

Université de Montréal

Effet des antirétroviraux sur la pathogénèse du VIH
Une étude par modélisation mathématique intégrant la cinétique
du virus, de l'immunité, du médicament, et le comportement
d'adhésion avec leurs variabilités interindividuelles

par
Steven Sanche

Faculté de pharmacie

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en Sciences pharmaceutiques

le 31 août 2018

© Steven Sanche, 2018

Résumé

Les traitements antirétroviraux actuels permettent à beaucoup de patients du VIH de maintenir leurs charges virales à de très faibles niveaux sur plusieurs décennies. Or, malgré ce succès scientifique, de nombreux problèmes persistent, et à ce jour, aucun traitement ne permet de venir à bout du virus. Une prise à vie d'antirétroviraux est donc nécessaire, impliquant ainsi des contraintes posologiques pour le patient, une potentielle atteinte à sa qualité de vie et un fardeau financier pour la société. À ces inconvénients s'ajoute le risque de développer de la résistance aux médicaments. Même si les traitements demeurent efficaces, la persistance du virus peut également causer des dommages aux différents tissus et organes de l'hôte.

En se basant sur des connaissances de pointe dans le domaine du VIH, cette thèse aborde ces problématiques par une approche de pharmacologie quantitative des systèmes, appuyée par des données cliniques. L'objectif principal fut d'informer les mécanismes sous-jacents au développement de la résistance et à la persistance du virus *in vivo*. Nous avons tiré profit d'un maximum d'information sur l'ensemble des composantes impliquées dans la réponse virologique aux médicaments. Les modèles que nous avons développés joignent différentes échelles d'information, allant de l'échelle moléculaire, à virale, puis cellulaire, jusqu'au niveau clinique. Nous avons ensuite évalué la cohérence d'hypothèses de causalité aux phénomènes étudiés en testant la capacité de ces modèles à expliquer et reproduire des données empiriques.

En premier lieu, nous avons développé un modèle visant à mieux comprendre l'ampleur du développement de la résistance chez les patients sous plusieurs traitements. Le modèle combine plusieurs composantes, dont la cinétique virale, l'immunité, la pharmacocinétique et pharmacodynamique, l'adhésion au médicament ainsi que leurs variabilités interindividuelles. Les prédictions du modèle *in silico* concordent avec les observations cliniques d'échec virologique pour les trois traitements considérés et qui font intervenir l'efavirenz, l'emtricitabine, le ténofovir, le darunavir et le ritonavir. Par cette approche intégrative, nous avons remédié à la lacune des modèles précédents qui sous-estimaient grandement le risque de résistance. Nos résultats soulignent le rôle important que joue la faible pénétration des médicaments au niveau des ganglions lymphatiques dans le développement de la résistance. Ce modèle se veut prometteur de son utilité dans la prédiction de réponse virologique en clinique.

Nous nous sommes ensuite intéressés au phénomène du déclin en différentes phases, de plus en plus lentes, des charges virales des patients sous traitement antirétroviral. Les causes sous-jacentes à ce phénomène restent encore obscures. Une divergence d'opinions sur le rôle de la faible pénétration tissulaire des médicaments quant à l'existence de ces phases, divise actuellement les efforts de recherche. Afin de mettre la lumière sur cette implication, nous avons ajouté à notre modèle intégratif des taux différents de pénétration tissulaire. Nos résultats indiquent que l'implication seule de la pénétration des médicaments dans l'explication des phases de déclin serait synonyme d'un grand risque de développement de résistance. Ces prédictions contredisent quantitativement la réalité observée (peu de résistance), nous faisant conclure que cette hypothèse ne peut vraisemblablement pas expliquer le phénomène en question.

La dernière partie de la thèse se penche sur la capacité de certains patients à maintenir de faibles charges virales après l'interruption d'un traitement prolongé. Nous avons revisité une corrélation rapportée entre la charge virale résiduelle et la durée de maintien post-interruption de charges faibles. L'interprétation de cette corrélation s'avère difficile, puisque la durée en question n'inclut pas seulement le temps de contrôle de la virémie, mais également le temps nécessaire à la charge virale d'atteindre un seuil de tolérance à partir de la charge virale résiduelle. En utilisant un modèle mécanistique et des techniques statistiques avancées, nous avons réussi à estimer la durée attendue de contrôle réel de la charge virale ainsi que la variabilité interindividuelle associée. Contrairement à l'interprétation directe de la corrélation rapportée dans la littérature, notre analyse révèle que la variabilité interindividuelle du temps de contrôle de la virémie n'est pas associée à la charge virale résiduelle.

L'approche *in silico* adoptée dans cette thèse s'inscrit dans l'effort global de ces dernières années visant à minimiser le fardeau humain et le coût financier dans le développement du médicament. L'ensemble de nos résultats de modélisation suggèrent qu'une meilleure pénétration dans les ganglions lymphatiques diminuerait le nombre de cas de résistance chez les patients non-adhérents. Cependant, ils indiquent qu'une telle amélioration aurait peu d'influence sur la vitesse de déclin des charges virales. Aussi, quelle que soit l'influence de la pénétration lymphatique sur la virémie résiduelle, son amélioration n'aurait pas d'impact sur la capacité des patients à contrôler leurs charges virales après avoir cessé les antirétroviraux.

Mots-clés : adhésion aux traitements; cellules latentes; contrôle post-traitement; distribution des médicaments; échecs virologiques; efficacité des traitements; extrapolation *in vitro* à *in vivo*; modélisation mathématique; persistance du VIH; pharmacologie quantitative des systèmes; rebonds virologiques; résistance aux traitements; traitements antirétroviraux hautement actifs (HAART); variabilité interindividuelle; VIH;

Abstract

Tremendous progress was made in treating people living with HIV. Nowadays, antiretroviral therapy usually allows patients to suppress viral loads for several years, if not decades. Despite this scientific achievement, chronic drug intake is usually necessary as no treatment can completely eradicate the virus. Patients under these conditions may experience constant side effects. Further, patients' tissues and organs may become damaged over time, as chronic immune activation is observed in most patients despite adequate drug intake and undetectable viremia. The number of treatment options can also become seriously limited over time if patients' viruses develop and accumulate drug resistance. These issues motivate current scientific efforts. Hopefully, results from these efforts may lead to improvements in patient's quality of life and lower the financial burden HIV imposes on society.

Using the most up-to-date knowledge in the field of HIV, this thesis addresses some of the above-mentioned issues through a quantitative systems pharmacology approach supported by clinical data. Our main objective was to inform the mechanisms underlying the development of resistance and the persistence of the virus *in vivo*. We used available information on the components involved in the virologic response to drugs. This allowed developing models that bridge multiple scales, going from molecular, to viral, then cellular and finally to the clinical level. By assessing the ensuing models' ability to explain and reproduce empirical data, we studied the consistency of hypotheses regarding the causality of studied phenomena.

First, we developed a model allowing to better understand the extent of drug resistance development in patients undergoing antiretroviral therapy. The model combines several components, including viral kinetics, immunity, pharmacokinetics and pharmacodynamics, drug adherence as well as their interindividual variability. Predictions originating from our *in silico* model are consistent with clinical observations of virologic failure for the three treatments that were considered consisting of efavirenz, emtricitabine, tenofovir, darunavir and ritonavir. Through this integrative approach, we have remedied previous models that largely underestimated the risk of resistance. Our results highlight the important role played by low lymph node drug penetration in the development of resistance. The model we developed is an added forward step toward the use of *in silico* methods in the prediction of virologic responses in HIV patients.

We then investigated the causes underlying the increasingly slow phases of viral decline observed in patients initiating antiretroviral therapy. Opinions differ as to the role played by low drug penetration in tissues in the existence of these phases, leading to a fragmentation in research efforts. To shed light on this issue, we additionally considered, in our integrative model, different rates of tissue penetration. Our results indicate that, if low drug penetration were the only cause of the decline in phases of the viral load, then the ensuing low drug exposure would put patients at a very high risk of developing drug resistance. Prediction of high risk quantitatively contradicts the observed reality (little to no resistance), leading us to conclude that low penetration is unlikely a fair explanation to the phases of viral decline.

The last part of the thesis examines the ability of some patients to maintain low viral loads after interrupting prolonged antiretroviral therapy. We revisited a reported correlation between viral load values at interruption (also called residual viremia) and the duration of time patients maintained low viral loads afterwards. The interpretation of this correlation proved to be challenging since this duration includes both a time of complete control of viremia and a time during which viremia grows to a low-level threshold once control is lost. Using a mechanistic model and advanced statistical techniques, we were able to estimate the expected duration of complete viremia control along with its interindividual variability. Contrary to a direct interpretation of the correlation reported in the literature, our analysis reveals that the time of viremia control is unlikely associated with patients' residual viremia.

In summary, the *in silico* approach adopted in this thesis is part of an overall effort aiming to minimize patient recruitment and the financial costs involved in drug development. Our models suggest that better drug penetration in lymph nodes would likely lead to a decreased risk of drug resistance in non-adherent patients. However, our results also suggest that such an improvement would have little influence on the rate of decline of viral loads. Further, and regardless of the influence lymphatic tissue drug penetration may have on residual viremia, this improvement would not favour viremia control after antiretroviral discontinuation.

Keywords : drug adherence; drug distribution; drug efficacy; drug individualization; drug resistance; highly active antiretroviral therapy (HAART); HIV; HIV persistence; *in vitro* to *in vivo* extrapolation; interindividual variability; latent cells; mathematical modeling; post-treatment control; quantitative systems pharmacology (QSP); viral rebound; virologic failure;

Table des matières

Résumé.....	i
Abstract.....	iv
Table des matières.....	vi
Liste des tableaux.....	ix
Liste des figures.....	x
Liste des sigles et abréviations.....	xiv
Remerciements.....	xvi
Chapitre 1. Introduction.....	1
1.1. Histoire du VIH.....	1
1.2. Incidence, prévalence et mortalité.....	4
1.3. Infection virale : anatomie, pathophysiologie et réponse immunitaire.....	7
1.3.1. Virologie de base.....	7
1.3.2. Réponse immunitaire à l'infection virale.....	9
1.3.3. VIH.....	14
1.4. Conséquences cliniques.....	19
1.5. Les antirétroviraux.....	20
1.6. Les biomarqueurs du VIH.....	22
1.7. Les défis scientifiques.....	25
1.7.1. La persistance du virus.....	25
1.7.2. Adhésion et fardeau posologique.....	27
1.7.3. La résistance aux antirétroviraux.....	29
1.8. Apport de la modélisation mécanistique.....	31
1.8.1. Utilité de l'approche.....	34
1.8.2. Principes de formulation d'un modèle.....	36
1.8.3. Représentation mathématique.....	38
1.9. Sujets et structure des prochains chapitres.....	46

Chapitre 2. A Mathematical Model to Predict HIV Virological Failure and Elucidate the Role of Lymph Node Drug Penetration.....	48
2.1. Pr��ambule	48
2.2. Abstract and keywords.....	49
2.3. Introduction.....	49
2.4. Methods.....	51
2.4.1. Antiretroviral treatment	51
2.4.2. The model and parameter values	52
2.4.3. Simulations	55
2.5. Results.....	57
2.6. Discussion	61
2.7. Supplementary Material.....	65
2.7.1. The initial model	65
2.7.2. Final model	70
2.7.3. Summary of parameter values	84
2.7.4. Algorithm for the final model.....	86
2.7.5. Validation of one assumption	89
Chapitre 3. Exploring an alternative explanation for the second phase of viral decay: infection of short-lived cells in a drug-limited compartment during HAART	91
3.1. Pr��ambule	91
3.2. Abstract and keywords.....	92
3.3. Introduction.....	93
3.4. Results.....	98
3.4.1. Does a model with a drug-limited compartment hosting new short-lived CD4 cell infections have the capacity to predict viral load decay following treatment initiation? .	98
3.4.2. What values of $f_{u,2}$ and ϕ_2 allow reproducing the first two phases of decay in a typical patient under efavirenz, tenofovir DF and emtricitabine?	100
3.4.3. What drug concentrations lead to an inhibition of 60% of the infection in the drug-limited compartment?	101

3.4.4. What is the expected risk of virologic failure over a long-term period if the less exposed compartment allowing the infection of new short-lived CD4 cells is responsible for the second phase of viral load decay?	103
3.5. Discussion	106
3.6. Methods.....	111
3.6.1. The model and simulations	111
3.6.2. Viral load data fitting.....	113
3.6.3. Identification of kIEFV, kITFV and kIFTC in the less drug-exposed compartment	113
3.6.4. Statistical analysis.....	114
Chapitre 4. How is baseline viremia linked to post-treatment control? Analysis of treatment interruption data based on a mechanistic model	115
4.1. Préambule	115
4.2. Abstract and keywords.....	116
4.3. Introduction.....	117
4.4. Methods.....	119
4.4.1. Empirical data	120
4.4.2. The model	121
4.4.3. Data fitting	123
4.5. Results.....	126
4.5.1. The base model ($f_v = 1$).....	127
4.6. Discussion	130
4.7. Supplementary Material.....	133
Chapitre 5. Discussion	136
5.1. Principaux messages	136
5.2. Limites et perspectives.....	140
5.2.1. Connaissances manquantes.....	140
5.2.2. Limites et amélioration des modèles visant la prédiction	144
5.3. Retour sur les études par modèles mécanistiques	146
Bibliographie.....	i

Liste des tableaux

Table 1.1 – Molécules actives notoires pour le traitement du VIH, leurs classes et abréviations	22
Table 1.2 – Les principaux biomarqueurs du VIH et leur rôle principal	25
Table 2.1 – Distribution of weight values associated with the probability of pill intake following a missed dose (PIM), per adherence level. These are used to obtain population prevalence estimates of virological outcomes.....	82
Table 2.2 – Plasma protein binding correction kp , lymph node penetration kl and variability in the susceptibility of wild-type virus in the population	86
Table 4.1 – Distribution of baseline viremia in the sample of patients	121
Table 4.2 – Number of participants by duration of the period of low-level viremia and by baseline viremia. (Li et al., 2016) Percentages were computed over patients having the same level of baseline viremia. A threshold of 200 copies per mL of plasma was used.....	121
Table 4.3 – Parameter maximum likelihood estimates for the four considered models.....	127
Table 4.4 – Empirical percentages (Zheng et al.) and model-based percentage of patients by baseline viremia categories	133

