# GBIO0033 - Advances in In Silico Medicine

Brandoit Julien - s200710

Academic Years 2024 - 2025

## Contents

# List of Figures

# List of Tables



# List of Boxes

# 1 Introduction

## 1.1 Context and importance of in silico medicine

In Silico Medicine represents a major breakthrough in the biomedical field, with promises that could profoundly transform research and development in healthcare. It harnesses the power of digital technologies to simulate complex biological processes, accelerate drug and medical device development, reduce costs, and limit reliance on animal models and human clinical trials. It also offers a unique opportunity for treatment personalization by integrating patient-specific data to tailor therapeutic approaches. These perspectives, combined with the ability to better understand fundamental biological mechanisms, pave the way for more effective, ethical, and sustainable medicine.

However, it is essential to clarify that these promises do not aim to replace in vivo and in vitro research, but rather to complement them. The goal is to provide additional evidence that, when integrated with results from experimental approaches, can improve and inform decision-making. In silico models offer a simulation environment where hypotheses can be posed, scenarios explored, and uncertainties reduced before proceeding to the more costly and complex stages of traditional trials.

A key aspect of this approach lies in the clear definition of a context of use. The goal is not to create models that perfectly represent the complexity of human biology – an unrealistic and unnecessary task, in my view. Instead, the objective is to design models that are "*wrong but useful*", suited to specific questions and designed to provide relevant answers within a defined framework. This pragmatic, utility-focused approach is crucial to maximizing the impact of digital models.

Thus, for In Silico Medicine to become a truly credible and integrated tool in healthcare decision-making, it is imperative to apply rigorous principles of Verification, Validation, and Uncertainty Quantification (VVUQ). While regulatory bodies have yet to define strict standards for the use of in silico models, the VVUQ methodology is an essential element in transforming these models into **real scientific evidence**. Verification and validation ensure that models are technically and scientifically robust, while uncertainty analysis quantifies margins of error and the predictive limits of the models. Without this quality assurance process, even the most sophisticated models may fail to be accepted as credible evidence by decision-makers, thus hindering their adoption in the decision-making process.

The development of digital models will only be valuable if they can actually be used in practice, particularly to inform therapeutic choices and medical strategies. Researchers must, therefore, focus not only on building increasingly accurate models but also take the time to perform the necessary VVUQ analyses to validate the usefulness of these models. This includes evaluating associated uncertainties, highlighting model limitations, and clearly documenting the context of use. Without this methodical approach, models risk remaining theoretical tools with no tangible impact on medical decisions.

Finally, the definition of the context of use should be seen as a central and preliminary element in the construction of any in silico model. Similar to clinical studies, the precise delimitation of the model's objective, parameters, and limitations should be done at the very beginning of any project. In an ideal world, this reflection should take place well before even the model's development to ensure it is designed to answer specific and useful questions within defined medical contexts. This process is indispensable to ensure that in silico models can provide relevant answers and ultimately contribute significantly to healthcare decision-making.

In this way, In Silico Medicine will not just be an innovative research tool, but a driving force for change in the biomedical field—assuming research continues to reach new heights of excellence. The path forward is full of challenges, but by adhering to this rigorous approach, it is possible to **shift perceptions**, **gain the trust of institutions**, and ultimately **establish In Silico Medicine as a key element in healthcare decision-making**.

## 1.2 Objectives of the report

This academic report, carried out as part of the course *GBIO0033 - Advances in In Silico Medicine*, has a dual objective: on one hand, a critical analysis of an existing in silico model, and on the other hand, a practical study aimed at applying and testing this model by conducting an "in silico" clinical trial. The

model under study is a *Viral Therapy* of the HIV-1 virus, proposed in [27]. The objective is to analyze the credibility and quality of this model through the lens of the principles of Verification, Validation, and Uncertainty Quantification (VVUQ), to determine to what extent it can provide useful and reliable results in a medical context.

The objectives of the report are as follows:

- **Presentation and contextualization of the viral therapy model**: The first part of the report will be dedicated to a detailed presentation of the in silico Viral Therapy model and the underlying biology. Special focus will be placed on the method used, including data acquisition and the analyses carried out in the reference article. This is done in Sections 2.1, 2.2.1, 2.2.3 and 2.2.4.

- **Critical analysis of the scientific article through the prism of VVUQ**: A critical analysis will be conducted to evaluate how the authors considered the principles of VVUQ in their work. This analysis will focus particularly on how the researchers sought to validate the credibility of the model, through verification, validation, and uncertainty quantification. The goal is not to judge whether the model perfectly represents biological reality, but to determine whether the researchers followed a rigorous and transparent approach to enhance the credibility of their results. This is presented in Section 2.3.

- **Credibility assessment**: The second part of the report will consist of applying this model to conduct an in silico study on HIV-1 therapy. This study will first consist of a rigorous methodological approach based on VVUQ, including:

  - Evaluation of Credibility: Identifying the research question and context of use, defining acceptability criteria, and performing a risk analysis.

  - Verification and Validation: Reviewing verification steps (platform and user) and available comparison data to validate the model's results.

  - Uncertainty Quantification: Selecting relevant parameters, justifying their choice, and using appropriate technology to quantify uncertainties.

  All of this is the whole point of Section 3.1.

- **Credibility Matrix**: A summary table of the credibility evaluation will be presented, in the spirit of [21]. This matrix is reported in Table 6, part of Section 3.1.5.

- **Application of the viral therapy model on a virtual population**: In addition to the VVUQ analysis, I will apply the Viral Therapy model to a virtual population. This application will involve simulating the model on a virtual group of patients. This will allow us to observe the results obtained with the model and assess the treatment's effectiveness in a controlled, simulated environment. The in silico trial is designed and conducted in Section 3.2.

- **Highlight of the limitations of the in silico trials**: Section 3.2.5 and 3.1.5 are really key parts of this report as they put emphasis on what are the main things that limit the medical impact of the model and of the in silico trial. The lack of data and the structure of the model are the main things that should explore if one wants to increase the range of applicability of the model.

In summary, this report aims to critically analyze the Viral Therapy model, apply it in a virtual trial, and assess its credibility and relevance in the context of HIV-1 therapy and this, as a 'toy' problem in the understanding of what is - or should be - *in silico medicine*.

# 2 Critical analysis of the article and the model

## 2.1 Biological and medical Context

HIV-1 (human immunodeficiency virus type 1) is a retrovirus responsible for the AIDS (acquired immunodeficiency syndrome) pandemic [15]. Discovered in the early 1980s, it primarily targets CD4+ T lymphocytes, key cells of the immune system, leading to their progressive destruction. This systematic attack weakens the body, making it vulnerable to various opportunistic infections and cancers.

HIV-1 is distinguished by its complex life cycle, which includes the integration of its genetic material into that of the host through an enzyme called reverse transcriptase. This process allows the virus to replicate efficiently while evading the immune defenses of the organism. Furthermore, the virus's high mutation rate contributes to its resistance to antiretroviral treatments, posing a major challenge in controlling the epidemic.

Clinically, HIV-1 progresses through several phases: the acute infection, where the virus replicates rapidly; the clinical latency phase, during which symptoms are minimal but the virus continues to spread; and finally AIDS, characterized by a severe collapse of the immune system. Without treatment, the progression to AIDS is generally inevitable, leading to severe complications and a fatal outcome.

Modern antiretroviral therapies, while effective in slowing the progression of the disease, do not completely eradicate HIV-1 due to the presence of latent viral reservoirs. This limitation has driven research into new approaches, including in silico models and alternative therapies, such as the use of genetically modified viruses to specifically target cells infected with HIV-1.

As a scientific, medical, and societal challenge, HIV-1 remains a priority area of biomedical research, with ongoing efforts to develop innovative solutions to better understand and treat this infection.

## 2.2 Analysis of the reference article

### 2.2.1 Introduction of the article

The article studied proposes an innovative mathematical model aimed at evaluating the effectiveness of a viral therapy in treating HIV-1. This approach relies on the use of genetically modified viruses, referred to as therapeutic viruses, to selectively target and eliminate HIV-1 infected cells. Unlike conventional antiretroviral treatments, this therapy not only aims to reduce the viral load but also seeks to partially restore CD4+ T lymphocyte populations, which are essential for immune response.

The model described in the article explores the dynamic interactions between five main elements: healthy CD4+ T cells, cells infected with HIV-1, cells doubly infected with both HIV-1 and the therapeutic virus, HIV-1 viral particles, and therapeutic virus particles. The dynamics of these components are described through a system of ordinary differential equations (ODEs).

The article itself derives its foundation from the paper [30], which is an *in vitro* study of the therapeutic virus, a product of genetic engineering and synthetic biology. The study [30] is frequently cited by [27], which is the article of interest presenting an in silico study of the therapeutic virus in an *in vivo* context. This work, therefore, represents an in silico trial of the virus as a potential therapy for HIV-1.

The main assumptions of the model are as follows:

- Only CD4+ T lymphocytes are considered as host cells for the HIV-1 virus. This assumption is supported by studies [7, 39, 25], which demonstrate that over 90% of HIV-1 replication occurs in these cells.

- The immune response to viral invasion is neglected. This is a strong assumption but is supported by the objectives of the model (which can be likened to a *context of use*, as discussed in detail in Section 3.1). Specifically, the model aims to simulate the biology at an early and long-term stage but not during intermediate stages. Although the immune response is acknowledged in the literature as relevant during HIV-1 infection [7], it is transient because the immune system, being a primary target of HIV-1, is significantly compromised in advanced stages of the infection.

- The parameters governing the dynamics are assumed to be constant, and no mutation mechanisms are incorporated.
- The model is deterministic, and biological variability is simplified to an average value.

  These last two assumptions are not justified by the original paper, except for the simplification of modeling, and no recommendations are provided for addressing these limitations.

The article is divided into sections in a typical manner, with an introduction to the ODEs of the model, a theoretical analysis of the system's stability, and a numerical study for different sets of parameters. The remainder of this subsection aims to present the model, the theoretical results, and the methods used for parameter estimation, following the original paper, and then to present the major results. Finally, subsection 2.3 will detail the gaps in the article from the perspective of VVUQ as described in the introduction and objectives.

### 2.2.2 Computational model

The model is a simple viral dynamics model as proposed by [22] and is summarized in box #1. The model is a relatively simple system of ODEs. The parameter values as presented in the article are summarized in box #2, and the initial conditions of the simulations are summarized in box #3.

**Regarding the analytical analysis**

The article focuses heavily on studying the equilibrium states of the system under different conditions (no infection, single infection, and double infection). The theoretical results on equilibria form an important basis for the results presented in the paper. Box #4 summarizes these equilibrium states and their conditions of existence. Attention can be paid to the link established between the model's initial conditions with treatment and the equilibrium states of the system under single infection. It can also be noted that the equilibrium states are *independent* of the initial conditions.

**Regarding the parameters**

The paper lacks transparency in parameter estimation, which significantly affects the credibility of the model. While the authors reference external studies to justify parameter values, they do not explicitly describe how these values were derived or fitted to data. This absence of detail makes it difficult to assess the reliability and biological relevance of the chosen parameters. For example, the infection rate of host cells by HIV-1 ($\beta$) is mentioned to be small enough to allow both infections to develop simultaneously, but the exact method or data source used to determine this value is not provided. Such ambiguity undermines confidence in the model's ability to represent real-world dynamics accurately.

Moreover, the paper does not adequately address biological variability across individuals or populations. Parameters such as the infection rate ($\beta$), production rate of recombinant viruses ($c$), and the removal rate of recombinant virus ($q$) are known to vary significantly among individuals. Ignoring this variability limits the model's applicability to heterogeneous populations. The assumption of uniform distributions for sensitivity analysis oversimplifies the inherent complexity of biological systems, where parameters often follow more nuanced distributions, such as Gamma or log-normal distributions.

Another critical gap is the absence of calibration against population-level data. The authors do not validate their parameter choices by comparing model outputs to observed data from diverse populations. This omission raises concerns about whether the selected parameter ranges are representative of real-world scenarios. For example, the production rate of recombinant viruses by double-infected cells ($c$) is based on in vitro experiments, but the specific details of these experiments are not discussed, leaving room for uncertainty about the biological relevance of the parameters.

Additionally, the paper fails to explore interactions between key parameters in sufficient depth. For instance, the interaction between the recombinant virus infection rate ($\alpha$) and the production rate of recombinant viruses ($c$) plays a crucial role in determining treatment efficacy. Neglecting these interactions oversimplifies the analysis and limits the model's predictive power.

Finally, the reliance on hypothetical values for certain parameters introduces additional uncertainty.

While the authors acknowledge the challenges of parameter estimation, they do not systematically justify or quantify the impact of these assumptions. This lack of rigor compromises the robustness of the model and highlights the need for future research to refine parameter estimation methods. Overall, addressing these gaps in parameter handling would significantly enhance the credibility and applicability of the model.

**Regarding the numerical methods**

The paper is relatively vague on the numerical methods used.

The software `MATLAB` is mentioned, and the authors indicate using a *"stiff"* solver. The use of a stiff solver is a good idea and adds credibility to the convergence of the results, but it is clear that the article does not provide enough detail on the subject.

Indeed, there are many stiff solvers with numerous hyperparameters. The methodology section of the paper really leaves much to be desired and does not allow for easy reproduction of the results.

The estimation of parameters based on the data is never really addressed in the paper. In practice, it only makes numerous references, so one must dive into these references to understand the origin and type of the data, as well as the methods used to determine the parameters.

Finally, there is no link to their implementation, which would provide an idea of how the model is practically implemented.

### 2.2.3 Main results

Two main analyses are conducted in the article. First, a study of the temporal dynamics in the case of *single-infection*, via numerical simulations, aiming to identify the infection peak and its intensity. These dynamic results are then compared to the dynamics in the case of *double infection*. Second, an equilibrium study is conducted for different parameter values in the case of *double infection*. These equilibrium results form the basis of the discussion on the potential of treatment using a recombinant virus.

**Dynamical analysis**

The study of temporal dynamics aims to compare the model subjected to a single infection of HIV-1 with the case of *double infection*, where the equilibrium state achieved during the single infection is perturbed by the introduction of the recombinant virus.

Figure 1 illustrates the dynamics of the *single infection*, allowing visualization of the *infection peak* reached towards the end of the first week of infection. The article studies the timing and intensity of this peak as functions of the initial conditions, such as viral load and host cell density. The main result shows that the timing of the peak is strongly influenced by the initial host cell density (a higher initial density accelerates the peak), while the initial viral load has minimal influence. Conversely, the intensity of the peak is significantly affected by both the initial viral load and host cell density. These results, already well-known in the literature, primarily validate the model by demonstrating its ability to reproduce findings widely supported by the scientific community in an *in vivo* context. The value ranges studied are supported by references based on *in vivo* data, ensuring that the ranges are "physiological." While the *single infection* is not the primary focus of the model, it serves as a *sanity check* and provides a foundation for comparisons with the dynamics of *double infection*. Consistent with analytical results, the equilibrium position of the system is independent of initial conditions.

Figure 2 illustrates the dynamics of *double infection* from the equilibrium state achieved during the *single infection*. In summary, this represents the dynamics observed during the treatment of an HIV-1-infected patient via injection of a recombinant virus. The article examines the *initial* transient behavior (within the first week) of the infection, focusing on the recovery of healthy host cells, rapid reduction in viral load, and the decline in the number of infected cells. Observations are made regarding the oscillatory but damped behavior of the transient period around the new equilibrium states. The oscillations in healthy host cell density are minimal compared to the more pronounced oscillations of other system components. Consistent with the conclusion that the model better represents short- and very long-term behavior rather

than medium-term dynamics, no analysis is conducted beyond the first week. Equilibrium positions are considered representative of long-term behavior. Viral load decreases by a factor of 12 during the first week, and healthy cell density increases by an order of magnitude.

A second simulation is analyzed, with a recombinant virus infection rate four times higher. The results on the dynamics of healthy host cell density during the first week are presented. The main result indicates that the recovery of a healthy cell population is faster with a higher infection rate, and the long-term intensity (equilibrium position) is also greater. The simulation is reproduced and presented in Figure 3. While no quantitative results on dynamics are provided for this second simulation, they could easily be extracted from the figures.

An additional aspect discussed but not shown in the article is the existence of parameter sets leading to stable limit cycles around the equilibrium values. The frequency and amplitude of these oscillations depend on the parameters.

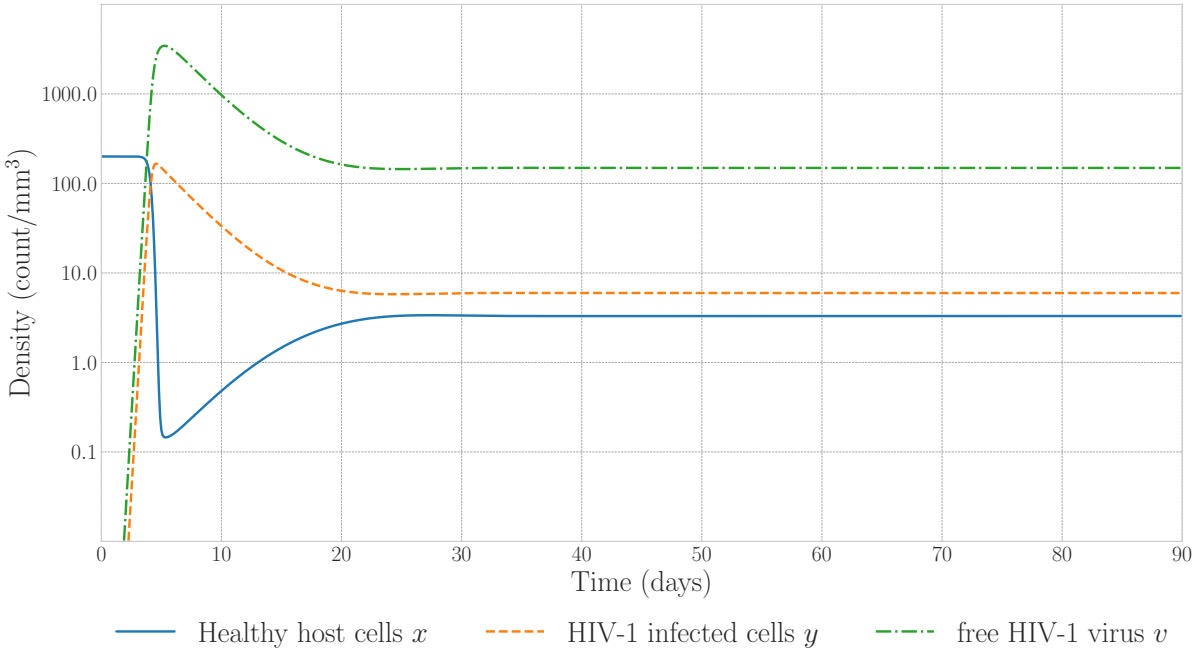


Figure 1: **Reproduction of Figure 2a from [27].** Temporal dynamics of a *single HIV-1 infection*. The infection peak is reached towards the end of the first week, with its timing primarily influenced by the initial host cell density and its intensity affected by both initial host cell density and viral load. These results validate the model by reproducing well-known *in vivo* findings. The equilibrium position of the system is independent of initial conditions, confirming analytical expectations.

