# HIV Infection Project

M442, Spring 2022

Due Friday April 8

## 1 Overview

Over the last twenty-five or so years mathematical models have proven an invaluable resource in the understanding and treatment of a wide variety of diseases. One of the most studied cases is that of *human immunodeficiency virus* (HIV). In this project, we will study the role mathematical models have played in understanding the dynamics of HIV infection, focusing especially on a critical study from 1996 (see reference [PNMLH]), and we will also develop our own models that describe the full progression of HIV infection from initiation to the onset of acquired immunodeficiency syndrome (AIDS) to remission after treatment.

## 2 Biological Aspects

The first step toward modeling a new phenomenon mathematically consists in learning as much as possible about the phenomenon. In this section, we will review a few concepts from biology, virology, and immunology that will be critical for the development of any reasonable model of HIV infection. A number of specialized terms arise, so a glossary has been included in Section 4.

### 2.1 Viral Infection

A viral particle (or *virion*) is essentially a renegade gene: a nucleic acid (DNA or RNA) enclosed in a protein shell (or *capsid*), about 15–25 nanometers[1] in diameter (see Figure 1). Unlike cells, virions are unable to reproduce independently and must rely on using living cells as hosts—which is why most biologists don't consider virions to be alive in the strictest sense of the word. Built for this purpose, their protein shells act like Trojan horses, tricking unsuspecting cells into thinking they're harmless, or even friendly. The typical MO of a virion is to attach itself to some particular type of cell (its host) and insert a copy of its DNA into that cell.[2] Then, when the host is stimulated to reproduce (also by the virus),

[1] One nanometer is $10^{-9}$ meters.

[2] Most of the action at this level is determined by the shapes of things, how amino acids fit together to make proteins, and how proteins fit together with other proteins (like three-dimensional puzzle pieces). The protein on the capsid of a virus fits in this way with certain cells in the host's body. In the case of HIV, the best fit is with T cells.

it produces a multitude of copies of the virus. For example, it will be important for this project that the HIV virus creates roughly 1000 copies of itself from a single host T cell.[3] This "budding" can take place slowly, sparing the host cell, or rapidly, destroying it.

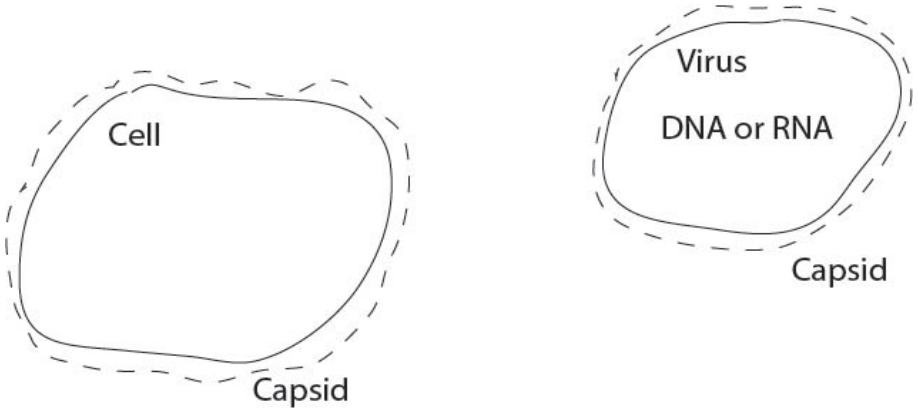


Figure 1: Cell and virus schematic.

A second type of virus, a *retrovirus,* carries a copy of its RNA (instead of its DNA) and must transcribe it into DNA before replicating. HIV is a retrovirus. This process of transcribing RNA into DNA is carried out by an enzyme called *reverse transcriptase* and adds a step to the virion's reproductive process. On first glance, this extra step may sound like an inefficiency that would mitigate the potency of HIV, but instead it turns out to be one of its strengths. The problem with retroviruses in general is that errors are sometimes made in the transcription of RNA into DNA, leading to inexact, or mutated, copies of the virus. With billions of virions being created every day, new strains are constantly emerging, and drugs that work on one version may not work on another. (Survival of the fittest over extremely short time intervals.)