Liste des figures

Figure 1.1 – Nombre de personnes vivant avec le VIH, et nombre de nouveaux diagnostics entre 1990 et 2016 à travers le monde.(UNAIDS, 2018b)	6
Figure 1.2 – Mortalités annuelles du SIDA et nombre de personnes dont on estime la mort évitée par les antirétroviraux entre 1990 et 2016 à travers le monde. (UNAIDS, 2018b)	6
Figure 1.4 – Schématisation du virus de l'immunodéficience humaine	14
Figure 1.5 – Cycle de reproduction du VIH : A) Liaison aux récepteurs; B) Fusion; C) Décapsidation; D) Transcription inverse; E) Intégration; F) Polymérisation de l'ARN viral; G) Traduction en protéines; H) Bourgeonnement; I) Maturation.	18
Figure 1.6 – Fonction de densité de la distribution log-normale. La distribution normale associée a une moyenne de 0 et un écart-type de 1.....	45
Figure 2.1– Simplified model algorithm and parameters: the special case of monotherapy for one patient with the assumption of only one major resistance mutation. Index k goes from 0 to the number of evaluated time steps. Index i takes a value of 1 for the wildtype strain and 2 for the resistant strain.	54
Figure 2.2 – Probability of virological outcomes as a function of adherence under efavirenz monotherapy. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented.....	58
Figure 2.3 –. Probability of virological success as a function of adherence under efavirenz, tenofovir disoproxil fumarate, and emtricitabine combination therapy. The model prediction for virological success is illustrated by the green bars and compared to a regression curve fitting clinical data (green line).(Dierynck et al., 2007)	58
Figure 2.4 – Probability of virological outcomes as a function of adherence under efavirenz monotherapy with complete lymph node penetration. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented.....	59
Figure 2.5 – Probability of virological success as a function of adherence under efavirenz, tenofovir disoproxil fumarate, and emtricitabine combination therapy: results from the model assuming complete lymph node penetration for all drugs. The model prediction for virological	

success is illustrated by the green bars and compared to a regression curve fitting clinical data (green line).(Gordon et al., 2015) 59

Figure 2.6 – Interindividual variability in the probability of virological outcome under efavirenz monotherapy as a function of adherence. The impact of model variability in adherence patterns and infection phase is displayed using whiskers on the graph. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented. 60

Figure 2.7 – Interindividual variability in the probability of virological success under efavirenz, tenofovir disoproxil fumarate, emtricitabine combination therapy as a function of adherence. The impact of model variability in adherence patterns and infection phase is displayed using whiskers on the graph. 61

Figure 2.8 – Model-based number of infected T-cells and number of plasma virions over time 90

Figure 3.1 – Hypotheses related to the decay of virion-producing cells after treatment initiation. A) Virion-producing cells are short-lived infected CD4 cells and all new cell infections are prevented due to high drug exposure. This translates into only one phase of viral decay; B) There are two types of virion-producing cells having half-lives of around 1 (short-lived) and 14 days (long-lived), respectively, and all new cell infections are prevented due to high drug exposure.(Perelson et al., 1997) This translated into two phases of viral decay C) Virion-producing cells are mostly short-lived infected CD4 cells located in two compartments, one with high (pink) and one with low (blue) drug exposure. The compartment with low drug exposure partially allows new cell infections, effectively leading to two phases of viral decay. The compartment is not associated to a specific tissue at this point, as its existence is hypothesized. 94

Figure 3.2 – Viral load data extracted from Perelson et al.(Perelson et al., 1997) (blue dots), model fit (black curve), and associated $f_{u,2}$ and ϕ_2 values. $f_{u,2}$ is the average fraction of total infection events not affected by the drugs and ϕ_2 is the fraction of the plasma viral load setpoint. Both parameters refer to the drug-limited compartment. 99

The parameter value $f_{u,2}=40\%$ was translated into concentration values of efavirenz, tenofovir and emtricitabine in the drug-limited compartment. In fact, the association between $f_{u,2}$ and

concentrations were mathematically described for many drugs using *in vitro* experiments.(L. Shen et al., 2008) The *in vivo* relationships can be deduced from these associations by assuming that intracellular concentrations dictate drug efficacy, independently of the medium in which the cells reside.(Sanche et al., 2017) The concentration inside the mononuclear cells of a compartment can be derived from extracellular plasma concentrations using two parameters,

..... 101

Figure 3.3 – Regression line (black) for the typical patient undertaking a treatment combining efavirenz, tenofovir DF and emtricitabine, model fit (red), and associated $f_{u,2}$ and ϕ_2 values. $f_{u,2}$ is the average fraction of total infection events not affected by the drugs and ϕ_2 is the fraction of maximum plasma viral load. Both parameters refer to the drug-limited compartment. Regression curve based on data from Karris et al.(Karris et al., 2017) 101

Figure 3.4 – Space of kl parameter values associated with efavirenz (EFV), tenofovir (TFV) and emtricitabine (FTC) leading to an inhibition of 55%-65% of infection events in the drug-limited compartment. Parameter kl represents the ratio of concentrations inside peripheral blood mononuclear cells to its equivalent in the hypothesized drug-limited compartment..... 103

Figure 3.5 – Simulation results for one random patient after 48 weeks of therapy with efavirenz (600mg), tenofovir DF (300mg) and emtricitabine (200mg) taken once daily. The model included a drug-limited compartment that allowed new short-lived cell infections. The virologic failure threshold was set at 400 copies/mL..... 105

Figure 3.6 – Boxplot of the risk of virologic failure as predicted by the model and obtained from 200 simulations, one per sampling point of drug penetration values (kl^{EFV} , kl^{TFV} , kl^{FTC}). For comparison purposes, the area shaded in darker blue is the confidence interval for the equivalent but observed in a real patient sample.(Gallant et al., 2006) 106

Figure 4.1 – Modeling the distinct viremia phases during the delay between treatment cessation and virologic failure observation (>200 copy/mL). The black line represents viremia. The dotted green line represents the occurrence of a rebound triggering event (RTE) separating the first phase (left) from the second phase (right). The dotted red line represents the virologic failure threshold. Baseline viremia was set at around 0.1 copy/mL in this example. 120

Figure 4.2 – Non-parametric empirical (grey line) and model-based (red line) survival curves for the duration of the period of low-level viremia (<200 copies/mL). The four model-based survival curves superimpose almost perfectly. 127

Figure 4.3 – Cumulative distribution of the time of occurrence of rebound triggering event (or tC, i) in a patient population, as predicted by model where $f(v) = 1$. Intervals with level of confidence of 95% are displayed for 0.1, 1, 10 and 100 days after treatment cessation. The x-axis is under a \log_{10} scale. 128

Figure 4.4 – Proportion of the patient population having an expected time of RTE ($\lambda i - 1$) below the values displayed on the x-axis, as predicted by the base model ($f(v) = 1$). From left to right, the colored dots represent the 10th, 25th, 50th, 75th, 90th and 97.5th percentiles. 129

Figure 4.5 – Probability that a specific patient experiences the rebound triggering event (RTE) before the time indicated on the x axis, as predicted by the base model. Each curve represents a patient-specific parameter value $\lambda i - 1$. The colors match those of Figure 4.4. 129

Figure 4.6 – Quantile-quantile plot of log-transformed empirical (Ioannidis et al., 2000) and model-based data. 134

Figure 4.7 – Cumulative distribution of the time to reach viremia threshold from the start of rebound in population of patients treated during the chronic infection phase 135

Liste des sigles et abréviations

Sigle ou abréviation	Description
3TC	Lamivudine
ABC	Abacavir
AIDS	Acquired immunodeficiency syndrome
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messager
ATV	Atazanavir
AZT	Azidothymidine
BCR	B-cell receptor
CCR5	C-C chemokine receptor 5
CD4	Cluster of differentiation 4
CD8	Cluster of differentiation 8
CDC	Centers for disease control and prevention
CMH	Complexes majeurs d'histocompatibilité
CMSP	Cellules mononucléées sanguines périphériques
cp	Copy
CXCR4	C-X-C chemokine receptor type 4
d4T	Stavudine
ddl	Didanosine
DLV	Delavirdine
DNA	Deoxyribonucleic acid
DRV	Darunavir
DTG	Dolutégravir
EFV	Efavirenz
ELISA	Enzyme-linked immunosorbent assay
Env	Envelope polyprotein
ETR	Étravirine
EVG	Elvitégravir
FDA	US Food and drug administration
FOS	Fosamprenavir
FTC	Emtricitabine
Gp41	Glycoprotein 41
Gp120	Glycoprotein 120
HAART	Highly active antiretroviral therapy
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen

Sigle ou abréviation	Description
IC ₅₀	Inhibitory concentration 50%
IDV	Indinavir
IIV	Inter-individual variability
INTI	Inhibiteur nucléotidique de la transcriptase inverse
INNTI	Inhibiteur non-nucléotidique de la transcriptase inverse
LPV	Lopinavir
MEMS	Medication event monitoring systems
MRV	Maraviroc
NK	Natural killer
NVP	Névirapine
PIM	Probability of dose intake following a missed dose
QSP	Quantitative systems pharmacology
RAL	Raltégravir
RCP	Réaction en chaîne par polymérase (RCP)
RNA	Ribonucleic acid
RNase	Ribonuclease
RPV	Rilpivirine
RRMM	Reconnaissance de motifs moléculaires
RTE	Rebound triggering event (RTE)
RTV	Ritonavir
SIDA	Syndrome d'immunodéficience acquise
SNP	Single nucleotide polymorphism
SQV	Saquinavir
T20	Enfuvirtide
TAHA	Traitements antirétroviraux hautement actifs
TCR	T-cell receptor
TDF	Ténofovir disoproxil fumarate
TFV	Ténofovir
TPV	Tipranavir
Vif	Viral infectivity factor
VIH	Virus de l'immunodéficience humaine
Vpr	Viral protein R
Vpu	Viral protein U
ZDV	Zidovudine

Remerciements

Cette thèse est le résultat d'un long cheminement à la fois personnel et professionnel où plusieurs ont joué un rôle important. En ordre chronologique, je tiens à remercier mes anciens collègues du centre de recherche de St. Mary. Tout particulièrement, j'aimerais remercier Jane McCusker qui a grandement contribué au perfectionnement de mes habiletés professionnelles, Éric Belzile qui a encouragé ma curiosité et mon développement, et Manon DeRaad qui a été d'un grand support. Je tiens également à remercier ma famille et mes proches d'avoir appuyé ma décision d'entreprendre cette aventure et d'avoir été à mon écoute lorsque d'importants défis ont obstrué mon parcours. À l'université, j'aimerais tout spécialement remercier la professeure Fahima Nekka qui a été, sans l'ombre d'un doute, celle qui a le plus contribué à mon épanouissement professionnel. Par ses encouragements, son exaltation devant les réussites, sa curiosité scientifique, sa critique toujours constructive et sa grande sensibilité, elle a fait croître le meilleur de moi et de l'ensemble de son groupe d'étudiants. Je remercie également le professeur Jun Li pour son regard critique et son souci du détail. Responsables de nombreux partages scientifiques extrêmement enrichissants, j'aimerais souligner mon appréciation pour la confiance et l'apport de Nancy Sheehan, Thibault Mesplède, ainsi que Mark Wainberg, ce dernier ayant été et étant toujours d'une grande inspiration. Finalement, je veux remercier mes nombreux collègues pour qui j'ai développé une amitié sincère : Abdullah Aljutayli, Guillaume Bonnefois, Morgan Craig, Fatéma Dodat, Soudeh Fakhari, Imad Kassem, Paul-Antoine Leboeuf, Cassandre Legault, Katia Mellal, Anaëlle Monfort, Jean-Michel Rabanel, Sara Soufsaf, Florence Véronneau-Veilleux, Xiaotian Wu et bien d'autres encore!

Chapitre 1. Introduction

La première partie de cette introduction a pour objectif de situer le lecteur dans le contexte historique du VIH et du SIDA. Elle rappelle l'urgence des premières années par la croissance rapide de l'épidémie et la gravité du pronostic, mais également la persévérance des scientifiques qui, malgré les nombreux obstacles, sont parvenus à développer des traitements antirétroviraux hautement actifs. L'arrivée de ces traitements a grandement amélioré l'espérance et la qualité de vie des patients atteints et coïncide avec un ralentissement important de l'épidémie. Malgré ce ralentissement, l'état de la situation épidémiologique actuelle, décrite subséquemment, souligne l'importance de maintenir les efforts de recherche.

1.1. Histoire du VIH

En 1981 dans l'état de la Californie, le journal « Morbidity and Mortality Weekly Report » (MMWR) fait état de cinq patients ayant souffert de pneumonie due à une infection au champignon *Pneumocystis jirovecii*. (Centers for Disease Control and Prevention, June 5, 1981) Cette infection fait partie des maladies opportunistes, c'est-à-dire des maladies qui affectent généralement des individus présentant un déficit immunitaire. La gravité des symptômes suggère que le déficit immunitaire est sévère, la pneumonie ayant causé la mort de deux des cinq patients mentionnés. (Centers for Disease Control and Prevention, June 5, 1981) Ce qui distingue les cinq patients est l'absence de critères normalement associés à un déficit immunitaire, soit l'âge avancé ou une fréquence élevée de maladies infectieuses dans l'historique médical. (Gottlieb, 1998) Le Dr. Michael Gottlieb, professeur assistant à l'université UCLA, a étudié trois de ces patients en 1980. Ceux-ci étaient également infectés par un champignon de la famille candida, associé à un défaut des lymphocytes T. (Gottlieb, 1998) Grâce à l'utilisation d'anticorps monoclonaux, l'équipe du Dr. Gottlieb a identifié une sous-population de lymphocytes T qui est particulièrement touchée, soit les cellules portant le marqueur CD4. (Gottlieb, 1998) Des analyses ultérieures ont démontré que tous les patients qui présentaient les signes d'infections susmentionnés avaient un important déficit en cellules CD4. (Gottlieb, 1998) En 1982, le nom de syndrome de l'immunodéficience acquise (SIDA) a été donné à cette affection du système immunitaire. (Centers for Disease Control and Prevention, September 24, 1982)

Les premiers individus souffrants du SIDA et qui ont été rapportés étaient des hommes américains. Des nourrissons et des femmes sont également affectés en 1982 et 1983.(Centers for Disease Control and Prevention, January 7, 1983; HIV.gov, 2016) Le nombre de personnes affectées croîtra rapidement aux États-Unis dans la décennie. Au total à la fin de l'année 1981, on estime à 270 le nombre de cas, dont 121 seraient décédés de complications associées au SIDA.(HIV.gov, 2016) Le bilan grimpe à 47 993 décès à la fin de l'année 1987. (HIV.gov, 2016) En 1992, le SIDA devient la première cause de décès chez les hommes âgés de 25 à 44 ans. (HIV.gov, 2016) D'autres pays sont également touchés, dont le Canada, où des premiers cas sont rapportés dès 1982.(Catie.ca, 2017) Dans le monde, le nombre de mortalités associées au SIDA augmente de manière exponentielle, pour atteindre plus d'un quart de millions de personnes vers 1990.(UNAIDS, 2018a)

La cause du SIDA a pu être identifiée en 1983.(Gottlieb, 1998) Au départ, la majorité des hommes affectés par le SIDA ont rapporté être homosexuels, ce qui suggère un lien entre le comportement sexuel et l'acquisition du syndrome.(Gottlieb et al., 1981) De plus, les premières femmes affectées étaient partenaires sexuelles d'hommes ayant des rapports sexuels avec d'autres hommes.(Centers for Disease Control and Prevention, January 7, 1983) Les nourrissons affectés avaient tous reçu une transfusion sanguine, de sorte qu'ils auraient pu avoir acquis le syndrome par du sang provenant d'individus infectés.(HIV.gov, 2016) Les chercheurs se sont donc initialement concentrés sur une maladie relativement commune dans la population homosexuelle, le cytomégalovirus.(Gottlieb et al., 1981) Ce virus a été détecté chez les premiers patients et engendre une diminution du nombre de cellules CD4.(Gottlieb et al., 1981) Malheureusement, la corrélation n'était pas indicatrice de causalité. Il s'avérera plutôt que l'infection au cytomégalovirus est une conséquence et non une cause de la dysfonction immunitaire.(Durack, 1981) Finalement, une équipe menée par la Dre. Barré-Sinoussi isola un nouveau virus à partir des cellules CD4 de ganglions lymphatiques des personnes affectées.(Barre-Sinoussi et al., 1983) On lui donne le nom initial de virus associé à une lymphadénopathie, mais le virus sera éventuellement rebaptisé virus de l'immunodéficience humaine, ou VIH.(HIV.gov, 2016) Le virus s'avérera responsable du développement du SIDA.

