A compound interest approach to HIV cure

Daniel B Reeves [1], Elizabeth R Duke [1,2], Martin Prlic [1,3],
Florian Hladik [1,2,4], Joshua T Schiffer [1,2,5]

July 11, 2016

[1] Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, 1100 Eastlake Ave, Seattle, WA 98109

[2] Department of Medicine, University of Washington, 1959 NE Pacific St, Seattle, WA 98105

[3] Department of Global Health, University of Washington, 1959 NE Pacific St, Seattle, WA 98105

[4] Departments of Obstetrics and Gynecology, University of Washington, 1959 NE Pacific St, Seattle, WA 98105

[5] Clinical Research Division, Fred Hutchinson Cancer Research Center, 1100 Eastlake Ave, Seattle, WA 98109

Contents

Abstract

In the era of antiretroviral therapy (ART), HIV-1 infection is no longer tantamount to early death. Yet the benefits of treatment are available only to those who can access, afford, and tolerate taking daily pills. True cure is challenged by HIV latency, the ability of integrated virus to persist within memory CD4$^+$T cells in a transcriptionally quiescent state and reactivate when ART is discontinued. Using a mathematical model of HIV dynamics, we demonstrate that treatment strategies offering modest but continual enhancement of reservoir clearance rates result in faster cure than abrupt, one-time reductions in reservoir size. We frame this concept in terms of compounding interest: small changes in interest rate drastically improve returns over time. On ART, latent cell proliferation rates are orders of magnitude larger than activation rates. Contingent on subtypes of cells that may make up the reservoir and their respective proliferation rates, our model predicts that coupling clinically available, anti-proliferative therapies with ART would result in functional cure within 2-10 years rather than many decades on ART alone.

1 Introduction

The most significant accomplishment in HIV medicine is the suppression of viral replication and prevention of AIDS with antiretroviral therapy (ART) [1]. However, HIV cure remains elusive due to viral latency, the ability of integrated virus to persist for decades within $CD4^+$T cells in a transcriptionally quiescent state. When ART is discontinued, latently infected cells soon reactivate, and virus rebounds [2, 3]. Attempts to cure HIV by eradicating the latent reservoir (summarized in [4]) have been unsuccessful except in one notable example [5].

To investigate persistence of the latent reservoir, we use a mathematical model that distinguishes between latently and actively infected cells. After demonstrating analytically that infection of new cells does not meaningfully contribute to sustaining the latent reservoir in ART-treated patients, we separate the observed clearance half-life of 44 months [2] into one positive component rate (proliferation) and two negative component rates (death and reactivation). The reactivation rate defines the rate at which memory $CD4^+$T cells transition to an activated phenotype in which HIV initiates replication. We demonstrate the impact of changes to any of these rates via an analogy with the compound interest formula in finance: small changes in interest rates result in large overall gains or losses over time. In the absence of an intervention, proliferation and death rates far exceed the reactivation rate [6]. Thus, reduction of the proliferation rate will deplete the latent pool much more rapidly than a comparable increase in the reactivation rate as in latency reversal therapy ('shock-and-kill'). In keeping with the financial analogy, we call the strategy of continuous reduction of the proliferation rate *compound interest cure*. We conclude by proposing strategies to test this concept and by addressing theoretical reasons for its failure, including latency in various T cell subsets with different proliferation rates and potential toxicities of anti-proliferative drugs.

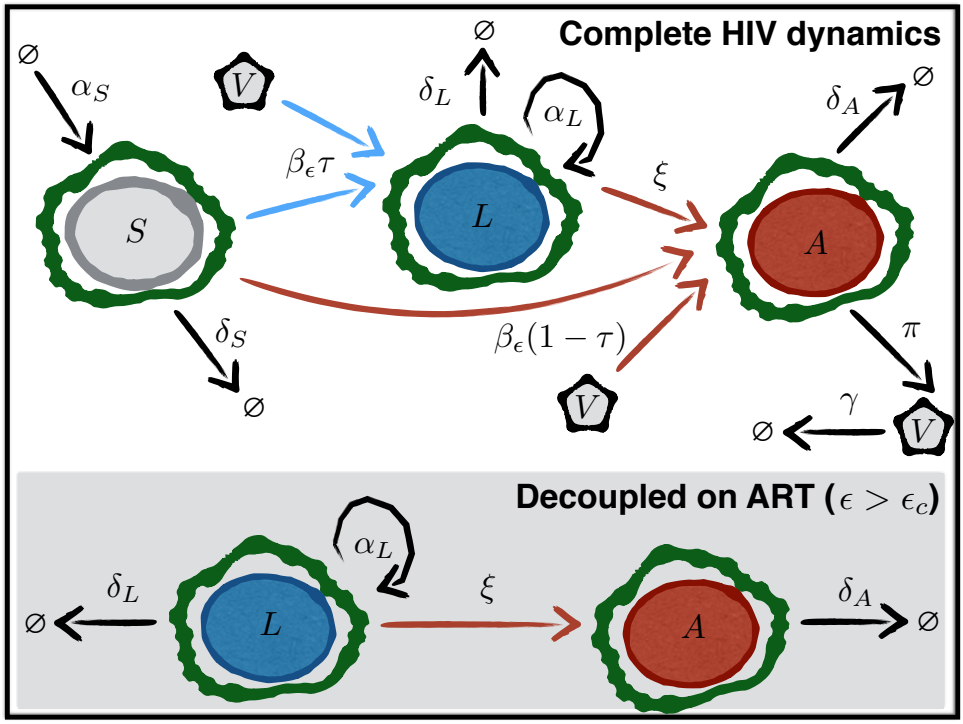

Figure 1: Schematics of models for HIV dynamics on and off ART. The top panel shows all possible transitions in the complete set of equations. The bottom shaded panel shows the available transitions for the decoupled dynamical equations when ART suppresses the virus. The parameters in the model are found in Table 1. Infection by HIV virus V of susceptible cells S mediated by ART of efficacy ϵ is given by β_ϵ. The probability of latent infection is τ. The rate of activation from the latently infected (L) to active state (A) is ξ. Cellular proliferations and deaths are determined by α and δ for each compartment.

2 Results

2.1 ART decouples latent pool dynamics from ongoing infection

To model the latent reservoir, we expand the infected compartment of the original HIV model [7] to account for the latently infected cells L and actively replicating, infected cells A [6, 8, 9, 10]. Our model is a set of ordinary differential equations (visualized schematically in the top panel of Fig. 1 and detailed in the Materials and Methods section). All parameters and their meanings are tabulated in Table 1 with further information in the Supporting Information (SI).

In particular, we adjust the infectivity parameter β to capture the effectiveness of ART therapy by defining $\beta_\epsilon = \beta(1-\epsilon)$ such that inert therapy implies $\epsilon = 0$, and perfectly effective therapy implies $\epsilon = 1$. If ART is perfectly effective, no virus is generated, and no new infections occur even when

cells reactivate from latency. Then, the dynamics of the latent pool can be considered on their own, decoupled from the dynamics of the other cell types, and the only mechanisms changing the latent cell pool size are cell proliferation, death, and activation (shown in the lower panel of Fig. 1).

Perfectly effective ART is not strictly necessary to consider the latent pool on its own. Using the tools of dynamical systems mathematics, we can derive an expression for the viral set-point and calculate the critical therapy efficacy ϵ_c in terms of the rates in Table 1 (Materials and Methods) [11]. When ART achieves the critical efficacy, viral production from reactivation infrequently creates another latent cell because the probability of latency is so low ($\tau \ll 1$). Therefore, we can consider the latent cells as their own system, decoupled from the other cell types. Put another way, the model does not preclude the possibility of some small, ongoing viral replication and infection, but rather demonstrates that this mechanism is unlikely to contribute to latent reservoir dynamics. [10, 12].

The mathematical analysis is pertinent to discussion of viral evolution and ART escape mutations. From a practical point of view, using parameters from Table 1, we find $\epsilon_c \sim 85\%$, considerably lower than empirically derived measures of ART potency [13, 14]. Thus, because true ART efficacy is likely much larger than the necessary minimum ϵ_c, we predict little *de novo* infection in patients suppressed on ART, consistent with the reported lack of viral evolution following years of ART [12, 15, 16]. We do not address the question of anatomic drug sanctuaries directly in the model. However, to have an impact on the reservoir, a sanctuary would have to contribute enough ongoing infection to drive the average ART efficacy below critical levels and thus to permanent viral rebound, which does not occur on suppressive ART.

Above the critical efficacy, we can solve for the decoupled time-evolution of the latent pool (see Materials and Methods):

$$L(t) = L_0 e^{\theta_L t} \tag{1}$$

in which we define $\theta_L = \alpha_L - \delta_L - \xi$ as the net clearance rate for the latent cells and L_0 the initial size of the reservoir. Eq 1 implies that the clearance rate of infected cells is a function of their proliferation rate α_L minus both death δ_L and activation ξ rates. Experimental measurements indicate an average latent cell half-life of 44 months ($\theta_L = -5.2 \times 10^{-4}$ per day) [2, 17] and an

Table 1: Parameters used in the HIV latency model

Param.	Value	Dimensions	Source	Meaning
θ_L	-5.2×10^{-4}	day^{-1}	[2, 17]	net latent T cell clearance rate on ART
α_{cm}	0.015	day^{-1}	[18, 19]	latent central memory T cell proliferation rate
δ_{cm}	0.0155	day^{-1}		latent central memory T cell death rate
α_{em}	0.047	day^{-1}	[18, 19]	latent effector memory T cell proliferation rate
δ_{em}	0.0475	day^{-1}		latent effector memory T cell death rate
α_n	0.002	day^{-1}	[18, 19]	latent naive T cell proliferation rate
δ_n	0.0025	day^{-1}		latent naive T cell death rate
ξ	5.7×10^{-5}	day^{-1}	[8]	activation rate
α_A	0	day^{-1}	[14]	active T cell proliferation rate
δ_A	1.0	day^{-1}	[14]	active T cell death rate
τ	10^{-4}	day^{-1}	[10]	probability of latency given infection
α_S	300	cells/(μL-day)	[20]	susceptible T cell growth rate
δ_S	0.2	day^{-1}	[14]	susceptible T cell death rate
β	10^{-4}	μL/(virus-day)	[6]	HIV infectivity off ART
β_ϵ	$\beta(1-\epsilon)$	μL/(virus-day)	[6]	HIV infectivity on ART (with efficacy $\epsilon \in [0,1]$)
π	10^3	virus/(cell-day)	[8]	viral production rate
γ	23	day^{-1}	[14]	viral clearance rate

average initial latent pool size L_0 of one-million cells [2]. A mathematical correspondence to the principle of continuous compound interest is now evident with L_0 as the principal investment and θ_L as the interest rate.

2.2 Existing thresholds are used to compare cure strategies

We use existing experimentally derived thresholds to compare potential cure therapies in the framework of our mathematical model. Hill *et al.* employed a stochastic model to estimate that a 2,000-fold reduction in the latent pool would result in HIV suppression off ART for a median time of one year. After a 10,000-fold reduction in latent cells, 50% of patients would remain functionally cured (defined by the authors as ART-free remission for at least 30 years) [8]. Pinkevych *et al.* inferred from patient data that decreasing the latent reservoir by 50-70-fold would lead to HIV remission in 50% of patients for one year [21]. Using the Pinkevych *et al.* results, we extrapolate a functional cure threshold as a 2,500-fold decrease in the reservoir size (Materials and Methods). In the following discussion, we consider all four thresholds—henceforth referred to as Hill 1-yr, Hill cure, Pinkevych 1-yr and Pinkevych cure.