A family of drugs has been created to take advantage of this extra step by interfering with the virus's reverse transcriptase. For obvious reasons, these drugs are called *reverse transcriptase inhibitors*.

## 2.2 Basic Immunology

One challenging aspect of HIV is that its primary host cells are precisely the $CD4^+$ T cells that control a person's immune system response. ("CD4" denotes a protein marker on the surface of the T cell, one that thinks the HIV virus is friendly; "T" refers to the thymus, where T cells mature after migrating from the bone marrow, where they are created.) During the course of most infections, the immune system plays the role of police officers swooping in to stop an assault. In HIV infection, the police officers are the ones being assaulted.

Typically, when a foreign substance (antigen) is introduced into the body, the body elicits an immune response in an attempt to purge the thing as quickly as possible. The antigen is first encountered by *macrophages*, cells that scavenge, engulf, and examine foreign particles,

[3]It will also be important that of these 1000 copies only about 10, on average, actually become active.

then present their findings to the $CD4^+$ T cells. These cells (i.e., the $CD4^+$ T cells), often referred to as *helper T cells*, decide whether or not a response is necessary. If so (which is to say, if the antigen turns out to be a pathogen), they typically respond in three ways: (1) they build up a command center by producing more of themselves; (2) they stimulate the production of $CD8^+$ T cells (or *killer T cells*), which hunt down infected cells and put them out of their misery (this is called *cellular immune response*); and (3) they stimuate B cells to produce *antibodies*, which are specifically engineered to destroy the pathogen at hand (this is called *humoral immune response*).

If the immune response is successful, certain cells of each type retain memory of the pathogen, and if it ever has the audacity to come around again, they'll be ready. This, of course, is the idea behind vaccinations. In particular, currently proposed HIV vaccinations consist of benign cells covered with the same protein that covers HIV. When exposed to these cells, the body learns to respond to this particular protein without having to actually fight off the virus.

## 2.3 HIV Infection inside the Cell

Once the virus has injected its RNA into a cell, a series of processes is initiated, leading to the ultimate *budding* mentioned above. As depicted in Figure 2, these processes include: (1) the reverse transcription of the viral RNA into DNA with the protein *reverse transcriptase*; (2) the integration of the viral DNA into the cellular DNA with the protein *integrase*; and (3) the separation of the resulting chain of virions with the protein *protease*.

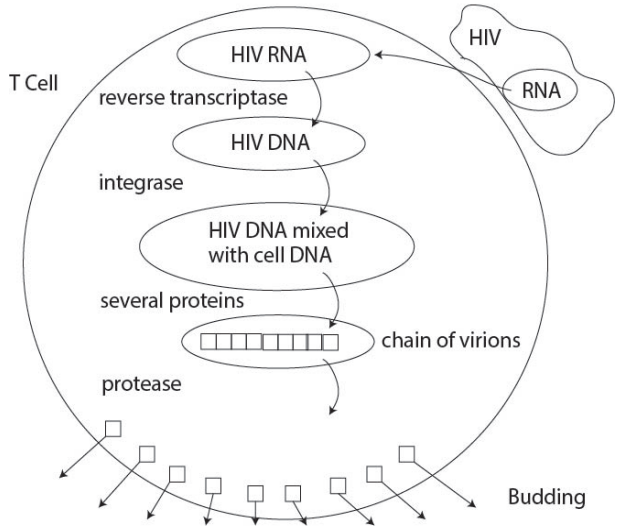


Figure 2: HIV progression inside a $CD4^+$ T cell.