Les besoins cliniques ont motivé l'élaboration d'expériences scientifiques pour étudier le virus. Heureusement, les expériences *in vitro* ont grandement été facilitées en 1983 lorsqu'un

laboratoire mené par le Dr. Robert Gallo réussit à faire croître le virus.(Popovic, Sarngadharan, Read, & Gallo, 1984) Ceci accélèrera l'élaboration de tests de dépistages et de molécules thérapeutiques. Le premier test de détection du VIH est commercialisé en 1985, faisant appel à un protocole ELISA, pour « enzyme-linked immunosorbent assay ».(HIV.gov, 2016) Le test de détection, très sensible, mais peu spécifique, a mené à de nombreux cas de faux négatifs.(Chappel, Wilson, & Dax, 2009) Heureusement, des tests plus spécifiques ont été développés vers la fin des années 80.(Alexander, 2016) Du côté thérapeutique, le premier antiviral approuvé par le US Food and Drug Administration (FDA) en 1987 est la zidovudine (AZT).(FDA, 2018,-a; Mitsuya & Broder, 1987)

Plusieurs complications émanent de l'utilisation de la zidovudine. Initialement, le traitement doit être administré par voie intraveineuse à des doses toxiques pour les patients.(R. E. Walker et al., 1988) De plus, de nombreux laboratoires ont confirmé la présence de virus résistant à la zidovudine, une résistance qui peut se développer rapidement suite au début des traitements.(Wainberg et al., 1991) L'article qui en fait le constat et écrit par le Dr. Mark Wainberg, l'un des coauteurs du premier article de cette thèse, motive le développement de nouvelles molécules pour traiter les nombreux patients pour qui la zidovudine n'arrive plus à contrôler l'infection dû au développement de résistance. Heureusement, plusieurs molécules seront élaborées dans les années qui suivent (voir la section intitulée « Les antirétroviraux »). Cependant, ces nouvelles molécules feront également face à des problèmes liés au développement de la résistance (voir la section intitulée « La résistance aux antirétroviraux »).

L'idée d'utiliser plusieurs antirétroviraux afin de pallier au problème de la résistance surgit tôt après le développement de différentes molécules pour traiter le VIH. En 1991, l'effet synergique résultant de l'utilisation de deux antirétroviraux (la zidovudine et la didanosine, nouvellement approuvée par la FDA) est démontré *in vitro*.(Dornsife et al., 1991) Ceci suggère que la combinaison de médicaments pourrait être une solution plus efficace pour traiter le VIH *in vivo*. En 1992, un nouvel antiviral, le zalcitabine, est utilisé chez les patients en combinaison avec la zidovudine.(McLeod & Hammer, 1992) La combinaison résulte en une efficacité accrue, mais modeste *in vivo*, en comparaison avec l'utilisation de zidovudine en monothérapie.(Collier et al., 1993) En 1995, un médicament appartenant à une nouvelle classe d'antiviral, un inhibiteur de la protéase nommé saquinavir, est approuvé par la FDA.(FDA, 2018,-b) En association avec

des inhibiteurs de la transcriptase inverse, l'impact clinique est particulièrement important : pour la première fois aux États-Unis, on constatera une régression marquée du taux de mortalité associé au SIDA dans les années qui suivent.(Centers for Disease Control and Prevention, June 3, 2011; Noble & Faulds, 1996) Ceci sonne le glas d'une nouvelle ère pour les patients séropositifs du VIH, celle des traitements antirétroviraux hautement actifs (TAHA).

1.2. Incidence, prévalence et mortalité

Cette section fait état de la progression de l'épidémie en Amérique du Nord suivi de l'équivalent au niveau mondial.

Aux États-Unis, l'incidence du VIH a atteint son maximum au milieu des années 1980 pour ensuite décroître au début des années 1990 et demeurer stable depuis 2010.(Centers for Disease Control and Prevention, June 1, 2001; Hall et al., 2017) De nos jours, elle se situe à environ 50 000 nouveaux cas par année.(Moore, 2011) On attribue la stabilité de l'incidence à plusieurs facteurs. Des traitements plus efficaces et des comportements sexuels plus sécuritaires diminuent le risque de transmission du VIH par contact sexuel.(Hall et al., 2017) En parallèle, l'amélioration des traitements entraîne une plus grande longévité chez les personnes atteintes, résultant en un bassin plus important de personnes pouvant transmettre la maladie au cours des années.(Hall et al., 2017) Également, une amélioration des diagnostics permet d'identifier une plus grande quantité de patients atteints, malgré qu'une proportion non négligeable vit toujours avec la maladie sans le savoir.(Hall et al., 2017) Au Canada, l'incidence du VIH suit une tendance similaire.(Catie.ca, 2017)

En ce qui a trait à la prévalence, on estimait à environ 1.3 million le nombre de personnes vivant avec le VIH en Amérique du Nord en 2012.(Fettig, Swaminathan, Murrill, & Kaplan, 2014) La maladie touche tout particulièrement les hommes ayant des relations sexuelles avec des hommes (prévalence d'environ 14.5% aux États-Unis et 14.9% au Canada selon UNAIDS)(UNAIDS, 2018b) et les usagers de drogues injectables (prévalence d'environ 3.6% aux États-Unis et 10.9% au Canada).(UNAIDS, 2018b) D'autres populations sont également considérées à plus haut risque, dont les prisonniers et les travailleurs du sexe.

La situation mondiale est décrite par la Figure 1.1 et la Figure 1.2. La première de ces figures représente la prévalence et l'incidence du VIH au niveau mondial entre 1990 et 2016 (Figure 1.1). On constate que l'épidémie du VIH est constamment ralentie depuis le développement des traitements antirétroviraux et grâce à la prévention (courbe en bleu). Malgré tout, environ 36,7 millions de personnes vivaient avec le VIH en 2016, dont 1,8 million nouveaux cas rapportés durant l'année. La Figure 1.2 représente l'évolution de la mortalité due au développement du SIDA entre 1990 et 2016 dans le monde. Plus spécifiquement, le nombre annuel de personnes ayant succombé à des complications liées au SIDA et le nombre de personnes dont le décès a été évité grâce à l'usage d'antirétroviraux y sont illustrés. On constate que le SIDA a été associé à environ 1 million de décès en 2016, mais un nombre similaire aurait été épargné par l'utilisation d'antirétroviraux.(UNAIDS, 2018b) Les pays les plus touchés par l'épidémie du VIH sont le Swaziland (27.2% de la population), le Botswana (21.9%), Lesotho (25%) et l'Afrique du Sud (18.9%).(UNAIDS, 2018b) De manière encourageante, on remarque une diminution constante du nombre annuel de décès depuis 2005. Le nombre de décès évités par le VIH est relativement stable depuis 2014. Même si beaucoup de personnes affectées (20.9 millions) ont eu recours à une forme de thérapie en Juin 2017, de l'amélioration quant à l'accès aux antirétroviraux est nécessaire au niveau mondial (seulement 56.9% des individus vivant avec le VIH) pour diminuer davantage le nombre de décès. (UNAIDS, 2018b)

En bref, l'ampleur de l'épidémie est toujours importante. Elle motive un effort scientifique soutenu afin de diminuer l'incidence du SIDA, d'augmenter la qualité de vie des personnes vivant avec le VIH partout dans le monde et de trouver une solution permettant d'éradiquer le virus chez ces derniers. Les progrès à venir nécessiteront certainement un niveau de compréhension élevé du virus et de son interaction avec l'hôte.

Figure 1.1 – Nombre de personnes vivant avec le VIH, et nombre de nouveaux diagnostics entre 1990 et 2016 à travers le monde.(UNAIDS, 2018b)

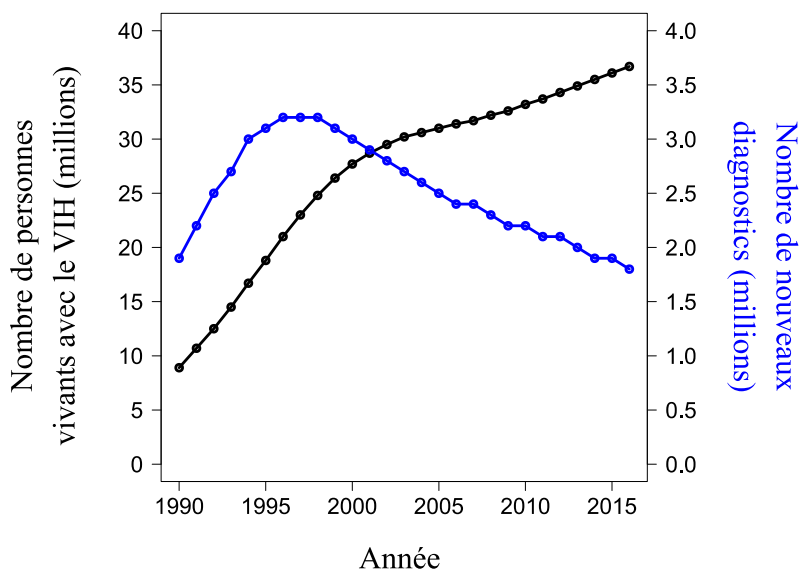
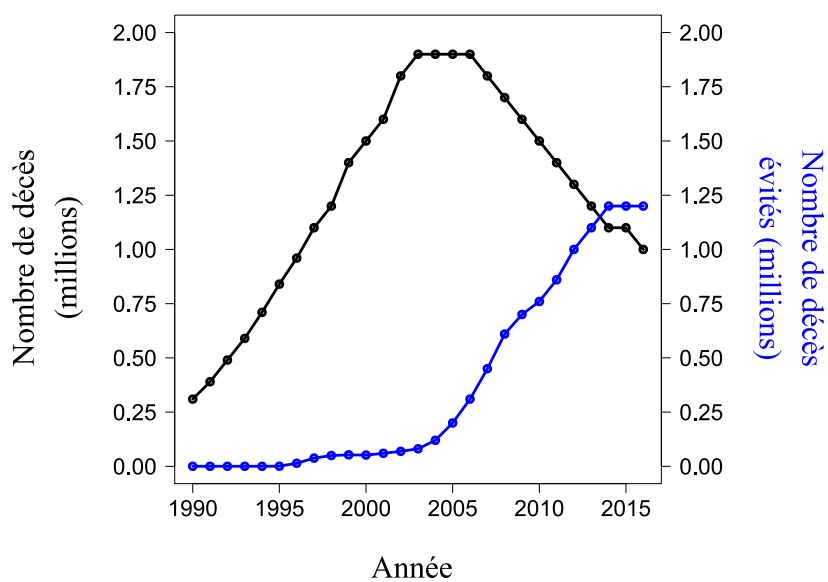


Figure 1.2 – Mortalités annuelles du SIDA et nombre de personnes dont on estime la mort évitée par les antirétroviraux entre 1990 et 2016 à travers le monde. (UNAIDS, 2018b)



1.3. Infection virale : anatomie, pathophysiologie et réponse immunitaire

Dans cette thèse, nous décrivons des modèles de l'interaction entre le VIH, le système immunitaire et les antirétroviraux. Ces modèles ont comme objectif de prédire comment progressent la taille et la composition de la population virale chez son hôte. Il s'avère donc utile de comprendre comment le virus se reproduit, évolue et est éliminé *in vivo* lorsque l'immunité et les médicaments exercent une pression sur la survie du virus. À cette fin, nous décrirons tout d'abord l'anatomie d'un virus et son cycle de reproduction. On insistera sur ce qui fait évoluer le virus au niveau génétique et affecte sa capacité à se reproduire dans diverses conditions. Il s'en suivra une description de la réponse immunitaire face à l'infection virale et son impact sur le virus. Afin de demeurer le plus général possible, on a séparé ce qui est commun à tous les virus de ce qui est caractéristique du VIH. L'action des antirétroviraux sur les diverses composantes du VIH sera discutée dans une autre section intitulée « Les antirétroviraux ».

Les prochaines sous-sections contiennent des éléments qui permettront à un non-initié de comprendre une part importante de la littérature sur le VIH. En ce sens, il n'est pas nécessaire de maîtriser tous les détails qui y sont mentionnés pour comprendre les modèles rapportés dans les chapitres qui suivent. Le lecteur peut, s'il le désire, se concentrer sur le nom et la fonction des lymphocytes T (sous-section « Réponse adaptée ») et le nom et la fonction des protéines virales mentionnées dans la Figure 1.5 concernant le cycle de reproduction (sous-section « VIH »).

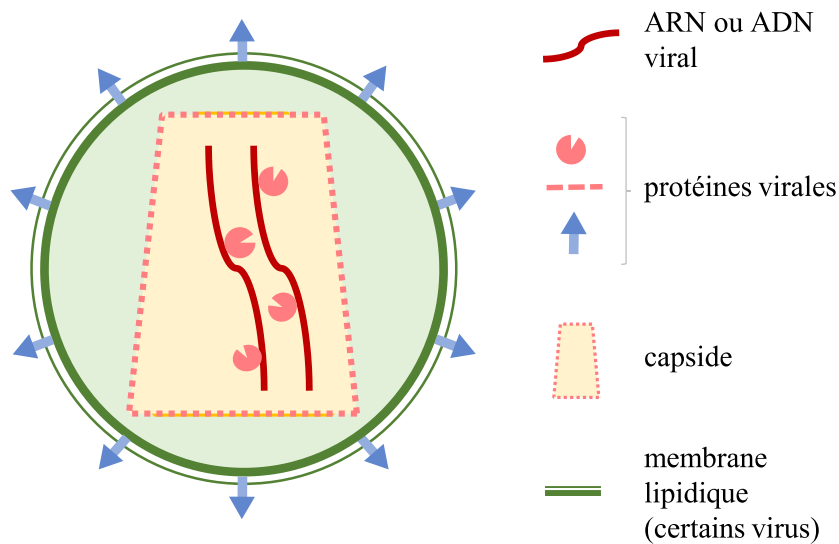
1.3.1. Virologie de base

Plusieurs éléments anatomiques et de la reproduction d'un virus sont communs à tous les virus. Cette section en résume les grandes lignes. (Dimmock, Easton, & Leppard, 2016) Afin de faciliter la compréhension, la Figure 1.3 schématise le virus. Tous les virus sont composés de matériel génétique (le génome du virus) entouré d'une enveloppe nommée capsid. La capsid composée de protéines a comme principale fonction de protéger le génome de la détérioration. Ce génome est composé soit exclusivement d'ADN ou bien d'ARN. L'ensemble du génome et de la capsid porte le nom de nucléocapsid. Il arrive que le nucléocapsid soit lui-même entouré

d'une enveloppe composée de protéines et de lipides. Dans la majorité des cas, cette enveloppe provient de la couche de phospholipides de la cellule ayant produit le virus. Les virus peuvent également contenir des protéines spécifiques au virus et/ou des composantes de la cellule hôte. Ces protéines exercent diverses fonctions qui facilitent ou protègent le cycle de reproduction du virus. La particule complète porte le nom de virion.