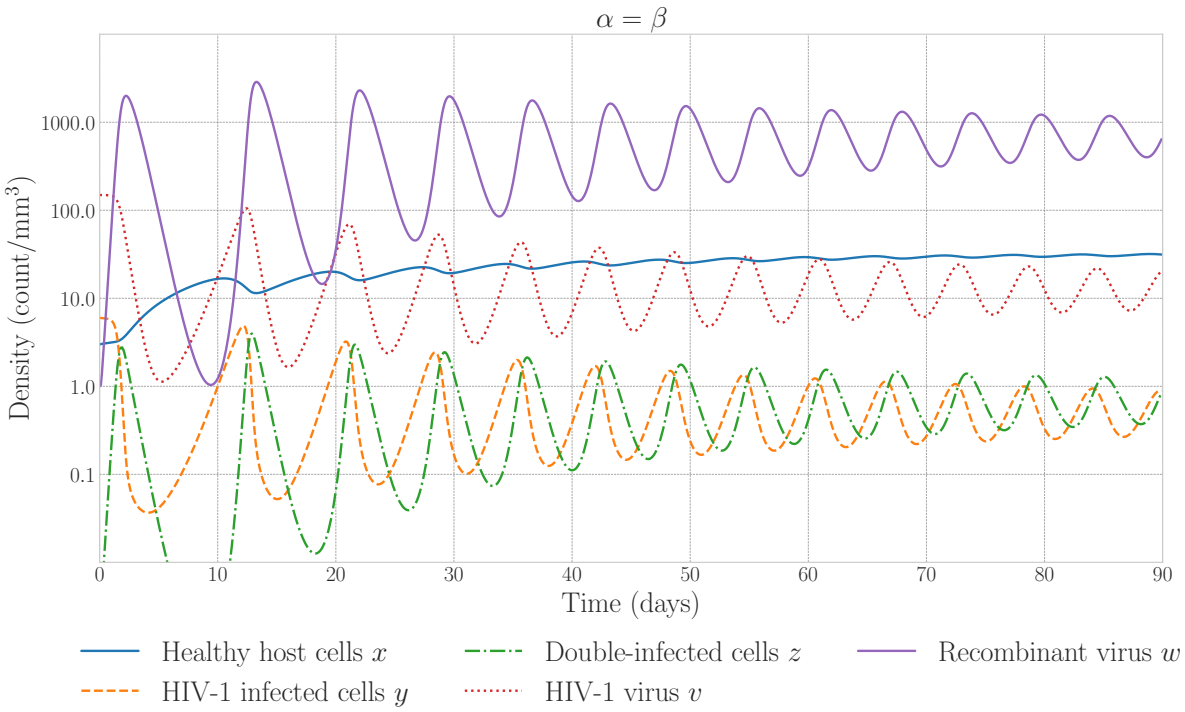


Figure 2: **Reproduction of Figure 2b from [27].** Dynamics of *double infection* starting from the equilibrium state of the *single infection*. This scenario models the treatment of an HIV-1-infected patient with a recombinant virus. The highly transient phase (first week) is characterized by the recovery of healthy host cells, a rapid decline in viral load, and oscillatory behavior of infected cells. The viral load decreases by a factor of 12, while the density of healthy cells increases by an order of magnitude. Long-term equilibrium positions indicate a sustained improvement in host cell population.

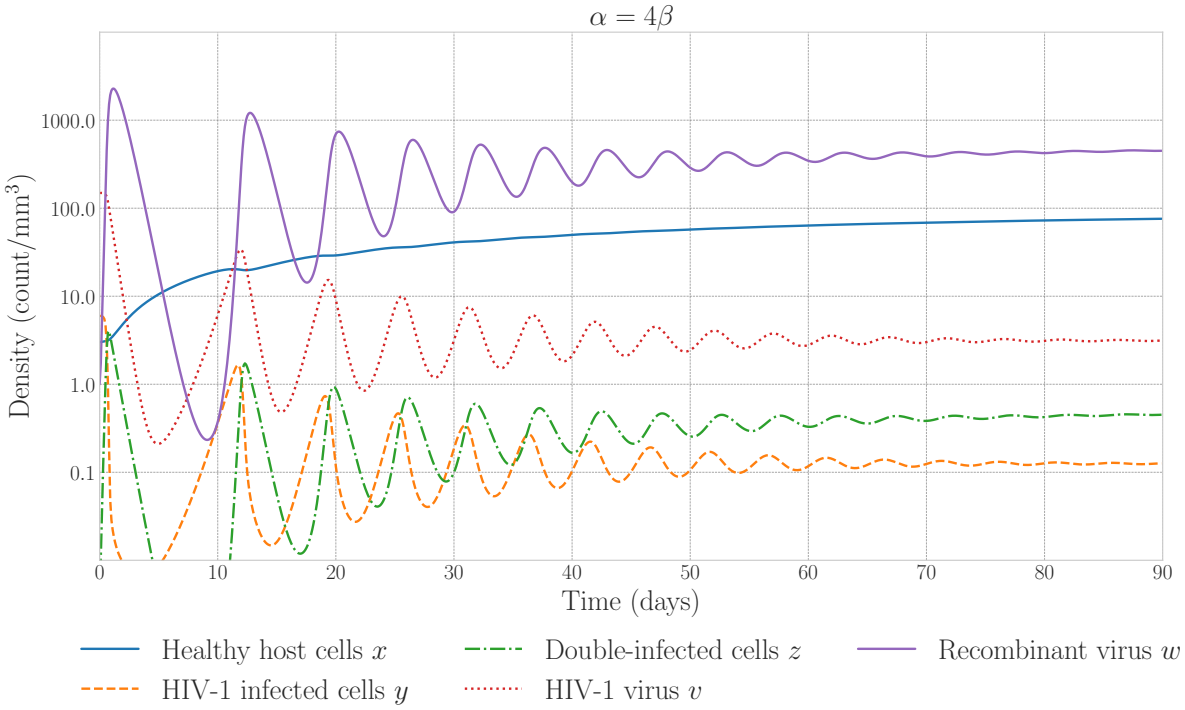


Figure 3: **Reproduction of Figure 2c from [27].** Dynamics of *double infection* with an increased recombinant virus infection rate ($\alpha = 4\beta$). A higher infection rate leads to a faster recovery of healthy host cells and a higher long-term equilibrium density. While no quantitative results are extracted from this simulation, the qualitative trends align with the expected impact of increased recombinant virus efficiency.

**Equilibrium analysis**

In addition to the dynamic study, Figures 2 and 3 introduce equilibrium analyses, showing recovery rates of 17% (compared to the same patient under no infection conditions) in Figure 2 and 44% in Figure 3. Surprisingly, these values are not compared to those in the case of *single infection* by HIV-1. After calculation, this value is 1.7%, providing insight into treatment effectiveness.

An extended equilibrium point analysis is performed, presented both as a sensitivity analysis of the model and as a study of its optimal performance, which can be associated with uncertainty quantification (UQ). Results are presented in a table that details the recovery percentage of healthy host cell density (compared to uninfected conditions) and the reduction in viral load (compared to the *single infection* case). This table is not reproduced here but can be accessed in the original article [27], *"**Table 2** - Effect of parameter values on the density of normal cells $x_i$, pathogen $v_i$, and system stability."*

The results from this table are challenging to interpret due to the conflation of sensitivity analysis (SA) and UQ. This issue is extensively discussed in Section 2.3, which critiques the article. The overall synthesis of these results indicates that substantial recovery of healthy cells and significant viral load reduction are achieved with a high reproduction rate of healthy cells ($\lambda$) and high recombinant virus production ($c$), particularly when the recombinant infection rate exceeds that of HIV-1 ($\alpha > \beta$). A treatment efficacy threshold is provided as a 1000-fold reduction in viral load. The table also distinguishes between stable equilibrium states and stable limit cycle cases by analyzing numerical solutions after 200 days of treatment and reporting amplitude values. Due to the model's nature, the limit cycle always encircles the equilibrium point, ensuring that amplitude ranges include the equilibrium value.

Additionally, a detailed analysis is performed for a population with fixed parameters—$d, \lambda, a, b, k, u$, and $q$ as defined in Box #2—and varying treatment designs (values of $\alpha$ and $c$). These results are partially reproduced in Figure 4 and are considered key findings of the paper. As previously emphasized, the recombinant virus's tunability through synthetic biology is a central strength, with its parameters being pivotal for therapy. Two key conclusions emerge: (1) the recombinant virus production rate, $c$, is the determining factor for treatment efficacy, with higher values leading to greater viral load reduction and healthy cell recovery; (2) for a given $c$, treatment is more effective when the ratio $\frac{\beta}{\alpha}$ is high, meaning the recombinant virus's infection rate surpasses that of HIV-1. As this ratio increases, the influence curves of $c$ become more pronounced. Limit cycles disappear for sufficiently large values of $c$, corresponding—although not explored in the article—to a subcritical Hopf bifurcation. The *in silico* results closely align with those obtained *in vitro* by [30].

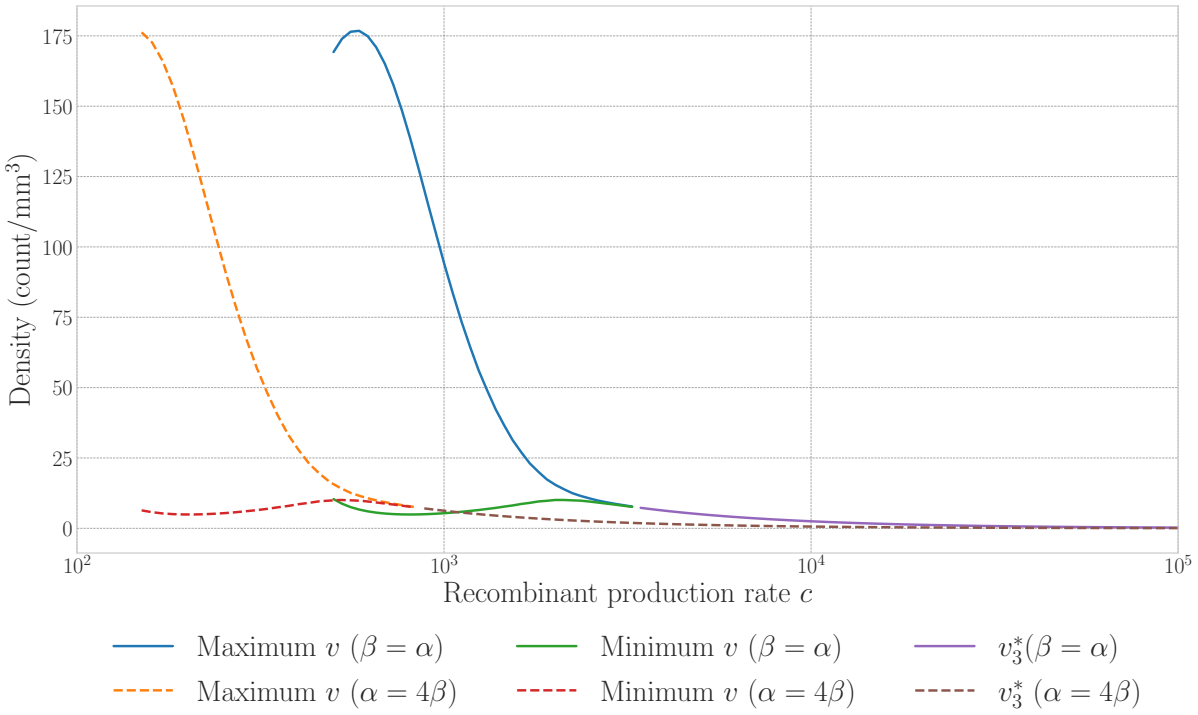


Figure 4: **Partial reproduction of Figure 3 from [27].** Analysis of treatment efficacy under varying values of $\alpha$ and $c$ while keeping other parameters fixed (as defined in Box #2). The study examines two cases: $\alpha = \beta$ and $\alpha = 4\beta$. Two main conclusions arise: (1) The recombinant virus production rate $c$ is the primary determinant of treatment success, with higher values leading to enhanced viral load reduction and healthy cell recovery. (2) For a given $c$, treatment is more effective when the ratio $\frac{\beta}{\alpha}$ is high, meaning the recombinant virus infects cells more efficiently than HIV-1. As this ratio increases, the effect of $c$ becomes more pronounced. Oscillations are assessed based on the solution behavior between $t = 150$ days and $t = 200$ days, with amplitudes below 0.1 considered negligible. Limit cycles vanish for sufficiently large $c$, potentially indicating a subcritical Hopf bifurcation.

**General discussion**

The article concludes by contrasting the viral therapy presented with traditional antiviral drug treatments. The authors emphasize that while viral therapy shows lower efficacy in reducing viral load and restoring healthy cell populations, *in silico* models and *in vitro* results from [30] support that once introduced, the recombinant virus persists in the system, exerting a lasting effect with a single treatment dose. The declining efficacy of antiviral drugs is contrasted with the stable nature of the *double infection* state.

The ability to manipulate the recombinant virus to achieve specific parameter values is a key strength of the results, presented as strong support for the *in silico* model's utility in preemptively determining optimal parameters.

Finally, scenarios leading to stable limit cycles are presented as treatment failures, indicating insufficient control over HIV-1 infection. Such predator-prey-like oscillatory behavior is considered undesirable from the patient's perspective.

### 2.2.4 Future perspectives

The paper outlines several future directions, aiming to improve recombinant virus engineering and enhance the presented model. Emphasis is placed on better parameter estimation and introducing mechanisms and interactions with the immune system, which becomes non-negligible once a partially restored healthy population is achieved.

## 2.3 What is missing?

This section highlights missing elements in [27] from the perspectives of VVUQ and *in silico* evidence. First, the paper's positive aspects are identified, showing alignment with VVUQ principles even if not explicitly stated. Subsequently, the missing and problematic elements that compromise the credibility of the *in silico* results are presented.

**Positive Aspects**

The paper clearly introduces initial assumptions and their limitations, proposing future directions to address them. Furthermore, the model is introduced as a tool to support *in vitro* results and optimize treatment design. This defines a ***Question of Interest*** and a ***Context of Use***. However, the *Context of Use* is not explicitly stated, which will be addressed in Sections 3.1 and 3.1.5.

Although the model is deterministic, an effort was made to present known parameter variability within the population.

The model is validated against *in vitro* predictions from [30]. It is also validated in a simplified case (referred to earlier as the *sanity check*), comparing against *in vivo* data.

Although verification is not explicitly described (neither for numerical simulation packages nor their usage), analytical results lend credibility, as simulations converge to equilibrium points. `MATLAB` is cited, and a stiff system solver is employed, supporting verification efforts.

While debatable, a *sensitivity analysis* (SA) is introduced. A positive aspect is its inclusion, despite shortcomings in execution, and to some extent in interpreting results.

Finally, clear *Quantities of Interest* are defined, specifically viral load reduction relative to *single infection* and recovery of healthy host cells relative to a healthy patient.

**Negative aspects and missing Elements**

The primary shortcoming is the approach to SA and UQ.

First, the UQ analysis is inadequate. Although parameter variation effects are presented, no clear quantification of result uncertainty stemming from population-level parameter variability is provided. The article demonstrates parameter-dependence but fails to quantify the uncertainty associated with biologically coherent probability distributions.

Regarding SA, while the article includes an SA through Table 2 in [27], presenting sensitivity of Quantities of Interest to parameter variations, this approach is insufficient. Parameters are not explicitly ranked or globally analyzed—results stem from a limited set of simulations. Given the simplicity and computational efficiency of the *in silico* model, more simulations should have been performed and presented. Furthermore, parameter distributions are not explicitly detailed. Most critically, the SA interpretation is flawed. The authors claim high treatment efficacy based on extreme parameter values, evaluated for a single individual with no parameter fluctuations. This claim is particularly concerning given the paper's acknowledgment of parameter estimation challenges and significant population variability. Proper calibration or identification of influential parameters via SA, alongside UQ for biological variability, is essential for model credibility.

Finally, while the article presents temporal dynamics analysis, these are not contextualized within a therapeutic framework. Additionally, without defining short- and long-term ranges consistent with model assumptions (e.g., neglecting the immune system), it is challenging to assess result coherence regarding immune system recovery potential.

**# 1: Model**

The model describes the dynamics of the following variables in 1 $mm^3$ of blood:

* The number of virus-free host cells (CD4+ T cells): $x$ [count/$mm^3$];
* The number of cells infected by the HIV-1 virus: $y$ [count/$mm^3$];
* The number of free HIV-1 viral particles in the blood plasma: $v$ [count/$mm^3$];
* The number of viral particles of the *recombinant* virus: $w$ [count/$mm^3$];
* The number of cells double-infected by HIV-1 and the recombinant virus: $z$ [count/$mm^3$];

Dynamics is encoded via the following system of ODEs (based on [22]):

$$\dot{x} = \lambda - dx - \beta xv \tag{1}$$
$$\dot{y} = \beta xv - ay - \alpha wy \tag{2}$$
$$\dot{z} = \alpha wy - bz \tag{3}$$
$$\dot{v} = ky - uv \tag{4}$$
$$\dot{w} = cz - qw \tag{5}$$

The parameters and their values are presented in Box #2.

These equations provide a mathematical representation of the main phenomena associated with viral infection. Specifically:

- Eq. (1) encodes a cell production term ($\lambda$), a cell death term ($-dx$), and the infection of cells by HIV-1 ($-\beta xv$).
- Eq. (2) represents the viral dynamics, including the infection of cells by HIV-1 ($\beta xv$), cell death ($-ay$), and the phenomenon of double infection by the recombinant virus ($-\alpha wy$).
- Eq. (3) similarly represents double infection ($\alpha wy$) and cell death ($-bz$).
- Eqs. (4, 5) describe the dynamics of viral particles (HIV-1 and recombinant virus, respectively), with terms for their synthesis by infected cells ($ky$ and $cz$) and their removal from the blood plasma ($-uv$ and $-qw$).

This model can be classified as a *'physics-based model'* on the in-silico spectrum, with a small *data-driven* component for parameter estimation. However, the model remains relatively simple and abstract regarding the underlying mechanisms, such as cell death, viral particle synthesis, and their elimination. This simplicity, along with its 'physics-based' nature, makes the model interpretable, especially concerning the relationship between parameter values and the underlying biology, as discussed in Box #2.

***This box reproduces the equations as presented in [27]*, equations (2)**.

## # 2: Model Parameters

The parameters of the model presented in Box # 1 are listed in the table below. Comments are provided to justify the values.

| Parameter | Description | Value ($\text{day}^{-1}$) |
|---|---|---|
| $d$ | Death rate of host cells (activated CD4+ T cells) | 0.01 |
| $\lambda$ | Production rate of new host cells | 2 cell/$\text{mm}^3$ |
| $\beta$ | Infection rate of host cells by HIV-1 | 0.004 $\text{mm}^3$/vir |
| $a$ | Death rate of HIV-1 infected cells | 0.33 $\text{day}^{-1}$ |
| $\alpha$ | Infection rate by recombinant virus | 0.004 $\text{mm}^3$/vir |
| $b$ | Death rate of double-infected cells | 2 |
| $k$ | HIV-1 production rate by a single infected cell | 50 vir/cell |
| $u$ | Removal rate of HIV-1 | 2 |
| $c$ | Production rate of recombinant viruses by double-infected cells | 2000 vir/cell |
| $q$ | Removal rate of recombinant virus | 2 |

**Comments:**

The comments below are paraphrased from the reference article.