## 2.4 Timeline for HIV Infection

Generally speaking, infection with HIV progresses on three time scales (see Figure 3): (1) during the first ten weeks or so after the virus is introduced into the body, the number of

(uninfected) helper T cells falls dramatically, while the number of virions spikes, as does the number of infected T cells;[4] (2) over a period ranging from 2 to as many as 20 years (and sometimes even longer) the number of virions and number of T cells roughly stabilize (small fluctuations persist), though with significantly fewer than the body's typical (uninfected) 1000 $mm^{-3}$ T cells (read: 1000 T cells per cubic millimeter); (3) in the final phase of the infection, a relatively short period of 1–3 years, the T cells once again crash, and the number of virions grows out of control. When the T-cell count falls below 200 $mm^{-3}$, the person is officially classified as having AIDS (Acquired Immune Deficiency Syndrome).

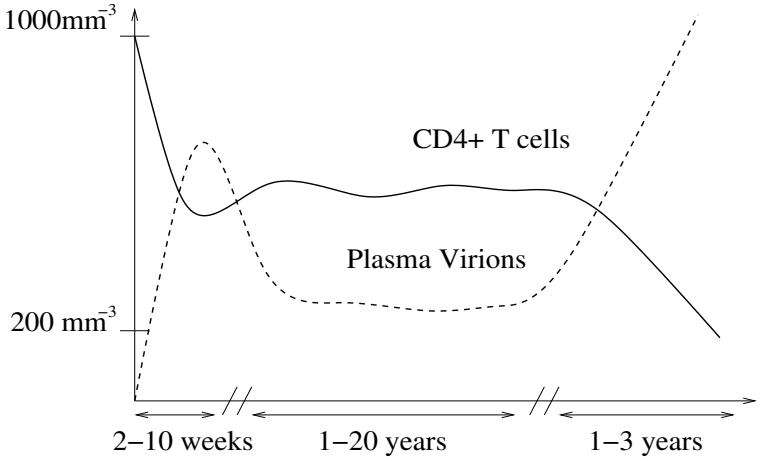


Figure 3: Timeline of HIV infection. The scale refers to the number of $CD4^+$ T cells. In practice, the initial virion spike can be much too large to fit on this figure ($\sim 10^6$).

### 2.5 Protease Inhibitors

The HIV treatments relevant to this project are called *protease inhibitors*. As the name suggests, this family of drugs blocks the efficacy of the protein protease described above, halting the viral budding process.

## 3 Assignments

For this project, your group will carry out two analyses. First, you will study a very crude (though certainly useful) mathematical model of HIV infection proposed in 1996. Second, you will build on this model to develop one of your own. We will begin our study of HIV infection by considering a critical advance in its understanding, made in 1996 by David Ho and his collaborators. (David Ho was named *Time* magazine's Man of the Year in 1996 for his contributions.)

1. Get a copy of the paper listed at the end of this assignment as reference [PNMLH], and also of the article beginning on p. 56 of the December 30, 1996 / January 6, 1997 copy of *Time* magazine. (Electronic copies of these articles are easy to find.)

[4] During this period the patient often exhibits symptoms similar to those of pneumonia.

2. In reference [PNMLH], the authors have two HIV models, equations (1)–(2) for infection in steady state, and equations (3)–(4)–(5) for the infection under treatment with the protease inhibitor Ritonavir. In terms of the population dynamics we have discussed in class, describe the processes modeled by each term in these equations.

3. Use equations (1)–(2) and equations (3)–(4)–(5) to derive the authors' expression in equation (6) for $V(t)$—total virion concentration. See the authors' footnote 15 if you want to check that you're on the right track after solving for $T^*(t)$.

4. The critical contribution these authors made in this paper was their determination of the virion production and clearance rate, which was previously considered much smaller (see, for example, the section entitled "A Team Effort" in *Time*'s article on David Ho). We will repeat this calculation for Patient 107, whose viral RNA counts are given graphically in the bottom left plot of the first figure in [PNMLH]. In tabulated form, these values are given in Table 1 of this assignment.[5] You can mimic the log plot the authors use with MATLAB's command, *semilogy*(). Use these data points and MATLAB's *lsqcurvefit()* to determine values $V_0$, $c$, and $\delta$ for Patient 107.[6] Your report should include a scatterplot of your data along with your best-fit curve.