Les virus ne peuvent se reproduire seuls, ils doivent utiliser la machinerie cellulaire de leur hôte pour y parvenir. Pour y avoir accès, le virus se lie tout d'abord à sa cellule cible. Ceci s'effectue via des protéines à la surface du virus qui se lient spécifiquement aux récepteurs à la surface de cette cellule (e.g. les protéines schématisées en bleu sur la Figure 1.3). Une fois lié, le virus entre dans la cellule par une série de mécanismes propres à celui-ci. Les virus qui sont

Figure 1.3 – Le virion et ses composantes



entourés d'une couche bilipidique verront celle-ci fusionner à la membrane de la cellule cible par une série d'interactions moléculaires. Une fois dans le cytoplasme de la cellule, la capsid qui est relâchée doit rendre son génome accessible afin d'entamer la production de nouveaux virions par un processus nommé décapsidation. La biosynthèse d'un virus s'effectue ensuite en plusieurs étapes impliquant la réplication du génome, la transcription du génome en ARN messager (ARNm) et la traduction de cet ARN en protéines virales. Les virus doivent utiliser

les ribosomes de l'hôte pour effectuer l'étape de traduction. Le génome est également copié plusieurs fois. Les génomes nouvellement produits peuvent eux-mêmes être répliqués et transcrits pour amplifier le processus. Une fois tous les éléments du virus produits, un processus d'assemblage permet d'unir les éléments de manière ordonnée afin de produire des particules virales. Le cycle de reproduction se termine par l'excrétion virale, i.e. le relâchement du virus hors de la cellule infectée. Parfois, un cycle de maturation du virus est nécessaire afin de rendre les protéines virales fonctionnelles. Le cycle complet prend de 5 à 15 heures en moyenne chez l'animal, mais il existe une grande variabilité entre les virus.

Les virus s'adaptent rapidement à leur environnement et la capacité d'adaptation dépend de la taille et de la nature de leur génome. En général chez les virus à ARN (comme le VIH), la synthèse d'ADN viral à partir de l'ARN n'utilise pas la machinerie cellulaire de l'hôte, mais une protéine qui lui est propre. Cette protéine ne possède pas de fonction de relecture, de sorte que de nombreuses erreurs d'encodage se produisent. (Bowman, Hu, & Pathak, 1998; Delviks-Frankenberry et al., 2011) On estime à une erreur d'encodage pour toutes les 10^3 ou 10^5 bases, ce qui équivaut à environ une erreur par génome produit pour un virus de taille moyenne. Pour les virus à ADN, la situation est différente, avec une erreur produite toutes les 10^9 à 10^{10} bases. Ainsi, comparés aux virus à ADN, les virus à ARN produisent beaucoup de virions non infectieux, ce qui donne lieu à une très large population de virions qui diffèrent par leurs séquences génétiques spécifiques. L'ensemble des souches virus qui composent la population virale correspond aux quasiespèces. L'adaptation des virus à des pressions environnementales (e.g. médicaments et système immunitaire) s'effectue alors par sélection darwinienne.

1.3.2. Réponse immunitaire à l'infection virale

Comme nous l'avons mentionné plus haut, le VIH s'attaque aux cellules du système immunitaire, le même système qui tente de le détruire. L'interaction qui en résulte est complexe. Afin de mieux comprendre celle-ci, décrivons tout d'abord la réponse immunitaire typique à une infection virale. (Dimmock et al., 2016)

1.3.2.1. Réponse innée

La réponse immunitaire se scinde en deux parties : la réponse innée et la réponse adaptée. La réponse innée n'est pas spécifique à un virus en particulier. Elle débute au niveau de la cellule infectée qui a détecté la présence du pathogène. Cette détection se fait par l'entremise de récepteurs de reconnaissance de motifs moléculaires (RRMM), qui se lient à des motifs moléculaires associés à la présence d'un pathogène comme de l'ADN dans le cytoplasme ou bien de l'ARN double brin. Une telle détection engendre une cascade de signalisation qui altère l'expression des gènes dans la cellule. Ces changements ont pour but de générer un milieu hostile à la réplication virale, de signaler aux cellules voisines la présence potentielle de pathogènes et d'attirer les cellules du système immunitaire. En général, la réponse innée se fait par la production d'interférons de type 1.(Ng, Mendoza, Garcia, & Oldstone, 2016) Ce sont des cytokines interagissant avec des récepteurs qui peuvent être situés à divers emplacements à même la cellule ou à l'extérieur de celle-ci. Entre autres, les interférons stimulent la transcription de gènes producteurs de molécules qui ont eux-mêmes des propriétés antivirales, qui signalent une détresse aux cellules du système immunitaire ou stimulent l'autophagie ou l'apoptose.(Schmeisser, Bekisz, & Zoon, 2014) Les facteurs de restriction sont de telles molécules dont la fonction est de nuire à la réplication virale. L'autophagie est le processus par lequel la cellule dégrade certaines de ses composantes cellulaires via la stimulation d'activité enzymatique, nuisant par le fait même à l'intégrité du matériel viral et à la machinerie cellulaire nécessaire à sa réplication.(Shibutani, Saitoh, Nowag, Munz, & Yoshimori, 2015) L'apoptose est synonyme de suicide cellulaire et est associée à la fragmentation du noyau et de son ADN.(Muppidi, Porter, & Siegel, 2004) L'apoptose peut également être engendrée par des lymphocytes NK (NK pour « Natural Killer », en anglais) qui reconnaissent les signes de détresse de la cellule susmentionnés.

Tous ces mécanismes de protection sont souvent insuffisants pour inhiber complètement l'activité virale. La capacité d'adaptation des virus est une des raisons expliquant sa survie malgré l'action de l'immunité innée. En effet, la plupart des virus pathogènes se sont adaptés de manière à pouvoir atténuer d'une manière ou d'une autre l'impact de la réponse aux interférons. Le VIH a d'ailleurs développé de nombreux mécanismes afin de pallier l'action du système immunitaire inné.(Altfeld & Gale, 2015; Jakobsen, Olagnier, & Hiscott, 2015) En général, ceci

peut s'effectuer par l'entremise de séquences d'ADN ou d'ARN viral ou par des protéines virales qui affectent la cascade de réponses aux interférons. À défaut d'enrayer l'infection, la réponse innée permet de gagner du temps avant que la réponse immunitaire adaptée, plus lente à répondre à l'infection, ne puisse intervenir.

1.3.2.2. Réponse adaptée

La réponse adaptée s'effectue par l'entremise d'interactions avec les complexes majeurs d'histocompatibilité (CMH). Ces protéines ont pour fonction de présenter des bouts de matériel biologique présent à l'intérieur d'une cellule, à sa surface. Les CMH de classe I (CMH-I, il en existe en effet plusieurs classes) sont présentes sur la très grande majorité des cellules du corps humain. Le matériel exposé peut provenir du soi ou bien être d'origine étrangère, auquel cas il portera le nom d'antigène. La réponse immunitaire innée provoque l'induction de l'expression de CMH-I chez la cellule infectée, augmentant la présentation d'antigènes à la surface de la cellule. Cette présentation est cruciale, car elle permet l'interaction avec des cellules du système immunitaire qui patrouillent le système circulatoire sanguin et lymphatique dans le but de détecter la présence de ces antigènes.

Les lymphocytes T et B font partie de ces cellules qui patrouillent le corps et ce sont les principales composantes de l'immunité adaptée. Ces cellules possèdent à leur surface des récepteurs (TCR, pour « T-cell receptor », ou BCR pour « B-cell receptor ») qui reconnaissent des parties d'antigènes nommés épitopes. La reconnaissance est spécifique, puisqu'un épitope ne peut être reconnu que par un sous-ensemble de lymphocytes T et B. Avant la première rencontre avec l'antigène, les lymphocytes T et B sont dits naïfs. (Pennock et al., 2013) Lorsque ces cellules entrent en contact avec l'épitope qui leur est spécifique et s'y lie suffisamment fortement, des changements intracellulaires et membranaires importants s'en suivent. Ces changements provoquent la migration de la cellule dans un des organes lymphatiques tels les ganglions lymphatiques ou les plaques de Peyer. Une fois située dans ces organes, la cellule se divisera pour produire un très grand nombre de clones par un processus intitulé expansion clonale. (Polonsky, Chain, & Friedman, 2016) Les clones réagissent au même épitope que la cellule originale. Certaines cellules produites deviendront des cellules mémoires, dont on verra l'importance plus loin. Cependant, la majorité des cellules issues de l'expansion clonale

deviennent plutôt activées ou effectrices. Dans cet état, les cellules immunitaires auront tendance à migrer à l'extérieur des organes lymphatiques à la recherche de l'épitope et une fois trouvé, à exercer leur fonction immunitaire spécifique. La nature de ces fonctions dépend du type de cellule qui a été activée.

Un premier type de cellules T est le lymphocyte T cytotoxique. Son rôle chez les patients séropositifs du VIH a été le sujet de plusieurs études.(Demers, Reuter, & Betts, 2013; B. Walker & McMichael, 2012) Les lymphocytes T cytotoxiques se distinguent par la présence de marqueurs CD8 à leur surface, d'où sa seconde appellation de cellules CD8. Les cellules qui sont issues de l'expansion clonale traquent les cellules qui sont infectées et qui présentent l'épitope reconnu. Si une telle rencontre se produit, les cellules CD8 activées peuvent s'attacher aux cellules infectées via leur récepteur TCR. Cette interaction provoque le relâchement d'un cocktail de molécules cytotoxiques (autrement dit, toxiques pour les cellules). Ces molécules ont pour but d'induire la mort de la cellule infectée, stoppant par le fait même la réplication du virus.

Les lymphocytes T auxiliaires constituent un deuxième type de cellules T et se distinguent par la présence de marqueurs CD4 à leur surface. Ces cellules sont les principales cibles du VIH. Contrairement aux TCR de cellules CD8 qui se lient aux CMH de type I chargés d'un épitope, les TCR des cellules CD4 ne peuvent se lier qu'aux CMH de type II. Ce type de CMH est exprimé par un sous-ensemble de cellules, dont les plus notoires sont les cellules dendritiques, les phagocytes et les cellules B.(Schraml & Reis e Sousa, 2015; P. Shen & Fillatreau, 2015; Varol, Mildner, & Jung, 2015) Les cellules dendritiques patrouillent normalement le corps à la recherche de protéines sécrétées par les cellules infectées, qu'elles absorbent par phagocytose et dégradent afin de présenter les antigènes résultants par leur CMH de type II. Si la protéine absorbée interagit avec un ou plusieurs RRMM appartenant aux cellules dendritiques, ces dernières cellules s'activeront et migreront vers les ganglions lymphatiques où elles pourront interagir avec toutes autres cellules dont les récepteurs sont compatibles avec les CMH de type II. Lorsque les CD4 rencontrent leur épitope présenté par un CMH de type II, elles procèdent à une expansion clonale. Les cellules issues se différencieront pour devenir des cellules mémoires ou bien des cellules effectrices dont il existe principalement les types I et II. Les cellules CD4 effectrices de type I ont pour fonction d'aider l'activité des cellules CD8 et

des macrophages, alors que les cellules effectrices de type II aident l'action des cellules B. L'aide en question se fait principalement par la sécrétion de cytokines (e.g. interleukines, interféron- γ) qui stimulent la production et l'activité de ces cellules, tout en promouvant la migration de celles-ci là où se situent les cellules infectées.

Les cellules B sont le sujet de nombreuses études en lien avec le VIH, notamment concernant le développement d'un vaccin.(Haynes et al., 2014; Mylvaganam, Silvestri, & Amara, 2015) Elles ont comme particularité de pouvoir produire des anticorps, aussi connus sous le nom d'immunoglobulines, des molécules pouvant se lier aux antigènes directement ou aux complexes formés par un CMH et l'antigène. La fonction des anticorps est variée, allant de la neutralisation de virus à l'induction de phagocytose ou de l'activation par agglomération du virus. Les cellules B sécrètent également des cytokines qui aident à la coordination de la réponse immunitaire.

L'action de l'ensemble du système est complexe et hautement coordonnée. La précision des communications est attribuable à la spécificité des récepteurs membranaires et la diversité des cytokines.(Chatenoud, 2008) Les cytokines contrôlent l'activation des cellules du système immunitaire, l'expansion, la différenciation des cellules qui permet à celles-ci de se spécialiser à certaines fonctions, et l'apoptose.(Chatenoud, 2008)

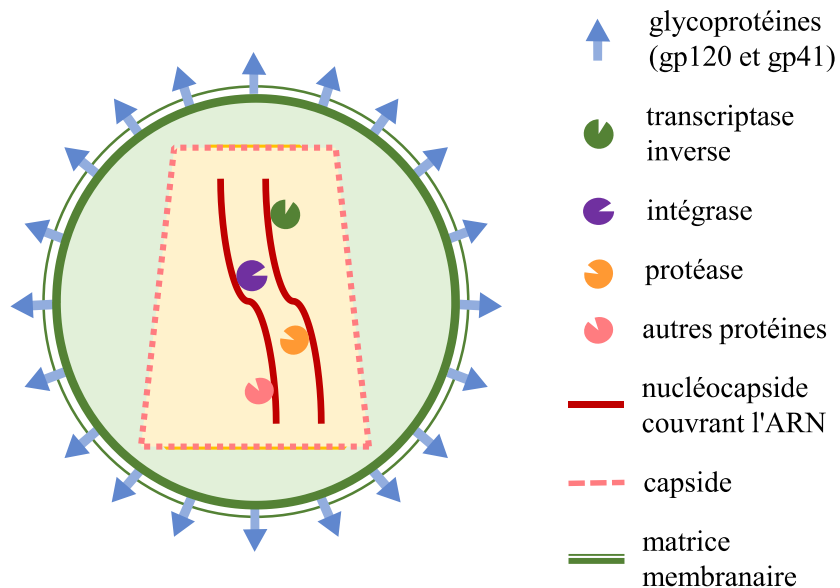
Le signal d'apoptose est particulièrement important lorsque l'infection virale se résorbe chez l'hôte. Le système immunitaire entre alors en phase de contraction.(McKinstry, Strutt, & Swain, 2010) Le chapitre 2 et 3 discutera du rôle que cette phase peut exercer sur le VIH. Durant cette phase, la très grande majorité des cellules effectrices meurent, laissant derrière quelques cellules, dont plusieurs cellules mémoires.(Buchholz, Schumacher, & Busch, 2016) Ces cellules mémoires permettent une réponse beaucoup plus rapide si le même virus vient à réinfecter l'hôte. En effet, si les cellules mémoires reconnaissent leur épitope, elles provoqueront une expansion clonale donnant rapidement lieu à une population importante et diversifiée de cellules effectrices.(Buchholz et al., 2016) Les cellules mémoires peuvent survivre plusieurs années, leur demi-vie étant estimée à environ 8 à 15 ans.(Pennock et al., 2013) Plus loin, on verra que les cellules mémoires ont un rôle important à jouer sur la persistance du virus du VIH chez leur hôte.(Chomont, DaFonseca, Vandergeeten, Ancuta, & Sekaly, 2011)

1.3.3. VIH

Dans cette thèse, nous nous concentrons sur le VIH-1, qui est le virus prédominant en Amérique et en Europe, par opposition au VIH-2 plus présent en Afrique. Un schéma de l'anatomie du VIH-1 est présenté à la Figure 1.4.(Dimmock et al., 2016) La section qui suit décrit ce virus en plus amples détails.(Dimmock et al., 2016)

Le VIH-1 est composé de deux brins simples d'ARN qui doivent tout d'abord être inversement transcrits en ADN avant de pouvoir produire de l'ARN qui sera traduit en protéines virales. Le virus fait conséquemment partie de la classe des rétrovirus. Le génome du virus est contenu dans un nucléocapside. Le nucléocapside est lié à plusieurs protéines qui permettent la réplication du virus, soit la transcriptase inverse, l'intégrase, la protéase et d'autres protéines dont on verra la fonction plus loin (RNase H, Vif, Vpu et Vpr). Le tout est protégé d'une capsid elle-même recouverte d'une couche de phospholipides. Les phospholipides sont traversés par les glycoprotéines gp41 (dans la membrane) et gp120 (attachés à gp41, en dehors de la membrane).