- The death rate of host cells ($d$) is based on the average lifespan of CD4+ T cells, which is approximately two years, translating to $d = 0.0014\,\text{day}^{-1}$. However, to get a better results from several model, much higher values are required, reference leads to $d = 0.01$.
- The production rate of host cells ($\lambda = 2\,\text{cell/mm}^3\text{/day}$) is assumed to be proportional to the equilibrium value of the host cell density, which is $X_0 = 1000\,\text{cell/mm}^3$ in healthy individuals. $\lambda$ is calculated as $\lambda = X_0 d = 10$. Assuming that 20% of new cells become activated, $\lambda$ is set to 2 .
- The infection rate of host cells by HIV-1 ($\beta$) is chosen to be small enough to allow both infections to develop simultaneously. This value is based on estimates from previous studies, with variations such as 0.00027, 0.00065, 0.0035, and an upper bound of 0.007 .
- The death rate of HIV-1 infected cells ($a$) is based on the observed lifespan of HIV-1 infected cells, which is approximately three days. Other estimates suggest values of 0.49 and $0.39\,\text{day}^{-1}$ .
- The infection rate by the recombinant virus ($\alpha$) is assumed to be identical to the infection rate of HIV-1 for simplicity. This assumption allows for the development of both infections.
- The recombinant virus is assumed to cause rapid cytopathic infection, leading to the death of double-infected cells within a short period. *In vitro* [30] reports $b = 2$.
- The HIV-1 production rate by a single infected cell ($k$) is estimated based on the total number of virions produced, which is approximately 140 to 180 per cell .
- The removal rate of HIV-1 ($u$) is based on the half-life of HIV-1 in plasma, which is approximately 12 hours, translating to $u = 2\,\text{day}^{-1}$. Others references lead to $u = 3$.
- The production rate of recombinant viruses by double-infected cells ($c$) is estimated based on in vitro experiments [30], where the total number of recombinant viruses produced per cell is approximately 3333. Here, a slightly lower value is used for simplicity.
- The removal rate of recombinant virus is assumed to be identical to the removal rate of HIV-1.

***This box is, in part, a reproduction of Table 1 from [27]. For an exhaustive list of references, the reader is directed to the original article.***

### # 3: Initial Condition

The initial conditions of the model, i.e., the values of $x|_{t=0}$, $y|_{t=0}$, $z|_{t=0}$, $v|_{t=0}$, and $w|_{t=0}$, with $t = 0$ representing the moment of the second infection, are determined as follows:

$$x|_{t=0} = \frac{au}{\beta k} \approx 3, \tag{6}$$

$$y|_{t=0} = \frac{\lambda}{a} - \frac{du}{\beta k} \approx 6, \tag{7}$$

$$v|_{t=0} = \frac{\lambda k}{au} - \frac{d}{\beta}, \approx 149, \tag{8}$$

The first three conditions correspond to the equilibrium values of the HIV-1 system in the case of a single infection;

$$z|_{t=0} = 0, \tag{9}$$

$$w|_{t=0} = 1, \tag{10}$$

The last two conditions reflect the start of the second infection (no double-infected cells initially) with a minimal amount of recombinant virus;

***This box is based on Fig. 2 from [27].***

**# 4: Equilibrium States of the System**

The equilibrium states of the system described by the set of ODEs and the parameters derived from [27], presented respectively in boxes #1 and #2, can be analyzed mathematically and constrained biologically. The equilibrium points correspond to the solutions of $\dot{\boldsymbol{x}}|_{\boldsymbol{x}=\boldsymbol{x}^*} = \mathbf{0}$ from a mathematical perspective, and are considered biologically valid if and only if $\boldsymbol{x}^* \geq \mathbf{0}$.

The study of these equilibrium points is of great interest for this model as they provide important insights. On one hand, they inform a biologically consistent initial state, such as the one described in box #3. On the other hand, they highlight the efficacy of the treatment being considered. This aspect is explored in detail in subsection 2.2.3 of this report.

The study of these equilibrium points yields the following theoretical results:
**Model in the complete absence of virus**

$$(x_1^*, y_1^*, z_1^*, v_1^*, w_1^*) = \left(\frac{\lambda}{d}, 0, 0, 0, 0\right) \tag{11}$$

This equilibrium state is unconditional and stable.
**Model with single infection [HIV-1 Only]**

$$(x_2^*, y_2^*, z_2^*, v_2^*, w_2^*) = \left(\frac{au}{\beta k}, \frac{\lambda}{a} - \frac{du}{\beta k}, 0, \frac{\lambda k}{au} - \frac{d}{\beta}, 0\right) \tag{12}$$

This equilibrium state is stable and should only be considered if $R_v = \dfrac{\beta\lambda k}{adu} > 1$, corresponding to $v_2^* > 0$. In other words, the single infection is stable if the invasion by HIV-1 is sufficiently strong. Such conditions are common for this type of model and are analyzed in depth in [22].
**Model with double infection [HIV-1 and recombinant virus]**

$$(x_3^*, y_3^*, z_3^*, v_3^*, w_3^*) = \left(\frac{\lambda}{d + \frac{\beta bkq}{\alpha cu}}, \frac{bq}{\alpha c}, \frac{qw_3^*}{c}, \frac{bkq}{\alpha cu}, \frac{\alpha\beta\lambda ck - \beta abkq - \alpha acdu}{\alpha\left(\beta bkq + \alpha cdu\right)}\right) \tag{13}$$

This equilibrium state is either stable or leads to an attractive limit cycle and should only be considered if $R_w = \dfrac{\alpha\lambda c}{abq}\dfrac{R_v - 1}{R_v} > 1$, corresponding to $w_3^* > 0$. In other words, the recombinant virus must also exhibit sufficiently strong invasion to persist permanently.

***This box is based on the results presented in Section 3 of [27].***

## 2.4 The impact of the paper

The reference paper [27] is directly cited by a total of 62 articles, which in turn are cited an average of 30 to 40 times. Among these articles, there are theoretical papers, notably on the analysis of limit cycles [9] or the stability of "predator-prey" type models [36] [22]. There are also many modeling papers that use the relatively simple model as a basis and add elements based on *in vitro* and *in vivo* observations.

Among these papers, there are models of cellular delays and cell cycles [8, 18, 2], and especially papers that address the most significant hypothesis and simplification of the model, namely the neglect of immune system recovery [1, 37].

Among the references of these references, there are animal trials on vaccines based on these recombinant viruses in the context of HIV-1 [5, 26].

I did not find any articles discussing clinical trials for these therapies.

It is difficult to determine the exact impact of a research paper, but it is clear that the idea and model presented by [27] have had an impact on scientific research. Although some aspects of their VVUQ approach and *in silico* medicine are open to criticism, the authors highlighted an essential building block that has been reused and improved by the scientific community. More recent papers are closer to what is expected for the credibility of results in *in silico* studies.

# 3 In silico trial based on the viral therapy model

While the previous section of this work was a presentation, analysis, and reproduction of the results from [27], this section proposes a completely original contribution.

This section aims to use the model presented in the boxes #(1, 2, 3, 4) as a "toy problem" to explore various aspects of *good simulation practice* [34]. More specifically, Section 3.1.3 will establish elements of *verification* and *validation*, Section 3.1.4 will conduct a sensitivity analysis, and Section 3.1.5 will synthesize all the elements to consider in a credibility matrix for conducting an in silico trial based on the model. Une approche *risk-based* est utilisée et introduite dans la Section 3.1.

Then, in Section 3.2, an in silico trial will be conducted on a virtual population, with a *control* population and two populations under two different treatments. These trials will be conducted under the framework of uncertainty quantification (UQ), and the uncertainty of biological parameters will be incorporated into the model to observe the uncertainty and results that can be obtained for the populations.

The theoretical results presented in #4 will be extensively used.

## 3.1 Credibility assessment

A *risk-based* approach will be used to determine the acceptability of verification and validation. If we later want to extend the *in silico* results beyond the *context of use* (CoU), it will be necessary to start over, in the sense that we will need to verify whether the verification, validation, and subsequent steps are still satisfactory within the new CoU.

### 3.1.1 The question of interest and the context of use

The *in silico* model aims to explore the following question of interest:

> **Question of Interest (QoI 1)**: Is viral therapy—that is, the use of the dual infection mechanism—via a recombinant virus a *possible* and *effective* approach in the context of HIV-1 infection—that is, a reduction in viral load and recovery of a healthy CD4+ cell population—and, if so, what are the important viral characteristics to engineer in the recombinant virus?

This can be broken down into sub-questions summarized as:

1. "Is this therapy possible?" Not from the perspective of virus engineering but rather from the perspective of: "Does the introduction of a recombinant virus impact the quantities of interest?"

2. "Is the therapy effective?" This part of the question of interest is actually very broad and will be the central question in final decision-making, for example, regarding the reimbursement of such therapy by healthcare systems. In practice, within the context of use that will be specified below, it is not yet necessary to address this question in such depth. We will limit ourselves to trying to establish whether a *sufficiently significant* impact on the quantities of interest is possible without precisely defining what "*sufficiently significant*" means.

   Indeed, as with all the results analyzed from the *in silico* model, these results must be put into perspective within the context of use—that is, for this sub-question, the in silico model is positioned before any clinical or even pre-clinical studies. The model plays a role in the research and design of the treatment (the recombinant virus) *in vitro*, and this is why we can limit our interpretation of the treatment's efficacy to having a *sufficiently significant* impact to be studied *in vitro*.

3. "What are the viral characteristics that are important to engineer correctly?" Thus, the in-silico model aims both to guide *in vitro* research in the quantities to explore (reducing the time and cost of *in vitro* research) and to guide experimental research more broadly on the techniques that are important to perfect.

   In the case of the recombinant virus, for example, if the model indicates that the infection rate ($\alpha$) is the most important and sensitive quantity for obtaining results and that this quantity must be

large, experimental research can focus its efforts on designing viruses with a high value of $\alpha$ in the laboratory (effectively reducing the design cost since the range of infection rate values to explore is narrowed by the *in silico* model). Broader fundamental and experimental research will have an indication that it may be important to understand what influences the infection rate.

We can add to the initial question of interest the following:

> **Question of Interest (QoI 2)**: What results can we expect to observe *in vivo* if such viral therapy were actually used on a heterogeneous population—that is, with individuals exhibiting variations in biological parameters?

This second question of interest is deliberately separated from the first because the associated *context of use* will be quite different. The goal is to illustrate the *risk-based* approach and its impact on the acceptability of the verification, validation, and subsequent steps.

The question of interest (whether we are talking about the first or the second) is centered around the following *quantities of interest*:

- The HIV-1 viral load $v$ and, more specifically, its reduction compared to a single HIV-1 infection scenario, over the long term. We are therefore interested in the quantity denoted as follows:

$$\texttt{Quantity of Interest 1 - Viral Load Reduction:} \quad \eta \approx \frac{v_3^*}{v_2^*} \tag{14}$$

  An *effective* treatment approaches a value of $\eta$ close to 0, while a treatment with no impact corresponds to $\eta = 1$. The symbols with an asterisk always refer to the equilibrium values of the system as studied by [27] and presented in the previous section, with the results summarized in box #4. The approximation symbol is used here because, in practice, equilibrium situations in the model can sometimes be limit cycles, characterized by oscillations of $v_3$ around $v_3^*$. More details on managing these oscillations will be provided in the following Section 3.2.

- The density of the healthy CD4+ population $x$ and, more specifically, its recovery compared to a single HIV-1 infection scenario, over the long term. We are therefore interested in the quantity denoted as follows:

$$\texttt{Quantity of Interest 2 - Recovery of a Healthy Population:} \quad \xi \approx \frac{x_3^*}{x_2^*} \tag{15}$$

  Note that, unlike [27], recovery is studied relative to the single HIV-1 infection case and not relative to a healthy individual ($\frac{x_3^*}{x_1^*}$). This choice is made because I consider it more informative about the treatment's efficacy, which is used for patients infected with HIV-1 and not for healthy patients. Of course, the ideal treatment would fully restore a healthy population so that $x_3^* \to x_1^*$, but from a perspective of exploring the most effective treatment (which aligns with the context of use), it is simpler to compare treatments using this metric since the variation in results due to variations in HIV-1 biological parameters is directly readable. We can consider that in very critical HIV-1 infection situations, the treatment's efficacy lies more in reducing the criticality of the situation than in the absolute results compared to a healthy situation.

  Finally, note that we can ensure, based on the model's structure, that regardless of the value of $\xi$, the value of $\frac{x_3^*}{x_1^*}$ is bounded by 100% [I provide a personal proof in Appendix A]. It is therefore not possible to obtain a treatment that would excessively increase the CD4+ population.

  An effective treatment increases the value of $\xi$, while a treatment with no effect gives $\xi = 0$.

Finally, the question of interest is embedded in a *context of use* for the *in silico* model. Thus, the context of use defines the framework in which the *in silico* results are used and determines, based on the associated *risks*, acceptability criteria for the various steps of model credibility assessment. I differentiate here two contexts of use, which will illustrate the *risk-based* approach.

**Context of Use (CoU 1) associated with QoI 1**: The mechanistic model aims to guide and accelerate experimental research on the design, through engineering and synthetic biology, of a recombinant virus as a therapy for HIV-1 infection. The model and *in silico* results play a role in the *research and development phase* of the treatment. The *in silico* model can be considered to have no impact on pre-clinical or clinical phases since the goal is to continue developing and testing recombinant viruses *in vitro*.

$\Longrightarrow$ **Decisions on the direction of *in vitro* research are informed by the *in silico* model.**

*In vitro* research and development should focus on recombinant viruses that perform well *in silico*.

**Context of Use (CoU 2) associated with QoI 2**: The mechanistic model aims to predict directly the *clinical outcomes* of viral therapy in a heterogeneous population of patients within the framework of clinical trials (phase II and phase III). Unlike CoU 1, where the model is used to guide experimental research, this context relies on *in silico* simulations to draw conclusions with significant impact on the efficacy and safety of the therapy in real-world conditions.

Clinical decisions based on the model would include:

- Approval or disapproval of the therapy for advanced clinical trials.
- Identification of target patient groups for whom the therapy is directly recommended.
- Assessment of risks related to inter-individual variability.

$\Longrightarrow$ **Decisions based on clinical data are replaced to save time and money, and decisions are directly informed by the *in silico* model.**

Obviously, we can already see that the second CoU seems much more critical than the first. The following sections will highlight the limitations of the *in silico* method in this context through acceptability criteria, an excessively high risk factor, and the model's excessive impact on decision-making.

### 3.1.2 Risk analysis and establishment of acceptability criteria

The clear definition of the QoI and CoU allows the model to be placed in a context and clearly defines what the model is being credited for. This means that if the model meets the acceptability criteria, it can be judged *sufficiently* credible and used "safely" within this CoU only. The "quantification" of what we mean by a "*sufficiently*" credible model is done through a *risk-based* approach.

Intuitively, a model operating in a riskier context requires much more substantial and stricter evidence than a model with low risks. In practice, two elements come into play in measuring the riskiness of a model's operation:

1. The model's weight in decision-making. A model that operates autonomously, making decisions entirely on its own, is riskier than a model whose *in silico* results complement other evidence—for example, *in vivo* and *in vitro* results—for decision-making. The term *regulatory impact* may sometimes be used to describe this aspect of risk.

2. The consequences related to decision-making, and more specifically, the severity of the consequences if a wrong decision were made. It is clear that the more severe and harmful the consequences of a wrong decision, the riskier the model's operating conditions. The term *decision consequence* may sometimes be used to describe this aspect of risk.

While CoU 1 and CoU 2 are not fundamentally different from a *technical* perspective for the model — in both cases, it is about observing the output based on the input — an observation through the *risk-based* lens places the two in high contrast.

Unfortunately, there is no standard for quantifying these aspects numerically. However, we can rely on risk assessment matrix, presented in Table 2. One way of determining which category a CoU falls into for decision consequence is to assess both the likelihood of a negative consequence and the severity of that consequence - in the event of a bad decision. For the impact on regulation, the category must highlight the extent to which the model is used to make the decision.

**Risk analysis for QoI 1 and its CoU 1**:

***Decision consequences***: The decision consequences for QoI 1 and CoU 1 are relatively low, as the model is primarily used to guide in vitro research. However, there is a financial risk for the laboratory if the model leads to unproductive research directions. For example, if the model suggests a treatment approach that fails in vitro, it could result in wasted time and resources. Nonetheless, these risks are manageable, as the model is intended to reduce the cost and time of experimental research by narrowing down the range of parameters to explore.

$\Longrightarrow$ **Very low**

***Regulatory impact***: The regulatory impact for QoI 1 and CoU 1 is minimal, as the model is not used for clinical decision-making. It is primarily a research tool, and its results are intended to inform in vitro experiments rather than regulatory approvals.

$\Longrightarrow$ **Very low**

***Combined risks associated with the model*** : The combined risks associated with the model in CoU 1 are low, as the model operates in a research context with minimal regulatory impact and manageable financial risks. The impact on patient health is very indirect.

$\Longrightarrow$ **Very low**

**Risk analysis for QoI 2 and its CoU 2**:

***Decision consequences***: The decision consequences for QoI 2 and CoU 2 are significantly higher than for QoI 1, as the model is used to predict clinical outcomes. If the model incorrectly predicts the effectiveness of the therapy, it could lead to inappropriate clinical decisions, such as approving a treatment that is ineffective or even harmful. This could have serious consequences for patient health and safety.

$\Longrightarrow$ **High to very high**

***Regulatory impact***: The regulatory impact for QoI 2 and CoU 2 is substantial, as the model is intended to replace or significantly inform clinical trials (phase II and III). Decisions based on the model could directly influence regulatory approvals, patient treatment protocols, and public health policies. If the model is inaccurate, it could lead to regulatory decisions that compromise patient safety or public health.

$\Longrightarrow$ **High to very high**

***Combined risks associated with the model*** : The combined risks associated with the model in CoU 2 are very high, as the model operates in a clinical context with significant regulatory impact and potentially severe consequences for patient health and safety. The stakes are much higher than in CoU 1, as the model is used to make decisions that directly affect human lives.

$\Longrightarrow$ **Very high**

Using the Table 2, we can directly identify that CoU 2 is far more critical than CoU 1, and this should be reflected in the *acceptability criteria*.

| **Regulatory Impact** / **Decision Consequences** | **Low** | **Medium** | **High** |
|---|---|---|---|
| **Low** | **Very Low** (CoU 1) | Low | Medium |
| **Medium** | Low | Medium | High |
| **High** | Medium | High | **Very High** (CoU 2) |

Table 2: Risk assessment matrix showing the interaction between decision consequences and regulatory impact. CoU refers to "Context of Use" and are presented in 3.1.3.

The risk-based approach aims to establish acceptability criteria in a consistent manner. This means that regulation and model credibility proofs do not need to be as strict for low-risk models as for high-risk models. Acceptability criteria establish a trade-off between the quality of the model's representation *in silico* (including the associated numerical methods) and its complexity, which directly determines the development cost of the model, including calibration, research, etc.

In summary, the acceptability criteria define what is meant by a "*sufficiently*" credible model.

