goicochea M, Dubé MP, et al. HIV viral kinetics and T cell dynamics in antiretroviral naïve persons starting an integrase strand transfer inhibitor and protease inhibitor regimen. HIV Clin Trials. 2017; 18(2):67–74. Epub 2017/01/31. https://doi.org/10.1080/15284336.2017.1282578 PMID: 28134057.

27. Gallant JE, DeJesus E, Arribas JR, Pozniak A, Gazzard B, Campo RE, et al. Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirenz for HIV. N Engl J Med. 2006; 354(3):251–60. Epub 2006/01/20. https://doi.org/10.1056/NEJMoa051871 PMID: 16421366.

28. Rosenbloom DI, Hill AL, Rabi SA, Siliciano RF, Nowak MA. Antiretroviral dynamics determines HIV evolution and predicts therapy outcome. Nat Med. 2012; 18(9):1378–85. Epub 2012/09/04. https://doi.org/10.1038/nm.2892 PMID: 22941277; PubMed Central PMCID: PMCPMC3490032.

29. Rohatgi A. WebPlotDigitizer v4.0 Austin, Texas, USA2012 [cited 2017 December 12]. Available from: https://automeris.io/WebPlotDigitizer.

30. Shen L, Peterson S, Sedaghat AR, McMahon MA, Callender M, Zhang H, et al. Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. Nat Med. 2008; 14(7):762–6. Epub 2008/06/17. https://doi.org/10.1038/nm1777 PMID: 18552857; PubMed Central PMCID: PMCPMC2743464.

31. Sanchez A, Cabrera S, Santos D, Valverde MP, Fuertes A, Dominguez-Gil A, et al. Population pharmacokinetic/pharmacogenetic model for optimization of efavirenz therapy in Caucasian HIV-infected patients. Antimicrob Agents Chemother. 2011; 55(11):5314–24. Epub 2011/09/08. https://doi.org/10.1128/AAC.00194-11 PMID: 21896912; PubMed Central PMCID: PMCPMC3195031.

32. Valade E, Treluyer JM, Bouazza N, Ghosn J, Foissac F, Benaboud S, et al. Population pharmacokinetics of emtricitabine in HIV-1-infected adult patients. Antimicrob Agents Chemother. 2014; 58(4):2256–61. Epub 2014/02/05. https://doi.org/10.1128/AAC.00258-13 PMID: 24492366; PubMed Central PMCID: PMCPMC4023733.

33. Baheti G, Kiser JJ, Havens PL, Fletcher CV. Plasma and intracellular population pharmacokinetic analysis of tenofovir in HIV-1-infected patients. Antimicrob Agents Chemother. 2011; 55(11):5294–9. Epub 2011/09/08. https://doi.org/10.1128/AAC.00194-11 PMID: 21896912; PubMed Central PMCID: PMCPMC3194996.

34. Acosta EP, Limoli KL, Trinh L, Parkin NT, King JR, Weidler JM, et al. Novel method to assess antiretroviral target trough concentrations using in vitro susceptibility data. Antimicrob Agents Chemother. 2012; 56(11):5938–45. Epub 2012/09/12. https://doi.org/10.1128/AAC.00691-12 PMID: 22964257; PubMed Central PMCID: PMCPMC3296892.

35. Jilek BL, Zarr M, Sampah ME, Rabi SA, Bullen CK, Lai J, et al. A quantitative basis for antiretroviral therapy for HIV-1 infection. Nat Med. 2012; 18(3):446–51. Epub 2012/02/22. https://doi.org/10.1038/nm.2649 PMID: 22344296; PubMed Central PMCID: PMCPMC3296892.

36. Ribeiro RM, Qin L, Chavez LL, Li D, Self SG, Perelson AS. Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. J Virol. 2010; 84(12):6096–102. Epub 2010/04/02. https://doi.org/10.1128/JVI.00127-10 PMID: 20357090; PubMed Central PMCID: PMCPMC2876646.

37. Ioannidis JP, Haviir DV, Tebas P, Hirsch MS, Collier AC, Richman DD. Dynamics of HIV-1 viral load rebound among patients with previous suppression of viral replication. AIDS. 2000; 14(11):1481–8. Epub 2000/09/13. PMID: 10953654.

38. Palmer S, Maldarelli F, Wiegand A, Bernstein B, Hanna GJ, Brun SC, et al. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. Proc Natl Acad Sci U S A. 2008;
39. Gilmore JB, Kelleher AD, Cooper DA, Murray JM. Explaining the determinants of first phase HIV decay dynamics through the effects of stage-dependent drug action. PLoS Comput Biol. 2013; 9(3): e1002971. Epub 2013/04/05. https://doi.org/10.1371/journal.pcbi.1002971 PMID: 23555209; PubMed Central PMCID: PMCPMC3610612.

40. Érdi P, Tóth J. Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models: Manchester University Press; 1989.

41. McKinstry KK, Strutt TM, Swain SL. Regulation of CD4+ T-cell contraction during pathogen challenge. Immunol Rev. 2010; 236:110–24. Epub 2010/07/20. https://doi.org/10.1111/j.1600-065X.2010.00921.x PMID: 20636812; PubMed Central PMCID: PMCPMC2908916.

42. Sun Y, Fu Y, Zhang Z, Tang T, Liu J, Ding H, et al. The investigation of CD4+ T-cell functions in primary HIV infection with antiretroviral therapy. Medicine (Baltimore). 2017; 96(28):e7430. Epub 2017/07/13. https://doi.org/10.1097/MD.0000000000007430 PMID: 28700479; PubMed Central PMCID: PMCPMC5515751.

43. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. J Virol. 2010; 84(19):9864–78. Epub 2010/07/28. https://doi.org/10.1128/JVI.00915-10 PMID: 20660205; PubMed Central PMCID: PMCPMC2937799.

44. Hockett RD, Kilby JM, Derdeyn CA, Saag MS, Sillers M, Squires K, et al. Constant mean viral copy number per infected cell in tissues regardless of high, low, or undetectable plasma HIV RNA. J Exp Med. 1999; 189(10):1545–54. Epub 1999/05/20. PMID: 10330433; PubMed Central PMCID: PMCPMC2193638.

45. Haase AT. Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues. Annu Rev Immunol. 1999; 17:625–56. Epub 1999/06/08. https://doi.org/10.1146/annurev.immunol.17.1.625 PMID: 10358770.

46. Markowitz M, Louie M, Hurley A, Sun E, Di Mascio M, Perelson AS, et al. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. J Virol. 2003; 77(8):5037–8. Epub 2003/03/29. https://doi.org/10.1128/JVI.77.8.5037-5038.2003 PMID: 12663814; PubMed Central PMCID: PMCPMC152136.

47. Blumm MR, Chittick GE, Begley JA, Zong J. Steady-state pharmacokinetics of emtricitabine and tenofovir disoproxil fumarate administered alone and in combination in healthy volunteers. J Clin Pharmacol. 2007; 47(6):751–9. Epub 2007/05/24. https://doi.org/10.1177/0091270007300951 PMID: 17519400.

48. Agresti A. Categorical data analysis: John Wiley & Sons; 2013.