Figure 1.4 – Schématisation du virus de l'immunodéficience humaine



Un schéma du cycle de reproduction est donné à la Figure 1.5. La première étape du cycle de reproduction est l'entrée dans la cellule cible.(Melikyan, 2014) La glycoprotéine gp120 à la surface du virus se lie en premier lieu au récepteur CD4, qui est exprimé par les lymphocytes

T auxiliaires et certains monocytes de la lignée des macrophages (Figure 1.5-A). Pour entrer dans la cellule, le virus doit utiliser un co-récepteur à la surface de celle-ci. Certains virus utilisent le co-récepteur CCR5 qui est exprimé par les macrophages et lymphocytes T, alors que d'autres virus utilisent le co-récepteur CXCR4, surtout exprimé par les lymphocytes T auxiliaires. Cette préférence de co-récepteur dépend de la composition de la glycoprotéine gp120. Ces liaisons permettent à la membrane du virus de fusionner à la membrane cellulaire et d'exposer la capsid au cytoplasme de la cellule (Figure 1.5-B). Par la suite, la capsid se désassemble dans le cytoplasme par un mécanisme peu connu, résultant en la libération de deux nucléocapsides et de protéines virales (Figure 1.5-C).

La deuxième étape du cycle de reproduction peut ensuite être entamée : la transcription inverse de l'ARN viral en ADN (Figure 1.5-D). Cette étape fait intervenir la transcriptase inverse qui exerce trois rôles : i) la transcription inverse d'ARN viral à de l'ADN viral, ii) la synthèse d'ADN à partir d'ADN et iii) la digestion de l'ARN des molécules ARN-ADN pour donner de l'ADN simple brin (aidée par la RNase H). (Menendez-Arias, Sebastian-Martin, & Alvarez, 2017) La transcription inverse donne lieu à de l'ADN double-brin qui contient l'ADN du virus (provirus) et une longue répétition de plus petites séquences d'ADN qui l'aidera à s'intégrer à l'ADN de la cellule.

Pour intégrer l'ADN de la cellule hôte, le provirus a tout d'abord besoin de migrer dans le noyau de la cellule (Figure 1.5-E). L'intégration du provirus qui s'en suit nécessite l'action d'une autre protéine : l'intégrase. (Andrake & Skalka, 2015) Le processus d'intégration est irréversible : le provirus intégré ne peut pas se détacher de l'ADN de l'hôte sans avoir recours à une quantité importante d'énergie. En général chez les rétrovirus, une cellule infectée contient de 1 à 20 copies de provirus intégré à son ADN.

Pour produire de nouveaux génomes et de nouvelles protéines virales, l'ADN proviral doit être transcrit en ARN viral (Figure 1.5-F). L'ARN-polymérase de la cellule hôte exerce cette fonction. (Griffiths et al., 2005) Le processus de transcription est pratiquement le même que pour fabriquer de l'ARN messager à partir du matériel génétique de l'hôte. Plusieurs copies de l'ARN viral peuvent être obtenues d'un seul provirus. L'expression des gènes du VIH-1 s'effectue en plusieurs étapes (Figure 1.5-G). (Dimmock et al., 2016) En premier lieu, un petit nombre d'ARN viral complet est produit. L'ARN est ensuite clivé de manière à produire de

l'ARNm codant pour plusieurs protéines dont la plus importante à cette étape est la protéine Tat. Une fois traduit par les ribosomes de l'hôte, Tat a pour fonction d'augmenter la rapidité à laquelle l'ARN-polymérase peut transcrire l'ADN viral. Ceci a pour conséquence de graduellement augmenter la concentration d'une autre protéine virale : Rev, qui a pour fonction d'empêcher le clivage de l'ARN du virus, permettant l'accumulation d'ARN complet et d'ARN clivés à des endroits précis dans le cytoplasme. C'est à cette étape que le génome du virus est amplifié, ce qui augmente la production du reste des protéines virales : Gag, Gag-Pol, Env, Vif, Vpu et Vpr. La polyprotéine Gag contient les éléments qui formeront éventuellement la capsid (p24) et le nucléocapsid (p7). Pol est une polyprotéine contenant la protéase, la transcriptase inverse et l'intégrase. Env, quant à elle, contient les glycoprotéines gp120 et gp41.

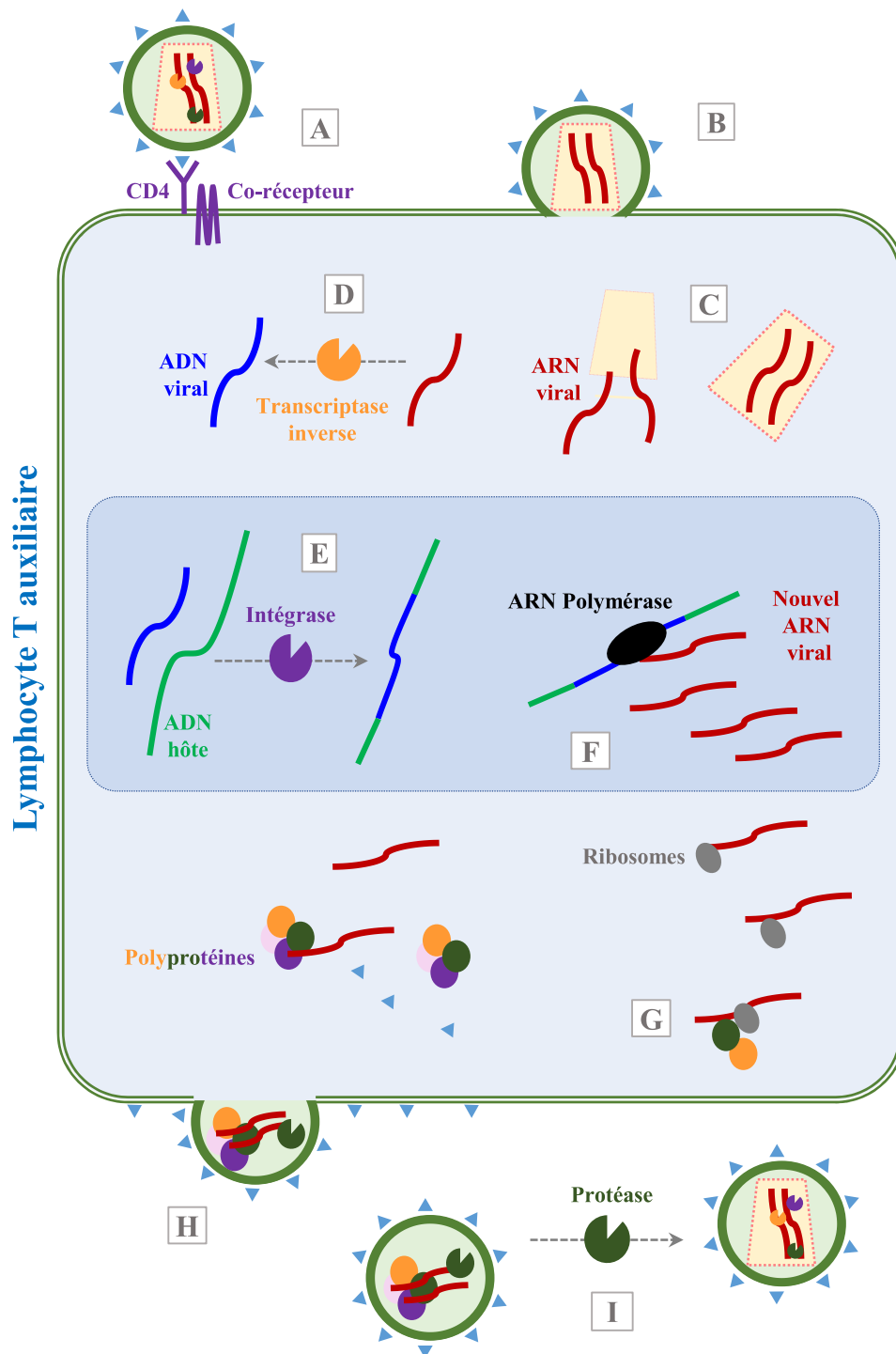
Tous les éléments du virus étant produits, il reste trois étapes pour obtenir des virions matures : i) l'assemblage des éléments, ii) le bourgeonnement et iii) la maturation. (Sundquist & Krausslich, 2012) Toutes les composantes du virus doivent au préalable avoir voyagé de leur lieu de fabrication à la paroi cellulaire. L'assemblage du virus est ensuite orchestré par la polyprotéine Gag qui se colle à la paroi cellulaire, et se lie aux protéines et à l'ARN viral. (Sundquist & Krausslich, 2012) Le bourgeonnement du virus, c'est-à-dire la formation d'une sphère de phospholipides à partir de la membrane de la cellule qui contient le matériel viral, est assuré par un processus de l'hôte (Figure 1.5-H). Le virus, encore immature, est relâché à l'extérieur de la cellule. Par la suite (et également un peu avant), la protéase clivera les polyprotéines, assurant la conformation de toutes les composantes virales finales dont la capsid (Figure 1.5-I). (Sundquist & Krausslich, 2012) Le cycle de réplication se termine par la production de plusieurs centaines de virions par cellule CD4 infectée.

D'autres protéines interviennent à différents moments du cycle de reproduction afin de faciliter les processus impliqués dans la réplication ou bien contrecarrer l'action du système immunitaire. (Dimmock et al., 2016; Moir, Chun, & Fauci, 2011) Les plus notables sont Vif, Vpr, Nef et Vpu. Vif est une protéine virale qui protège le génome du VIH de l'action de APOBEC3G, protéine de l'hôte qui tente de provoquer des mutations qui empêcheraient le virus de fabriquer des protéines fonctionnelles. Vpr a plusieurs fonctions dont celui de faciliter le transport de l'ADN viral dans le noyau. L'action de Nef, quant à elle, est d'induire une cascade de signalisation qui veut imiter l'activation d'un lymphocyte et ainsi induire la transcription de

l'ADN viral en ARN d'un lymphocyte infecté. Vpu cause la dégradation du récepteur CD4 de la cellule infectée, réduisant la réponse immunitaire adaptée dont l'activation des lymphocytes T cytotoxiques.

En résumé, le virus du VIH agit de manière complexe et efficace pour se reproduire dans les cellules exprimant le récepteur CD4. Le virus infecte préférentiellement les lymphocytes T qui sont en état actif; les changements encourus dans la cellule une fois celle-ci active la rendent beaucoup plus vulnérable à son infection. Les lymphocytes T auxiliaires qui sont au repos (inactifs) expriment abondamment un facteur de restriction (SAMHD1) dont le rôle est de dégrader une molécule clé permettant la complétion de la transcription inverse, expliquant la difficulté du virus à infecter les cellules CD4 au repos.(Baldauf et al., 2012; Sze, Olganier, Lin, van Grevenynghe, & Hiscott, 2013) En ce qui a trait aux lymphocytes T en état d'activation, une infection résulte généralement en une mort cellulaire après environ 48 heures, dont 24 heures de production de virions.(Holmes, Zhang, & Bieniasz, 2015) La cause de la mort cellulaire est incertaine, mais elle fait possiblement intervenir la pyroptose,(Doitsh et al., 2014) un mécanisme induit par la réponse immunitaire innée provoquant l'éclatement de la cellule et le relâchement de matériel qui promeuvent la réponse immunitaire.(Fink & Cookson, 2005)

Figure 1.5 – Cycle de reproduction du VIH : A) Liaison aux récepteurs; B) Fusion; C) Décapsidation; D) Transcription inverse; E) Intégration; F) Polymérisation de l'ARN viral; G) Traduction en protéines; H) Bourgeonnement; I) Maturation.



1.4. Conséquences cliniques

Le premier stade de l'infection au VIH est l'infection primaire.(Cohen, Shaw, McMichael, & Haynes, 2011) Durant ce stade, le patient peut manifester des symptômes se rapprochant de ceux de la grippe. Cet épisode symptomatique se manifeste habituellement quelques semaines après l'infection, et ne dure que quelques jours. Initialement durant cette période, le nombre de lymphocytes T auxiliaires décroît dû à la destruction de ceux-ci alors que le nombre de lymphocytes T cytotoxiques à épitope VIH augmente. Par la suite, un équilibre dynamique est atteint entre les lymphocytes T cytotoxiques, dont le nombre et l'activation dépendent de l'action des lymphocytes T auxiliaires, ces lymphocytes T auxiliaires, dont le nombre et l'activation dépend de l'activité des lymphocytes T cytotoxiques et de l'implication générale du système immunitaire, et la production de nouveaux virions. La quantité de virions décelés dans le sang une fois cet équilibre atteint porte le nom de « set point », en anglais. On croit que ce nombre de virions dépend du ratio de lymphocytes CD8+/CD4+, mais également de la capacité d'adaptation du virus qui modifie constamment son génome pour contrecarrer l'action du système immunitaire (e.g. peut résulter en changement d'épitopes reconnus par le système immunitaire).(Asquith, Edwards, Lipsitch, & McLean, 2006; Bar et al., 2012) En anglais, ce dernier processus porte le nom de « viral escape ».