The following section aims to establish acceptability criteria for the two CoUs introduced earlier. These criteria are defined *a priori* to the model testing. This means the model must be refined until it meets these criteria, and not that the criteria are adjusted until the model is deemed credible. The acceptability criteria primarily concern the verification and validation steps. They are built directly upon the risk analysis.

**Acceptance Criteria for CoU 1 and QoI 1**:

$\Longrightarrow$ *Very low-risk CoU.*

- Calculation verification and user-related verification should include a convergence analysis of the solver's dynamics. Given the low risk, this analysis can be qualitative and based on the theoretical equilibrium analysis. Specifically, all solutions must be bounded. This will also confirm that the model is correctly implemented with respect to the reference paper [27]. In particular, the implementation must be able to qualitatively reproduce Figures 2a, 2b, and 2c from [27].
- The numerical tools used must demonstrate a degree of robustness. This can be proven through *Software Quality Assurance* (SQA) or by a thorough convergence analysis of the error.
- Additional credibility can be provided via a versioning system, such as Git, with established sanity checks. However, this is not mandatory and represents only an additional positive point for long-term credibility.
- To validate the model, it must be able to reproduce known *in vivo* results for single HIV-1 infections.
- To validate the model for use in *in vitro* virus design, it must be able to reproduce experimentally obtained results [30]. This reproduction must be at least qualitative, and a measure of observed differences must be provided.
- Ideally, a sensitivity analysis will be conducted to highlight the critical calibration elements for further analysis and posterior updates with future experimental results.

* Validation should ideally focus on the quantities of interest $\xi$ and $\eta$. If experimental data for these quantities are unavailable, other quantities may be used for validation and serve as evidence.

* The goal being to guide *in vitro* research, the model must be updated as scientific discoveries and experiments progress.

By ensuring that these verification and validation criteria are met, the model can be considered sufficiently credible for guiding *in vitro* research in CoU 1, which carries very low risk. While not all criteria need to be strictly quantitative, they provide a foundation for ensuring the model remains useful and reliable for its intended purpose.

**Acceptance Criteria for CoU 2 and QoI 2**:

$\Longrightarrow$ *Very high-risk CoU.*

- Numerical methods (ODE solvers, integration methods, etc.) must be thoroughly verified. This includes ensuring the solver converges to expected results and that approximation errors remain within acceptable bounds. Solution convergence will be verified using standard test cases, comparing results at different step sizes, and ensuring that errors decrease with smaller step sizes.
  Since CoU 2 has high regulatory and decision-making implications, numerical verification must be **quantitative** and supported by a formal error analysis. Furthermore, error bounds must be provided for a wide range of the model's most sensitive parameters.
- Verification steps must be conducted on different architectures to ensure the pipeline's execution is independent of the computational environment. A documentation of

software dependencies (libraries, operating system versions, etc.) must be provided.

- The model must reproduce known theoretical equilibrium results (such as those in Box #4) with high precision.
- Model verification must also ensure no unintended changes occur between model versions, especially during updates or modifications. This can be verified through a version control system like Git, including sanity-check scripts to ensure consistent outcomes after updates. A versioning system **must** be implemented to ensure long-term credibility.
- The model may only be used after updates following extensive testing of all sanity checks, ensuring strictly bounded error levels.
- A complete documentation of the model and its numerical pipeline must be provided.
- The model **must** be able to explicitly quantify uncertainties associated with its outputs. Additionally, a sensitivity analysis **must** be conducted to provide *quantitative* insights on model calibration. Parameters must be estimated to ensure the target population is well-represented, including potential outliers.
- The model must be validated against *in vivo* data from clinical trials addressing similar questions, primarily Phase I (and Phase II depending on the therapy's development and market introduction stage). *In vivo* sanity checks must also be identified and validated, such as the HIV-1 infection dynamics without therapy.
- The risks associated with interactions between the recombinant virus, HIV-1, and the immune system must be quantified. Strong model assumptions must be validated against *in vivo* data.
- Given the decision-making weight and associated risks for accelerating therapy market introduction, discussions with **regulatory bodies** must take place to identify potential additional criteria and establish quantitative error thresholds during validation and verification phases.
  * Validation must focus on the quantities of interest $\xi$ and $\eta$.
  * If applicable, post-decision monitoring must be implemented, and results should be compared with the model to identify and document potential differences.

In summary, the high risks associated with CoU 2 necessitate strict and quantitative acceptability criteria. Comparison with *in vivo* data becomes paramount.

It is essential to emphasize that for high-risk cases, close collaboration with regulatory authorities (such as the FDA or EMA) is critical throughout the validation process to ensure compliance with standards and identify additional criteria that may emerge based on the therapeutic product's specifics.

The distinction in precision requirements for CoU 2, which are significantly more stringent than for CoU 1, is evident. High-risk scenarios demand quantitative and extensive measures, broad validation with *in vivo* data, more robust verification, etc.

These acceptability criteria must all be met for the model to be considered credible for its intended CoU. If the model is credible, it can be used to provide answers (or elements of answers) to the QoI.

**Important Comment**

**Two CoUs were presented, and the acceptability criteria for these CoUs were established based on the risk analysis for both.** In practice, only CoU 1 is analyzed in the continuation of this work, and credibility assessment is performed solely for this CoU. This is due, on one hand, to "practical" reasons, as it is just a "toy problem"; but more importantly—and this point will be discussed extensively in Section 3.1.5 of this report—because the necessary *in vivo* data for the verification and validation steps of CoU 2 are unavailable.

### 3.1.3 Verification and Validation

The verification step aims to explore and provide evidence that the numerical methods used to generate the results behave correctly. This involves, for example, ensuring that the ODE system solver is stable, that the numerical solution behaves qualitatively as expected, and that the error introduced by the numerical approximation is well controlled. In addition to this software-related verification, aspects

of verification can also be proposed regarding how the user has implemented the numerical equations. Verification is therefore independent of the model's quality in representing reality but focuses rather on the quality of the numerical methods and their use in representing the model.

On the other hand, the validation step aims to assess the model's quality in representing a certain biological reality. Thus, we compare the model's results and dynamics with known experimental results.

Both verification and validation are necessary to provide credibility to the model. Analyzing the results of these two steps is critical and must be done within the *context of use*. Indeed, the goal is not to obtain an exact numerical solution or to create a model that is an exact representation of reality—if such a thing is even possible—but rather to demonstrate that the results are acceptable for the model's intended use.

Here, we address verification and validation of CoU 1, associated with QoI 1, and in light of the acceptability criteria established in the previous subsection.

**Verification**

The model is implemented in Python, specifically in the following version: `Python 3.12.7 | packaged by Anaconda, Inc. | (main, Oct 4 2024, 13:27:36) [GCC 11.2.0] on linux`. The following packages are used for the various computational steps, including solving the dynamics (ODE system), data manipulation, and figure generation. These are essential because the evaluation criteria largely rely on qualitative analyses.

| **Package** | **Version** | **Usage** |
|---|---|---|
| Numpy [10] | 1.26.4 | Used for numerical computations and data manipulation. The default precision (float64, as described in the Numpy documentation) is employed. It was verified that the machine used supports this precision. It is also used for generating parameter samples from given distributions (UQ). |
| Scipy [35] | 1.14.1 | Used for model simulation via the `solve_ivp` method with a solver designed for stiff systems, specifically "BDF" [31]. Default parameters are used except for the `rtol` that is constrained to $10^{-6}$, and solutions are evaluated at every $dt = 0.1$ days, corresponding to the maximum time step. |
| Matplotlib [12] | 3.9.2 | Used for creating figures, including those used for visual inspections during verification and validation stages. |
| SAlib [13] | 1.4.7 | Used for conducting sensitivity analysis (SA) of the model. |

Table 3: List of packages and software used for the *in silico* approach of the model.

Figures 1, 2, 3 and 4, presented earlier in Section 2 of this work, are reproductions of Figures 2a, 2b, and 2c from [27]. A visual inspection confirms that the implementation used in this work produces results qualitatively similar to [27].

It can thus be considered that user-related verification is achieved—that is, the equations have been correctly transcribed, and parameters have been correctly implemented. Furthermore, the model can be considered verified from a computational accuracy perspective, as the solutions do not diverge, and the theoretical equilibrium presented in #4 is properly reached.

The libraries used are widely adopted within the scientific community working in Python. These libraries operate under version control systems similar to Git and are subject to numerous benchmark sanity checks for each stable release. The open-source nature of these libraries also allows for code inspection

and community verification. These aspects constitute what is referred to as SQA.

Additionally, Figure 5 empirically confirms that the solver converges correctly. It is known that the equilibrium position is independent of the initial conditions. The experiment illustrated in this figure consists of 500 simulations performed with the same set of parameters (and thus the same theoretical equilibrium position) but with different initial conditions sampled from a normal distribution centered on the equilibrium value, with bounds extending up to 250% of the equilibrium value. This wide range demonstrates convergence for even extreme initial conditions. Specifically:

$$\boldsymbol{x}|_{t=0} \sim \max\left\{\mathcal{U}(\boldsymbol{x}^* \pm 125\%\boldsymbol{x}^*), 10^{-6} \text{ or } 1(\text{for } x \text{ and } y)\ \right\}, \tag{16}$$

No initial value is negative or zero. If the initial number of healthy cells or single-infected cells is below 1, we set it to 1. We observe that convergence occurs regardless of the initial conditions. Figure 6 represents the distribution of relative error compared to the theoretical equilibrium value at times $T = 90$ days and $T = 180$ days. Only the error on $v_3$ is shown for visualization purpose. That is :

$$e_{v_3}(t = T) = \frac{v_3(t = T) - v_3^*}{v_3^*} \tag{17}$$

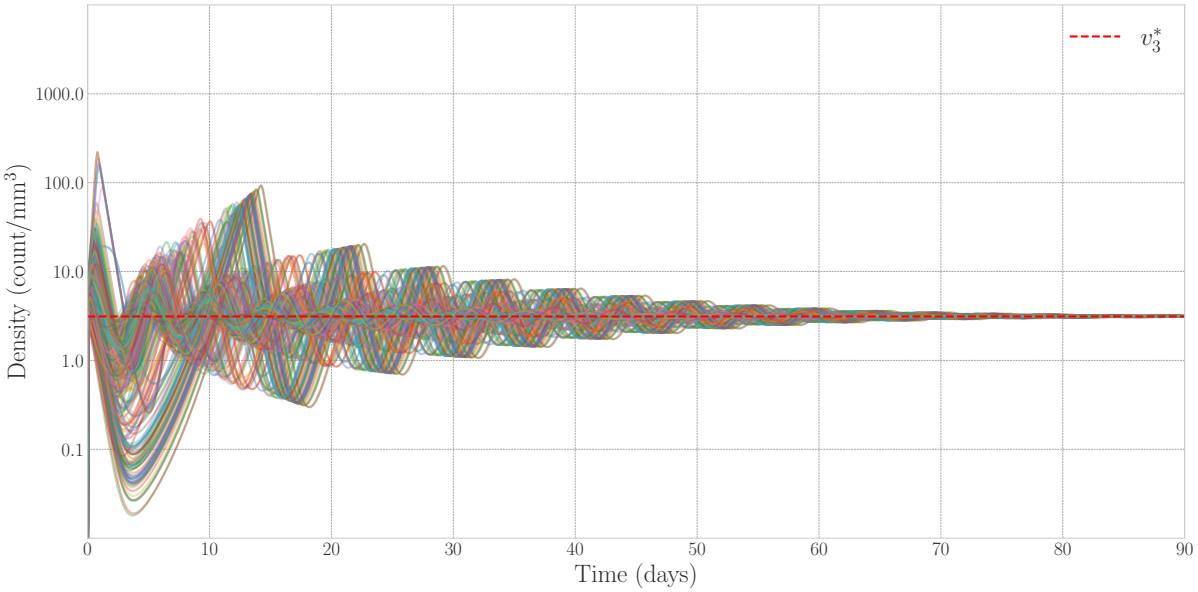


Figure 5: Convergence of the solver from different initial conditions. N = 500 simulations are carried out in the case $\alpha = 4\beta$ and whose initial conditions are sampled from a uniform distribution: $\boldsymbol{x}|_{t=0} \sim \max\left\{\mathcal{U}(\boldsymbol{x}^* \pm 125\%\boldsymbol{x}^*), 10^{-6} \text{ or } 1(\text{for } x \text{ and } y)\ \right\}$. Qualitatively, convergence can be observed since all simulations are bounded. Fig. 6 quantifies the errors of those simulations.

It is clear that the relative error is low and decreases with simulation time across all simulations and more importantly, remains bounded. The error is approximately 0.1% at 90 days and 0.08% at 180 days of simulation. Additionally, the error is centered around 0, indicating that the numerical method does not introduce a bias with respect to the equilibrium value.

One can also argue that the implementation is verified, since it gives similar results to those of [27], which are carried out on a different machine and using different numerical tools (`MATLAB`).

$\Longrightarrow$ The *in silico* model meets all acceptability criteria for verification;

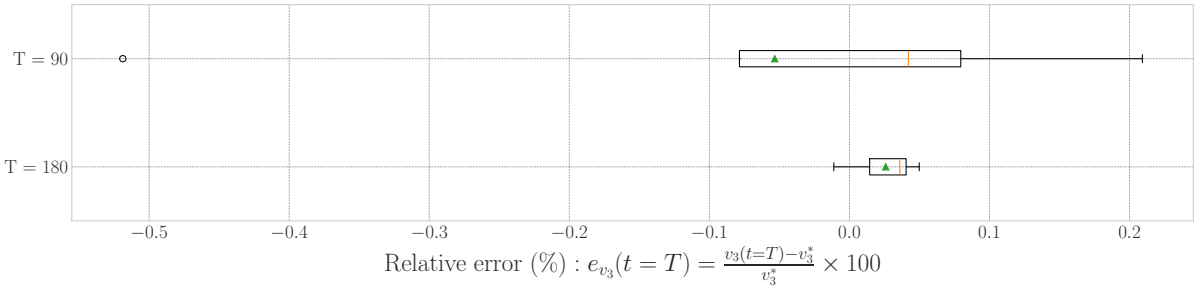


Figure 6: Distribution of the relative error for $v_3$ at $T = 90$ days and $T = 180$ days of simulation. The relative error is calculated as $e_{v_3}(t = T) = \frac{v_3(t=T)-v_3^*}{v_3^*}$, with a small error ( 0.1% at 90 days, 0.08% at 180 days). The error remains bounded and is centered around 0, indicating no bias in the numerical method.

**Validation**

The acceptability criteria for validation require comparing the model to *in vivo* data for the sanity check of the single HIV-1 infection, and to *in vitro* data for the complete model that includes double infection. The quantities of interest are $\xi$ and $\eta$. Following the approach of [27] and consulting the data cited in the referenced works ([30, 24, 22]), the comparison data are summarized in Table 4.

Firstly, *in vivo*-based validation compares the temporal position of the infection peak during single infection and the equilibrium value of healthy cells. More specifically, [24] reports a peak occurring within the first few weeks of infection (between 1 and 4 weeks) after HIV-1 infection, and stabilization of the healthy population at a level approximately two orders of magnitude lower than the uninfected state. Figure 1 validates these results for the model, showing a 60-fold reduction in the healthy population and a peak occurring around the 10th day post-infection. Notably, the position and magnitude of the peak are highly sensitive to infection parameters and initial conditions—this sensitivity is also confirmed *in vivo*. Thus, *in vivo* validation reinforces the necessity of sensitivity analysis (SA) for effective model calibration. The behavior and dynamics of single infection in the double-infection model qualitatively align with established single-infection dynamics, serving as a robust *sanity check* based on *in vivo* data. Additionally, [22] corroborates the theoretical results and highlights parameter sensitivity, which the model replicates. *In vitro* data [30] also show the peak occurring between days 10 and 15 post-infection.

Secondly, *in vitro*-based validation is critical as it directly relates to double infection, the main focus of the model. Furthermore, the model's context of use (CoU) specifically targets *in vitro* designs. Beyond validation, differences observed must also be highlighted. [30] reports a reduction in the percentage of single-infected cells from approximately 80% to about 10%. Figure 7 illustrates these percentages in our model. While quantitative differences exist, the magnitudes are consistent, and the model qualitatively reproduces the reduction relative to single infection. Notably, the observed error primarily pertains to the single-infection case, where the article reports about 80% of cells as HIV-1 positive, whereas the model estimates around 70%. Considering the CoU focuses on double infection rather than single infection and given the low-risk context not requiring strict quantitative reproduction, the model can confidently be deemed validated for these results. [30] also reports a viral load reduction of approximately $\eta \approx 0.6\%$ (Figure 7A). However, our model predicts a more conservative reduction of $\eta \approx 8\%$, representing a significant difference. Based on the reference study, it is possible to achieve a viral load reduction of $\eta \approx 0.1\%$, closer to experimental results, using parameter values $\alpha = 4\beta$, $c = 10000$, and $q = 0.5$. In practice, the authors of [27] explain the complexity of parameter measurement, suggesting this as an avenue for future research. Following this philosophy, the model is validated against *in vitro* data, but continuous adaptation and calibration will be essential as future experiments refine parameter estimates.

$\Longrightarrow$ The *in silico* model satisfies all acceptability criteria for validation;

| **Source of Data** | **Type of Data** | **Comparable Quantities** |
|---|---|---|
| Widely accepted in the scientific community, the mechanism and dynamics of single infection are presented in [22, 24]. The *in vivo* data cited in these articles originate from [17]. Additional *in vivo* data sources are available in [22, 24]. | *In vivo* data (large samples of patients from various studies and heterogeneous populations). Specifically, [17] refers to a controlled clinical trial, where the data used are from the *placebo* arm consisting of $n = 38$ patients with weekly measurements over six months. | Since the references focus on single infection, quantities $\xi$ and $\eta$ are not meaningful.<br><br>[22] highlights key parameters influencing equilibrium states and validates the theoretical results of equilibrium positions for single infection presented in #4. However, [22] does not present numerical simulations for the single-infection system, so qualitative comparisons are not possible.<br>$\Longrightarrow$ Equilibrium positions for the single-infection case.<br><br>[24, 17] illustrate the temporal dynamics of the infection peak and the reduction of healthy cell populations, particularly $\frac{x_2^*}{x_1^*}$.<br>$\Longrightarrow$ Dynamics of the infection peak for the single-infection case. |
| The original study introducing the construction of a recombinant virus targeting HIV-1-infected cells [30] is based on *in vitro* experiments. The mechanism of double infection was tested and reported. | *In vitro* data from 96-well plate experiments with duplicates for each experiment, and sometimes triplicates. | Quantities $\xi$ and $\eta$ are not directly reported in *in vitro* experiments. However, a "dynamic" value of $\eta$ can be estimated—where "dynamic" implies that the ratio is calculated based on the measured free infectious virions over time rather than equilibrium values. Additionally, the percentage of cells in a single-infection state after introducing the recombinant virus into the system, i.e., $\frac{y}{x+y+z}$, can be compared. |

Table 4: Source and type of data used to validate the *in silico* model, along with the quantities compared.