5. Use your results from Item 4 to determine the number of virions produced per day in the entire body during the course of HIV infection. (Read the third column on p. 1583 of [PNMLH] carefully.) Explain the significance of this number. In particular, explain why this changed the way we view the dynamics of HIV infection.

6. Use your results from Item 4 to determine values for (1) the average life-span of a virion; (2) the average life-span of an infected T-cell (productively infected T-cell); and (3) the average viral generation time.

7. Use your results from Item 4 and the data from Table 1 of [PNMLH] to find an equilibrium point $(T_e, T_e^*, V_e)$. **Note on units.** In Table 1 of [PNMLH] the authors give T-cell counts per mm$^3$ and virion counts multiplied by $10^3$ per ml. For example, Patient 107 has $77 \times 10^3$ virions per milliliter. But this is exactly 77 virions per mm$^3$ (i.e., 1000 mm$^3$ = 1 ml).

8. When developing mathematical models, researchers often try to find a balance between strategically simplified models that are relatively straightforward to analyze and more sophisticated models that might be cumbersome to work with. While the model employed in [PNMLH] is simple enough to analyze exactly, it fails to capture the full dynamics of HIV infection (as depicted qualitatively in Figure 3 of this assignment).

[5] These values were taken directly from the plot, and so are certainly only an estimation of the exact values, which the authors annoyingly do not provide in the article. You should expect your values of $c$, $\delta$, and $V_0$ to have the same orders of magnitude as the corresponding values from the article, but you should not expect the values to be exactly the same. Also, the values given in Table 1 are viral counts per mm$^3$, whereas the values in the first figure of the article are RNA counts per mL. The viral count is half the RNA count, and the count per mm$^3$ is $10^{-3}$ the count per mL. See also the note on units in Item 7.

[6] Our standard methods for obtaining initial parameter values don't work here, so you'll need to be creative.

Using the discussion of HIV progression given in class and in this assignment (as well as any outside material you find appropriate, such as the discussion in [PNMLH] and the article from *Time*),[7] develop your own model of HIV infection that captures the first two phases of infection shown in Figure 3 of this assignment. (Notice that you're only studying the progression of infection here, *not treatment*, so your starting point is (1)–(2), not (3)–(4)–(5).) Here's a good place to begin: Notice that in [PNMLH], equations (1)–(2), the number of uninfected T cells is assumed to remain constant. The idea there was that during the steady-state period of the infection (labeled as 1–20 years in Figure 3 of this assignment), T cells remain nearly constant. It's clear from Figure 3 of this assignment, however, that if you want to model the full dynamics of HIV infection, the number of uninfected T cells will have to become a variable that changes with time. Begin by adding a third equation to (1)–(2) in which uninfected T cells are created by the body and killed by the virus.

9. Specify the units (not dimensions) of each parameter in your model from Item 8. Use your results from Item 4, and the two equilibriuim points $(T_e, T_e^*, V_e)$ and $(1000, 0, 0)$ to obtain values for the parameters in your model.

10. For each equilibrium point in your model, determine whether the point is stable or unstable, and explain the significance in either case.

11. Using your model from Item 8 and your parameter values from Item 9, create a stacked plot with uninfected T cells as a function of time on top and virions as a function of time on bottom. Take a long enough interval of time so that the first phase from Figure 3 is entirely included and the second phase has started. As initial values, you should take 1000 $\text{mm}^{-3}$ uninfected T cells, 0 $\text{mm}^{-3}$ infected T cells, and .001 $\text{mm}^{-3}$ virions. Depending on your model, the numerics can be delicate for this part, so you may need to increase accuracy by reducing the values of AbsTol and/or RelTol.