Ce stade atteint d'équilibre dynamique est connu sous le terme d'infection chronique et est majoritairement asymptomatique, avec seulement la présence de ganglions lymphatiques enflés due à un état d'inflammation chronique.(Dimmock et al., 2016) La durée moyenne de cette période est de 10 ans, mais de grandes variations existent entre les individus. Sans traitement, le patient peut entrer dans un second stade, où de nouveaux symptômes se manifestent. Ces symptômes sont caractéristiques d'un affaiblissement du système immunitaire : perte de poids et apparition d'infections opportunistes mineures (fièvre, candidose, diarrhées). Le dernier stade de l'infection est le SIDA, où les patients manifestent des infections plus graves (myopathie, démence, cancers). Ce qui distingue les phases chroniques, intermédiaire et le SIDA est la quantité de lymphocytes T auxiliaires : >500, 200-500 et <200 cellules/ μ L, respectivement.

D'autres complications peuvent se manifester à un moment ou à un autre de l'infection, principalement dû à un état d'activation chronique du système immunitaire. Cette activation implique de l'inflammation qui favorise la destruction des tissus. Ainsi, les porteurs du VIH ont un plus grand risque de souffrir de maladies cardiovasculaires, neurologiques, du foie, du rein et des os.(Deeks, Tracy, & Douek, 2013)

1.5. Les antirétroviraux

La Table 1.1 dresse une liste des principales molécules actives approuvées par la FDA pour traiter le VIH.(FDA, 2018) Il existe plusieurs classes de ces médicaments antirétroviraux (antirétroviral : contre les rétrovirus), auxquelles on peut associer une des étapes du cycle de reproduction du VIH et un mode d'action spécifique (voir la Figure 1.5 pour un rappel).

Tel que mentionné ci-haut, la zidovudine (AZT) a été le premier médicament à indication VIH à être approuvée par la FDA.(HIV.gov, 2016) Ce médicament est un analogue nucléosidique qui inhibe l'action de la transcriptase inverse (INTI). Plus précisément, il se lie au site d'action de la transcriptase inverse et insère ladite molécule à la chaîne d'ADN viral qu'il tente de produire (Figure 1.5-D).(Dimmock et al., 2016) Or, les analogues nucléosidiques n'ont pas la terminaison nécessaire pour permettre à d'autres nucléotides de s'insérer dans la séquence d'ADN, causant la terminaison prématurée de celle-ci. (Dimmock et al., 2016) L'AZT avait comme grand défaut d'être aussi substrat, mais à moindre mesure, de l'ADN polymérase de cellules de l'hôte, d'où sa grande toxicité. Depuis, une grande quantité d'analogues nucléosidiques ont été approuvés par la FDA.

Il y a également les inhibiteurs non-nucléotidiques de la transcriptase inverse. Le premier médicament de cette classe à être approuvé fut le névirapine (NVP). Contrairement aux INTI, ces médicaments ne se lient pas au site catalytique de la transcriptase inverse, mais se lient à l'extérieur de ce site, ce qui a pour conséquence de changer la conformation de la transcriptase inverse et d'empêcher son bon fonctionnement. (Dimmock et al., 2016) Un exemple notoire de molécule INNTI est l'efavirenz (EFZ).

D'autres protéines virales sont la cible d'antirétroviraux. On note les inhibiteurs de la protéase qui, comme la nomination le sous-entend, se lient à la protéase virale de manière à

bloquer un site de clivage de protéines, et d'empêcher la maturation des virions, étape qui les rendent infectieux (Figure 1.5-I). (Dimmock et al., 2016) Un exemple notoire est le darunavir (DRV). Les inhibiteurs de l'intégrase, quant à eux, se lient au site de pré-intégration de l'intégrase et empêchent le transfert de l'ADN viral à l'ADN de l'hôte (Figure 1.5-E). Un exemple est le raltégravir (RAL).(FDA, 2018)

Finalement, certains antirétroviraux empêchent l'étape de fusion de la membrane du virus à celle de la cellule hôte (Figure 1.5-B). On note l'enfuvirtide (T20) qui se lie à gp41 et empêche la formation d'un pore qui permettrait à la capsid virale d'entrer dans la cellule hôte.(Greenberg & Cammack, 2004) Une autre molécule qui agit à l'étape d'entrée du virus est le maraviroc (MVC) qui se lie au CCR5 et l'empêche d'interagir avec la glycoprotéine virale gp120 (Figure 1.5-A).(Biswas, Tambussi, & Lazzarin, 2007)

Malheureusement, certains génomes du virus pourraient produire des protéines qui se lient moins facilement aux antirétroviraux. Les virions qui possèdent ces génomes arrivent à se reproduire malgré la présence d'antirétroviraux. Pour éviter cette situation, il est courant d'utiliser des combinaisons de médicaments de différents types. Pour pouvoir continuer à se reproduire, il faut que le virus puisse produire des protéines qui résistent à l'ensemble des médicaments utilisés, ce qui est moins probable, mais malheureusement possible (voir section « Les défis scientifiques »).

Bien que les médicaments mentionnés soient principalement destinés aux personnes déjà atteintes du VIH, ils ne se limitent pas à cette seule population. Ils peuvent également être utilisés en prophylaxie, puisqu'il a été démontré que les séronégatifs prenant le ténofovir en combinaison avec l'emtricitabine ont un risque significativement plus bas d'infections par transmission horizontale (infection entre individus qui ne sont pas parents).(Krakower & Mayer, 2015) De plus, certains antirétroviraux diminuent grandement le risque de transmission verticale (transmission de la mère à l'enfant dont elle était enceinte), ce qui ajoute à leur utilité.(Rosa et al., 2015)

Table 1.1 – Molécules actives notoires pour le traitement du VIH, leurs classes et abréviations

Classe	Abréviation	Nom de la molécule active
Inhibiteur nucléosidique de la transcriptase inverse	AZT	Azidothymidine
	ddl	Didanosine
	FTC	Emtricitabine
	3TC	Lamivudine
	ABC	Abacavir
	d4T	Stavudine
	TFV*	Ténofovir
	ZDV	Zidovudine
Inhibiteur non-nucléosidique de la transcriptase inverse	DLV	Delavirdine
	EFV	Efavirenz
	ETR	Étravirine
	NVP	Névirapine
	RPV	Rilpivirine
Inhibiteur de la protéase	ATV	Atazanavir
	DRV	Darunavir
	FOS	Fosamprenavir
	IDV	Indinavir
	LPV	Lopinavir
	RTV	Ritonavir
	SQV	Saquinavir
	TPV	Tipranavir
Inhibiteur de fusion	T20	Enfuvirtide
Antagoniste du co-récepteur CCR5	MRV	Maraviroc
Inhibiteur de l'intégrase	DTG	Dolutégravir
	EVG	Elvitégravir
	RAL	Raltégravir

* Dans cette thèse, on fait souvent référence à une de ses formulations: ténofovir disoproxil fumarate (TDF)

1.6. Les biomarqueurs du VIH

Il existe plusieurs marqueurs biologiques (biomarqueurs) associés à l'ampleur de l'infection au VIH et son impact sur l'hôte. Une description de ces marqueurs qui sont le plus souvent mentionnés dans la littérature scientifique est détaillée ici. Par la suite, un résumé est fourni sous forme de tableau (voir Table 1.2).

Le biomarqueur majeur représente la concentration d'ARN viral dans le plasma, aussi nommé charge virale plasmatique ou virémie.(Pasternak, Lukashov, & Berkhout, 2013) La charge virale est mesurée par le nombre de copies d'ARN par millilitre de plasma et peut largement varier chez les individus infectés, allant de moins d'une copie à plusieurs millions de copies/mL.(Rosenbloom, Hill, Rabi, Siliciano, & Nowak, 2012) Cette mesure se veut représentative du nombre de virions libres dans le plasma. Comme il existe deux copies d'ARN par virion, on juge que la concentration de virions libres dans le plasma équivaut à environ la moitié de la charge virale. D'un point de vue mécanistique, la charge virale résulte d'une dynamique d'entrée et d'élimination des virions du plasma. Or, l'entrée dans le compartiment plasmatique à partir du tissu lymphatique serait rapide et l'élimination des virions à partir du compartiment plasmatique ou des tissus lymphatiques secondaires serait encore plus rapide.(De Boer, Ribeiro, & Perelson, 2010; Ramratnam et al., 1999) Ainsi, on considère régulièrement que la charge virale est proportionnelle au nombre de cellules productrices de virions au moment où le sang a été prélevé.(Pasternak et al., 2013)

D'un point de vue technique, on mesure la charge virale par un échantillon de sang duquel le plasma est extrait par centrifugation. L'ARN polymérase amplifie l'ARN viral de manière à pouvoir la quantifier dans l'échantillon de plasma.(Pasternak et al., 2013) Le processus d'amplification porte le nom de réaction en chaîne par polymérase (RCP). La limite de détection a grandement évolué pour atteindre une capacité de détection d'une seule copie d'ARN par mL.(Palmer et al., 2003) Ainsi, par exemple, sur un échantillon plasmatique de 8 mL, la limite de détection atteint environ 0.2 copie/mL (seul autour de 55% du plasma est utilisé pour quantifier l'ARN, alors que le reste de l'échantillon est utilisé comme contrôle).(Zheng et al., 2013) Il est rare de nécessiter un tel niveau de précision au niveau clinique. Dans ce contexte, on utilise plutôt des techniques dont la limite de détection est d'environ 20 copies/mL. En général, un des objectifs thérapeutiques est le maintien de la charge virale en dessous de 20 ou alternativement 50 copies/mL, auquel cas on estime que la charge virologique est supprimée ou, de manière équivalente, qu'il y a suppression virologique, et on conclut au succès virologique. Un échec virologique représente la situation opposée.

Une seconde mesure d'importance est la quantité d'ADN viral extraite des cellules mononucléées sanguines périphériques (CMSP), principalement composées de lymphocytes et

de monocytes (précurseurs de macrophages circulant dans le sang).(Riedhammer, Halbritter, & Weissert, 2016) Toute forme d'ADN du VIH peut être détectée de ces cellules, que l'ADN soit intégré ou pas aux chromosomes de l'hôte. La mesure est généralement rapportée en nombre de copies par millions de CMSP.(Rouzioux & Avettand-Fenoel, 2018) Elle est obtenue en isolant les CMSP d'un échantillon sanguin du patient, en lysant les cellules et en quantifiant l'ADN par RCP.(Massanella & Richman, 2016) La mesure serait représentative du nombre de cellules qui sont capables de produire des virions, chez l'hôte.(Massanella & Richman, 2016) Il est intéressant de noter que les concentrations intracellulaires d'ARN viral peuvent également être mesurées chez les CMSP.(Pasternak et al., 2013) L'interprétation de ce marqueur différerait de celle de la charge virale plasmatique, indiquant plutôt un nombre de cellules en fin de cycle de réplication.(Pasternak et al., 2013)

Afin de quantifier l'impact du VIH sur le système immunitaire de l'hôte, on rapporte également la taille des populations de lymphocytes T. Une attention particulière est attribuée aux lymphocytes T auxiliaires (CD4) qui sont les principales cibles du VIH, et aux lymphocytes T cytotoxiques (CD8) qui ont pour rôle de causer la destruction des cellules infectées. Les concentrations de cellules sont mesurées à partir d'un échantillon sanguin en cellule/ μ L ou cellule/ mm^3 .(Serrano-Villar et al., 2014) La concentration de lymphocytes T auxiliaires permet d'évaluer la fatigue immunitaire et le risque d'infection opportuniste.(Ford, Meintjes, Vitoria, Greene, & Chiller, 2017) Le ratio du nombre de lymphocytes T cytotoxiques sur celui des lymphocytes T auxiliaires est également d'intérêt puisqu'associé à la mortalité chez les séropositifs du VIH.(Serrano-Villar et al., 2014) De plus en plus, la littérature rapporte le dénombrement de sous-populations de lymphocytes T auxiliaires ou cytotoxiques et leur influence sur l'évolution de la maladie.(Geginat et al., 2013; Mittrucker, Visekruna, & Huber, 2014) Ainsi on s'intéresse au nombre de lymphocytes naïfs, effecteurs et mémoires, et également au nombre de lymphocytes ayant des fonctions très spécialisées (e.g. régulation de population de lymphocytes, implication dans la reconnaissance du soi, etc). Ces cellules se distinguent principalement par la prépondérance de marqueurs de surface spécifiques.(Delves, Martin, Burton, & Roitt, 2017)

Table 1.2 – Les principaux biomarqueurs du VIH et leur rôle principal

Biomarqueur	Rôle
Charge virale ou concentration plasmatique d'ARN viral	Mesure associée à l'activité virale au moment de la collecte de plasma, i.e. la quantité de nouveaux virions produits
Concentration d'ADN viral	Représentative du nombre de cellules infectées qui sont productrices de virions ou en état de latence
Concentration de lymphocytes T auxiliaires	Indicatrice de l'impact cumulé du VIH sur le système immunitaire
Concentration de lymphocytes T cytotoxiques	Indicatrice du risque de complications associées au VIH

1.7. Les défis scientifiques

Dans cette section, on décrit plusieurs problématiques associées au traitement du VIH. On y discute particulièrement les sujets qui sont abordés au cours de la présente thèse, i.e. la persistance du virus, la résistance aux antirétroviraux, l'adhésion et la diminution du fardeau posologique. D'autres thèmes d'intérêt ne sont pas couverts ici puisque largement discutés dans la littérature, autant dans le domaine des vaccins et les difficultés relatives d'en produire pour éradiquer le virus, la détection tardive, les méthodes de prévention, l'accès aux médicaments et la stigmatisation sociale.(Chambers et al., 2015; Okwundu, Uthman, & Okoromah, 2012; Palmisano & Vella, 2011; Rosenberg, Pilcher, Busch, & Cohen, 2015; Shapiro, 2014; Whaley & Mayer, 2014)

1.7.1. La persistance du virus

Pour la plupart des patients, les traitements antirétroviraux hautement actifs doivent être pris de manière continue afin de diminuer les charges virales ou de les maintenir à de faibles

niveaux. Chez les patients initiant un tel traitement, la diminution de la charge virale s'effectue en plusieurs phases distinctes.(Costiniuk & Jenabian, 2015) Au cours des premiers jours, les patients voient celle-ci diminuer très rapidement, laissant présager une suppression rapide de la charge virale. Durant cette période, la demi-vie de réduction est d'environ 24 heures.(Cardozo et al., 2017) Malheureusement, cette cadence n'est pas maintenue. Environ une semaine après le début des traitements, la demi-vie de réduction augmente à 2 semaines, ralentissant significativement la décroissance de charge virale.(Perelson et al., 1997) Après ces deux premières phases, on constatera une nouvelle phase où la demi-vie augmentera encore, cette fois à environ 39 semaines, suivi d'une dernière phase où la charge virale demeure pratiquement constante.(Palmer et al., 2008) D'un point de vue mécanistique, les causes de l'observation de ces phases demeurent incertaines. On élaborera sur ce sujet au chapitre 3.