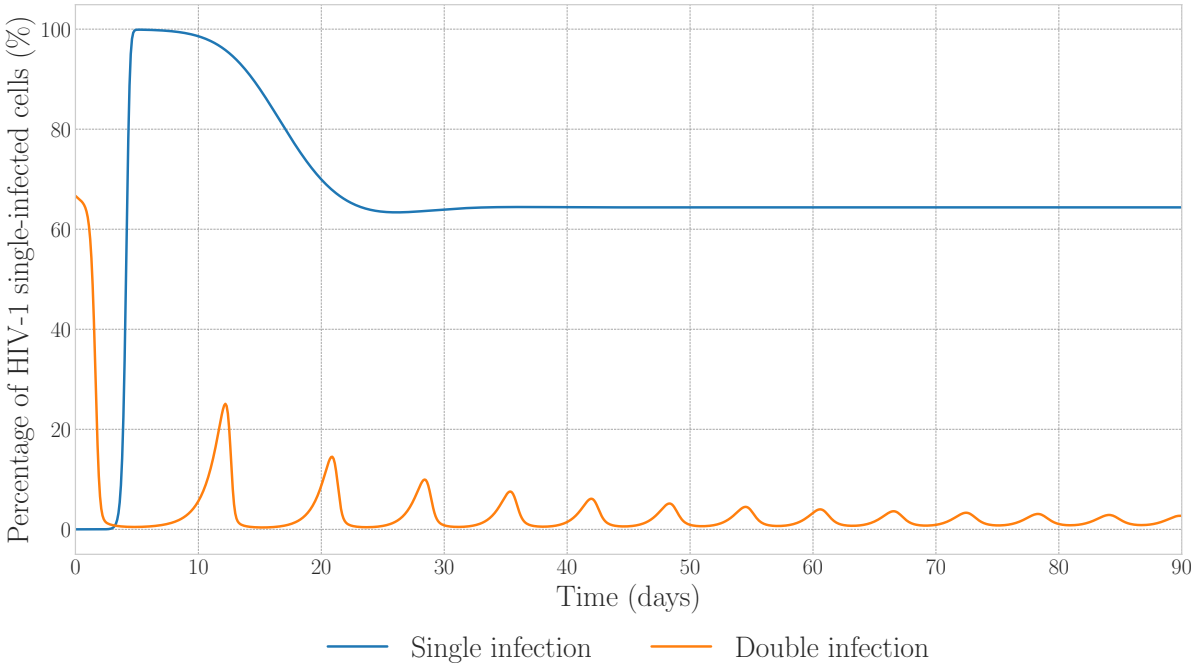


Figure 7: **Reproduction of Figure 7A from [30].** The figure illustrates the percentage of HIV-1-infected cells before and after recombinant virus treatment. The model estimates an initial percentage of approximately 70% of infected cells in the single-infection case, compared to 80% reported in [30]. Following double infection, the model predicts a reduction to about 10%, qualitatively aligning with experimental results. While the initial single-infection percentage exhibits some quantitative differences, the model successfully captures the relative reduction in HIV-1-positive cells, supporting validation. It is noteworthy that in the reference study, the oscillations are less visible, likely due to the sampling frequency *in vitro*.

### 3.1.4 Sensitivity analysis

#### Methodology

Global sensitivity analysis (SA) is conducted in the context of variance-based SA, specifically by computing first-order Sobol indices [32] as well as total-order indices introduced by Saltelli et al. [29]. These indices assess the influence of different model parameters on the variance of the model outputs. The `SAlib` library is used for computing these indices, providing a robust and flexible implementation suited for this type of analysis.

The interpretation of this SA can be summarized as follows:

- The **first-order Sobol indices** quantify the contribution of an individual parameter to the total variance of the model outputs. In other words, they measure the direct impact of a parameter on the model results, independently of its interactions with other parameters.
- The **total-order indices of Saltelli** account not only for the direct effect of a parameter on output variance but also for its interactions with all other model parameters. These indices thus quantify the global influence of a parameter, including its interactions.

The methodology used for this analysis is outlined below:

1. **Definition of input parameters:** A range of values is defined for each model parameter based on information available in the literature.
2. **Sampling plan construction:** A sampling matrix is generated using a quasi-Monte Carlo method adapted for Sobol index computation. The Saltelli method [29] is used here to ensure an efficient and accurate estimation of sensitivity indices.
3. **Simulation execution:** Each combination of parameters from the sampling plan is simulated using the *in silico* model. The model outputs include the quantities of interest, notably $\xi$ and $\eta$, which are then used to compute sensitivity indices.
4. **Results analysis:** First-order Sobol indices and total-order indices are computed using the `SAlib` library. The results are visualized to identify the parameters with the most significant influence on the quantities of interest.

Finally, the results of the global SA are presented in Figure 8. This analysis highlights the critical parameters requiring careful calibration and provides valuable insights for guiding research and optimizing the design of the recombinant virus. Additionally, it helps identify parameters for which reducing uncertainty could significantly impact the accuracy of model predictions.

#### Parameter distributions

The Saltelli sampling method constructs a quasi-random sequence adapted for sensitivity analysis based on Sobol indices. In this study, model parameters are assumed to follow uniform distributions within variation ranges defined from the literature. These ranges represent the uncertainties associated with each parameter and their possible biological variability.

The parameter ranges are summarized in Table 5, specifying the minimum and maximum values used for sampling. These bounds were chosen to cover a realistic spectrum of the studied biological scenarios. When searching the literature, the priority order for data sources is *in vivo* > *ex vivo* > *in vitro* > *in silico* > *hypothesis*. The "*hypothesis*" source corresponds to cases where no data is available in the literature. In this case, the distribution bounds are centered around the value given by #2, with $\min_i = \theta_i \times 0.1$ (conservative) and $\max_i = \theta_i \times 10$ (optimistic).

A uniform distribution was used for each parameter, as it does not require any prior assumptions about the probability of each value and ensures that the support covers only physiological (positive) values. This approach is suitable for exploratory scenarios and guarantees a balanced coverage of the parameter space.

It is important to note that a preprocessing step is performed before evaluating the sensitivity indices. Specifically, only stable double-infection scenarios are retained, meaning that the criteria $R_v > 1$ and $R_w > 1$ must be satisfied. The interest of this filtering is to focus on steady-state double infections ($R_w > 1$) (rather than merely transient ones) in a population affected by AIDS ($R_v > 1$). Parameter sets that do not meet these criteria are re-sampled.

| Parameter | Min value | Max value | References |
|---|---|---|---|
| $d$ | 0.001 | 0.1 | *In vivo* values are around $d \approx 0.0014$ [19]. In practice, modeling requires a higher value for this type of model [22] to be validated against both *in vitro* and *in vivo* data, as shown by [33]. The value range is chosen under a *hypothesis* that aligns with the *in silico* modeling of [27], whose simulations are validated using *in vitro* data. |
| $\lambda$ | 1 | 3 | The value range is based on a combination of *in vivo* and *in vitro* data [24, 17]. The variability reflects both differences in the percentage of active CD4+ cells after their generation and the existing variability in CD4+ density in blood plasma. It should be noted that the latter, although sourced from *in vivo* data, is established based on healthy individuals. |
| $\beta$ | 0.0027 | 0.007 | Many sources provide data on the HIV-1 infection rate from *in vivo* and *in vitro* experiments [24, 33, 20, 16]. The value range corresponds to the extreme values found in the literature. |
| $a$ | 0.3 | 0.5 | The initially used value of $a = 0.33$ [24] is already quite conservative. The literature also suggests values of $a = 0.49$ and $a = 0.39$ [33, 23]. The range is constructed based on these *in vivo* and *in vitro* data to encompass the extremes. |
| $\alpha$ | 0.001 | 0.16 | *Hypothesis* centered around the estimated *in vitro* value from [30]. The range is intentionally widened, as engineering allows for a broad variability of achievable values. |
| $b$ | 0.2 | 20 | *Hypothesis* centered around the estimated *in vitro* value from [30]. |
| $k$ | 30 | 120 | Slightly extended range, but it includes the extremes based on *in vitro* data suggesting $k \approx 46$ and $k \approx 60$ [6, 4]. The upper bound for a highly acute infection (high $k$ value) is significantly extended. |
| $u$ | 1 | 3 | The value range is derived from the extreme values reported in [33, 23] and estimated *in vivo*. |
| $c$ | 1000 | 10000 | The range is very broad and not centered around the reference value $c = 2000$. *In vitro* data suggest $c \approx 3333$, and for an effective viral therapy, the model might explore higher values of $c$ [30]. Furthermore, for high values of $c$, the oscillatory behavior disappears, which could be desirable in the design of the therapy. |
| $q$ | 1 | 3 | No available data, but the *hypothesis*-based range is reduced. In practice, it is likely that $q \approx u$, as there is little control over clearance in blood plasma. |

Table 5: Parameter ranges used for the global sensitivity analysis. Each parameter is assumed to follow a uniform distribution, with the bounds reported in this table. When multiple data sources are available, the priority order is *in vivo* > *ex vivo* > *in vitro* > *in silico* > *hypothesis*.

**Simulation and computation of indices**

Each sample generated from the defined ranges is evaluated using the *in silico* model. The quantities of interest, $\xi$ and $\eta$, are computed for each combination of parameters, providing the necessary data for sensitivity analysis. For SA, oscillations are not considered, and only the theoretical equilibrium value is analyzed. Even though this is not ideal, the philosophy of SA remains to identify a ranking and measure the importance of calibrating different parameters. At this stage, it is not yet a question of evaluating the performance of a treatment on a heterogeneous population. However, it is important to highlight the limitations this may introduce in SA. The main limitation is that the purely quantitative aspect is probably imprecise—though this is, in any case, dependent on the distributions associated with each parameter. Given the limited knowledge of their true distributions, the quantitative aspect remains uncertain. Nevertheless, this should not significantly impact the qualitative aspect of SA.

First-order and total Sobol indices are computed using the `SAlib` library. This library provides a robust implementation of the methods proposed by Saltelli et al. [29], enabling the calculation of both direct and global contributions of parameters to the variance of the model outputs.

The complete process follows these steps:

1. Generation of parameter samples using the Saltelli method.

2. Estimation of $\xi$ and $\eta$ values according to their definitions given in Eq. (15) and Eq. (14) and based on #4 for equilibrium values. Note that <u>all</u> equilibrium values are adjusted (i.e., those for single infection and double infection) to ensure consistency for a patient initially infected by HIV-1 and later undergoing viral therapy with the recombinant virus.

3. Analysis of results to obtain Sobol indices using estimators (18) and (19), including:

   - First-order indices ($S_i$), representing the direct contributions of parameters.

   - Total indices ($S_{T_i}$), accounting for parameter interactions.

$$S_i^Y = \frac{\mathbb{V}_{\theta_i}\left[\mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[Y|\theta_i]\right]}{\mathbb{V}[Y]} \quad \xrightarrow{\text{Monte Carlo Estimator [29]}} \quad \frac{\frac{1}{N}\sum_{k=1}^{N}\left[Y_{B,k}\left(Y_{A_i,k}-Y_{A,k}\right)\right]}{\frac{1}{N-1}\sum_{k=1}^{N}\left(Y_{A,k}-\bar{Y}_A\right)^2} \tag{18}$$

$$S_{T_i}^Y = 1 - \frac{\mathbb{V}_{\boldsymbol{\theta}_{\sim i}}\left[\mathbb{E}_{\boldsymbol{\theta}_i}[Y|\boldsymbol{\theta}_{\sim i}]\right]}{\mathbb{V}[Y]} \quad \xrightarrow{\text{Monte Carlo Estimator [29]}} \quad \frac{\frac{1}{2}\frac{1}{N}\sum_{k=1}^{N}\left[\left(Y_{A,k}-Y_{A_i,k}\right)^2\right]}{\frac{1}{N-1}\sum_{k=1}^{N}\left(Y_{A,k}-\bar{Y}_A\right)^2} \tag{19}$$

where the following notations are used:

- $Y$: Model output *in silico*, corresponding to one of the quantities of interest, such as $\xi$ or $\eta$.

- $\theta$: Vector of model parameters. Each component $\theta_i$ represents a specific parameter, while $\boldsymbol{\theta}_{\sim i}$ represents the vector of parameters excluding $\theta_i$ (all other parameters).

- $\mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[Y|\theta_i]$: Conditional expectation of $Y$ when parameter $\theta_i$ is fixed, with other parameters $\boldsymbol{\theta}_{\sim i}$ varying according to their distribution.

- $\mathbb{V}[Y]$: Total variance of the output $Y$.

- $\mathbb{V}_{\theta_i}\left[\mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[Y|\theta_i]\right]$: Variance of the conditional expectation, reflecting the direct contribution of $\theta_i$ to the total variance of $Y$.

- $\mathbb{V}_{\boldsymbol{\theta}_{\sim i}}\left[\mathbb{E}_{\boldsymbol{\theta}_i}[Y|\boldsymbol{\theta}_{\sim i}]\right]$: Remaining variance after fixing all parameters except $\theta_i$, used to quantify interactions between $\theta_i$ and other parameters.

- $N$: Number of samples generated in the Monte Carlo estimation of Sobol indices.

- $Y_{A,k}$, $Y_{B,k}$, $Y_{A_i,k}$: Model outputs corresponding to different parameter combinations obtained via sampling matrices $A$, $B$, and $A_i$, following the Saltelli method.

- $\bar{Y}_A$: Mean of outputs $Y_{A,k}$ over the $N$ samples, used for total variance calculation.

**Results**

Since the SA method used does not require simulation via ODE resolution for each parameter set, a large number of samples can be utilized. The presented results are based on $N = 2^{17} = 131072$ individuals. Fig. 8 highlights the results of SA.

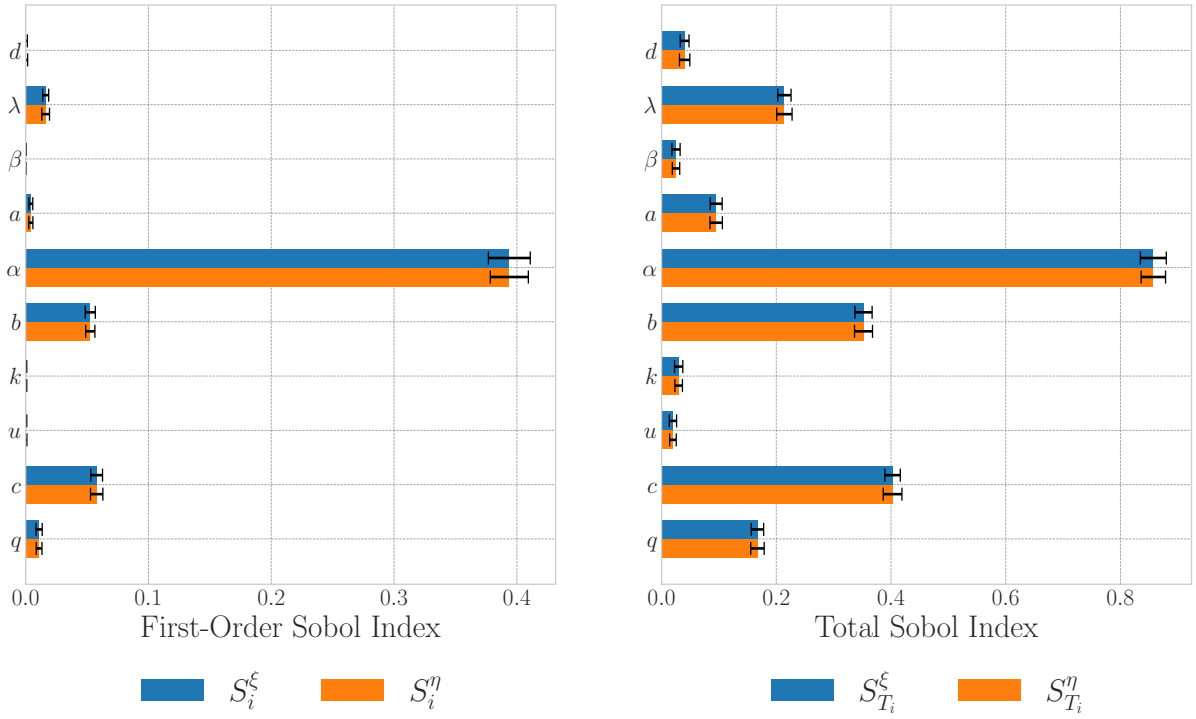


Figure 8: Global sensitivity analysis of the model. The figure shows the first-order (left) and total (right) Sobol indices for each model parameter, computed separately for the quantities of interest $\xi$ and $\eta$. The vertical axis lists the model parameters, while the horizontal axis indicates each parameter's contribution to output variance. Error bars represent confidence intervals of the estimates. A noticeable difference between first-order and total sensitivity highlights the importance of parameter interactions.

The sensitivity analysis clearly shows that the model is more sensitive to some parameters than others. A difference between first-order and total sensitivity is also observed.

On one hand, the most influential parameters are $\alpha$, $b$, $c$, and $\lambda$. Additionally, $d$, $q$, and $a$ are of intermediate importance—mostly through interactions with other parameters. From a biological perspective, this SA provides two insights:

1. The parameters $\alpha$, $b$, and $c$ are exclusively related to the properties of the recombinant virus. Their strong influence on output variability suggests a priority in correctly designing these properties in the laboratory. The SA does not necessarily indicate the need for precise estimation but rather emphasizes the importance of a robust design through genetic engineering.

2. The parameter $\lambda$ as well as $d$ and $a$ (likely $u \sim q$ as well) are associated with the HIV-1 virus and cannot be manipulated in the laboratory. SA thus informs us of the importance of *calibrating* these parameters and estimating them accurately to best characterize their *in vivo* distribution. Poor calibration could lead to underestimating or overestimating the potential and expected effects of viral therapy.

The error estimates from the `SAlib` library provide confidence in the convergence of SA results.

These results are highly significant as they guide researchers on which parameters should be estimated with high precision and inform decision-making regarding the credibility of results, despite known challenges in parameter estimation.

Thus, it would likely be crucial to improve the estimation of *production rate of new host cells*, *death rate of host cells*, *death rate of HIV-1 infected cells*, and *removal rate of recombinant virus* before using the model for decision-making. Meanwhile, during the design phase, the focus should be on *infection rate by recombinant virus*, *death rate of double-infected cells*, and *production rate of recombinant viruses by double-infected cells*, especially if the process is highly costly.

### 3.1.5 Credibility matrix and limitations

This section aims to establish a *credibility matrix* (CM) as introduced by [21]. This CM should, on its own, provide enough information to assess the credibility of the model and therefore must present the QoI, the CoU, the main points of the risk-based analysis, the resulting acceptance criteria, and the key elements from the VV and SA steps that provide a (positive or negative) response to the acceptance criteria. The CM serves as an identity card for the model's credibility and, in my opinion, is essential for transitioning models from theory to practice and for ensuring effective communication with institutions such as the FDA or EMA, which make decisions in the medical field.

It is obviously very difficult to imagine a CM that is adaptable to all situations and models, but the set of elements introduced by [21] in these CMs seems necessary for the credibility of a model.

The credibility matrix related to CoU 1 is presented in Table 6 and constitutes a good summary of everything that has been presented in the rest of Section 3.1. It is because we are able to present such a CM that Section 3.2 can be approached with "confidence".