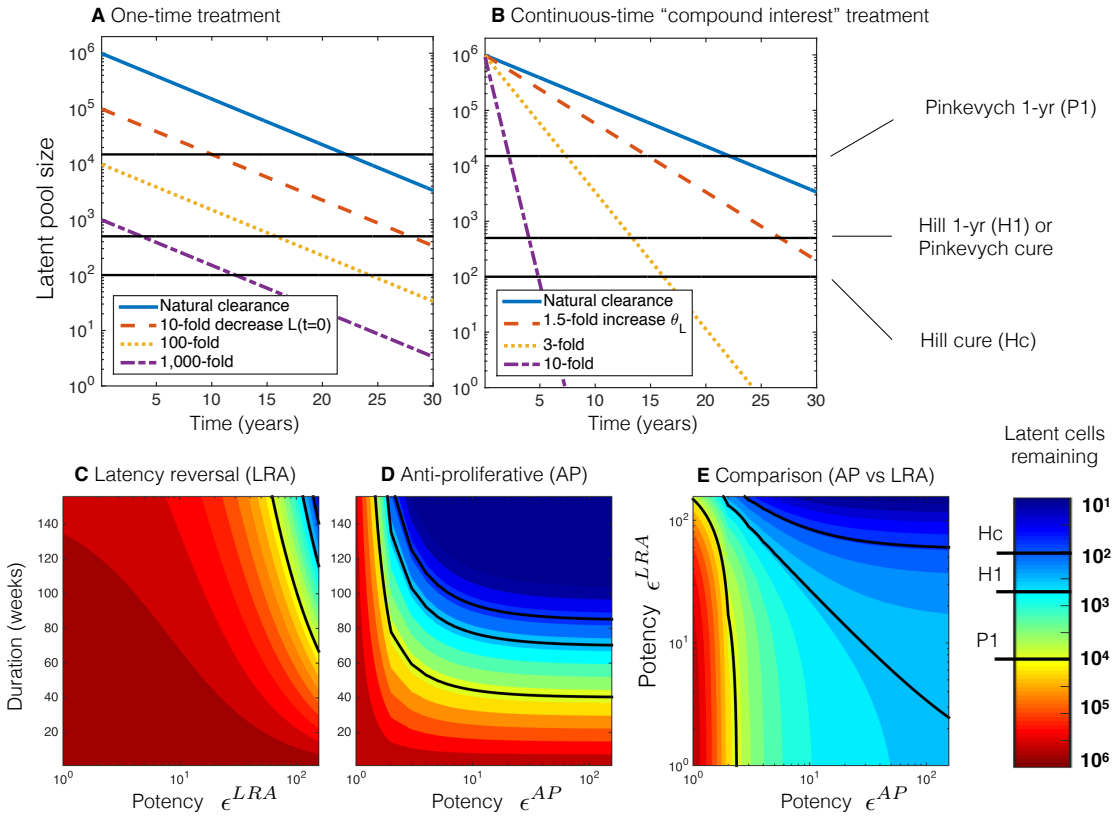

Figure 2: Simulated comparisons of latent reservoir eradication strategies on standard antiretroviral (ART) treatment. The thresholds (shown by black lines) are discussed in the text. A) One-time therapeutic reductions of the latent pool (L_0). B) Continuous therapeutic increases in the clearance rate (θ_L). Relatively small decreases in rates produces markedly faster times to cure than much larger decreases in the initial reservoir size. C-E) Latency reversal agent (LRA) and Anti-proliferative (AP) therapies are given continuously for durations of weeks with potencies given by ϵ^{LRA}, and ϵ^{AP} respectively. The color bar is consistent between panels, and thresholds of cure are shown as solid black lines both on plots and on the color bar. C) Latency reversing agent therapy (LRA) administered alone requires years and potencies above 100 to achieve the cure thresholds. D) Anti-proliferative therapies (AP) administered alone lead to cure thresholds in 1-2 years provided potency is greater than 2-3. E) LRA and AP therapies are administered concurrently, and the reduction in the latent pool is measured at 70 weeks. Because the proliferation rate is naturally greater than the reactivation rate, increasing the AP potency has a much stronger effect than increasing the LRA potency.

2.3 Sustained mild effects on clearance rate deplete the reservoir more rapidly than large, one-time reductions in the reservoir

The HIV cure strategy most extensively tested in humans is 'shock-and-kill' therapy: small molecules (called latency reversing agents or LRAs) activate latent cells to replicate and express HIV proteins, allowing immune clearance while ART prevents further infection [4, 22]. Other strategies in

development include therapeutic vaccines [23], viral delivery of DNA cleavage enzymes [24], and transplantation of modified HIV-resistant cells [25] informed by the 'Berlin patient' who was cured following a similar protocol [5]. Some of these therapies may manifest as one-time reductions in the number of latent cells. We simulate such instantaneous decreases using Eq. 1 in Fig. 2A. With θ_L constant and a 100-fold reduction in L_0, the Pinkevych 1-yr threshold is immediately satisfied, but the Hill 1-yr and Pinkevych cure still require 15 years of ART. Hill cure requires a 1000-fold reduction and over 10 subsequent years of ART.

In comparison to one-time treatments, continuous-time interventions are more promising. In Fig. 2B, relatively small changes in θ_L in Eq. 1 lead to significant changes in the time to cure. On ART alone, the reservoir is depleted slowly with estimated cure occurring at roughly 70 years [2]. However, just a 3-fold increase in clearance rate achieves Hill-cure in fewer than 20 years. If a 10-fold sustained increase is possible, Hill cure requires only 5 years. Analogous to the 'miracle of compound interest,' increasing the clearance rate for an extended duration produces profound latency reduction.

2.4 Smaller reductions in proliferation rate achieve more rapid reservoir depletion than comparable relative increases in reactivation rate

Latency reversing therapy can be modeled by Eq. 1 if treatment is assumed to be a continuous-time multiplication of the reactivation parameter (ξ). Simulations of the latent pool at various potencies and therapy durations in Fig. 2C indicate both Hill and Pinkevych cure thresholds require more than a 150-fold multiplication of ξ sustained for two or three years, respectively.

Experimental evidence [15, 19, 26, 27] and theoretical analysis [6] demonstrate that proliferation of latently infected cells is critical for reservoir persistence. While two anti-proliferative drugs, mycophenolate mofetil (MMF) and hydroxyurea, have been used in the past as adjuncts to ART to further limit replication [28, 29, 30], anti-proliferative strategies have not been tested specifically to treat HIV latency [31]. We now justify mathematically how such therapies could have curative potential.

The latent cell proliferation rate is considerably larger than the reactivation rate ($\alpha_L \gg \xi$). Thus,

anti-proliferative therapies would clear the reservoir faster than equivalently potent latency reversal strategies. When the reservoir of $CD4^+$T cells harboring replication competent HIV is assumed to consist only of central memory cells (T_{cm}), a 10-fold reduction in $\alpha_{L(cm)}$ reaches Pinkevych 1-yr, Hill 1-yr, Pinkevych cure, and Hill cure in 0.8, 1.6, 1.6, and 1.8 years, respectively (Fig. 2D).

The improvement in cure time in Fig. 2D (as opposed to equivalent 10-fold reduction of θ_L in Fig. 2B) is possible because decreasing the proliferation rate means the net clearance rate approaches the death rate. The potency is relatively unimportant beyond reducing the proliferation rate by a factor of ten because the underlying death rate δ_{cm} is actually the bound on the clearance rate. Fig. 2D also demonstrates substantial asymmetry of the contours over its $y = x$ axis, conveying that outcomes improve more by extending duration than by equivalent increases in potency. This can also be demonstrated mathematically (see SI). Fig. 2E further illustrates that the relative impact of anti-proliferative therapy is much greater than that of latency reversal therapy when the two therapies are given concurrently for just over a year.

2.5 Heterogeneity in reservoir cell types may necessitate prolonged anti-proliferative therapy

Recent studies indicate that the reservoir is heterogeneous, consisting of $CD4^+$central memory (T_{cm}), naïve (T_n), effector memory (T_{em}), and stem cell-like memory (T_{scm}) T cells with composition differing dramatically among patients [19, 32, 33]. This heterogeneity suggests the potential for differing responses to anti-proliferative agents. T_{cm} homeostatic proliferation rates exceed the rates for $T_n$but lag behind antigen-driven T_{em} turnover rates (Table 1). T_{scm} are assumed to proliferate at the same frequency as T_n in our model based on similar properties. We simulate possible reservoir profiles in Fig. 3A-C. At least 7 years of treatment is needed for Pinkevych functional cure if slowly proliferating cells (T_n and/or T_{scm}) comprise more than 20% of the reservoir. In contrast, an increased proportion of T_{em} has no clinically meaningful impact on time to cure. Slowly proliferating cells are predicted to comprise the entirety of the reservoir within two years of 10-fold anti-proliferative treatment regardless of initial percentage T_n or T_{scm} (Fig. 3D&E). Therefore, anti-proliferative strategies may face a challenge akin to the cancer stem cell paradox, whereby only the rapidly proliferating tumor cells are quickly expunged with chemotherapy. For

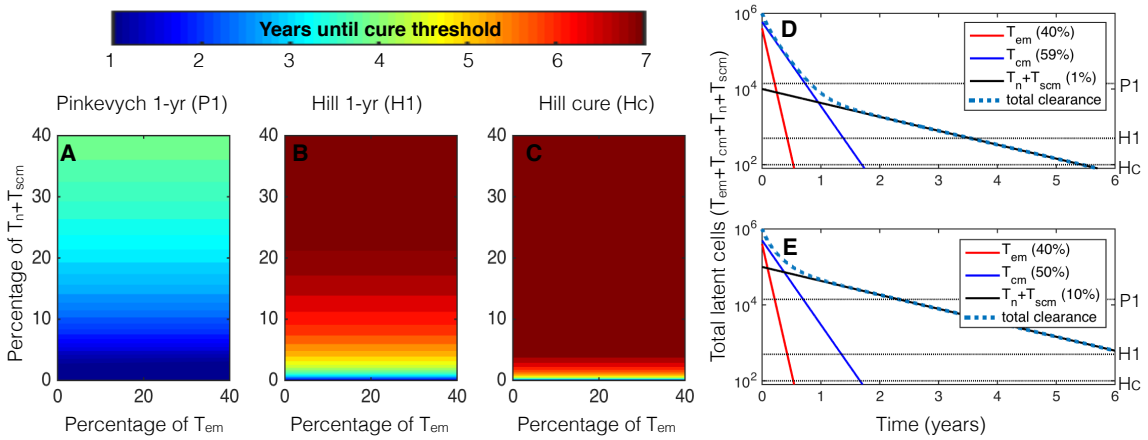

Figure 3: Simulated comparisons of anti-proliferative therapies on standard antiretroviral therapy (ART) assuming variable reservoir composition; see proliferation and death rates in Table 1. The potency of the therapy is $\epsilon^{AP} = 0.1$ (i.e., each cell type i has proliferation rate equal to $\alpha_i/10$ with $i \in [em, cm, n]$). Plausible initial compositions of the reservoir ($L_i(0)$) are taken from experimental measurements [19, 32, 33]. It is assumed that the HIV reactivation rate ξ is equivalent across all reservoir subsets. A-C) Plots of times to therapeutic landmarks on long-term ART and anti-proliferative therapy with heterogeneous reservoir compositions consisting of effector memory (T_{em}), central memory (T_{cm}), and naïve plus stem cell-like memory (T_n+T_{scm}) CD4$^+$T cells. T_{em} and T_n+T_{scm} percentages are shown with the remaining cells representing T_{cm}. Times to one-year remission and functional cure are extremely sensitive to percentage of T_n+T_{scm} but not percentage of T_{em}. D&E) Continuous 10-fold therapeutic decreases in all proliferation rates (α_i) result in Hill 1-yr/Pinkevych cure in D) 3.5 years assuming T_n+T_{scm}=1% and E) 6 years assuming T_n+T_{scm}=10%. The reservoir is predicted to become T_n+T_{scm} dominant within 2 years under both assumptions, thus providing an indicator to gauge the success of anti-proliferative therapy in a potential experiment.

example, tyrosine kinase inhibitors suppress proliferation of cancer cells in chronic myelogenous leukemia (CML). While many patients achieve 'undetectable minimal residual disease,' some patients relapse to pre-therapy levels of disease following cessation of therapy—perhaps due to slowly proliferating residual cancer cells [34].