La dernière phase est caractérisée par une virémie résiduelle, c'est-à-dire une charge virale relativement basse, voire indécélable.(L. Shen & Siliciano, 2008) Cette virémie résiduelle exhibe une grande variabilité interindividuelle, et varie de moins de 0.2 à quelques dizaines de copies/mL.(Zheng et al., 2013) La présence d'ARN viral laisse sous-entendre que le virus continue à se répliquer durant cette quatrième phase.(L. Shen & Siliciano, 2008) En effet, comme la clairance de virions est considérée comme un processus rapide, on s'attendrait à ce que tout virion détecté soit un nouveau virion. À ce jour, la source de ces virions demeure incertaine.(Sahu, 2015)

Une explication possible aux phénomènes susmentionnés serait la faible exposition de certains tissus aux antirétroviraux, permettant ainsi l'infection de nouvelles cellules et la production de nouveaux virions dans ces tissus.(Murray et al., 2007) Cette hypothèse est appuyée par l'observation de différences importantes de concentrations d'antirétroviraux entre les compartiments physiologiques.(Fletcher et al., 2014; Huang et al., 2016; Yilmaz, Price, & Gisslen, 2012) Plus particulièrement, une étude rapporte de très faibles concentrations à l'intérieur des monocytes des ganglions lymphatiques comparativement aux monocytes du sang périphérique, et ce pour de nombreuses molécules actives.(Fletcher et al., 2014) Plus de détails sur cette étude et sur les hypothèses qu'elle permet de soulever sont fournis aux chapitres 2 et 3.

Après l'atteinte d'une charge virale très basse, il serait tentant de recommander un arrêt des traitements chez les patients. Hélas, même si la virémie résiduelle est indécélable, un arrêt des traitements résulte presque toujours en un retour à des charges virales élevées.(Li et al., 2016) Pour cette raison, on croit qu'un réservoir du virus subsiste chez l'ensemble des patients et peut provoquer la résurgence de l'infection.(Eisele & Siliciano, 2012) Ce réservoir aurait comme principale composante des cellules infectées latentes, c'est-à-dire des cellules infectées qui ne produisent pas de virions, qui évadent la détection par le système immunitaire et dont la production de virion peut être enclenchée par réactivation.(Sengupta & Siliciano, 2018) Ces cellules latentes seraient majoritairement des lymphocytes T auxiliaires de type mémoire qui contiendraient de l'ADN viral.(G. Q. Lee & Lichterfeld, 2016) Les mécanismes sous-jacents à la formation, à l'activation ou à l'élimination de ces cellules demeurent relativement inconnus.(Sengupta & Siliciano, 2018) Ainsi, les antirétroviraux actuels ne permettent pas une destruction totale du virus. La recherche se poursuit activement pour mettre sur pied des moyens de traiter définitivement les patients ou pour prévenir de nouveaux cas par l'élaboration de vaccins efficaces. Les pièces manquantes dans le puzzle de la dynamique du virus constituent l'étape limitante dans l'aboutissement de ces efforts.

1.7.2. Adhésion et fardeau posologique

Le terme « adhésion » arbore plusieurs définitions dans la littérature scientifique.(Vrijens et al., 2012) Dans cet ouvrage, l'adhésion au régime posologique représente le niveau de concordance entre la façon dont le patient prend ses médicaments et le régime posologique prescrit. Elle est souvent mesurée par le ratio du nombre de doses prises durant une période donnée sur le nombre de doses prescrites pour cette même période (niveau général d'adhésion).

Dans ce qui suit, nous préciserons les difficultés qu'engendre la prise continue de plusieurs médicaments chez les patients du VIH (fardeau posologique). Afin de maintenir la suppression virale (en rappel, ceci correspond à une charge virale sous la limite des 20 ou 50 copies/mL) et éviter le développement de résistance, un niveau général d'adhésion d'au moins 95% est habituellement recommandé.(Osterberg & Blaschke, 2005) Cet état de discipline doit être maintenu durant tout le reste de la vie des patients atteints, sans quoi ceux-ci risquent de

nuire à leur système immunitaire à long terme. Or, de nombreux facteurs peuvent influencer l'adhésion du patient à son traitement, tel le nombre de pilules qu'il doit prendre, l'occurrence d'effets secondaires ou bien le fardeau social et financier.(Chesney, Morin, & Sherr, 2000)

Le nombre de pilules que prennent les séropositifs du VIH est particulièrement grand aux États-Unis. Dans une étude sur 151 sujets, le nombre médian de pilules prises était de 8, dont environ 3 pilules directement pour le VIH (étendue de 2 à 5) et environ 6 pilules pour des conditions autres que le VIH (étendue de 3 à 12.5).(Zhou et al., 2014) Heureusement, des efforts sont menés afin de réduire ce fardeau de polymédication. Ainsi, une seule pilule peut maintenant être prise pour un traitement antirétroviral complet combinant efavirenz, tenofovir disoproxil fumarate et emtricitabine.(Goicoechea & Best, 2007) Il en est de même pour un traitement combinant elvitegravir, cobicistat, emtricitabine et tenofovir disoproxil fumarate(Perry, 2014). Plusieurs autres pilules tout-en-un existent ou sont en voie de développement.(Caplan, Daar, & Corado, 2018)

Le deuxième facteur d'ordre pharmacologique est l'occurrence d'effets secondaires qui peuvent grandement affecter la qualité de vie des séropositifs du VIH et, par conséquent, leur niveau d'adhésion. Bien que ces effets secondaires ne soient pas létaux dans la très grande majorité des cas, ils peuvent être accablants pour les patients.(Margolis, Heverling, Pham, & Stolbach, 2014) La toxicité des antirétroviraux peut toucher plusieurs systèmes physiologiques : gastro-intestinal (e.g. diarrhée et vomissements), le système nerveux central (e.g. difficultés de concentrations ou à maintenir l'équilibre), sanguin (e.g. anémies) et peut aussi affecter plusieurs organes à long terme, incluant les reins (e.g. dysfonction), le foie et le cœur (e.g. infarctus du myocarde).(Hawkins, 2010) Si le patient souffre de comorbidité, la polypharmacie et les interactions médicamenteuses sous-jacentes peuvent également poser un problème majeur.(Reust, 2011)

Une lueur d'espoir existe afin d'alléger davantage le fardeau posologique, car une minorité de patients arrivent à maintenir de faibles charges virales et à préserver un nombre élevé de lymphocytes T auxiliaires (>200 cellules/ μ L) sans avoir recours aux antirétroviraux. On distingue plusieurs types de ces patients, dont : i) les contrôleurs élites qui constituent un très petit groupe de patients (0.2 à 0.5%) et qui maintiennent des charges virales en deçà de 50 copies/mL sans avoir eu recours à la médication; ii) les contrôleurs virologiques qui

maintiennent des charges virales entre 200 et 2000 copies/mL sans avoir eu recours à la médication et iii) les contrôleurs post-traitement qui maintiennent de faibles charges virales (<50 copies/mL) après avoir interrompu un traitement antirétroviral.(Hocqueloux et al., 2010; Okulicz & Lambotte, 2011) On peut également mentionner le groupe des non-progresseurs à long terme constitué de patients aptes à maintenir un nombre élevé de lymphocytes T auxiliaires sans prise médicamenteuse. Les causes de ces différentes conditions médicales avantageuses sont nébuleuses, mais font l'objet de nombreuses études.(Gonzalo-Gil, Ikediobi, & Sutton, 2017) Il y a un grand espoir que les informations issues de ces études mènent concrètement à des molécules et stratégies thérapeutiques permettant de diminuer le fardeau posologique des patients.

1.7.3. La résistance aux antirétroviraux

Il est malheureusement à noter que le risque de développer de la résistance aux antirétroviraux demeure un danger réel. Cliniquement, il n'est pas rare que les cliniciens rencontrent des patients dont le virus a accumulé assez de mutations pour résister à la plupart des traitements.

Les mécanismes par lesquels le virus acquiert de la résistance aux traitements sont bien connus. Premièrement, un virus de type sauvage (i.e. susceptible aux antirétroviraux) peut générer des virus dont le génome diffère du sien par mutation. Ces génomes peuvent produire des protéines qui se lient plus difficilement aux antirétroviraux, nuisant ainsi à leur potentiel d'inhibition. Plusieurs mécanismes sont associés à une grande fréquence de mutation du virus du VIH. Entre autres, la transcriptase inverse effectue beaucoup d'erreurs de transcriptions en substituant des codons ou en permettant la recombinaison de génomes différents.(Menendez-Arias, 2002; Roberts, Bebenek, & Kunkel, 1988; Schlub et al., 2014) Le taux de mutation a précédemment été estimé à 3×10^{-5} mutations par paires de bases par cycle de réplication.(Mansky & Temin, 1995) Ceci correspond à plus d'une erreur de codage par génome produit et environ une erreur par transcription de génome pour la protéine de la transcriptase inverse seule. Plus récemment, des estimations de taux de mutations encore plus grandes ont été rapportées *in vivo* (4.1×10^{-3} mutations par base par cycle de réplication).(Cuevas, Geller, Garijo, Lopez-Aldeguer, & Sanjuan, 2015)D'autres facteurs ont été mis en cause dans

l'acquisition de mutations, telles les protéines antivirales de l'hôte qui en augmente la fréquence (e.g. APOBEC3G).(Cuevas et al., 2015)

Certaines mutations augmentent la résistance du virus aux antirétroviraux. La résistance associée à une mutation est souvent rapportée par le rapport des concentrations inhibant 50% la réplication du virus entre la souche résistante et la souche sauvage qui est susceptible à l'antirétroviral.(Sampah, Shen, Jilek, & Siliciano, 2011) La plupart du temps, les mutations conférant une résistance à un antirétroviral confèrent également un coût d'adaptation.(Sampah et al., 2011) Ceci est logique si l'antirétroviral a démontré son efficacité sur les quasiespèces sauvages qui constituent le bassin de virus arrivant le mieux à se répliquer sans médication. En effet, le coût de l'adaptation du virus représente une perte d'efficacité du virus vis-à-vis de l'infection de cellules ou de la production de virions infectieux en absence d'exposition à l'antirétroviral.(Goudsmit, de Ronde, de Rooij, & de Boer, 1997) Ceci est normalement synonyme d'une plus faible prévalence des virions résistants dans la population virale de l'hôte si ce dernier n'a pas été exposé aux antirétroviraux. Une exposition intermédiaire aux antirétroviraux peut, cependant, exercer une pression de sélection. En effet, une exposition intermédiaire peut nuire fortement à la réplication de la souche sauvage, mais beaucoup moins à celle de la souche résistante.(Rosenbloom et al., 2012) Ceci résulte en une amplification de la population virale résistante et peut rendre, ultimement, l'antirétroviral inefficace à des doses non toxiques. (Rosenbloom et al., 2012) Qui plus est, le virus déjà résistant à un antirétroviral peut acquérir de nouvelles mutations par le même processus que susmentionné. Ainsi, le virus peut devenir résistant à plusieurs antirétroviraux et rendre l'action de combinaisons d'antirétroviraux inefficace.(Hill, Rosenbloom, & Nowak, 2012; Pennings, 2013)

Le problème de la résistance motive le développement de nouvelles molécules ou de nouveaux régimes posologiques qui seraient efficaces malgré l'accumulation de certaines résistances. Le chapitre 2 élaborera sur un problème associé au développement de nouveaux régimes posologiques, i.e. le fait qu'il soit difficile de prévoir le résultat virologique d'une combinaison d'antirétroviraux qui a peu souvent été le sujet d'études cliniques.

1.8. Apport de la modélisation mécanistique

L'approche adoptée dans cette thèse est celle de la modélisation mécanistique dans le cadre de la pharmacologie des systèmes (QSP en anglais, pour « quantitative systems pharmacology »). Elle s'intéresse aux mécanismes biologiques et pharmacologiques donnant lieu à l'effet des médicaments. L'approche en question diffère substantiellement de celles communément rapportées dans les journaux scientifiques portant sur la pharmacologie. Afin de mieux comprendre ce que la modélisation mécanistique peut apporter à ce domaine, nous décrivons en premier lieu les qualités et lacunes des approches plus communes.

Si on y tirait au hasard dans la littérature pharmacologique un article ayant comme objectif de mieux comprendre un phénomène donné, il est probable que l'approche qui y est rapportée soit la suivante : i) des données empiriques ont été recueillies selon un protocole expérimental dans le but de tester une hypothèse de recherche, et ii) des méthodes statistiques ont été appliquées à ces données afin d'en soutirer des valeurs de paramètres ou d'évaluer la force des liens entre variables et facteurs, tout en estimant les degrés d'incertitude associés. Habituellement dans ce type d'études, on tente de limiter toute influence provenant de suppositions préalables dans l'analyse statistique afin que ce soit majoritairement les données expérimentales qui dictent les résultats de la recherche. Si toute analyse statistique fait tout de même appel à certaines suppositions, cette approche veut en minimiser le nombre, augmentant le niveau de crédibilité de l'interprétation basée sur ces résultats. Or, ce gain en crédibilité a également un prix.

En fait, il est rare que de l'information supplémentaire provenant d'études indépendantes serve à même l'analyse statistique qui a été effectuée puisque l'inclusion d'information nécessite de faire appel à des suppositions supplémentaires. Par exemple, il arrive rarement qu'un paramètre estimé dans une étude indépendante serve dans une procédure de régression appliquée sur d'autres données. Des exceptions existent, notamment dans le contexte des études qui ont fait appel à des techniques bayésiennes basées sur des distributions de valeurs de paramètres *a priori*. (Gelman, Carlin, Stern, & Rubin, 1995) Il y a également la méta-analyse qui regroupe et analyse des données de plusieurs études portant sur le même sujet. Par contre et plus fréquemment, si des données autres que celles collectées au cours de l'étude ont été

considérées, elles ont servi à motiver les objectifs (e.g. étude clinique motivée par des résultats suggérant l'efficacité d'une molécule *in vitro*) ou sont utilisées afin d'ajouter à l'interprétation des résultats. À défaut de faire intervenir des données d'études indépendantes, ce sont les données collectées à même l'étude qui servent à estimer les valeurs de tous les paramètres associés aux relations considérées (e.g. paramètres de régression). Si on considère comme recommandation statistique générale le besoin de 10 observations par paramètre estimé afin d'éviter qu'un modèle ne surajuste les données,(Harrell, 2014) on comprend qu'il peut s'avérer souhaitable que tout modèle considéré soit simple. Ainsi, sous cette approche, un nombre restreint d'interactions entre les éléments d'un système peut être étudié à la fois.