#### Limitations

We can highlight the limitations identified by the credibility analysis of the model. There are two main limitations for the model. The first is intrinsic to the model itself and concerns the simplification of reality. The neglect of the recovery of immune system functions is a critical point of the model and could significantly hinder the applicability of the results. In practice, either it would need to be confirmed via *in vivo* experiments, or alternatively, *ex vivo* or *in vitro* experiments to ensure that this assumption does not overly compromise the quality of the predictions and that they remain sufficiently accurate for the CoU. Alternatively, the model could be *complexified* as suggested in certain references cited in Section 2.4. In other words, there is a *lack of data* to validate the model *in vivo*. The second limitation is independent of the model and concerns the estimation of parameters. The sensitivity analysis shows a strong sensitivity to certain parameters related to the population of treated individuals rather than the treatment itself. It is crucial to ensure, based on *in vivo* measurements, that the values and distributions used are representative of the target population. Ultimately, these two issues can be addressed in the same way, i.e., by implementing data collection *in vivo*. However, it should be noted that refining parameter measurements would be easier from an ethical standpoint compared to testing the model's validity *in vivo*, as the latter would require injecting the recombinant virus into individuals. Therefore, a two-step approach should be followed: first, refine the parameters, and subsequently, for example, with the support of *in silico* evidence and other types of proof, conduct *in vivo* trials.

Table 6: ***Credibility Matrix for the HIV-1 Viral Therapy Model***

| **Credibility Matrix (3 pages)** | |
|---|---|
| **Drug** | Viral therapy as an approach to the treatment of HIV-1. This therapy involves using a recombinant virus to induce a double infection mechanism that competes with HIV-1, potentially reducing viral load and promoting immune recovery. |
| **Type of model** | The model is a mechanistic dynamic system based on a set of ordinary differential equations. It represents interactions between (1) healthy CD4+ T cells, (2) HIV-1-infected cells and (3) recombinant virus-infected cells. |
| **Scientific Question(s) of interest (QoI)** | The main research question is whether viral therapy, via a recombinant virus, is a feasible and effective approach in HIV-1 treatment.<br>• Does the introduction of a recombinant virus significantly impact the key quantities of interest (HIV-1 viral load and CD4+ cell recovery)?<br>• To what extent is the therapy effective, and how does it compare to a single HIV-1 infection scenario?<br>• What viral characteristics should be optimized in recombinant virus design to maximize therapeutic impact?<br>• How robust are these effects in the presence of biological variability across different patients? |
| **Context of use (CoU)** | The model is used to guide in vitro research and development for engineering recombinant viruses targeting HIV-1. The context of use (CoU 1) is strictly preclinical and does not extend to regulatory decision-making for clinical applications. The model assists researchers in:<br>• Prioritizing experimental studies by identifying the most promising recombinant virus candidates.<br>• Informing decisions on which viral characteristics should be further investigated in laboratory settings.<br>• Providing theoretical insights into viral interactions that would otherwise require extensive in vitro experimentation.<br>This model is **not** intended for direct clinical predictions, regulatory approvals, or patient-specific decision-making. |
| **Regulatory impact** | Minimal regulatory impact: The model is a research tool rather than a decision-making system for clinical trials or regulatory approval. It does not replace human trials and is not intended for clinical predictions. Regulatory agencies (e.g., FDA, EMA) would not use this model for direct evaluation of treatment efficacy or safety. |
| **Risk-based analysis of decision consequence** | A risk assessment matrix classifies this model under a **low-risk** category:<br>• **Decision consequences:** The primary risk is financial and operational (e.g., wasted experimental resources if model predictions are incorrect). There is no direct impact on patient safety.<br>• **Regulatory impact:** The model does not contribute to clinical decision-making, meaning regulatory risks are negligible.<br>**Overall risk level: Low**. |
| **Model-informed decision** | The model supports in vitro research by providing:<br>• Recommendations on which viral characteristics to prioritize for recombinant virus design.<br>• Identification of key biological parameters requiring precise calibration for experimental success.<br>• Theoretical validation for observed in vitro effects, helping refine experimental protocols.<br>Importantly, the model does **not** inform clinical or regulatory decisions. It is purely a preclinical research tool aimed at optimizing viral therapy development. |

| **Credibility Matrix** (continued, 3 pages) | |
|---|---|
| **Acceptability criteria (Precision level)** | The model must meet the following criteria to be deemed credible for its intended use. These criteria ensure the reliability of the model by establishing rigorous verification, validation, and uncertainty quantification processes:<br>• **Verification:** Ensuring numerical correctness and theoretical consistency.<br>– Solver convergence analysis must confirm numerical stability by demonstrating bounded solutions and adherence to equilibrium conditions.<br>– The model must correctly implement theoretical equilibrium results for HIV-1 infection dynamics as described in the reference literature [27]. Specifically, it should qualitatively reproduce Figures 2a, 2b, and 2c from [27]. This would ensure that the *user* has correctly implemented the ODEs system.<br>– Computation must be validated against a reference implementation, ensuring that the model behaves consistently with established benchmarks.<br>– The numerical tools used should demonstrate robustness, either through *Software Quality Assurance* (SQA) processes or a thorough convergence analysis of numerical errors.<br>– Additional credibility may be provided through a versioning system (e.g., Git) with established sanity checks, though this is not mandatory.<br>• **Validation:** Ensuring biological and experimental relevance.<br>– The model must be capable of reproducing known *in vivo* HIV-1 infection dynamics (*sanity-check*), including but not limited to:<br>∗ Timing of viral load peak.<br>∗ CD4+ cell depletion trends over time.<br>– The model must qualitatively replicate *in vitro* experimental results for recombinant virus therapy, ensuring its applicability in guiding laboratory research.<br>– Any discrepancies between model predictions and experimental data must be systematically analyzed and justified in terms of parameter estimation uncertainty. This includes explaining deviations based on known biological variability and measurement limitations.<br>– Validation should prioritize the quantities of interest, denoted as $\xi$ and $\eta$. If direct experimental data for these variables are unavailable, surrogate measures should be used, with appropriate justification.<br>– Given that the model aims to support *in vitro* research, it must be continuously updated and evaluated in response to new scientific discoveries and experimental findings (such as the *in vitro* experiments).<br>• **Sensitivity Analysis:** Ensuring robustness of model predictions.<br>– A detailed sensitivity analysis must be conducted to identify the key parameters that significantly influence model outcomes. This analysis should highlight the most critical calibration elements for future refinements.<br>– The model must explicitly characterize parameter uncertainty, with particular emphasis on:<br>∗ The recombinant virus infection rate.<br>∗ CD4+ cell recovery dynamics post-infection.<br>The parameter distributions must be justified and as close as possible as the known *in vivo* distributions. If no data are available in the literature, a sufficiently large range of values must be used to ensure that extreme values are accounted for in the global SA.<br>All choices must be documented and justified in terms of the CoU. |

| **Credibility Matrix** (continued, 3 pages) | |
|---|---|
| **Credibility activities results** | All the details can be found in Section 3.1. All the acceptability criteria have been met.<br>$\Longrightarrow$ The model has been verified.<br>$\Longrightarrow$ The model has been validated.<br>$\Longrightarrow$ Sensitivity analysis has been conducted.<br>In particular,<br>The solver's convergence has been confirmed, ensuring numerical stability. The model qualitatively reproduces Figures 2a, 2b, and 2c from reference [27], ensuring consistency with HIV-1 infection dynamics. Validation against a reference implementation shows behavior consistent with expectations.<br>The model correctly reproduces HIV-1 *in vivo* infection dynamics, including the viral load peak and CD4+ cell depletion of [24]. It also aligns with *in vitro* experiments on recombinant virus therapy. Some quantitative discrepancies are observed. The differences between model predictions and experimental data are explained by uncertainties in biological parameters. Special attention is given to the quantities of interest $\xi$ and $\eta$, with surrogate measures used when direct data are unavailable.<br>The global sensitivity analysis, based on $N = 131072$ samples, reveals that the most influential parameters are $\alpha$, $b$, $c$, and $\lambda$. Certain parameters ($d$, $q$, $a$) play an important role through their interactions. These results emphasize the necessity of a robust design for recombinant virus properties and precise calibration of parameters related to HIV-1. Estimates from `SAlib` confirm the reliability of sensitivity indices, guiding model improvement before its use in decision-making.<br>**To ensure reproducibility of the results, a *SEED* is enforced, and the code is available open-source on GitHub**[1] |

[1] https://github.com/julienbrandoit/GBIO0033—HIV-1-viral-therapy-VVUQ-and-in-silico-trials.git

## 3.2 Design of the in silico trial

Section 3.1 aimed to explore the credibility of the model and explicitly define its context of use within a specific question of interest. The results of this credibility analysis are synthesized in the credibility matrix 3.1.5, reinforcing our confidence in using the model.

This section of the report aims to address the question of interest using the *in silico* model by conducting an *in silico clinical trial*.

### 3.2.1 Specific objectives of the trial

The objective of this *in silico clinical trial* is, of course, to answer the QoI. In practice, we aim to determine the outcomes of viral therapy on a *heterogeneous* population (i.e., characterized by a distribution of parameters) to identify various aspects:

1. The effectiveness of the treatment with respect to the quantities of interest $\xi$ and $\eta$;
2. The identification of treatments leading to oscillatory behaviors, which are considered undesirable;
3. Whether there exist subpopulations that react fundamentally differently to the treatments;
4. ...

The *in silico clinical trial* is thus approached through the lens of uncertainty quantification (UQ) and represents a method that could guide the *in vitro* design decisions of the recombinant virus.

Two different treatments will be compared after being tested *in silico*. One treatment is defined by $c = 5000$, $\alpha = 0.004$, and $b = 2$, while the other is defined by $c = 2500$, $\alpha = 0.016$, and $b = 2$. We can see here that a trade-off has been deliberately chosen: in practice, both alpha and c should be large. We're in a situation where only one of the two can be increased. These values are consistent with what is currently feasible in *in vitro* design [30]. Additionally, we assume that the individuals receiving treatment have been infected with HIV-1 for a sufficiently long time and thus initially present at the equilibrium state of single infection (Box #4) with $R_v > 1$. A selection process is carried out to retain only those patients for whom the treatment is theoretically stable in dual infection, meaning $R_w > 1$. This choice aligns with the potential objective of such a therapy if it were to become commercially available.

### 3.2.2 Digital population

The advantage of conducting *in silico* trials is the ability to study a relatively large population without ethical concerns and to test the treatment on highly heterogeneous populations, i.e., in an environment and sample that are much less controlled than in *in vivo* clinical trials, since, obviously, no *in silico* patient can suffer negative consequences.

A central element in designing an *in silico* clinical trial is the design of the *population*. In practice, this means characterizing a population using probability distributions over different parameters and then sampling a certain number of individuals. To obtain results as close as possible to those that would be observed *in vivo*, calibrating these distributions is the cornerstone of a well-conducted *in silico* trial.

A literature review is necessary, and the *validation* and SA steps already conducted in Sections 3.1.3 and 3.1.4 can assist in this regard. In practice, the value range used for SA already represented the strongest physiological *prior* that could be constructed based on the literature review. Table 5, developed for SA, thus also serves as the starting point for designing our population.

However, two modifications are made and justified below:

* Only the parameters $d$, $\lambda$, $\beta$, $a$, $k$, and $u$ are sampled. These parameters are related to individual physiology and thus define a true *in vivo* population. The parameters $\alpha$, $c$, and $b$ are <u>not</u> random variables; they define the treatment applied to the population. We assume that genetic engineering enables the construction of recombinant viruses with *exact* precision in these parameters. This aligns with the CoU. It would be valuable to quantify the uncertainty associated with these parameters if viral therapy progresses to preclinical or clinical phases. Finally, $u = q$ since we assume

that viral clearance is non-specific. In summary, a population of size $N$ can be described by a parameter matrix $[P]$ and a treatment applied to this population by a vector $\boldsymbol{T}$.

$$[P] = \begin{bmatrix} \boldsymbol{d} & \boldsymbol{\lambda} & \boldsymbol{\beta} & \boldsymbol{a} & \boldsymbol{k} & \boldsymbol{u} = \boldsymbol{q} \end{bmatrix} \in \mathbb{R}_+^{N\times 5}, \quad [P] \text{ being a matrix of random variables;}$$
$$\boldsymbol{T} = \begin{bmatrix} \alpha & c & b \end{bmatrix} \in \mathbb{R}_+^3, \quad \boldsymbol{T} \text{ being a vector of fixed and known parameters;}$$

* While uniform distributions were initially chosen for SA, these distributions likely do not represent biological reality accurately. In practice, uniform distributions are *bounded*, and while it is certain that biological parameters do not take extremely large values, it is more plausible to say that such values are just very improbable rather than inherently bounded by an underlying mechanism. Moreover, uniform distributions assume that all values within the interval are equally probable, which contradicts the widely observed principles of homeostasis in biology. In reality, a population's distribution is expected to be concentrated around the mean. For all these reasons, Gamma distributions are chosen for the population parameters. The Gamma distribution has a strictly positive support, aligning with biological priors, and allows for representation of outliers.

The parameters of these distributions are selected such that the first and second moments (mean and variance) match those of the uniform distributions presented in Table 5. Since the bounds were initially chosen as extreme values, the standard deviation is reduced by a factor of 0.8 before moment matching. The mathematical details are provided in Appendix C. Figure 9 visualizes the distributions used for each parameter.

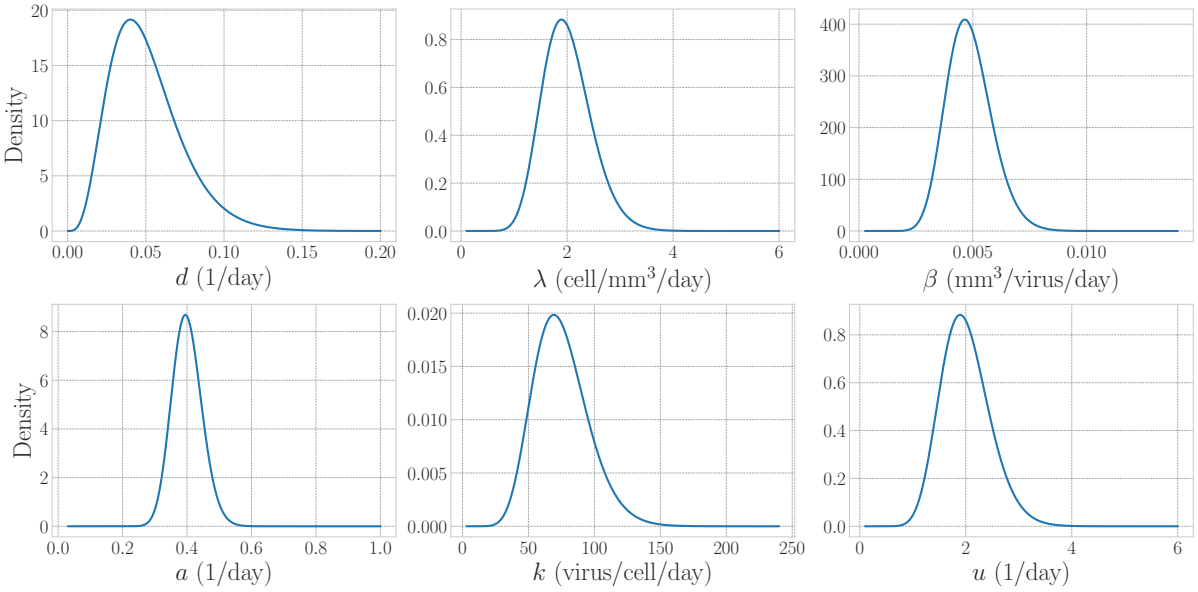


Figure 9: Visualization of the distributions used for each parameter in the digital population. The Gamma distributions are fitted to match the first and second moments of the uniform distributions from Table 5, with a rescaling factor of $f = 0.8$ applied to the standard deviation.

Two populations of size $N = 2500$ are sampled. Among these 5000 patients, all meet the criterion. This is a significant result as it suggests that both treatments could have broad applications, meaning they would not need to be specific to certain cases of HIV-1 infection.

$$\texttt{population 1:} \quad N \ \texttt{individuals} \xrightarrow{R_w>1} \texttt{100\% of patients can be effectively treated.}$$
$$\texttt{population 2:} \quad N \ \texttt{individuals} \xrightarrow{R_w>1} \texttt{100\% of patients can be effectively treated.}$$

Thus, the entirety of both populations can undergo the *in silico* clinical trial and follow the experimental protocol described in the next section.

### 3.2.3 Experimental protocol

Similar to an *in vivo* clinical trial, an *in silico* clinical trial must be detailed with a precise protocol clearly defined in advance. These protocols must be designed to support the QoI, ensuring statistical robustness while adhering to ethical principles.

For an *in silico* trial, the ethical principles differ from those in *in vivo* studies. Here, the focus is not on minimizing risks for the test population while balancing the necessity of results, but rather on two key aspects: first, ensuring rigor in the credibility assessment—something we consider to be effectively achieved here—and second, ensuring that the *in silico* results are well-categorized and representative of the target *in vivo* population. This requires a rigorous inspection of the data used, calibration methods, and so on. In practice, it is crucial to be aware of biases present in the data. For example, representation biases concerning women, minority populations, or individuals from non-Western regions are *omnipresent* in data science and particularly in medicine [38, 28, 11, 14].

For our *in silico* clinical trial, the challenge extends beyond the representation of *in vivo* data, as a significant portion of the parameters have not been estimated *in vivo*. This could be the most substantial point of criticism and will be discussed in detail in Section 3.2.5.

The experimental protocol is identical for both treatments. Each individual is simulated over a period of 250 days, with initial conditions for $x$, $y$, and $v$ set at equilibrium under single infection. On day 0 of the treatment, no double-infected cells are present ($z = 0$), and we inject the equivalent of 1 recombinant virus per $mm^3$ of blood plasma ($w = 1$). The injection is performed as a *bolus*, meaning in a single dose. Afterward, the individuals undergo no further interventions, and only the intrinsic system dynamics drive the treatment effects.

Starting from day 225, daily measurements of $v$ and $x$ are assumed to be taken. The feasibility of this protocol *in vivo* warrants consideration. Measuring $v$ experimentally is evidently possible, and measuring $x$ is also feasible [3]. Each patient is then classified into one of two categories: *oscillatory response* or *stable response*. A response is considered stable if the amplitude of oscillations measured in $x$ is less than 1. The choice of $x$ is based on the premise that the therapy aims to restore a *stable* immune system. If the primary goal were to reduce viral load, $v$ could also be used as a criterion. In practice, both are often targeted in HIV-1 therapies.

The values $\eta_{\min}$, $\eta_{\max}$, $\xi_{\min}$, and $\xi_{\max}$ are reported for oscillatory responses, while the values of $\eta$ and $\xi$ are reported for stable responses. The values of $\eta$ and $\xi$ are based on the measurements taken on the last day—i.e., day 250.

### 3.2.4 Results

There are many ways to present the results of clinical trials, each with its advantages and disadvantages. Figure 10 provides a view of the results, highlighting the uncertainty and variability of these results. A similar scale for the values being compared is enforced for simpler analysis. Numerical values are reported in Table 7. Combined with Figure 11, we can affirm that the response to both treatments is unimodal for both $\xi$ and $\eta$. This means that if subpopulations exist, they do not respond fundamentally differently, but rather the response is more or less pronounced. This is consistent with the theoretical analysis and the existence of a unique equilibrium point. Finally, Figure 12 illustrates some trajectories for each of the populations, where we have separated individuals with an oscillatory response from those with a stable response, for each population. This figure is not very useful for interpreting the results but confirms that the simulations appear to behave as expected—when compared to Section 3.1.