12. According to your model, determine the total number of T-cells created in the body per day during Phase II of infection.

13. The most difficult dynamic to capture in the progression of HIV infection is the eventual collapse of T cells and the spike of virions in the final period of infection. One possible explanation for the first aspect of this behavior is that T cell production increases during infection to fight the virus, but cannot continue indefinitely at this new level and eventually gives out. Incorporate this idea by changing your model so that the production of T cells begins to slow down during phase II (when your model is near equilibrium). Compute the length of time between the initiation of T cell failure and the onset of AIDS (when the T cell count drops below 200), and create a stacked plot beginning at the time you began decreasing T cell production and continuing until the onset of AIDS. While T cell production fatigue can explain the eventual collapse of the immune system, it does not necessarily account for the final spike of virions. A

[7] Another useful resource along these lines is the *Science* magazine special HIV issue, volume 280, June 19, 1998.

possible explanation for this final spike of virions is that HIV virions can also replicate in cells other than T cells. We will not try to model this effect.

14. Incorporate medication by Ritonavir into your model from Item 13 in the following way: combine your model with (3)–(4)–(5) in [PNMLH] to create a full model (with four equations) of HIV infection and treatment. Assume that once medication has been administered, T cell production returns to full capacity, and that all virions created after medication—whether created in T cells or elsewhere—are non-infectious. Taking T cell collapse to occur during Phase II (as described in Item 13), and supposing medication is administered at the onset of AIDS, determine the time required until the T cell count returns to 900 $\text{mm}^{-3}$ (90%).

| Times (in days) | Viral counts per $\text{mm}^{-3}$ |
|---|---|
| 1/12 | 119 |
| 1/6 | 137 |
| 1/4 | 111 |
| 1/2 | 119 |
| 3/4 | 89 |
| 1 | 103 |
| 5/4 | 89 |
| 3/2 | 58 |
| 7/4 | 67 |
| 2 | 72 |
| 3 | 30 |
| 4 | 25 |
| 5 | 16 |
| 6 | 8 |
| 7 | 6 |

Table 1: Table of viral counts and times for Patient 107 in [PNMLH].

# 4 Glossary

- **B cell**: Any lymphocyte not matured in the thymus. (cf. T cells.)
- **Amino Acids**: Any of a large group of organic acids containing a carboxyl group, COOH and an amino group $NH_2$. These guys link together to form proteins. The human body requires 20 to form its proteins: 1. alanine, 2. phenylalanine, 3. arginine, 4. asparagine, 5. aspartic acid (sold commercially as Nutrasweet), 6. cysteine, 7. glycine, 8. glutamine, 9. glutamic acid (in MSG), 10. histidine, 11. leucine, 12. isoleucine, 13. lysine, 14. methionine, 15. proline, 16. serine, 17. threonine, 18. tryptophan (in, for example, turkey; remember those Thanksgiving naps?), 19. tyrosine, and 20. valine.

- **Antibodies**: Oh, come on, I know you remember this much. But just in case: Protein molecules shaped to fit foreign proteins that are typically carried on the surface of invading germs (antigens).

- **Antigen**: A protein, toxin, or other substance of high molecular weight, to which the body reacts by producing antibodies.

- **Budding**: The process through which a multitude of virions emerge from their host cell, killing it.

- **Capsid**: The protein shell enwrapping a virion.

- **CD$4^+$ T cells** (helper T cells): Leukocytes that serve as the command center for the immune system, and one of HIV's favorite host targets.

- **CD$8^+$T cells** (killer T cells): Leukocytes that, in the event of *cellular immune response*, attack and kill friendly cells that have been infected.

- **Cellular Immune Response**: The immune system response to a pathogen in which killer T cells are created.