The uncertainty in the reservoir composition tempers the results in Fig. 2. On the other hand, our model assumes that the HIV reactivation rate ξ is equivalent across all CD4$^+$T cell reservoir subsets. It is biologically plausible, though unproven, that HIV rarely or never reactivates from resting T_n or T_{scm}. Under this assumption, more rapid functional cure akin to Fig. 3D would be expected because theoretically this would allow cessation of ART despite long-term persistence of T_n harboring latent HIV.

2.6 Reservoir size, reservoir composition, and efficacy of anti-proliferation agents predict time to HIV cure

We performed a single parameter sensitivity analysis (Fig. S3) and identified that the latency probability τ, ART efficacy above the critical efficacy ϵ, and reactivation rate of HIV ξ did not vary the kinetics of the HIV reservoir using concurrent ART and anti-proliferative therapy. Alternatively, low initial reservoir size L_0, low percentage of naïve cells in the reservoir $L_n(0)/L_0$, and high reservoir clearance rate $-\theta_L$ decreased time to cure.

We also performed an uncertainty analysis in which multiple therapeutic parameters were randomly selected from distributions and generated a wide range of therapeutic outcomes among individual simulated patients (Fig. S4A). No single parameter was completely predictive of time until Pinkevych 1-yr or Hill cure (Fig. S4B), and ART efficacy again had no impact on time to cure above the critical efficacy. Low initial reservoir size, low percentage of naïve cells in the reservoir, and high anti-proliferative therapy efficacy were the most predictive factors of rapid reservoir clearance (Fig. S4C).

2.7 Model output is congruent with available clinical data

Chapuis *et al.* treated 8 ART-suppressed, HIV-infected patients with 24 weeks of mycophenolate mofetil (MMF). They measured $Ki67^+CD4^+$T cells before and after MMF treatment as a marker of anti-proliferative effect and found the $Ki67^+CD4^+$T cells had decreased 2.2-fold. Incorporating that reduction in latent cell proliferation rate α_L into our model, we find an average reduction in the latent reservoir of 77% using the assumptions in Fig. 2. Chapuis *et al.* also estimated an average relative reduction in reservoir size by quantitative viral outgrowth assay in 6 of those patients and found a 1 to 2 log reduction in infectious units per million (IUPM) in 5 of those patients and a statistically significant reduction in 3 [29].

Garciá *et al.* treated 9 virally suppressed HIV patients with MMF for 17 weeks before ART was discontinued. $Ki67^+CD4^+$T cell measurements demonstrated an anti-proliferative response in 6 of 9. The median time to viral rebound (6-12 weeks) was longer, and average viral set-point was lower (by ≈ 1 log) than in the control cohort and in the 3 patients with no pharmacologic response

to MMF [28]. Our model estimates a 1-2 order of magnitude reduction in reservoir size after 17 weeks of treatment, which in turn would correspond with a median of a 4-8 week delay in HIV reactivation, consistent with their findings (see SI for parameter selections) [8].

3 Discussion

Several authors have studied similar systems of equations to represent the latency dynamics of HIV [6, 8, 9, 10]. Here, we focus on cure strategies and contrast their respective times to functional cure. We demonstrate that minor reductions in memory $CD4^+$T cell proliferation rate would exhibit powerful reductions in the latent pool when therapy duration is extended over time. We call this proposed strategy *compound interest cure* due to the correspondence with financial modeling.

Our results are relevant because the only HIV cure strategy currently being tested in humans—latency reversal therapy ('shock-and-kill')—does not capitalize on the advantages of a compound interest strategy. Promising latency reversing agents are typically dosed over short time-frames based on concern for toxicity. Furthermore, even if these agents exert a large relative impact on the reactivation rate of memory $CD4^+$T cells, we predict the reduction in the reservoir will be insignificant given the small size of the natural reactivation rate (orders of magnitude smaller than the proliferation and death rates).

The potential of the compound interest approach is enhanced by the existence of licensed medications that limit T cell proliferation, including mycophenolate mofetil (MMF) and azathioprine. In line with our suggestion that duration is more important than potency, these drugs are dosed over time periods of months or years. MMF is commonly used to treat several rheumatologic conditions and to prevent rejection in solid organ transplant. The most frequent side effects reported are gastrointestinal symptoms and increased risk of infection though the latter risk is obscured by concurrent use of high-dose glucocorticoids [35]. MMF has been given to a small number of HIV patients suppressed on ART for 17 [28] and 24 weeks [29]. No opportunistic infections or adverse events were observed in either study, and $CD4^+$T cell counts did not decrease significantly during therapy. We hypothesize that whereas MMF decreases proliferation of existing $CD4^+$T cells, it does not suppress bone marrow replenishment of these cells (see SI). Finally, MMF did not coun-

teract the effects of ART [28, 29], and thus we do not expect viral drug resistance to develop or ongoing viral evolution to occur on anti-proliferative therapy. However, despite these reassuring findings, future studies of HIV infected patients on anti-proliferative agents will require extremely close monitoring for drug toxicity and immunosuppression.

Our model suggests that slowly proliferating cells in the reservoir could present a barrier to rapid eradication of latently HIV-infected cells. Additional limitations could include insufficient drug delivery to anatomic sanctuaries, certain cellular subsets that are unaffected by treatment, and cytokine-driven feedback mechanisms that compensate for decreased proliferation. These challenges might be countered by combining anti-proliferative agents with other cure therapies. Avoidance of nucleoside and nucleotide reverse transcriptase inhibitors, which may enhance homeostatic proliferation, could provide an important adjunctive benefit [31, 36, 37].

The anti-proliferative approach is attractive because it is readily testable without the considerable research and development expenditures required for other HIV cure strategies. Anti-proliferative approaches require minimal potency relative to latency reversal agents, and T cell anti-proliferative medications are well studied mainstays of organ rejection prevention. Therefore, we propose trials with anti-proliferative agents as an important next step in the HIV cure agenda.

4 Methods

4.1 Decoupling the latent reservoir dynamics

HIV dynamics have been described mathematically in many fashions [38]. We describe a model (schematic in Fig. 1) that includes susceptible and latently or actively infected T cells. We follow the concentrations [cells/μL] of susceptible CD4$^+$T cells S, latently infected cells L, actively infected cells A, and plasma viral load V [virions/μL]. The system of ordinary differential equations tracks

the states and is expressed (using the over-dot to denote derivative in time)

$$\begin{aligned}\dot{S} &= \alpha_S - \delta_S S - \beta_\epsilon SV \\ \dot{L} &= \alpha_L L + \tau\beta_\epsilon SV - \delta_L L - \xi L \\ \dot{A} &= (1-\tau)\beta_\epsilon SV - \delta_A A + \xi L \\ \dot{V} &= \pi A - \gamma V,\end{aligned} \tag{2}$$

with α_S [cells/μL-day] as the constant growth rate of susceptible cells, δ_S [1/day] as the death rate of susceptible cells, and $\beta_\epsilon = (1-\epsilon)\beta$ [μL/virus-day] as the therapy-dependent infectivity with ϵ [unitless] the therapy efficacy ranging from 0 (meaning no therapy) to 1 (meaning perfect therapy). α_L and δ_L [1/day] are the proliferation and death rates of latent cells. The death rate of activated cells is δ_A, and the proliferation of activated cells $\alpha_A \approx 0$ is likely negligible [39]. τ [unitless] is the probability of becoming latent given infection, and ξ [1/day] is the activation rate from latent to actively infected cells. The viral production rate is π [virions/cell-day], mechanistically describing the aggregate rate of constant leakage and burst upon cell death. γ [1/day] is the clearance rate of virus. The values of the rates are defined in Table 1.

Equilibrium solutions to the set Eq. 2 can be calculated by solving the set when all $\dot{S}, \dot{L}, \dot{A}, \dot{V} = 0$. See SI for this calculation and stability analysis. Notably, the viral 'set-point,' or the equilibrium value for non-zero virus and infected cells, is

$$V^* = \frac{\pi\alpha_S f_L}{\gamma\delta_A} - \frac{\delta_S}{\beta_\epsilon} \tag{3}$$

where we have defined the dimensionless latency factor $f_L = 1 - \tau(\xi/\theta_L + 1)$ to encapsulate all latent pool dynamics. Solving this equation, the critical drug efficacy ϵ_c can be derived such that the viral equilibrium is unstable ($V^* < 0$):

$$\epsilon_c = 1 - \frac{\delta_S\gamma\delta_A}{\beta\pi\alpha_S f_L}. \tag{4}$$

If it is assumed that ART also affects π in Eq. S10, i.e. $\pi \to \pi(1-\epsilon)$ this decreases the critical efficacy (SI). Above the critical efficacy reactivation from latency is the only mechanism that produces

virus. After a reactivation, a transient small viral load exists, but the small probability (τ) of new latent infections from this small amount of virus allows us to ignore any increase in latent cells. Thus, the latent state can be considered independently, and the transition from latency to activation represents another pathway to removal from the latent pool. We express the dynamics of latency decoupled from virus with the two-equation system as shown in the shaded bottom panel of Fig. 1:

$$\begin{aligned} \dot{L} &= \alpha_L L - \delta_L L - \xi L \\ \dot{A} &= -\delta_A A + \xi L. \end{aligned} \tag{5}$$

A similar model has been studied recently in Conway and Perelson [10] who also show that transitions back to latent cells from replicating active cells are negligible. In another manner, we demonstrate the ability to ignore transitions back from active cells in simulation (Fig. S1&S2). The two equation system is linear, and we solve for the latent dynamics as

$$L = L_0 e^{\theta_L t} \tag{6}$$

where the initial number of latent cells is L_0 and the total clearance rate $\theta_L = \alpha_L - \delta_L - \xi$ for the latent cells corresponds with the experimentally known net clearance rate. We use the correspondence of Eq. 6 to the formula for continuous compound interest throughout the paper.

4.2 Modeling multiple T cell subsets

It has been demonstrated that several subsets of CD4$^+$T cells may contribute to the reservoir. We reviewed existing data to find the rates of proliferation for each subset as well as estimates of the total fractional makeup of the reservoir [19, 32, 33]. The typical groupings are central memory (T_{cm}), effector memory (T_{em}), and naïve (T_n) cells. Deuterated glucose measurements by Macallan *et al.* found turnover rates as 4.7% per day for T_{em}compared with 1.5% per day for T_{cm} and only 0.2% per day for T_n[18]. Transitional memory T cells (T_{tm}) have also been described to represent a transition from $T_{cm}$to T_{em}[19]. For the model we consider $T_{tm}$to have equivalent proliferation rates to central memory cells. Similarly, we characterize the recently described stem cell-like memory

$CD4^+$T cells (T_{scm}) as T_n given their likely slow turnover rate [33]. Chomont *et al.* determined the percentage of each cellular subset from 31 aviremic individuals [19]. They find mean contributions of T_{cm}+T_{tm} to be 86%, whereas T_{em} are 13% and Tn roughly 1-2% though there is a wide range of values amongst patients. Buzon *et al.* report larger percentages of naiv̈e cells up to 15% and include T_{scm} at 24%, T_{em} at 26%, and T_{cm} at 22% [33].

To model the heterogeneity, we make the latent pool $L \to L_i$ a vector with a dimension i for each T cell subset. The differential equation for the decoupled latent dynamics is then

$$\dot{L}_i = \theta_{ij} L_j. \tag{7}$$

We ignore transitions among subsets, as the natural composition of the reservoir is reasonably stable over time [19]. The solution then with diagonal $\theta_{ii} = \alpha_i - \delta_i - \xi$ is

$$L = \sum_i L_i(0) \exp[(\alpha_i - \delta_i - \xi)t]. \tag{8}$$

The percentages discussed in the preceding paragraph determine the initial makeup of the reservoir $L_i(0)$. We simulate the dynamics of a heterogeneous reservoir with 10% potent anti-proliferative therapy $\epsilon^{AP} = 10$ in Fig. 3 using

$$L = \sum_i L_i(0) \exp[(\alpha_i/10 - \delta_i - \xi)t] \tag{9}$$

in which the index $i \in [cm, em, n]$ indicates the T cell subsets with rates from Table 1.