Referring to Figure 10 and Table 7, we observe that the treatment applied to population 2 yields more satisfying results. Indeed, it leads to both a smaller amount of oscillatory response while also reducing the viral load and achieving a greater recovery of a healthy CD4+ population. **Therefore, this treatment is more effective across all established criteria.** Thus, the informed decision via the model is that it is preferable to explore *in vitro* this second recombinant virus compared to the first in order to study whether it is feasible and behaves similarly to the *in silico* simulations.

Focusing only on this second population—the observations are similar for the first—one can also highlight that there indeed exist at least two subpopulations. Part of the population leads to an oscillatory

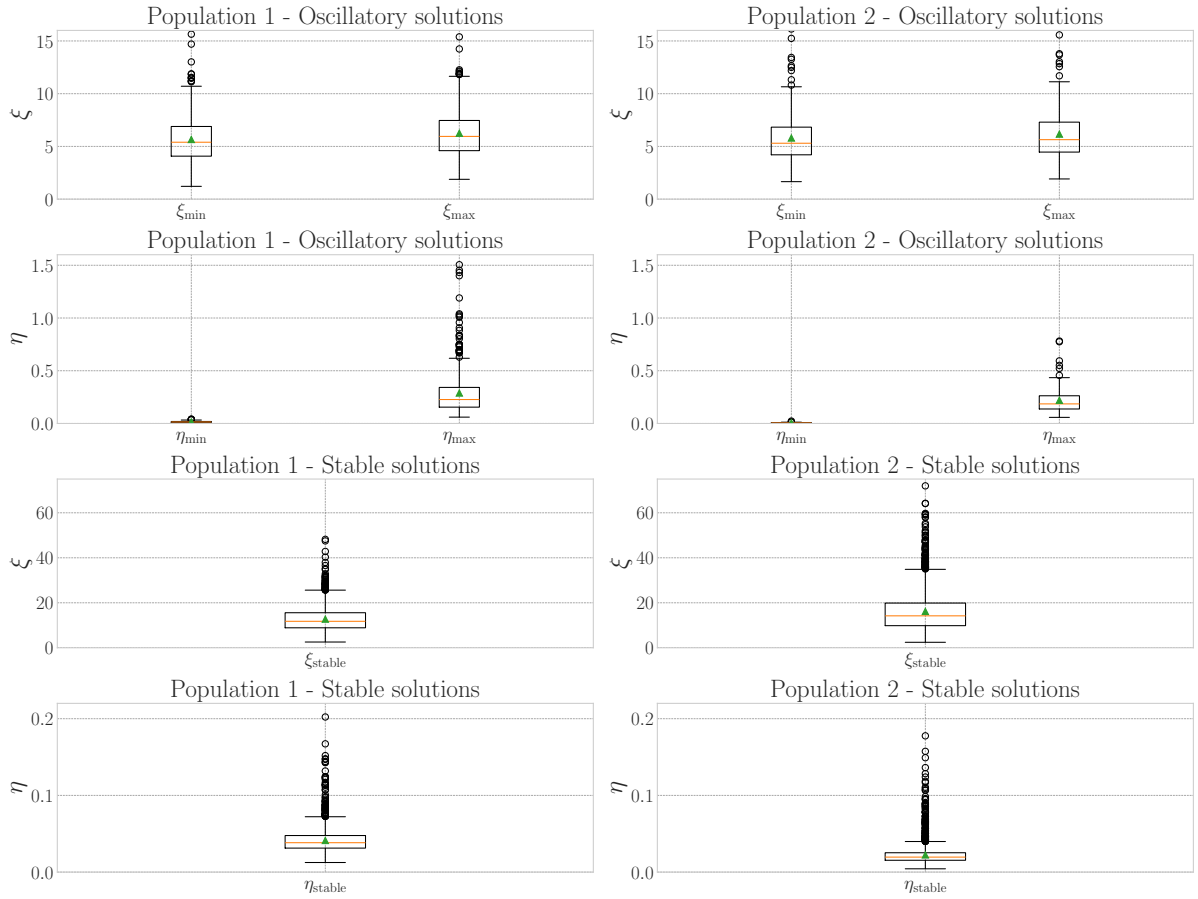


Figure 10: Boxplots of *in silico* clinical trial results, highlighting uncertainty and variability in the outcomes. The scale is forced for a direct comparison. On the left, population 1 treated with viral therapy characterized by $\boldsymbol{T}_1 = \begin{bmatrix}5000 & 0.004 & 2\end{bmatrix}$; on the right, population 2 is treated with $\boldsymbol{T}_2 = \begin{bmatrix}2500 & 0.016 & 2\end{bmatrix}$. The orange line indicates the median, and the green triangle represents the mean.

response, while another leads to a stable response. In addition to being a type of response that we wish to avoid, oscillatory responses are characterized by a *lower efficacy* of the treatment. One can compare the mean, median, and standard deviations from Table 7 or qualitatively observe Figure 12.

This result obtained *in silico* is extremely valuable as it may indicate a possible limit to the therapeutic value of this viral therapy. This limit should, of course, be studied in more depth, and this would involve determining strict criteria on the oscillations. In practice, one should study whether such behavior is *truly* problematic. One observation that can already be highlighted is that the oscillatory response sometimes leads to a *higher viral load* than before the treatment, as the value of $\eta_{\max}$ is, for some individuals, $> 1$. Thus, the perturbation of the system by the recombinant virus introduces a peak in HIV-1 infection similar to that observed during the initial infection.

A strong point of this *in silico* study, in addition to having identified the best of the two treatments and the existence of subpopulations, would be the ability to highlight what differentiates these two subpopulations. If these parameters are easily measurable *in vivo*, it could guide the decision-making of medical personnel in prescribing and managing viral therapy for HIV-1 patients.

An additional study is conducted by examining the distribution of parameters leading to a stable solution under therapy 2 and those leading to an oscillatory response. Figure 13 illustrates these distributions.

Statistical tests should be conducted to draw conclusions, but they are beyond the scope of this report. A qualitative analysis is performed. It is observed that the most important parameter for determining

the type of response is $u$, or the viral clearance. This quantity should be easily measurable *in vivo*, which is a positive aspect [23]. The parameters $\lambda$ and $d$ also appear to be important, but to a lesser degree.

### 3.2.5 Limitation of the *in silico* trials

The main limitation is, of course, the lack of *in vivo* data, as already discussed in Section 3.1.5. This limitation makes the quantification of the results complicated. In practice, however, the results can still be used qualitatively to guide *in vitro* research, with a perspective on the CoU. The ethical aspect of biases in the data cannot be addressed, and this is, in my opinion, a very negative element.

However, the *in silico* trial has highlighted fundamental results such as the existence of subpopulations and the fact that the treatment could be effective. Identifying what differentiates the subpopulations offers perspectives for research. One can imagine prescribing the therapy only if the value of $u$ is sufficiently low, or even decreasing the value of $u$ per medication.

Of course, the biological limits of the model are also a major cause for concern. In particular, there is a real need to assess whether the neglect of the immune system can still produce results consistent with biological reality once the recombinant virus has been introduced. The elements presented in 2.4 support this.

Although illustrated on a "toy problem", we clearly understand the interest and power of *in silico* trials, especially when accompanied by a credibility assessment.

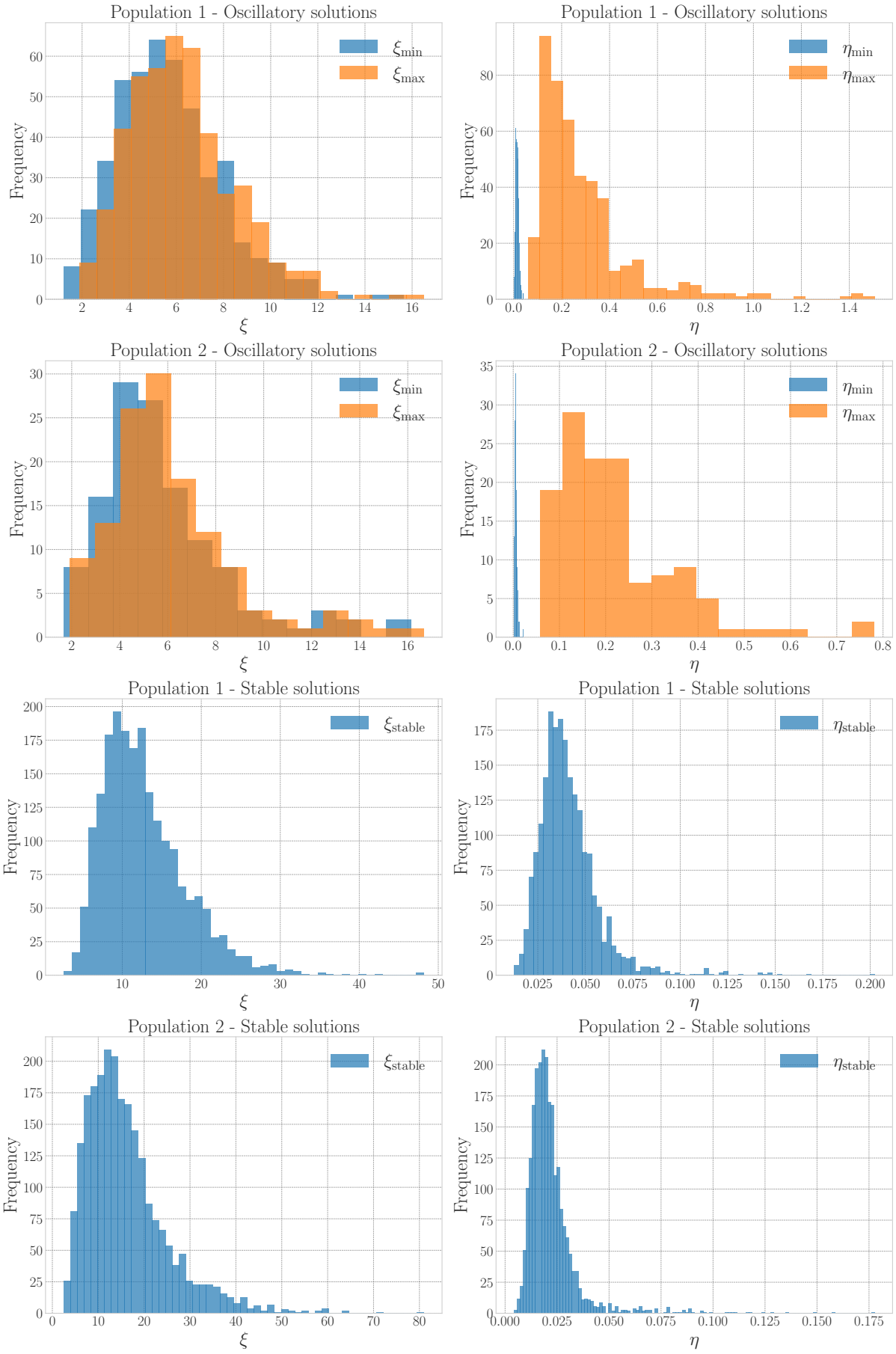


Figure 11: Histograms of the clinical trial results, showing the unimodal distribution of responses for both $\xi$ and $\eta$.

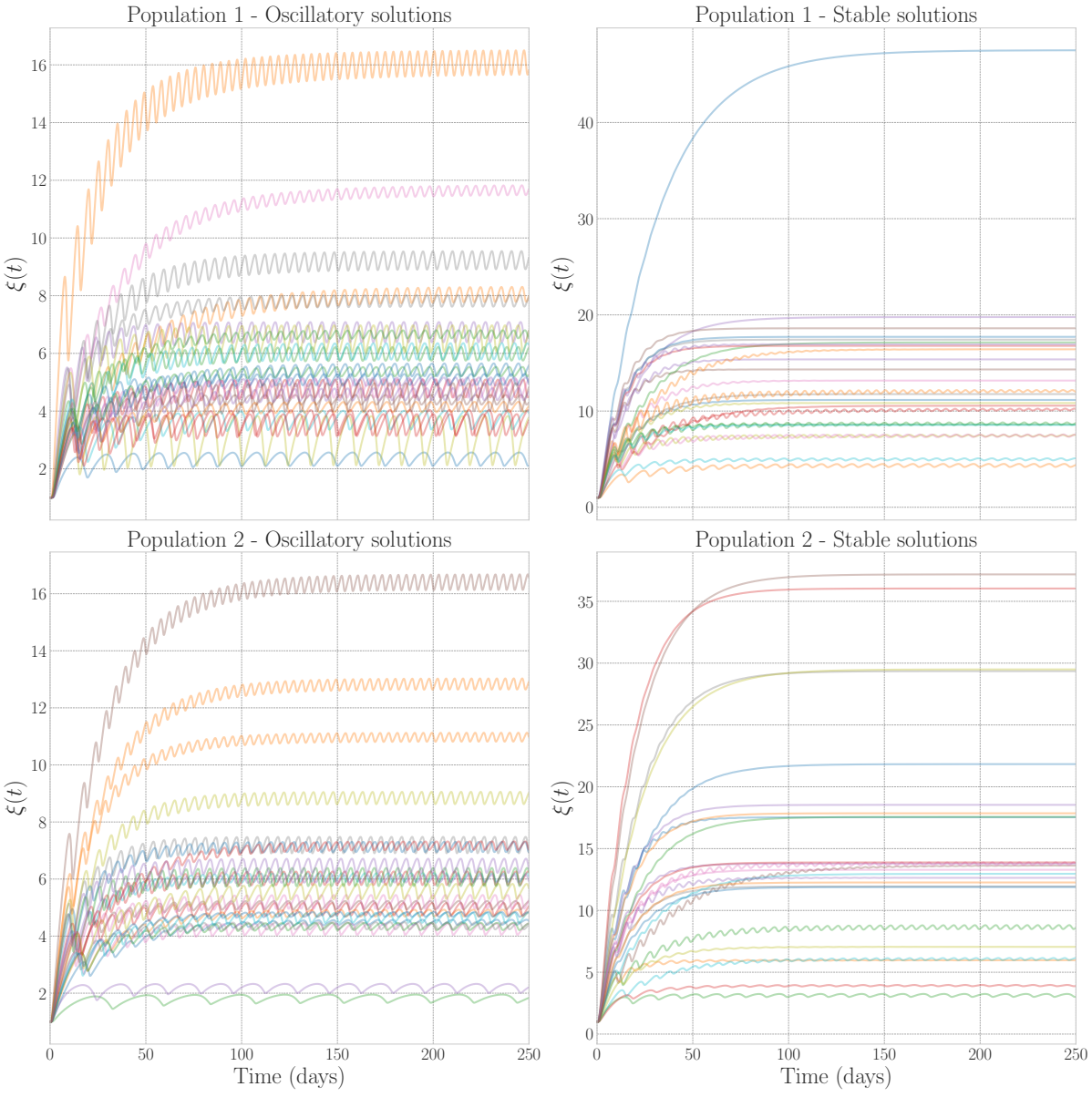


Figure 12: Illustration of trajectories for each population, distinguishing individuals with an oscillatory response and those with a stable response. 25 individuals per population are reported.

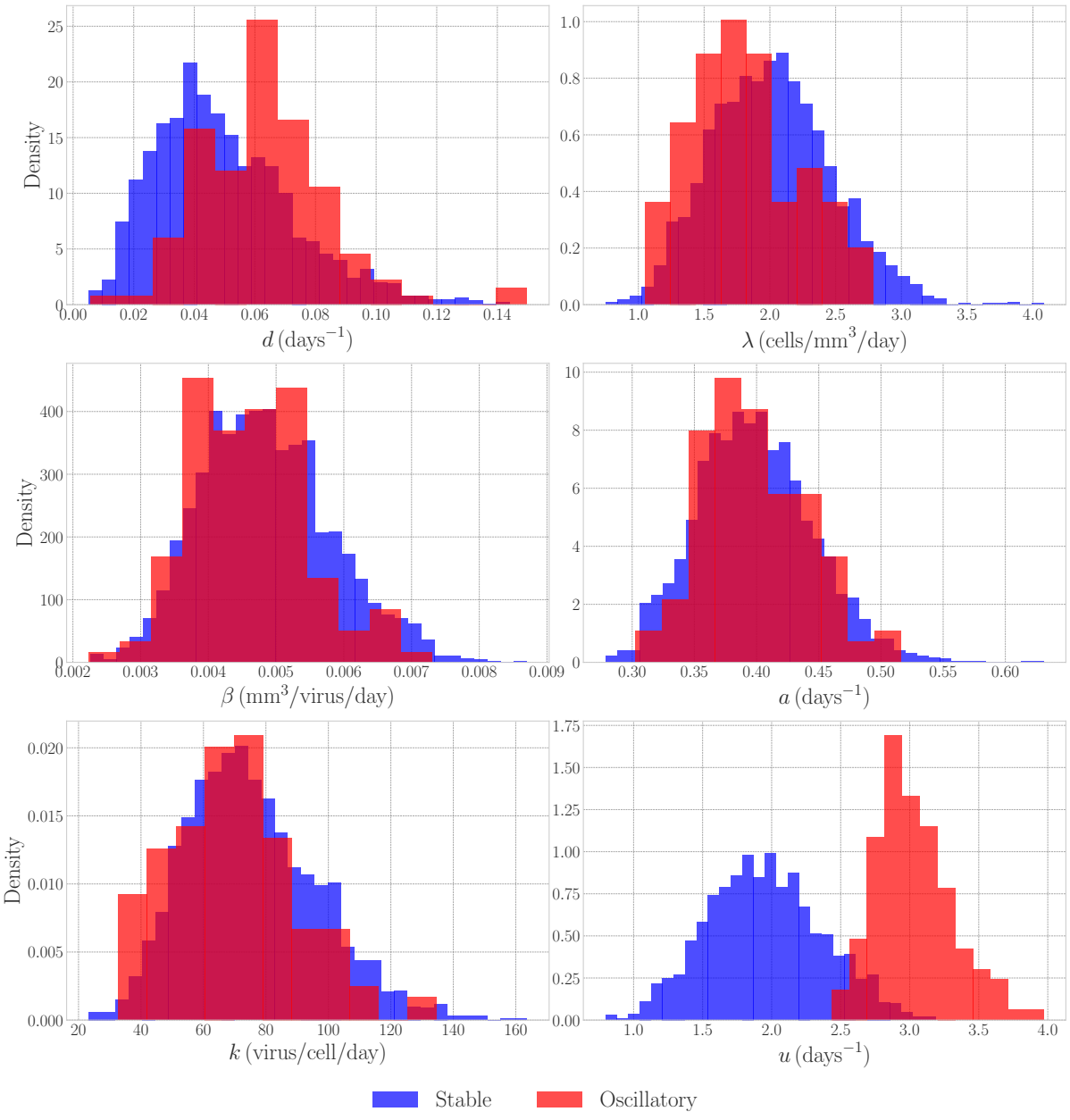


Figure 13: Distribution of parameters leading to a stable solution under therapy 2 and those leading to an oscillatory response. This figure illustrates how different parameter values influence the system's response. Viral clearance, $u$, is the parameter that seems to distinguish the two sub-populations.

| **Parameter** | **Population 1** | **Population 2** |
|---|---|---|
| Candidates | 2500 | 2500 |
| Retained ($R_v > 1, R_w > 1$) | 2500 | 2500 |
| Stable Response | 2046 | 2371 |
| Oscillatory Response | 454 | 129 |
| **Oscillatory Solutions** | | |
| $\boldsymbol{\xi_{\min}}$ | | |
| Mean | 5.66 | 5.80 |
| Standard Deviation | 2.23 | 2.70 |
| Min | 1.21 | 1.66 |
| Median | 5.40 | 5.30 |
| Max | 16.50 | 16.68 |
| $\boldsymbol{\xi_{\max}}$ | | |
| Mean | 6.25 | 6.17 |
| Standard Deviation | 2.25 | 2.72 |
| Max | 16.50 | 16.68 |
| Median | 5.95 | 5.64 |
| $\boldsymbol{\eta_{\min}}$ | | |
| Mean | 0.015 | 0.006 |
| Standard Deviation | 0.007 | 0.003 |
| Min | 0.002 | 0.002 |
| Median | 0.014 | 0.005 |
| Max | 1.51 | 0.78 |
| $\boldsymbol{\eta_{\max}}$ | | |
| Mean | 0.29 | 0.22 |
| Standard Deviation | 0.21 | 0.13 |
| Max | 1.51 | 0.78 |
| Median | 0.23 | 0.19 |
| **Stable Solutions** | | |
| $\xi_{\min}$ (Stable) | 2.56 | 2.44 |
| $\xi_{\max}$ (Stable) | 48.20 | 81.01 |
| $\eta_{\min}$ (Stable) | 0.013 | 0.004 |
| $\eta_{\max}$ (Stable) | 0.202 | 0.178 |
| $\xi_{\text{mean}}$ (Stable) | 12.75 | 16.19 |
| $\xi_{\text{std}}$ (Stable) | 5.39 | 9.13 |
| $\xi_{\text{median}}$ (Stable) | 11.77 | 14.20 |
| $\eta_{\text{mean}}$ (Stable) | 0.041 | 0.022 |
| $\eta_{\text{std}}$ (Stable) | 0.017 | 0.014 |
| $\eta_{\text{median}}$ (Stable) | 0.038 | 0.020 |

Table 7: Summary of results for both populations. All values are reported with two decimal precision. Stable solutions refer to the computed $\xi$ and $\eta$ statistics for stable responses. An effective treatment increases the value of $\xi$, decreases the value of $\eta$, and does not present oscillatory characteristics in the long term. Population 1 is treated with viral therapy characterized by $\boldsymbol{T}_1 = \begin{bmatrix}5000 & 0.004 & 2\end{bmatrix}$. Population 2 is treated with $\boldsymbol{T}_2 = \begin{bmatrix}2500 & 0.016 & 2\end{bmatrix}$.