- **Chromosome:** A chain of genes. Each human cell (except red blood cells) contains 46 chromosomes (22 nearly identical pairs, plus two sex-determining), each composed of hundreds or thousands of genes linked end to end. Each series of chromosomes contains the entirety of the person's DNA.

- **DNA** (deoxyribonucleic acid): A nucleic acid bound in double helical chains by hydrogen bonds between bases, forming the basic material of genes. The chains consist of four molecules: adenine, thymine, guanine, and cytosine, commonly referred to as A, T, G, and C. The fit together A–T and G–C, but no other way.

- **Enzyme**: A protein that acts as an organic catalyst to speed up specific chemical reactions.

- **Gene**: A collection of DNA (or, more specifically, a segment along the strand of DNA) whose purpose in life is to create exactly one type of protein. (Recently, this notion has come under suspicion, as there appear to be something like 30,000 genes in the human genome and roughly 100,000 types of proteins at work in the human body. (Though these numbers are debated.) An accessible account of this debate appears in [C].)

- **Humoral Immune Response**: The immune system response to a pathogen in which B cells are stimulated to create antibodies appropriate for killing the pathogen.

- **Leukocyte** (white blood cell): Any of the small, colorless nucleated cells in the blood, lympth, and tissues, which are important in the body's defenses against infection. T cells and B cells are both leukocytes.

- **Lymphocyte** (white blood cell in lymphatic tissue): A particular type of leukocyte formed in lymphatic tissue.

- **Macrophages**: Any of the various phagocytic cells in lymphatic tissue, bone marrow, etc.; these guys serve as lookouts for the immune system and, on occasion, as hit men.
- **mRNA** (messenger RNA): A version of RNA that serves as a copy of a DNA strand and can be used to transport that copy outside of the nucleus where it can be copied by ribosomes.
- **Pathogen**: Any agent, especially a microorganism, able to cause disease.
- **Phagocyte**: Any cell, especially a leukocyte, that ingests and destroys other cells, microorganisms, or other foreign matter in the blood and tissues.
- **Protease**: An enzyme carried by HIV that cuts the long line of proteins HIV forms in its hosts into individual proteins that make new HIV particles.
- **Protease Inhibitor**: Any drug that inhibits the ability of protease to do its job, and thus limits the reproductive capacity of HIV.
- **Protein**: Any of a large class of nitrogeneous substances consisting of a complex union of amino acids and containing carbon, hydrogen, nitrogen, oxygen, frequently sulfur, and sometimes phosphorous, iron, iodine, or other elements. A particular protein is distinguished by the specific sequence of amino acids that appears in its structure.
- **Reverse Transciptase**: The enzyme that transcribes RNA into DNA. (Called "reverse" because, as you may recall, the typical cellular process if for DNA to be copied onto RNA.)
- **Reverse Transcriptase Inhibitor**: Any drug that inhibits the enzyme reverse transcriptase from transcribing RNA into DNA.
- **RNA** (ribonucleic acid): a nucleic acid essential to all cells, playing a key role in the copy and transfer of DNA; see also RNA–polymerase and mRNA.
- **RNA–polymerase**: A verson of RNA that "unzips" the double helix of a DNA molecule and copies the DNA into mRNA.
- **Virus**: Escaped genes or fragments of genetic code, virus particles (or virions) typically consist of DNA or, less commonly, RNA surrounded by a protein shell called a capsid. The shell works like a Trojan horse, tricking cells into thinking it's friendly.

# References


[ARA] The AIDS Research Alliance of America website: http://www.aidsresearch.org.

[C] B. Commoner, *Unraveling the DNA Myth: the spurious foundation of genetic engineering*, Harper's **304** (2002), no. 1821, 39–47.

[HNPCLM] D. D. Ho, A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz, *Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection*, Nature **373** (1995) 123–126.

[IAPAC] The International Association of Physicians in AIDS Care web site: http://www.iapac.org.

[PNMLH] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho, *HIV-1 Dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time*, Science **271** (1996) 1582–1586.