4.3 Important model parameters

The parameters that affect our model most critically are those that contribute to the total clearance of latent cells θ_L. The half-life of latently infected cells has been determined experimentally using the quantitative viral outgrowth assay [2, 17]. We use a proliferation rate for α_L obtained using *in vivo* labeling of $CD4^+$T cells with deuterated glucose [18]. Importantly, these estimates depend on the length of the labeling period and the model used to interpret the measurements [40]. Thus, we perform uncertainty and sensitivity analyses (see SI) to explore a larger range of values. In

Fig. 2, we assume that the majority of latently infected $CD4^+T$ cells are in fact central memory T cells (T_{cm}), and thus $\alpha_L = \alpha_{cm} = 0.015$ per day. The importance of reservoir proliferation has been supported experimentally [19, 16] and predicted theoretically [6, 41]. In [6], α_L parameter values were chosen between 0.011 and 0.03 per day, consistent with our model. The activation rate ξ is several orders of magnitude smaller than the proliferation rate. Luo *et al.* estimated HIV model parameters from structured treatment interruptions in the AutoVac trial [14]. Using those estimates, Hill *et al.* found that on average, 57 $CD4^+T$ cells per day transition from latency to the activated state and assuming 1 million latent cells then, $\xi = 5.7 \times 10^{-5}$ per day [8]. We calculate the death rate for each of i cell types δ_i in Table 1, using the respective proliferation, activation and net clearance rates [18]. The remaining parameters are necessary here only to calculate the critical efficacy. They are tabulated in Table 1 for completeness and are discussed in further detail in the SI.

5 Acknowledgments

We thank Keith Jerome for his reading of the manuscript, the VIDD faculty initiative at the Fred Hutch, and the NIH (grant R01 AI116292 to FH, and 1DP2DE023321-01 to MP) for funding.

Some results in the main body of the paper were stated without full justification. Extended calculations and results from models that are secondary to the main model of the paper are included in this supplementary information document. The first section (§1) expounds upon the basic latency model and addresses some possible extensions. The second (§2) is a discussion of how model parameters were obtained from the literature. The third and final section (§3) contains both a local and global uncertainty and sensitivity analysis for parameters in the model as found in §2. We compare our theoretical treatment using anti-proliferative agents in conjunction with ART with previous experimental results, and discuss how anti-proliferative therapy can deplete HIV-infected $CD4^+T$ cells without depleting all $CD4^+T$ cells from the body.

6 Appendix I: Additional model discussion and calculations

6.1 HIV dynamical equations including latency and therapy

The $SLAV$ model for HIV latency is briefly described in the main body and schematized in Fig 1. Our system of equations captures the dynamics of the concentrations [cells/μL] of susceptible CD4$^+$ T cells, latently infected cells L, actively infected cells A, and viral concentration V [virions/μL]. Experimentally derived values of all parameters are found in Table 1 of the main body. The system is

$$\begin{aligned}
\dot{S} &= \alpha_S - \delta_S S - \beta_\epsilon SV \\
\dot{L} &= \theta_L L + \tau\beta_\epsilon SV \\
\dot{A} &= (1-\tau)\beta_\epsilon SV - \delta_A A + \xi L \\
\dot{V} &= \pi A - \gamma V
\end{aligned} \tag{S10}$$

where the over-dot denotes time derivative. We use the variables described in Table S1 throughout the paper.

6.2 Equilibrium solutions leading to decoupled equations for the latent pool

Equilibrium solutions (denoted by the asterisk) to the set of ODEs Eq. S10 can be calculated by setting $\dot{S}, \dot{L}, \dot{A}, \dot{V} = 0$. The viral free equilibrium has the solution $S^* = \alpha_S/\delta_S$, $L^* = 0$,$A^* = 0$, and $V^* = 0$. We assume that the model begins at this equilibrium such that $S(0) = S^*$. To calculate the non-trivial equilibrium, referred to as the viral set-point equilibrium, we begin by solving the fourth equation for A^* and multiply by δ_A

$$\delta_A A^* = \frac{\delta_A \gamma V^*}{\pi}$$

then identify $\delta_A A^*$ from the third equation to get

$$\frac{\delta_A \gamma V^*}{\pi} = (1-\tau)\beta_\epsilon S^* V^* + \xi L^*$$

Parameter	Description
α_S	Constant growth rate of susceptible cells due to thymic production
δ_S	Death rate of susceptible cells
$\beta_\epsilon = (1-\epsilon^{ART})\beta$	ART dependent infectivity
ϵ^{ART}	ART efficacy from 0 meaning no therapy to 1 meaning perfect therapy
α_L	Proliferation rate of latent cells
α_i	Proliferation rate of latent cells of subset $i \in [cm, em, n]$ to account for different T cell phenotypes
L_0	Initial number in the latent pool
$L_i(0)/L_0$	Initial fraction in the latent pool of subset i or the fraction of new infections of type i (expected to remain constant over time)
ϵ^{AP}	Anti-proliferative potency (acts as α_L/ϵ^{AP})
τ	Probability of becoming latent given infection
δ_L	Death rate of latent cells
ξ	Activation rate from latent to actively infected cell
ϵ^{LRA}	Latency reactivation agent potency (acts as $\xi \cdot \epsilon^{LRA}$)
π	Virus production rate, due to constant leakage or burst upon cell death
γ	Clearance rate of virus
$\theta_L = \alpha_L - \delta_L - \xi$	Net clearance rate of the latent reservoir, assumed the same between T cell subtypes

Table S2: List of parameters and descriptions. Dimensions and values are found in Table 1 of the body; confidence intervals are included in Table S7

then solving the second equation for L^* leads to

$$\frac{\delta_A \gamma V^*}{\pi} = (1-\tau)\beta_\epsilon S^* V^* + \xi \frac{\tau \beta_\epsilon S^* V^*}{-\theta_L}$$

so that we can factor and cancel V^*

$$\frac{\delta_A \gamma}{\pi} = \left[\frac{-\xi\tau}{\theta_L} + (1-\tau)\right] \beta_\epsilon S^*$$

and rewrite the bracketed term as $f_L = 1 - (1 + \xi/\theta_L)\tau$ so that we can solve for the viral set-point equilibrium value of the susceptible cells

$$S^* = \frac{\gamma\delta_A}{\beta_\epsilon \pi f_L}.$$

From here, we solve for the viral set-point equilibrium viral load concentration,

$$V^* = \frac{\alpha_S}{\beta_\epsilon S^*} - \frac{\delta_S}{\beta_\epsilon}$$

and the other equilibria follow, leading to the set of equilibrium solutions:

$$\begin{aligned} S^* &= \frac{\gamma\delta_A}{\beta_\epsilon \pi f_L} \\ L^* &= \frac{\tau}{\theta_L}\left[\frac{\gamma\delta_S\delta_A}{\beta_\epsilon \pi f_L} - \alpha_S\right] \\ A^* &= \frac{\alpha_S f_L}{\delta_A} - \frac{\gamma\delta_S}{\beta_\epsilon\pi} \\ V^* &= \frac{\alpha_S \pi f_L}{\gamma\delta_A} - \frac{\delta_S}{\beta_\epsilon} \end{aligned} \tag{S11}$$

where $f_L = 1 - (1 + \xi/\theta_L)\tau$. We refer to this factor as the latency factor because it encapsulates all the latent dynamics. When $\tau \ll 1$, we have $f_L \sim 1$ which is true for parameter values from the literature (see Table 1).

We can also calculate the basic reproductive ratio $\mathcal{R}_0^{ART}$ of HIV on ART quickly using the equation for the viral equilibrium. When the basic reproductive number is less than 1, the viral equilibrium point is unstable. We can rewrite

$$V^* = \frac{\delta_S}{\beta_\epsilon}\left(\frac{\alpha_S\beta_\epsilon\pi f_L}{\gamma\delta_A\delta_S} - 1\right)$$

and thus identify the combination of parameters

$$\mathcal{R}_0^{ART} = \frac{\alpha_S\beta_\epsilon\pi f_L}{\gamma\delta_A\delta_S}, \tag{S12}$$

where we include the superscript to remind the reader that the value is not the natural value,

but the value with β_ϵ. This expression for the basic reproductive number can be checked using the next-generation matrix method for example [42]. This expression is fairly independent of the activation rate. For example, if LRA treatment is not very potent

$$f_L =\approx 1-\tau,$$

and if LRA treatment provides a huge increase in the activation parameter ($\epsilon^{LRA}\xi > \alpha_L - \delta_L$), aside from the point when they are actually equal and the function is undefined, the expression approaches (using the Binomial expansion $[1-x]^n \approx 1 - nx + \ldots$)

$$f_L = 1-(1+\xi/\theta_L)\tau \approx 1+\tau\frac{\alpha_L-\delta_L}{\epsilon^{LRA}\xi} \approx 1,$$

because $\alpha_L - \delta_L$ is a negative number, as activation increases

$$\mathcal{R}_0^{ART} \approx \frac{\alpha_S\beta_\epsilon\pi}{\gamma\delta_A\delta_S}$$

which is the typical value for the viral dynamics model without a latent compartment. This is understandable because when reactivation is instantaneous, the latent model acts as a model without latency. This is excellent for clearance of the reservoir, but it is unclear how possible it is to have such a potency.

6.3 Critical drug efficacy

The critical drug efficacy ϵ_c is then calculated for the situation where $V^* < 0$ [6]. This leads us to

$$\epsilon_c > 1-\frac{\delta_S\gamma\delta_A}{\beta\alpha_S\pi f_L} \tag{S13}$$

which is precisely the drug efficacy for which $\mathcal{R}_0^{ART} < 1$. Thus when therapeutic effectiveness is above this threshold, the virus will decay and never reach a stable set point. We could also consider a therapeutic effect in the parameter that controls viral production from active cells. This could be implemented by changing $\pi \to \pi_\epsilon = \pi(1-\epsilon)$ for example. If this choice is made instead of the change in the infectivity the equilibrium analysis is identical. If both parameters are changed we

arrive at the second critical efficacy

$$\epsilon_c^{(2)} > 1 - \sqrt{\frac{\delta_S \gamma \delta_A}{\beta \alpha_S \pi f_L}} \tag{S14}$$

which is in general lower than the previous calculation (i.e., $\epsilon_c^{(2)} < \epsilon_c$), the critical efficacy provided by a drug only affecting the infectivity.

6.4 Stability analysis

By linearizing our system of equations, we can comment on the stability of the equilibrium points. We begin by defining our state variables as a vector $\mathbf{x} = [S, L, A, V]^T$ such that we can express the system of ODEs as $\mathbf{F}(\mathbf{x}) = \frac{d}{dt}\mathbf{x}$. We have calculated the equilibrium solutions above, and thus can Taylor expand our function around the equilibrium:

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}^*) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{\mathbf{x}^*} (\mathbf{x} - \mathbf{x}^*) + \left.\frac{\partial^2 \mathbf{F}}{\partial \mathbf{x}^2}\right|_{\mathbf{x}^*} (\mathbf{x} - \mathbf{x}^*)^2 + \dots$$

By construction the first term is zero, because it is the equilibrium point, and calling a small $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^* \ll 1$, then terms of $\mathcal{O}(\Delta\mathbf{x}^2)$ are negligible. Because the derivative function is linear and using $\mathbf{x} = \mathbf{x}^* + \Delta\mathbf{x}$ we can rewrite

$$\mathbf{F}(\mathbf{x}) = \frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}^*}{dt} + \frac{d\Delta\mathbf{x}}{dt}$$

and because $\mathbf{x}^*$ is not a function of time we have finally the linear equation

$$\frac{d}{dt}\Delta\mathbf{x} \approx \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{\mathbf{x}^*} \Delta\mathbf{x} = \mathbf{J}_{\mathbf{x}^*}\Delta\mathbf{x}$$

where the matrix of derivatives ($\frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \partial_i F_j$ where both $i, j \in [1, 4]$) is referred to as the Jacobian. The eigenvalues of the Jacobian describe how perturbations near equilibrium behave. Notably, if all eigenvalues λ_j of the Jacobian have negative real components $Re(\lambda_j) < 0$, perturbations decay back to equilibrium and the equilibrium is deemed stable. Importantly, a stable set-point equilibrium only exists when $\epsilon < \epsilon_c$. As mentioned in the body, for realistic parameters the critical efficacy is

$\epsilon_c \sim 85\%$. Above the critical efficacy of ART, no stable viral set-point equilibrium exists, and the viral-free equilibrium becomes stable.