# 4 Conclusion

In this report, we have critically analyzed an in silico model of viral therapy for HIV-1, focusing on its credibility, validation, and practical application. Our investigation was guided by the principles of Verification, Validation, and Uncertainty Quantification (VVUQ), with the overarching goal of assessing the model's potential to inform therapeutic development and decision-making. The report is structured around two main objectives: a critical analysis of the original article and the application of the model to conduct an in silico clinical trial.

## Critical analysis of the original article

The article [27] introduces a mathematical model that evaluates the effectiveness of recombinant virus therapy for HIV-1. The model considers the interactions between healthy CD4+ T cells, HIV-1-infected cells, and recombinant virus-infected cells. Through a rigorous analysis, we identified several strengths and weaknesses of the model from the perspective of VVUQ.

### Strengths

The model clearly outlines its assumptions and limitations, providing a solid foundation for future improvements. It validates against both in vitro and in vivo data, ensuring biological relevance. The equilibrium analysis provides important insights into the efficacy of the recombinant virus therapy. The model's deterministic nature simplifies complex biological interactions, enhancing interpretability.

### Weaknesses

The sensitivity analysis lacks quantification of uncertainty, particularly in the context of biological variability across different patients. The article does not fully address the regulatory impact and decision-making risks associated with the model. The model's assumptions, such as neglecting the immune response, limit its applicability to certain stages of HIV-1 infection. The parameter estimation challenges and significant population variability are not adequately addressed in the sensitivity analysis.

Despite these limitations, the model remains a valuable tool for guiding in vitro research and informing the design of recombinant viruses. The insights gained from the equilibrium analysis and sensitivity analysis can help researchers prioritize experimental studies and refine viral characteristics for optimal therapeutic impact.

## Application of the model in an in silico clinical trial

To further explore the model's utility, we conducted an in silico clinical trial on a virtual population, simulating two different treatments characterized by distinct recombinant virus parameters. The trial followed a rigorous methodological approach based on VVUQ, including risk analysis, verification, validation, and uncertainty quantification.

### Key findings

Treatment characterized by higher recombinant virus production ($c$=5000) and lower infection rate ($\alpha$=0.004) led to more pronounced oscillatory responses, while treatment with lower production ($c$=2500) and higher infection rate ($\alpha$=0.016) yielded more stable responses. The second treatment resulted in greater viral load reduction and healthier CD4+ cell recovery, making it more effective across all established criteria. The existence of subpopulations responding differently to the treatment was identified, with oscillatory responses being less favorable due to higher viral loads in some cases. The parameter $u$ (viral clearance) emerged as a key determinant of the response type, suggesting its potential as a **biomarker** for *personalized therapy.*

## Reflection on the results and limitations

The in silico trial underscores the complexity of viral therapy and the importance of considering biological variability. While the model provides valuable insights, several limitations must be acknowledged:

The lack of comprehensive in vivo data reduces the precise quantification of results and the full validation of the model. Ethical biases in the data, such as underrepresentation of certain demographic groups, could affect the generalizability of the findings. The model's deterministic nature and simplifications, such as neglecting the immune response, limit its applicability to certain stages of HIV-1 infection.

Despite these limitations, the in silico trial demonstrates the potential of in silico medicine to guide therapeutic development and inform decision-making. By identifying key parameters and subpopulations, the model can facilitate more targeted and effective in vitro research, ultimately contributing to the advancement of HIV-1 therapy.

## Future directions

Future work should focus on refining the model by incorporating additional biological mechanisms, such as the immune response, and expanding the parameter estimation to include more diverse and comprehensive in vivo data. Collaboration with regulatory authorities is essential to establish additional criteria and thresholds for validation and verification, ensuring the model's credibility and utility in clinical decision-making. The exploration of the sub-populations can be an interesting field of research.

In summary, this report has demonstrated the value of in silico medicine as a tool for accelerating research and development in HIV-1 therapy. By adhering to rigorous VVUQ principles, we have enhanced the credibility and applicability of the model, paving the way for more informed and effective therapeutic strategies.

# References


[1] Nigar Ali, Gul Zaman, and Ali Saleh Alshomrani. "Optimal control strategy of HIV-1 epidemic model for recombinant virus". In: *Cogent Mathematics* 4.1 (2017), p. 1293468.

[2] Nigar Ali et al. "The Effects of Time Lag and Cure Rate on the Global Dynamics of HIV-1 Model". In: *BioMed Research International* 2017.1 (2017), p. 8094947.

[3] Marc Bulterys et al. "Rapid HIV-1 Testing During Labor: A Multicenter Study". In: *JAMA* 292.2 (2004), pp. 219–223. DOI: 10.1001/jama.292.2.219.

[4] D. S. Dimitrov et al. "Quantitation of human immunodeficiency virus type 1 infection kinetics". In: *J. Virol.* 67 (1993), p. 2182.

[5] S. Finke and K.-K. Conzelmann. "Recombinant Rhabdoviruses: Vectors for Vaccine Development and Gene Therapy". In: *The World of Rhabdoviruses*. Ed. by Zhen F. Fu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 165–200. ISBN: 978-3-540-27485-8. DOI: 10.1007/3-540-27485-5_8.

[6] A. T. Haase. "Population biology of HIV-1 infection: viral and CD4+ T cell demography and dynamics in lymphatic tissues". In: *Ann. Rev. Immunol.* 17 (1999), p. 625.

[7] Ashley T. Haase. "Population Biology of HIV-1 Infection: Viral and CD4+ T Cell Demographics and Dynamics in Lymphatic Tissues". In: *Annual Review of Immunology* 17 (1999), pp. 625–656. DOI: 10.1146/annurev.immunol.17.1.625.

[8] M Hadjiandreou, Raul Conejeros, and Vassilis S Vassiliadis. "Towards a long-term model construction for the dynamic simulation of HIV infection". In: *Mathematical Biosciences & Engineering* 4.3 (2007), pp. 489–504.

[9] Maoan Han and Pei Yu. *Normal forms, Melnikov functions and bifurcations of limit cycles*. Vol. 181. Springer, 2012.

[10] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[11] Anita Holdcroft. "Gender bias in research: how does it affect evidence-based medicine?" In: *Journal of the Royal Society of Medicine* 100.1 (2007), pp. 2–3. DOI: 10.1177/014107680710000102.

[12] J. D. Hunter. "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[13] Takuya Iwanaga, William Usher, and Jonathan Herman. "Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses". In: *Socio-Environmental Systems Modelling* 4 (May 2022), p. 18155. DOI: 10.18174/sesmo.18155.

[14] Terri D. Keville. "The Invisible Woman: Gender Bias in Medical Research". In: *Women's Rights Law Reporter* 15 (1993-1994), p. 123.

[15] Wilhelm Kirch. *Encyclopedia of Public Health*. New York: Springer, 2008, pp. 676–677. ISBN: 978-1-4020-5613-0.

[16] S. P. Layne, J. L. Spouge, and M. Dembo. "Quantifying the infectivity of human immunodeficiency virus". In: *Proc. Nat. Acad. Sci. USA* 86 (1989), p. 4644.

[17] Sabine Kinloch-de Loës et al. "A Controlled Trial of Zidovudine in Primary Human Immunodeficiency Virus Infection". In: *New England Journal of Medicine* 333.7 (1995), pp. 408–413. DOI: 10.1056/NEJM199508173330702.

[18] Cuifang Lv and Zhaohui Yuan. "Stability analysis of delay differential equation models of HIV-1 therapy for fighting a virus with another virus". In: *Journal of Mathematical Analysis and Applications* 352.2 (2009), pp. 672–683.

[19] C. A. Michie et al. "Lifespan of human lymphocyte subsets defined by CD45 isoforms". In: *Nature* 360.6401 (1992), pp. 264–265. DOI: 10.1038/360264a0.

[20] J. M. Murray et al. "A model of primary HIV-1 infection". In: *Mathematical Biosciences* 154.2 (1998), pp. 57–85. DOI: 10.1016/s0025-5564(98)10046-9.

[21] Flora T. Musuamba et al. "Scientific and regulatory evaluation of mechanistic in silico drug and disease models in drug development: Building model credibility". In: *CPT: Pharmacometrics & Systems Pharmacology* 10.8 (2021), pp. 804–825. DOI: https://doi.org/10.1002/psp4.12669.

[22] Martin A. Nowak and Charles R. Bangham. "Population Dynamics of Immune Responses to Persistent Viruses". In: *Science* 272.5258 (Apr. 1996), pp. 74–79. DOI: 10.1126/science.272.5258.74.

[23] A. S. Perelson et al. "HIV-1 dynamics in vivo: virion clearance rate, infected cells life-span, and viral generation time". In: *Science* 271 (1996), p. 1582.

[24] Andrew N. Phillips. "Reduction of HIV concentration during acute infection: independence from a specific immune response". In: *Science (New York, N.Y.)* 271.5248 (1996), pp. 497–499. DOI: 10.1126/science.271.5248.497.

[25] Ted Pierson, John McArthur, and Robert F. Siliciano. "Reservoirs for HIV-1: Mechanisms for Viral Persistence in the Presence of Antiviral Immune Responses and Antiretroviral Therapy". In: *Annual Review of Immunology* 18 (2000), pp. 665–708. DOI: 10.1146/annurev.immunol.18.1.665.

[26] E. Ramsburg et al. "Highly effective control of an AIDS virus challenge in macaques by using vesicular stomatitis virus and modified vaccinia virus Ankara vaccine vectors in a single-boost protocol". In: *Journal of Virology* 78.7 (2004), pp. 3930–3940. DOI: 10.1128/JVI.78.7.3930-3940.2004.

[27] Tomás Revilla and Ginés García-Ramos. "Fighting a virus with a virus: a dynamic model for HIV-1 therapy". In: *Mathematical Biosciences* 185.2 (Oct. 2003), pp. 191–203. DOI: 10.1016/s0025-5564(03)00091-9.

[28] Maria T. Ruiz and Lois M. Verbrugge. "A two way view of gender bias in medicine". In: *Journal of Epidemiology and Community Health* 51.2 (1997), pp. 106–109. DOI: 10.1136/jech.51.2.106.

[29] A. Saltelli et al. *Global Sensitivity Analysis. The Primer*. Chichester, UK: John Wiley & Sons, 2008.

[30] Matthias J Schnell et al. "Construction of a Novel Virus That Targets HIV-1-Infected Cells and Controls HIV-1 Infection". In: *Cell* 90.5 (1997), pp. 849–857. ISSN: 0092-8674. DOI: https://doi.org/10.1016/S0092-8674(00)80350-5.

[31] L. F. Shampine and M. W. Reichelt. "The MATLAB ODE Suite". In: *SIAM J. Sci. Comput.* 18.1 (Jan. 1997), pp. 1–22.

[32] I. M. Sobol. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates". In: *Math. Comput. Simulat.* 55.1-3 (2001), pp. 271–280. DOI: 10.1016/S0378-4754(00)00270-6.

[33] M. A. Stafford et al. "Modeling plasma virus concentration during primary HIV infection". In: *J. Theor. Biol.* 203 (2000), p. 285.

[34] Marco Viceconti and Luca Emili, eds. *Toward Good Simulation Practice: Best Practices for the Use of Computational Modelling and Simulation in the Regulatory Process of Biomedical Products*. Springer Nature, 2024.

[35] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[36] Zhiping Wang and Rui Xu. "Stability and Hopf bifurcation in a viral infection model with nonlinear incidence rate and delayed immune response". In: *Communications in Nonlinear Science and Numerical Simulation* 17.2 (2012), pp. 964–978.

[37] Pei Yu, Jianing Huang, and Jiao Jiang. "Dynamics of an HIV-1 infection model with cell mediated immunity". In: *Communications in Nonlinear Science and Numerical Simulation* 19.10 (2014), pp. 3827–3844.

[38] Anneke Zanting et al. "The 'exotic other' in medical curricula: Rethinking cultural diversity in course manuals". In: *Medical Teacher* 42.7 (2020), pp. 791–798. DOI: 10.1080/0142159X.2020.1736534.

[39] Zhenqiang Zhang et al. "Sexual Transmission and Propagation of SIV and HIV in Resting and Activated CD4+ T Cells". In: *Science* 286.5443 (Nov. 1999), pp. 1353–1357. DOI: 10.1126/science.286.5443.1353.

# A Proof of the Maximum Recovery Bound in CD4+ Cells

Here, I demonstrate that regardless of the treatment's effectiveness, the model's structure ensures that the healthy CD4+ population is bounded by its physiological value. This demonstration is critical for evaluating the treatment's effectiveness in the context of the disease rather than the healthy state. According to the model, it is not possible for a treatment to lead to *hyperlymphocytosis*, a condition characteristic of certain leukemias or autoimmune diseases.

### Proof:

We have

$$\frac{x_3^*}{x_1^*} = \frac{\left(\frac{\lambda}{d + \frac{\beta bkq}{\alpha cu}}\right)}{\left(\frac{\lambda}{d}\right)} \tag{20}$$

$$= \frac{d}{d + \frac{\beta bkq}{\alpha cu}} < 100\% \tag{21}$$

Since $\frac{\beta bkq}{\alpha cu} \geq 0$ (the parameters have physiological relevance within the domain of positive real numbers) and due to the infectivity condition of HIV-1: $R_v > 1 \implies \beta \neq 0$, it follows that $\frac{\beta bkq}{\alpha cu} > 0$.

# B Numerical Aspect - Numpy Precision

The precision used for calculations is the default precision defined in Numpy [10], namely `float64`. This precision corresponds to a double-precision representation (64 bits) and is sufficient for the needs of this work. A verification was performed to ensure that this precision is supported by the machine used. The parameters associated with `float64` are detailed in Table 8. This information ensures the robustness and consistency of numerical calculations.

| **Parameter** | **Value** |
|---|---|
| Precision (`precision`) | 15 significant digits |
| Resolution (`resolution`) | $1.0000000000000001 \times 10^{-15}$ |
| Smallest exponent (`minexp`) | -1022 |
| Largest exponent (`maxexp`) | 1024 |
| Smallest normal value (`tiny`) | $2.2250738585072014 \times 10^{-308}$ |
| Smallest subnormal value (`smallest_subnormal`) | $4.9406564584124654 \times 10^{-324}$ |
| Largest representable value (`max`) | $1.7976931348623157 \times 10^{308}$ |
| Machine epsilon (`eps`) | $2.2204460492503131 \times 10^{-16}$ |
| Negative epsilon (`epsneg`) | $1.1102230246251565 \times 10^{-16}$ |

Table 8: Numerical parameters for the `float64` precision used in Numpy.

# C Matching of First and Second Moments Between a Gamma Distribution and a Uniform Distribution

The distribution of population parameters for *in silico* trials is chosen as a Gamma distribution, whose first and second moments match those of the uniform distributions detailed in the literature (see Table 5), after applying a rescaling factor of 0.8 to the standard deviation.

In practice, this matching can be achieved as detailed below.

Consider a uniform distribution with bounds $a < b$. The expectation and variance of this distribution are given by:

$$\mathbb{E}[\mathcal{U}(a;b)] = \frac{a+b}{2} \tag{22}$$

$$\mathbb{V}[\mathcal{U}(a;b)] = \frac{(b-a)^2}{12} \tag{23}$$

Now, consider a Gamma distribution parameterized by its shape parameter $k$ and scale parameter $\theta$, such that its probability density function (pdf) is given by:

$$f_{X\sim\text{Gamma}[k,\theta]}(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{\frac{-x}{\theta}} \tag{24}$$

Its expectation and variance are given by:

$$\mathbb{E}[\text{Gamma}(k;\theta)] = k\theta \tag{25}$$

$$\mathbb{V}[\text{Gamma}(k;\theta)] = k\theta^2 \tag{26}$$

Thus, if we denote $f = 0.8$ as the rescaling factor for the standard deviation, we can compute $k$ and $\theta$ from $a$ and $b$ using:

$$k = \frac{3(a+b)^2}{f^2(b-a)^2} \tag{27}$$

$$\theta = \frac{f^2(b-a)^2}{6(a+b)} \tag{28}$$

# D About using ChatGPT and other conversational agents

To do this work, I used ChatGPT and other similar conversational agents. In line with the faculty's expectations, this appendix defines the framework and extent of my use of these tools.

The agents used are ChatGPT, DeepL and Github Copilot. ChatGPT and Github Copilot were used for some pieces of code. In practice, they were mainly used to create the graphs and achieve a consistent style in the code. They were not used to write the report directly.

DeepL was used to translate certain sections originally written in French into English.

Finally, ChatGPT was used once the first version of the report was completed with the following prompt:

*You're an expert in the field of in silico medicine. What do you think of this academic report? Do you see any stylistic or content weaknesses?*

The response is then used to rework the report. The overall result was that some sentences were too long and some paragraphs needed to be reorganized.