Understanding how perturbations to the viral-free equilibrium behave gives insight into what happens when a small bit of virus is introduced. We evaluate the Jacobian at the viral-free equilibrium $\mathbf{x}^* = [\alpha_S/\delta_S, 0, 0, 0]^T$, leaving

$$\mathbf{J}_{\mathbf{x}^*} = \begin{pmatrix} -\delta_S & 0 & 0 & 0 \\ 0 & \theta_L & 0 & \tau\beta_\epsilon\alpha_S/\delta_S \\ 0 & \xi & -\delta_A & (1-\tau)\beta_\epsilon\alpha_S/\delta_S \\ 0 & 0 & \pi & -\gamma \end{pmatrix}.$$

We can see immediately that one of the eigenvalues is $-\delta_S$, and by noting that $\tau\beta_\epsilon\alpha_S/\delta_S$ is very small relative to the other rates, we see that another eigenvalue is θ_L. In fact, all eigenvalues (when calculated numerically) above the critical efficacy have negative real parts, so we know the equilibrium is stable. Furthermore, the timescale to return to equilibrium t^* is due to the largest (least negative) eigenvalue by $t^* = 1/\max\lambda_j$ [43]. Numerically, it turns out that this timescale is on the order of $1/\theta_L$, demonstrating that for the equilibrium to truly return to virus-free, all the latent cells must be cleared.

The mathematical analysis gives the condition for true eradication, and it is possible to make a heuristic argument for the dynamics of the latent cells. Soon after initiation of ART, virus only exists when long lived latent cells are activated, and these small activation events are quickly removed by ART. Furthermore, the few infections that do recur also have a minute probability of becoming latent (due to the magnitude of τ). All of this results in the observation that ongoing infection is ignorable in the latent reservoir [10]. We can now focus on the active and latent cells in the two equation system as shown and visualized in the bottom of Fig. 1 of the paper.

$$\begin{aligned} \dot{L} &= \alpha_L L - \delta_L L - \xi L \\ \dot{A} &= \xi L - \delta_A A \end{aligned} \tag{S15}$$

6.5 Solving the decoupled equations: the compound interest formula

The system Eq. S15 is linear and trivially solvable. Solving the equation using the initial number of latent cells L_0 yields

$$L = L_0 e^{(\alpha_L - \delta_L - \xi)t}$$

We will now define the total clearance rate $\theta_L = \alpha_L - \delta_L - \xi$ for the latent cells. This definition is important to make correspondence with the experimentally known parameter of the net clearance rate first measured by Siliciano *et al.* [2]. We have now

$$L = L_0 e^{\theta_L t} \tag{S16}$$

It is clear that when $\theta_L < 0$ the latent cell pool is cleared exponentially. Equivalently, if the latent cell proliferation rate α_L is greater than the sum of the reactivation rate ξ and the latent cell death rate δ_L, the latent reservoir L grows indefinitely. We call the exponential decay expression the **compound interest model** because of the correspondence with financial modeling. The principal investment is analogous to the initial size of the latent pool, and the interest rate analogous to the clearance rate of the latent pool.

The differential equation for the active cells is

$$\dot{A} + \delta_A A = \xi L_0 e^{\theta_L t}$$

with solution that can be computed by multiplying both sides by $\exp(\delta_A t)$ and identifying the total derivative $\partial_t[A \exp(\delta_A t)]$. We have then

$$A = \frac{\xi L_0}{\delta_A + \theta_L}\left(e^{\theta_L t} - e^{-\delta_A t}\right) + A_0 e^{-\delta_A t}$$

A plot of the trajectory of the latent pool with the drug efficacy above and below the critical efficacy is shown in Fig. S4. As expected by the stability analysis, above the critical efficacy the decoupled solution (the compound interest formula) matches the full solution perfectly. The full solution is achieved by solving the set of equations Eq. S10 using Matlab's `ode23s` solver.

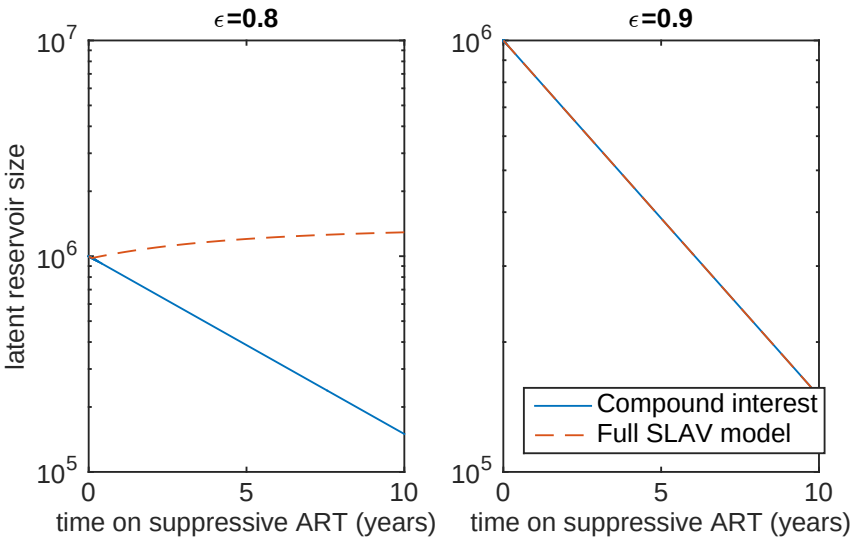

Figure S4: Plots of the latent pool clearance below and above the critical efficacy $\epsilon_c \sim 85\%$. Above the critical value, the compound interest formula $L = L_0 \exp(\theta_L t)$ matches the full ODE solutions exactly.

6.6 Impact of duration and rates on clearance time

It was claimed and demonstrated graphically in Fig. 2D of the body that it is better to increase the duration of therapy to deplete the latent pool than to increase the potency of the therapy equivalently. We can show this formally as well. If we examine the remaining fraction of latent cells over time

$$L/L_0 = \exp\left[(\alpha_L - \delta_L - \xi)t\right]$$

we can study the effect of multiplying any rate by a factor r or multiplying the duration of time by factor d. The ratio of the percent remaining after doing either of these multiplications tells us which procedure is more valuable. For example if we multiply the reactivation rate only

$$\frac{L^{(r)}}{L^{(d)}} = \exp\left[t(\alpha_L - \delta_L - r\xi - d(\alpha_L - \delta_L - \xi))\right]$$

then for the rate changed $L^{(r)}$ to be smaller than the duration changed $L^{(d)}$ we must have

$$\alpha_L - \delta_L - r\xi - d(\alpha_L - \delta_L - \xi) = (1-d)(\alpha_L - \delta_L) + \xi(d - r) < 0.$$

This means that the factor multiplying the rate must obey

$$r > \frac{(\alpha_L - \delta_L)(1-d) + \xi d}{\xi}$$

and note if we set $d = r$, we find that $d < 1$ because $(\alpha_L - \delta_L)$ is a negative number. Therefore, the decrease in the latent pool due to multiplying the duration of therapy will always be larger than an equivalent multiplication of the rate (unless the $r < 1$, which is a nonsensical proposition equivalent to making therapy less effective).

6.7 The minimal impact of allowing active to latent transitions

In some modeling works a transition from active to latent cells has been included [6, 44]. If we add that term to the model as a rate ϕ the equations for the latent and active pool (decoupled from the virus and susceptible as before) are

$$\begin{aligned} \dot{L} &= \alpha_L L - \delta_L L - \xi L + \phi A \\ \dot{A} &= \xi L - \delta_A A - \phi A. \end{aligned} \tag{S17}$$

These equations can be solved with typical linear algebraic methods, but the eigenvalue/eigenvector solutions are not intuitively enlightening. Instead we plot the solutions (of the "latency reversion" model) for varying values of ϕ in Fig. S5. The solution is compared with the compound interest model, and only for unnaturally large values of ϕ do the solutions deviate from the compound interest formula. For example, in Ref. [44], the largest values for the re-latent rate (fitting to patient data) is 0.078, well below the values that appear different from the compound interest model. In these simulations we use $L_0 = 10^6$ and $A_0 = 3 \times 10^5$ as in Hill *et al.* 2014. Only for large ϕ and at early points in time do we see any impact of the active cells. At that point, the active cells almost instantly transition to latency before the active cells decay. Then in the long time limit, the clearance is identical to the typical exponential decay. The result is that the solution is approximately $L = (L_0 + \phi A_0)\exp(\theta_L t)$. If A_0 is assumed to be larger the error in the model occurs for smaller latency reversion rate, but still above the expectation for ϕ. Indeed, this is not so much a problem for the model, because the initial conditions can be adjusted accordingly.

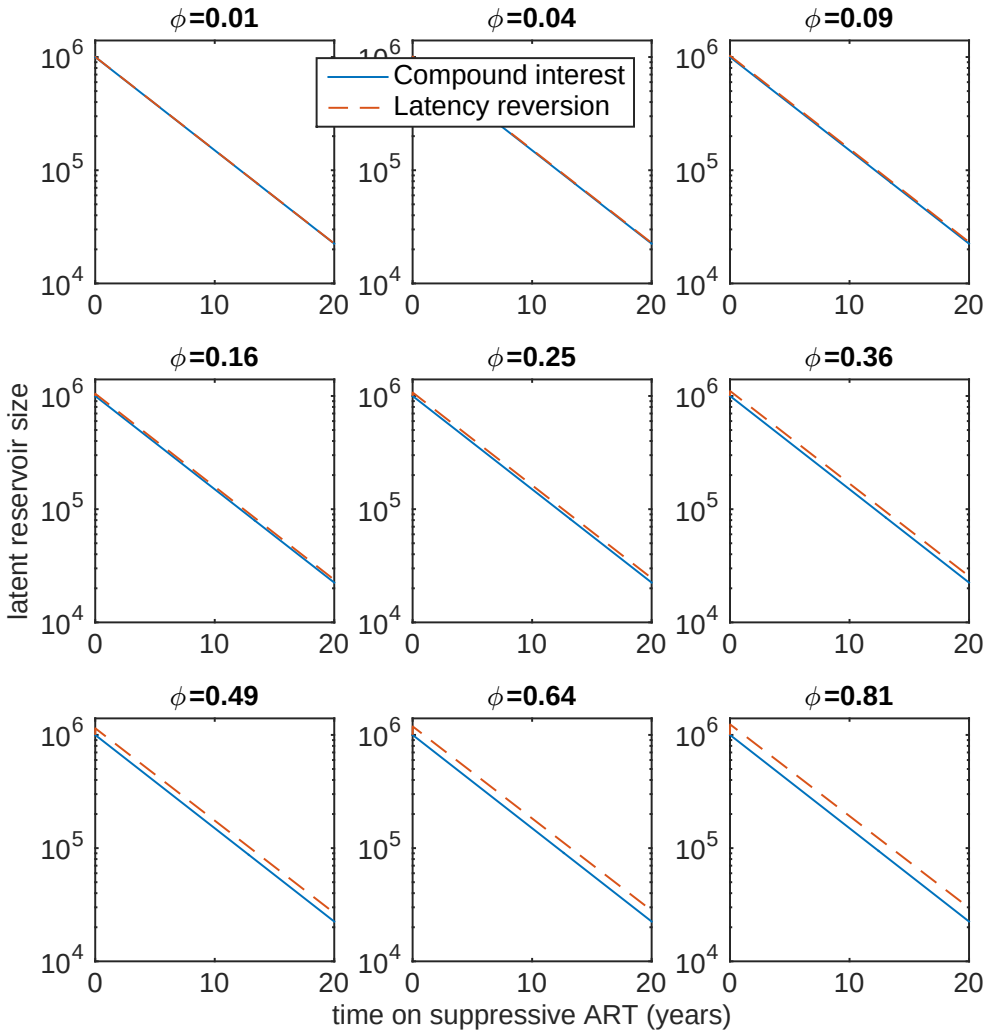

Figure S5: Solutions of Eq. S18 compared to the compound interest model $L = L_0 \exp(\theta_L t)$. Only for unrealistically large values of the return to latency rate ϕ does the solution deviate from the exponential decay.

6.8 Composition of the reservoir: modeling T-cell subsets

It has been demonstrated that several subsets of $CD4^+$ T cells may contribute to the reservoir [19, 32, 33]. We reviewed the existing data to find the rates of proliferation for each subset, as well as estimates of the total fractional makeup of the reservoir.

The typical T cell phenotypes considered to make up the reservoir are central memory T_{cm}, effector memory T_{em}, and naïve T_n cells. Transitional memory T_{tm} have also been described to represent a

transition from $T_{cm} \rightarrow T_{em}$. For the model we consider T_{tm} to have equivalent proliferation rates to T_{cm}. Similarly, we characterize newly recognized stem-cell-like memory CD4$^+$T cells (T_{scm}) as T_n given their likely slow turnover rate. Of note, these are conservative estimates that would not favor anti-proliferative therapy.

We include the T cell diversity into the decoupled model by breaking the differential equation for the latent pool down into three differential equations, one for each subtype L_i with $i \in [cm, em, n]$. We ignore transitions between types because on average the composition of the reservoir is reasonably stable over time and thus our model is the system

$$\dot{L}_i = \theta_i L_i. \tag{S18}$$

The total number of latent cells is the sum of the subset numbers, $L = \sum_i L_i$. We only use three types mentioned above (including stem-cell-like and transitional memory cells into naïve and central memory sets respectively). Solutions for the total latent pool when decoupled on ART are thus linear combinations

$$L(t) = \sum_i L_i(0) e^{\theta_i t} \tag{S19}$$

where $\theta_i = \alpha_i - \delta_i - \xi$, $L_i(0)$ are the initial numbers of each subtype, and because we have assumed there are no couplings between types.

As in Fig. 3 in the main body, simulations assume the same original net clearance rate and activation rates, but different proliferation rates α_i and thus different death rates δ_i. When the proliferation rate is reduced using an anti-proliferative therapy, the clearance rates are different among types. The initial conditions for each type $L_i(0)$ are inferred from Chomont *et al.* [19], though we do simulate larger values than this to cover the worst case scenario of many naïve cells as in Buzon *et al.* [33].

7 Appendix II: Discussion of model parameters

7.1 The components of θ_L: proliferation, death, and activation

Our most critical parameter values are those that describe the proliferation, death, and activation rates of the latent cell pool–the sum of which is the the net clearance rate of the latent cell pool θ_L. We estimate the proliferation rate α_L from Macallan *et al.* who used *in vivo* labeling with deuterated glucose to measure turnover in memory T cells in humans and found that on average, 1.5% of CD45R0$^+$CCR7$^+$ T cells (central memory) proliferate daily, whereas 4.7% of CD45R0$^+$CCR7$^-$ T cells (effector memory) and 0.2% of CD45R0$^-$CCR7$^+$ T cells (naïve) proliferate per day, corresponding with exponential growth rates of $\alpha_i = 0.015$, 0.047, and 0.002, respectively [18]. We initially describe the model assuming that latently infected CD4$^+$ T cells are in fact mostly central memory T cells but then go on to generalize the results to different fractional makeups of the reservoir.

Despite the fact that in the past, most mathematical models of HIV latency have not included this proliferation parameter α_L, Rong *et al.* demonstrated that incorporating a homeostatic proliferation rate allowed them to "describe the multi-phasic viral decline following initiation of antiretroviral treatment and maintain both low-level persistent viremia and the latent reservoir during therapy." In their models, α_L parameter values are chosen between 0.011 and 0.03 per day, consistent with our value of 0.015 [6].

The activation rate ξ is several orders of magnitude smaller than the proliferation rate. Using parameters calculated by Luo *et al.* from structured treatment interruptions, Hill *et al.* found that on average, 57 CD4$^+$T cells per day transition from latency to the activated state [14, 8]. Assuming that the reservoir contains 1 million cells on average, we find $\xi = 5.7 \times 10^{-5}$ is the rate of reactivation of a single cell per day [8], i.e. 57 cells/(10^6 cells $\times$ day) = 5.7 $\times 10^{-5}$ per day.

The half-life estimate of the latent pool from Siliciano *et al.*'s quantitative viral outgrowth assay yields the total clearance rate of latent cells, θ_L = -5.2 $\times 10^{-4}$ per day [2]. This result was corroborated closely by Crooks *et al.* [17]. Thus, the clearance, proliferation, and activations rates are derived from human experiments. However, the death rate is calculated. The total clearance

rate of the latent pool is the death and activation rates subtracted from the proliferation rate. A rearrangement of this equation (using central memory cell values) gives the daily death rate $\delta_L = \alpha_L - \xi - \theta_L = 0.0155$.

A large body of work has been reviewed in de Boer and Perelson on the variety of experimental techniques employed to measure the turnover of T cells [40]. The modeling and experimental challenges are discussed, and we include the reference here to emphasize how much uncertainty in the parameters indeed exists. The uncertainty and sensitivity analysis in §3 are included to demonstrate our awareness of the uncertainty and provide theoretical ranges for plausible outcomes given the varying measurements of these rates.

7.2 Fractional makeup of the reservoir

Chomont *et al.* determined the percentage of each cellular subset from 31 aviremic individuals. They found mean contributions of $T_{cm} + T_{tm} \approx 85\%$, whereas $T_{em} \approx 13\%$ and $T_n \approx 2\%$ [19]. Buzon *et al.* 2014 found higher percentages of naïve and stem-cell-like memory cells, up to 40% [33]. Thus in the model inclusive of diversity we set the initial conditions of each subtype proportional to their fractional makeup, e.g. $L_i(0) = [85, 13, 2]$ for the example above.

7.3 Susceptible and activated cell dynamics: α_S, δ_S, α_A, δ_A

The production of CD4$^+$T cells from the bone marrow and thymus is described by α_S; and δ_S is the rate of susceptible T cell death. In several of the early HIV modeling papers, a value of 10 per μL-day was estimated for the production rate[39], [45] with δ_S estimated at 0.02.

Luo *et al.* used a Bayesian Markov-Chain Monte-Carlo method to estimate HIV model parameters [14]. They used data from 10 patients who underwent a series of 30-day ART treatment interruptions with viral loads taken three times weekly following interruptions and then weekly following initiation of treatment. Each patient underwent 3-5 interruption/treatment cycles. They estimated $\alpha_S = 295$, $\delta_S = 0.18$, and $\delta_A = 1$. Huang *et al.* also use Bayesian methods (Markov Chain Monte Carlo) and fit their model to data from [46], an AIDS clinical trial comparing dosing regimens for indinavir and ritonavir [47]. This model also incorporated adherence, drug concentrations, and

drug susceptibilities [47]. They find $\alpha_S = 98.1$, $\delta_S = 0.08$, and $\delta_A = 0.37$. Note in Table S3 that the ratio of α_S to δ_S varies between 500 - 1639 with the more recent experiments in relative agreement near 1500. Thus, we chose $\alpha_S = 300$, $\delta_S = 0.2$ to reflect this ratio.

The death rate of productively infected cells δ_A was initially thought to be 0.24 [39]. Later, Perelson *et al.* find $\delta_A = 0.5$ based on frequent sampling of 5 patients after giving ritonavir monotherapy [7]. Using a fixed viral clearance rate $\gamma = 23$ from [48], Markowitz *et al.* determined $\delta_A = 1$ based on potent antiretroviral therapy with lopinavir-ritonavir, tenofovir, lamivudine, and efavirenz in 5 chronically-infected patients [49]. We chose $\delta_A = 1.0$ to reflect more recent experimental findings. We assign $\alpha_A = 0$, as the relative rate of proliferation of actively infected cells likely occurs at neglible rates compared to the death rate of these cells [39].

Parameter	Perelson [39]	Huang [47]	Luo [14]	Markowitz [49]	Units
α_S	10	98.1	295	-	per μL-day
δ_S	0.02	0.08	0.18	-	per day
α_S/δ_S	500	1226	1639	-	per day
δ_A	0.24	0.37	1.0	1.0	per day

Table S3: Proliferation and death parameters for susceptible and actively infected cells estimated in prior experiments and mathematical models.

7.4 Estimating the infectivity β

Perelson *et al.* calculate the infectivity of a virus β with Smoluchowski's formula for the diffusion-limited rate constant for two spherical particles to estimate the probability of cell-virus contact, multiplied by *in vitro* experimental values for the probability of viral attachment to cells and consequent infectivity and notes "one might safely assume that $k1$ is a strain-dependent parameter that can vary greatly," where $k1$ is the parameter used to represent infectivity in that model. They find $\beta = 2.4 \times 10^{-5} \mu$L per day [39]. This value is often referenced in other works. Using Bayesian methods, both Luo *et al.* and Huang *et al.* estimate parameter values for β [14, 47] as $3.9{\times}10^{-3}$ and $1.7{\times}10^{-5}$, respectively. Given that the range of these values covers two orders of magnitude, we chose a value between these extremes: 10^{-4}.

Parameter	Perelson [39]	Huang [47]	Luo [14]	Units
β	2.4×10^{-5}	1.7×10^{-5}	3.9×10^{-3}	μL per day

Table S4: Infectivity as estimated in prior experiments and mathematical models.

7.5 Viral dynamics

The viral 'burst rate' π is the amount of virus a single actively infected cell emits in a day (roughly its lifetime). Rong and Perelson note that π is problematic, as its experimental value is under question and affects the values of many of the other parameters [6]. Haase *et al.* used radioactively-labeled RNA probes and quantitative image analysis to determine the number of viral particles per mononuclear cell in biopsies from fixed lymph tissue and found a mean value of 74 [50]. Using quantitative, competitive, real-time PCR (QT-RT- PCR) to measure the mean viral RNA copy number per infected cell from fresh-frozen cervical lymph nodes from 9 HIV patients with varied viral loads, Hockett *et al.* found $\pi = 10^{3.6} = 4 \times 10^3$ [51]. Both of these estimates were obtained experimentally rather than derived from models; however, they do not necessarily reflect the number of copies produced in a day per cell or in a cell's lifetime, rather the amount that was being produced at the instant the experiments were performed. In any case, the value of π appears to be in the neighborhood of 10^3, which is the value we will adopt. For the viral clearance rate γ, we use Ramratnam *et al.*'s estimate $\gamma = 23$, obtained from viral load measurements taken over 5 days before, during, and after apheresis in 4 patients assuming a constant rate of viral production [48].

Parameter	Perelson [39]	Huang [47]	Luo [14]	Units
π	1200	976	5.9×10^3	copies/cell-day
γ	2.4	3.06	18.8	per day
π/γ	500	319	314	copies/cell

Table S5: Viral burst and clearance rates as estimated in prior experiments and mathematical models. Note: [14] used the data from [48] but used the geometric mean rather than the arithmetic mean, i.e. 18.8 rather than 23.

The 'latent cell fraction' τ is the rate at which newly infected cells join the latent cell pool, whereas $1-\tau$ is the rate at which they join the actively infected pool. There are few estimates available for this parameter, and the estimates vary across a wide range. Despite this, the choice of τ within the given range does not affect our cure estimates or our estimate of the critical epsilon (ϵ_c). Because

their estimates are based on modern measurements of the reservoir, we chose Conway and Perelson's value $\tau = 10^{-4}$, which is the upper bound on their estimates ranging from 10^{-7} to 10^{-4} [10].

Parameter	Callaway [45]	Jones [52]	Conway [10]	Units
τ	10^{-6}	10^{-3}	10^{-4}	unitless

Table S6: Latency fraction as estimated in prior experiments and mathematical models.

8 Appendix III: Further discussion of anti-proliferative therapy from a mathematical and clinical standpoint

8.1 Complete parameter table, sensitivity analysis and uncertainty analysis

The previous section reported how parameter values were found from the literature. Here we summarize the results and include associated confidence intervals or ranges found in the literature where parameter distributions are unspecified.

Working within the intervals given in Table S7 we first conduct a local uncertainty analysis [53] that shows how anti-proliferative therapy coupled with ART depletes the latent reservoir. First, we fix each parameters to its typical value and vary one parameter at a time within its bounds as given in Table S7 and with the constraint that $\mathcal{R}_0^{ART} < 1$. The results are shown in Fig. S6. We see that the variables affecting the basic reproductive number (here for example τ, ϵ^{ART}) have no effect on the time to cure. As we have shown, if $\mathcal{R}_0^{ART} < 1$, the dynamics of the latent reservoir effectively decouple from the viral dynamics. The original number of cells in the reservoir has a large impact on the time to cure. Additionally, as discussed in the main body, the makeup of the reservoir impacts the clearance: a higher percentage of naïve cells slows decay and extends the time to cure. Also, the value of the clearance rate changes time to cure as expected by the compound interest framework. However, the expected range of this value is not large, so our results are not drastically changed by this variation.

We also complete a global uncertainty and sensitivity analysis in which all variables for the model are drawn from uniform distributions based on the ranges in Table S7 and sampled using Latin Hypercube sampling [53]. Any variable whose range spans more than one order of magnitude is

Param.	Value	(CI)/[Literature Range]	Dimensions	Source
α_S	300	[10, 760]	cells per μL-day	[14, 39, 47]
δ_S	0.2	[0.02, 0.45]	per day	[14, 39, 47]
α_{cm}	0.015	(0.01, 0.02)	per day	[18]
δ_{cm}	0.0155*		per day	calculated*
α_{em}	0.047	(0.038, 0.057)	per day	[18]
δ_{em}	0.0475*		per day	calculated*
α_n	0.002	(0.0015, 0.0023)	per day	[18]
δ_n	0.0025*		per day	calculated*
δ_A	1.0	$(0.8, 1.2)$	per day	[14, 49]
L_0	10^6	$[10^4, 10^8]$	cells	[2]
$L_{cm}(0)$	$0.6L_0$	$[0.2, 0.8]L_0$	cells	[19, 33]
$L_{em}(0)$	$0.4L_0$	$[0.2, 0.8]L_0$	cells	[19, 33]
$L_n(0)$	$0.02L_0$	$[0, 0.1]L_0$	cells	[19, 33]
ξ	5.7×10^{-5}	$(5.4, 6.0) \times 10^{-5}$	per day	[8]
θ_L	-5.2×10^{-4}	$-(2, 8.4) \times 10^{-4}$	per day	[2]
β	1×10^{-4}	$[0.016, 6] \times 10^{-3}$	μL/copy-day	[14, 39, 47]
τ	10^{-4}	$[10^{-6}, 10^{-3}]$	unitless	[10, 52, 45]
π	10^3	[20, 5340]	copies/cell-day	[14, 39, 47]
γ	23	[9, 36]	per day	[48]

Table S7: A summary of all parameters used in our simulations. $^*\delta_L$ is back-calculated from known α_L, θ_L, and ξ. 95% confidence intervals are given in parentheses () where applicable from experimental data. Otherwise, the range is taken from our literature search or from the ranges given in the cited works not assumed to be normally distributed; these values are given in square brackets [].

sampled to achieve even sampling on a logarithmic scale. The simulations are carried out in Matlab using `lhsdesign` and `ode23s`. We note that a stochastic model is not necessary because we are studying decay dynamics with large numbers of cells as opposed to rebound dynamics.

The simulations from the global uncertainty analysis are used in a sensitivity analysis to correlate each parameter with time to cure on ART and anti-proliferative therapy. Time to cure is defined by the number of latent cells as compared to the thresholds described in the main body. We calculate correlation with the the Pearson correlation coefficient $\rho_{T,p}$, the ratio of the covariance of the time to cure T with the parameter of interest p normalized by each's standard deviation:

$$\rho_{T,p} = \frac{\langle (T - \langle T \rangle)(p - \langle p \rangle) \rangle}{[(\langle T^2 \rangle - \langle T \rangle^2)(\langle p^2 \rangle - \langle p \rangle^2)]^{1/2}},$$

where $\langle \cdot \rangle$ indicates the expectation value. The results of the simulations are shown in Fig. S7.

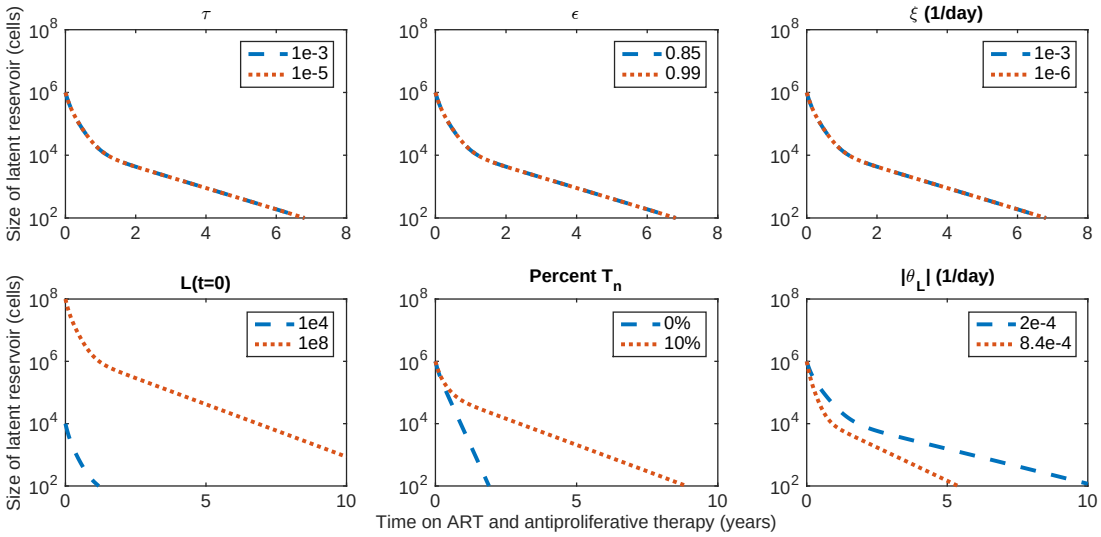

Figure S6: Local uncertainty analysis of the complete model, Eq. S1 where ART is coupled with anti-proliferative therapy (potency $\epsilon^{AP} = 5$) to analyze the variability of solutions and time to cure. Note that the model includes varying phenotypes of T cells with varying proliferation rates as in Table S7. Each parameter that is varied is listed above the panel with the range of variability shown in the panel legend, and all other values are constant as in the Value column of Table S7.

One thousand simulations are carried out and only the parameter ranges that obey $\mathcal{R}_0^{ART} < 1$ are considered. The variables are shown to correlate positively, negatively, or not at all with time to cure. As expected, as the initial size of the latent pool increases, the time to cure goes up, and as the anti-proliferative therapy potency increases, the time to cure goes down. The percentage of the latent reservoir that are naïve T cells delays the time to cure while a higher magnitude decay rate hastens cure. The probability of latency and $\mathcal{R}_0^{ART}$ do not affect time to cure strongly.

8.2 Comparison with latent pool reduction from Chapuis *et al.* 2000

Mycophenolate mofetil (MMF) is a clinically approved and commonly used drug for organ transplantation because it selectively inhibits the proliferation of lymphocytes. In 2000, Chapuis *et al.* administered MMF along with ART to 6 HIV-infected patients for 24 weeks. Three study participants demonstrated a 1-2 order of magnitude reduction in reservoir size by quantitative viral outgrowth assay. The authors concluded that MMF for 'the treatment of HIV infection deserves further investigation in controlled clinical trials'. In Fig. 2b of their paper, the potency of anti-proliferative medicine (ϵ^{AP}) follows a dose-response, but can exceed 10 which is well above the

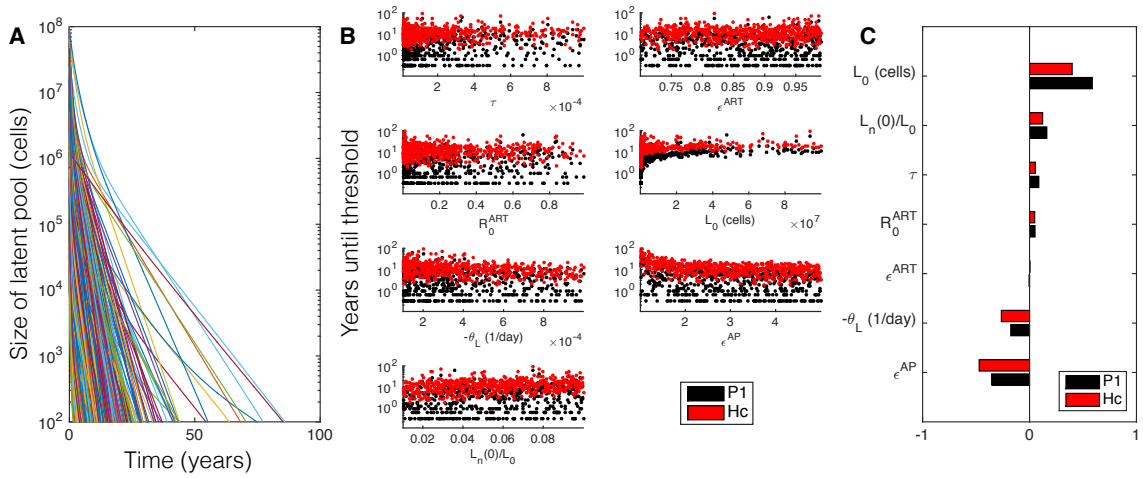

Figure S7: Global uncertainty and sensitivity analysis using the ranges of parameters from Table S7. A) The 1,000 simulations are shown to give an idea of the variance over the whole parameter range. B) The time until each cure threshold, Pinkevych 1 yr and Hill cure, are calculated as the time when the latent reservoir contains fewer than 20,000 and 200 cells respectively. In some cases cures are achieved within months, in others, many years. C) Pearson correlation coefficients indicate which variables correlate with time to cure.

saturation point where duration is more important than potency in Fig. 2D of our paper. Assuming an $\epsilon^{AP} > 3$, our model predicts a 1 to 2 order of magnitude decrease in the reservoir after 24 weeks of MMF, which is in keeping with Fig. 6a of their paper.

We suspect that the anti-proliferative drugs are not successful in all patients, so that the 3 patients who had no reservoir reduction may not have responded well to MMF. The suspicion is corroborated in the next subsection by the fact that the authors of a separate MMF study had to rearrange their clinical cohorts to deal with participants who received MMF but showed no anti-proliferative effect.

8.3 Comparison with time to rebound from García *et al.* 2004

In 2004, García *et al.* also assessed the effect of MMF in treatment interruption studies of ART [28]. In the study, seventeen HIV patients received ART for a year. They were then randomized into two groups, the control group that remained on ART only, and the experimental group that also received MMF for 17 weeks. After this time, ART was interrupted in both groups and viral rebound was measured. The benchmark to compare the two groups was to be the number of

individuals who maintained a viral load set-point below 200 copies/mL. 5/9 in the experimental group and 1/6 in the control group met this standard, though this was not statistically significant. The MMF recipients were subsequently reclassified into a group of the 6 that successfully inhibited proliferation (as measured by an *in vitro* assay) and a group of 3 with no clear inhibition. The viral load set-point measurement was indeed significantly different between these two groups, as was the time to rebound. The authors provide two mechanisms by which they expect MMF to control HIV infection: 1) antiviral in that MMF may deplete a substrate necessary for reverse transcription, and 2) immunologic in that the number of target cells is reduced.

Previous studies allow us to estimate the possible decrease in reservoir size due to MMF or similarly potent anti-proliferative therapy in general. Examining the patient data of García *et al.* (Fig. 1 of Ref. [28]), we see that the times to achieving a viral load above the typical detectable limit (10-50 copies per mL) is 1-4 weeks for the control group and 6-12 weeks for the MMF responder group. We can compare this experimental data with several theoretical predictions for the fold decrease in the latent reservoir. Using the results of Pinkevych *et al.*, a median time-to-detection of roughly 7 weeks (see Fig. 5B of Ref. [21]) corresponds to a 7-fold decrease in the latent reservoir. Using the results of Hill *et al.* the same median time-to-detection of roughly 7 weeks (see Fig. 4 of Ref. [8]) corresponds to a 50-fold reduction in the latent reservoir. The result of Hill matches well with our calculation. Because MMF was given to patients in the study for 17 weeks, and we assume the potency of MMF is above a factor of 2 ($\epsilon^{AP} > 2$) [29], our simulations suggest that the size of the latent reservoir would be reduced by approximately a factor of 40 (visualized in Fig. 1D in the main body).

8.4 Bone marrow/thymic production of T cells and the safety of MMF as an anti-proliferative therapy

There is some debate about how the susceptible cells are sustained [54, 55, 56]. In our model we assumed that bone marrow or thymic production is the mechanism that produces T cells. However, density-dependent growth could also contribute to the production of T cells. To account for this possibility, we and other modelers consider this effect to be grouped into the density-dependent death rate δ_S. Thus, assuming that anti-proliferative drugs have no impact on thymic/bone marrow

production, the total number of T cells should stay roughly constant. To be precise, the equilibrium solution for the susceptible cells $S^* = \alpha_S/\delta_S$ depends only on bone marrow/thymic production and death rate. Perhaps more important than our theoretical perspective, experimental results from studies discussed in the previous 2 subsections show that the total number of $CD4^+$ and $CD8^+$ T cells stayed roughly constant even as the number of proliferating cells decreased on MMF [29, 28].

Beyond the stability of $CD4^+$T cell counts, the safety of MMF should be vetted in effective dosages in HIV patients. Again, in the Chapuis and Garciái studies, no opportunistic infections or adverse events were observed, but these were performed in small populations for relatively limited durations [29, 28]. On the other hand, MMF has been given in many patients at higher dosages and for much longer durations in autoimmune disease and solid organ transplantation though much of this data is difficult to interpret given the usual, concurrent administration of steroids, calcineurin-inhibitors, and/or cyclosporine [35]. However, Kriss *et al.* report retrospectively on 23 liver transplant patients who were transitioned to MMF monotherapy after a median of 2 years on calcineurin-inhibitors +/- MMF. Following conversion to MMF monotherapy, patients were followed for an average of 4 years. During that time, one patient discontinued MMF due to persistent diarrhea; one required dose-reduction due to anemia; one developed zoster with resolution on acyclovir; and one developed post-transplant lymphoproliferative disorder (PTLD) but achieved complete remission with MMF dose-reduction and R-CHOP chemotherapy. Thus, the incidence of adverse events over 4 years was small; but if the case of PTLD was caused by the MMF, the consequences are too severe to merit treatment in HIV patients controlled on ART. The median dose in this trial was 2g/d, the same dose given in [29] and twice the dose given in [28]. Questions of causation and relatedness to dosage remain to be answered, but we propose worthy to be attempted with low doses and careful monitoring.

As a final note, it is also possible that the number of divisions resulting in cell proliferation is finite. This upper bound on rounds of proliferation is called the "Hayflick limit." This limit is unlikely to apply to our model, as destructive telomere shortening is not as relevant for T cells as other cell types. See Ref. [57] for further discussion and references.

References

[1] Richman DD et al. The challenge of finding a cure for HIV infection. *Science*, 323(5919):1304–1307, 2009.

[2] Siliciano JD et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nature Med*, 9(6):727–728, 2003.

[3] Finzi D et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature Med*, 5(5):512–517, 1999.

[4] Martin AR and Siliciano RF. Progress toward HIV eradication: Case reports, current efforts, and the challenges associated with cure. *Ann Rev Med*, 67(1):011514–023043, 2016.

[5] Hütter G et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *NEJM*, 360(7):692–698, 2009.

[6] Rong L and Perelson AS. Modeling latently infected cell activation: Viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Comp Bio*, 5(10):e1000533, 2009.

[7] Perelson AS, Neumann AU, Markowitz M, Leonard JM, and Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science (New York)*, 271(5255):1582–1586, 1996.

[8] Hill AL, Rosenbloom DS, Fu F, Nowak MA, and Siliciano RF. Predicting the outcomes of treatment to eradicate the latent reservoir for HIV-1. *Proc Nat Acad Sci*, 111(43):15597–15597, 2014.

[9] Rouzine IM, Weinberger AD, and Weinberger LS. An evolutionary role for HIV latency in enhancing viral transmission. *Cell*, 160(5):1002–12, 2015.

[10] Conway JM and Perelson AS. Residual Viremia in Treated HIV+ Individuals. *PLoS Comp Bio*, 12(1):e1004677, 2016.

[11] Kürten KE and Castiglione F. A dynamical model of b–t cell regulation. *Intl J Mod Phys C*, 12(03):367–375, 2001.

[12] Palmer S et al. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proc Nat Acad Sci*, 105(10):3879–3884, 2008.

[13] Shen L et al. A critical subset model provides a conceptual basis for the high antiviral activity of major HIV drugs. *Science Trans Med*, 3(91):91ra63, 2011.

[14] Luo R, Piovoso MJ, Martinez-Picado J, and Zurakowski R. HIV model parameter estimates from interruption trial data including drug efficacy and reservoir dynamics. *PLoS ONE*, 7(7), 2012.

[15] Bull ME et al. Monotypic HIV-1 genotypes across the uterine cervix and in blood suggest proliferation of cells with provirus. *JVI*, 83(12):6020, 2009.

[16] von Stockenstrom S et al. Longitudinal genetic characterization reveals that cell proliferation maintains a persistent hiv-1 DNA pool during effective HIV therapy. *JID*, 1, 2015.

[17] Crooks AM et al. Precise quantitation of the latent HIV-1 reservoir: implications for eradication strategies. *JID*, 212(9):1361–1365, 2015.

[18] Macallan DC et al. Rapid turnover of effector-memory CD4+ T cells in healthy humans. *J Exp Med*, 200(2):255–260, 2004.

[19] Chomont N et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature Med*, 15(8):893–900, 2009.

[20] Huang Y, Wu H, and Acosta E. Hierarchical Bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors. *Biomed J*, 52(4):470–486, 2010.

[21] Pinkevych M et al. HIV Reactivation from Latency after Treatment Interruption Occurs on Average Every 5-8 Days? Implications for HIV Remission. *PLoS Pathog*, 11(7):e1005000, 2015.

[22] Archin NM and Margolis DM. Emerging strategies to deplete the HIV reservoir. *Curr Opin Infect Dis*, 27(1):29–35, 2014.

[23] Fuller DH et. al. Therapeutic DNA vaccine induces broad T cell responses in the gut and sustained protection from viral rebound and AIDS in SIV-infected rhesus macaques. *PloS ONE*, 7(3):e33715, 2012.

[24] Aubert M et al. Successful targeting and disruption of an integrated reporter lentivirus using the engineered homing endonuclease Y2 I-AniI. *PloS ONE*, 6(2):e16825, 2011.

[25] Peterson CW, Younan P, Jerome KR, and Kiem H-P. Combinatorial anti-HIV gene therapy: using a multipronged approach to reach beyond HAART. *Gene Ther*, 20(7):695–702, 2013.

[26] Wagner TA et al. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science*, 345(6196):570–573, 2014.

[27] Maldarelli F et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*, 345(6193):179–83, jul 2014.

[28] García F et al. Effect of mycophenolate mofetil on immune response and plasma and lymphatic tissue viral load during and after interruption of highly active antiretroviral therapy for patients with chronic HIV infection: a randomized pilot study. *JAIDS*, 36(3):823–30, 2004.

[29] Chapuis A et al. Effects of mycophenolic acid on human immunodeficiency virus infection in vitro and in vivo. *Nat Med*, 6(7):762–768, 2000.

[30] Zala C, Rouleau D, and Montaner JS. Role of hydroxyurea in treatment of disease due to human immunodeficiency virus infection. *Clin Inf Dis*, 30 Suppl 2:S143–50, 2000.

[31] Hladik F. A new perspective on HIV cure. *F1000*, 77:0–4, 2014.

[32] Jaafoura S et al. Progressive contraction of the latent HIV reservoir around a core of less-differentiated CD4+ memory T cells. *Nature Comm*, 5:5407, 2014.

[33] Buzon MJ et al. HIV-1 persistence in CD4+ T cells with stem cell-like properties. *Nature Med*, 20(2):139–142, 2014.

[34] Ross D et al. Safety and efficacy of imatinib cessation for CML patients with stable undetectable minimal residual disease: results from the TWISTER study. *Blood*, 122(4):515–522, 2013.

[35] Mok CC. Mycophenolate mofetil for lupus nephritis: an update. *Exp Rev Clin Immunol*, 11(12):1353–1364, 2015.

[36] Hladik F et al. Mucosal effects of tenofovir 1% gel. *eLife*, 4:1–21, 2015.

[37] Margolis DA et al. Cabotegravir plus rilpivirine, once a day, after induction with cabotegravir plus nucleoside reverse transcriptase inhibitors in antiretroviral-naive adults with HIV-1 infection (LATTE): a randomised, phase 2b, dose-ranging trial. *Lancet*, 15(10):1145–1155, 2015.

[38] Perelson AS and Nelson PW. Mathematical Analysis of HIV-I : Dynamics in Vivo. *SIAM Rev*, 41(1):3–44, 2009.

[39] Perelson AS, Kirschner DE, and De Boer R. Dynamics of HIV infection of CD4+ T cells. *Math Biosci*, 114(1):81–125, 1993.

[40] De Boer RJ and Perelson AS. Quantifying T lymphocyte turnover. *J Theor Biol*, 327:45–87, 2013.

[41] Kim H and Perelson AS. Viral and latent reservoir persistence in hiv-1–infected patients on therapy. *PLoS Comp Biol*, 2(10):e135, 2006.

[42] Diekmann O, Heesterbeek JP, and Roberts MG. The construction of next-generation matrices for compartmental epidemic models. *J Roc Soc Interface*, 7:873–885, 2010.

[43] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.

[44] Ribeiro RM, Mohri H, Ho DD, and Perelson AS. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proc Nat Acad Sci*, 99(24):15572–15577, 2002.

[45] Callaway DS and Perelson AS. HIV-1 infection and low steady state viral loads. *Bull Math Bio*, 64(1):29–64, 2002.

[46] Acosta EP et al. Comparison of two indinavir/ritonavir regimens in the treatment of HIV-infected individuals. *JAIDS*, 37(3):1358–1366, 2004.

[47] Huang Y, Liu D, and Wu H. Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biomet*, 62(2):413–423, 2006.

[48] Ramratnam B et al. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet*, 354(9192):1782–1785, 1999.

[49] Markowitz M et al. A novel antiviral intervention results in more accurate assessment of HIV-1 replication dynamics and t-cell decay in vivo. *JVI*, 77(8):5037–5038, 2003.

[50] Haase AT et al. Quantitative image analysis of HIV-1 infection in lymphoid tissue. *Science*, 274(5289):985–989, 1996.

[51] Hockett RD et al. Constant mean viral copy number per infected cell in tissues regardless of high, low, or undetectable plasma HIV RNA. *J Exp Med*, 189(10):1545–1554, 1999.

[52] Jones LE and Perelson AS. Transient viremia, plasma viral load, and reservoir replenishment in HIV-infected patients on antiretroviral therapy. *JAIDS*, 45(5):483–93, 2007.

[53] Marino S, Hogue IB, Ray CJ, and Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J Theor Biol*, 254(1):178–196, 2008.

[54] Macallan DC et al. Measurement and modeling of human t cell kinetics. *Euro J Immunol*, 33(8):2316–2326, 2003.

[55] Douek DC et al. Assessment of thymic output in adults after haematopoietic stem-cell transplantation and prediction of T-cell reconstitution. *Lancet*, 355(9218):1875–1881, 2000.

[56] Franco JM et al. T-cell repopulation and thymic volume in HIV-1-infected adult patients after highly active antiretroviral therapy. *Blood*, 99(10):3702–3706, may 2002.

[57] De Boer RJ and Noest AJ. T cell renewal rates, telomerase, and telomere length shortening. *J Immunol*, 160(12):5832–5837, 1998.