Or, force est de constater que beaucoup de maladies ne sont pas le résultat de défauts d'un seul gène ou bien de simples interactions entre l'hôte et le pathogène, mais bien le résultat d'interactions biologiques complexes.(Schneider & Klabunde, 2013) L'ajout d'agents exogènes dans l'équation complique d'autant plus ces interactions. Par le fait même, l'effet global des agents exogènes est difficile à prédire via la considération d'une seule étude qui utilise l'approche susmentionnée. Il en revient donc aux scientifiques de la pharmacologie d'effectuer un travail d'intégration des diverses études et éléments de connaissance répartis dans la littérature pour assembler un « modèle global » d'un phénomène, très souvent parsemés de suppositions rationnelles afin de pallier au manque de disponibilité de certaines informations.(Craig, Gonzalez-Sales, Li, & Nekka, 2016) C'est à partir de ce modèle que de nouvelles hypothèses sont formulées et que d'importantes décisions scientifiques sont prises. Par exemple, un modèle de persistance du virus du VIH chez les patients sous thérapie implique principalement l'activation de cellules infectées latentes,(Sengupta & Siliciano, 2018) ces cellules ayant été observées chez les patients après plusieurs années de thérapie(Chun et al., 1995) et ayant pu produire des virions infectieux *in vitro*.(Rosenbloom et al., 2012) On suppose que ces cellules, si activées *in vivo*, seraient majoritairement éliminées par les lymphocytes T cytotoxiques de l'hôte avant qu'elles ne puissent produire de nouvelles cellules infectées latentes. Ce modèle prédit une diminution marquée de la taille du réservoir de telles cellules par l'administration d'agents d'activation chez l'hôte. Des études ont été menées dans le but de tester cette hypothèse par de nouvelles expériences.(Archin et al., 2012; Sogaard et al., 2015) Les résultats obtenus jusqu'à maintenant via cette approche sont décevants au niveau

quantitatif : on note une faible diminution de la taille du réservoir, et on en comprend mal la raison.(Sengupta & Siliciano, 2018) Le modèle, tel que nous l'avons décrit, est-il trop simple?

Ci-haut, c'est la prédiction du modèle imaginé par le scientifique qui motive sa décision de mener une expérience pour étudier l'effet des agents d'activations des cellules *in vivo*. Or, la précision des prédictions de l'humain quant au comportement d'un système décroît rapidement avec le nombre de processus qui interagissent.(Miller, 1956) Sans outil additionnel pour guider les nouvelles hypothèses et les décisions scientifiques qui découlent de modèles imaginés, il arrive que celles-ci soient erronées, et ces erreurs sont d'autant plus fréquentes que le système à l'étude est complexe. C'est à ce niveau que l'utilité de la modélisation mécanistique atteint son apogée, jugée comme « notre meilleure arme contre le comportement contre-intuitif de systèmes complexes » (traduction libre)(Phair, 2014)

Contrairement à l'approche plus classique, celle par modélisation mécanistique se veut intégrative. Elle consiste en la représentation de l'ensemble des processus qui dictent le comportement d'un système. Pour fin d'illustration, nous pouvons prendre l'exemple de la résistance aux antirétroviraux chez les patients du VIH, qui est le sujet du prochain chapitre. Le modèle que nous avons développé est une représentation mathématique intégrant l'ensemble des processus qui dictent le développement et l'évolution temporelle de la résistance *in vivo*. Ce modèle reproduit la concentration de médicaments aux endroits où le virus se réplique dans le corps et intègre la probabilité de mutation du virus, l'impact de ces mutations sur sa réplication (coût d'adaptation) et sur sa susceptibilité aux différents médicaments, etc. Afin de représenter l'ensemble des processus impliqués, nous avons intégré dans le modèle un ensemble d'informations quantitatives provenant de la littérature que nous avons complété par des déductions rationnelles et des hypothèses pour pallier à l'indisponibilité de certaines informations.

Un modèle mécanistique représente en général chaque processus par des liens de causalité. Un exemple détaillé est fourni dans la section intitulée « Supplementary material » du prochain chapitre. Les liens entre variables sont exprimés par des fonctions paramétriques ou d'algorithmes informatiques. Les paramètres du modèle peuvent être directement associés à une mesure physiologique (e.g. la clairance intrinsèque du médicament), ou être de nature purement

empirique pour des raisons pratiques (e.g. valeurs d'un paramètre de régression linéaire basées sur les données de deux variables).

La qualité du modèle est analysée en se basant sur les prédictions qui en découlent (e.g. prédiction de la charge virale d'un patient au cours du temps) et qui s'obtiennent par des méthodes analytiques ou numériques.(Collins, 1990; Perko, 2013) Ces prédictions peuvent être qualitatives en décrivant le comportement du système modélisé (e.g. observations d'oscillations au cours du temps de la charge virale d'un patient) ou numérique (e.g. valeurs de charges virales). Dans ce dernier cas, le modèle doit être entièrement défini, c'est-à-dire que toutes les valeurs de paramètres du modèle doivent être uniquement déterminées. Dans ce qui suit, nous discuterons de l'utilité et des limites associées à l'analyse de modèles mécanistiques.

1.8.1. Utilité de l'approche

Un modèle est évalué par le degré de concordance entre les prédictions de celui-ci et les observations empiriques. D'un point de vue interprétatif, un degré de concordance élevé signifie que le modèle (l'ensemble des processus modélisés, suppositions rationnelles et hypothèses) forme une explication plausible aux données observées. Une concordance élevée ne signifie pas que le modèle est la seule représentation raisonnable de la réalité puisque plusieurs modèles peuvent potentiellement expliquer les données observées avec une précision similaire.

On évalue la valeur du modèle par son degré de fiabilité. Cette approche partage des similarités avec celles inspirées de la philosophie bayésienne. (Berger et al., 2001; Kirk, Thorne, & Stumpf, 2013) La logique bayésienne stipule que le niveau de confiance ou degré de fiabilité associé à un modèle dépend du niveau de confiance initial du scientifique par rapport à ce modèle (incluant donc la subjectivité du scientifique) et de la probabilité d'observation des données empiriques si le modèle représente parfaitement la réalité.(Berger et al., 2001) Toujours selon cette conceptualisation, la quantité de données que le modèle peut prédire avec précision augmente le degré de fiabilité du modèle. Dans le cas de la modélisation mécanistique, il est possible de tenir compte du degré d'incertitude de nos croyances, tel que pour l'approche bayésienne, mais ceci n'est pas usage courant. Une autre différence notable entre l'approche bayésienne et l'approche par modélisation mécanistique est que la première attribue généralement une valeur de probabilité *a posteriori* d'un modèle parmi plusieurs modèles, alors

que l'approche par modélisation mécanistique évalue généralement un seul modèle sans avoir recours à une mesure particulière. Plutôt, on utilisera directement la capacité du modèle à répliquer des données observées afin d'évaluer si ce modèle est une explication plausible au phénomène modélisé. Une règle non écrite est que, en ordre croissant d'importance, un modèle est d'autant plus fiable s'il prédit des données : 1) qui ont partiellement servi à son élaboration, 2) qui n'ont pas servi à son élaboration, mais proviennent d'études similaires, 3) qui n'ont pas servi à son élaboration et proviennent d'études où les conditions expérimentales diffèrent. Autrement dit, plus le modèle arrive à décrire correctement des observations du phénomène modélisé, et ce dans des situations diverses, plus le niveau de confiance en la capacité du modèle à effectuer adéquatement d'autres prédictions augmente. À l'inverse, si le phénomène a donné lieu à des données empiriques discordantes avec les prédictions du modèle, celui-ci est jugé moins fiable. Une grande discordance entre données empiriques et prédictions suggère que ce modèle n'est pas une représentation adéquate de la réalité, i.e. au moins un des éléments du modèle est inapproprié. D'un point de vue rationnel, il est vraisemblable que les éléments en cause sont ceux qui ont un faible appui scientifique, et donc prioritairement les hypothèses intégrées au modèle.

Cette philosophie donne lieu à plusieurs utilités à la modélisation mécanistique :

- 1) **Identifier des explications plausibles à un phénomène mal compris.** D'un point de vue méthodologique, ceci revient à établir un modèle du phénomène où certains éléments sont bien établis alors que d'autres sont encore à découvrir. L'idée est de trouver des formulations plausibles pour les éléments inconnus de sorte que les prédictions du modèle concordent avec des données empiriques provenant d'observations du phénomène. La formulation ainsi obtenue peut inspirer des explications mécanistiques appuyées par les données empiriques. On verra des exemples aux chapitres 3 et 4.
- 2) **Évaluer la fiabilité d'hypothèses déjà formulées.** D'un point de vue méthodologique, ceci revient à établir un modèle du phénomène qui intègre l'hypothèse en question et de démontrer qu'il existe (ou pas) des valeurs de paramètres justifiées ou raisonnables qui permettent aux prédictions du modèle de concorder avec des données empiriques. Selon la logique étalée plus haut, le niveau

de confiance du modèle augmente significativement si celui-ci arrive à prédire adéquatement des données qui sont indépendantes. Inversement, si les éléments du modèle trouvent un appui scientifique à l'exception de l'hypothèse intégrée et qu'il y a une grande discordance entre prédiction et observation, la fiabilité de l'hypothèse diminue. On utilise cette approche dans les chapitres 2 et 3.

- 3) **Explorer le comportement du système modélisé dans des conditions qui n'ont pas été étudiées expérimentalement.** Si le modèle est jugé fiable, il peut être utilisé pour appuyer des décisions scientifiques basées sur des prédictions de situations encore inexplorées par des observations empiriques, mais où les mécanismes sous-jacents demeurent sensiblement les mêmes. Plus la décision scientifique a de l'importance (e.g. décision clinique), plus le niveau de fiabilité du modèle doit être élevé. On mentionnera la possibilité d'utilisation d'un modèle pour les décisions cliniques au chapitre 2.

Ainsi, l'utilité de l'approche par modélisation mécanistique est complémentaire à celle couramment utilisée et qui fait plus strictement appel à la statistique. Si la dernière approche permet de tester des hypothèses de recherche, la première aide à les suggérer et à faire un premier tri de celles-ci grâce à la comparaison entre données empiriques et prédictions de modèles intégrant ces hypothèses.

Il est important de noter que le parallèle entre statistiques et modèles mécanistiques élaboré plus haut concerne principalement les méthodes paramétriques (modèles avec paramètres). Les méthodes statistiques non-paramétriques peuvent également être très pratiques afin de représenter la complexité de données émanant d'un système (e.g. méthodes de l'apprentissage machine). Par contre et tout comme pour les méthodes paramétriques, les méthodes non-paramétriques ont un apport complémentaire à l'approche par modélisation mécanistique, ces premières permettant habituellement une grande précision des prédictions, mais sont souvent difficiles à interpréter et nécessitent une grande quantité de données.

1.8.2. Principes de formulation d'un modèle

La formulation d'un modèle implique plusieurs choix de la part du modélisateur. Le modélisateur doit premièrement identifier l'ensemble des processus qui ont une influence sur la

manifestation d'un phénomène étudié. Si certains de ces processus sont incertains ou inconnus selon la littérature scientifique, des hypothèses sont formulées afin d'avoir un portrait complet du système impliqué. Tous les processus identifiés seront incorporés au modèle. La représentation mathématique de ces processus résulte d'un choix de la part du modélisateur, ce choix étant majoritairement dicté par les objectifs de l'étude, la disponibilité et la fiabilité de données quantitatives (e.g. données de recherche, données cliniques, estimations de paramètres) permettant de caractériser les processus modélisés et un souci de parcimonie.

Afin d'illustrer ce propos, supposons que notre objectif est de prédire l'évolution temporelle de la charge virale d'un patient en fonction de son usage d'antirétroviraux. La charge virale étant la variable que nous désirons reproduire, le système modélisé devra représenter l'ensemble des processus ayant un impact significatif sur celle-ci. Comme l'objectif est de connaître l'impact de l'usage d'antirétroviraux, les processus qui lient le médicament à son effet devront être modélisés avec une assez grande précision. La disponibilité de données quantitatives fiables sur les différents antirétroviraux (e.g. clairance du médicament et sa variabilité entre les individus), permet de modéliser fidèlement les concentrations plasmatiques de ces médicaments en fonction de leur utilisation. Le système immunitaire peut également influencer l'évolution de la charge virale, quoiqu'à moindre mesure. L'information est moins riche dans la littérature quant à l'évolution temporelle du nombre de lymphocytes T cytotoxiques et auxiliaires actifs chez les patients. Les besoins de l'étude n'obligent pas une modélisation précise du phénomène. Par ailleurs, l'action du système immunitaire sera modélisée en utilisant un paramètre constant propre à chaque patient. Cette approximation peut être justifiée par la capacité du modèle à reproduire les données de charges virales de patients ayant cessé leur médication. Sans cette approximation, des suppositions supplémentaires plus difficilement justifiables auraient été nécessaires relativement à l'activité des cellules immunitaires.

La formulation d'un modèle résulte donc d'un équilibre entre précision et approximation. Une grande complexité ou bien la présence d'éléments non justifiés nuisent à la fiabilité d'un modèle. Heureusement, plusieurs processus biologiques peuvent être représentés par des équations relativement simples et ayant démontré leur pouvoir de prédiction dans diverses situations. Également et comme l'a mentionné le grand mathématicien, physicien et

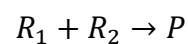
statisticien Pierre-Simon Laplace : « Tous les effets de la nature ne sont que résultats mathématiques d'un petit nombre de lois immuables ».(marquis de Laplace, 1847)

1.8.3. Représentation mathématique

Un modèle est constitué d'expressions mathématiques qui impliquent un ensemble de variables dont le rôle est de décrire l'état du système modélisé. Par exemple, dans un modèle du diabète, l'état du système d'intérêt à un moment t pourrait être caractérisé par la concentration d'insuline (variable 1) et le taux de sucre sanguin (variable 2). Les processus biologiques du système causent des changements aux valeurs des variables (e.g. métabolisme de l'insuline). L'ampleur des changements attendus à ces valeurs dépend généralement de l'état du système et/ou de ces états antérieurs (e.g. la quantité d'insuline éliminée qui dépend de la concentration d'insuline). Ceci explique pourquoi les processus sont souvent modélisés sur des intervalles de temps infinitésimaux dt , où les valeurs de variables peuvent être considérées constantes et l'ampleur du changement de la valeur d'une variable peut ainsi être caractérisée. Lorsque les variables modélisées sont continues ($\vec{c}(t)$) et que l'issue des changements sur des intervalles de temps infinitésimaux dt est unique (processus déterministe), les modèles sont généralement exprimés par un système d'équations différentielles, i.e. $\frac{d\vec{c}}{dt} = f(\vec{c}, t)$. Lorsque l'ampleur du changement est incertaine (processus stochastique), celle-ci est souvent exprimée comme une variable aléatoire dont les paramètres de distribution sont déterminés par une fonction $f(\vec{c}, t)$.

1.8.3.1. Forme mathématique de processus biologiques déterministes communs

Beaucoup de processus biologiques déterministes peuvent être décrits par des lois simples dont l'une des plus importantes est la loi d'action de masse. Initialement, cette loi a été utilisée pour décrire la quantité de produits dans un volume fixe en fonction du temps, ce volume contenant une solution homogène de deux réactants.(Lund, 1965) Cette loi stipule que la vitesse de création d'un produit P à un temps t est proportionnelle à la multiplication des quantités de réactants R_1 et R_2 . En d'autres mots, la réaction chimique décrite par:



peut s'exprimer par le système d'équations différentielles :(Lund, 1965)

$$\frac{dP}{dt} = \alpha R_1 R_2$$
$$\frac{dR_1}{dt} = \frac{dR_2}{dt} = -\alpha R_1 R_2$$

où α est une constante. Du point de vue de la physique, le système représente ce qui se produit lorsque deux types de molécules sont présents de manière homogène dans un volume délimité, s'y déplacent et entrent en contact les uns avec les autres de manière aléatoire, et une proportion de ces rencontres résulte en une transformation des molécules. Or, d'autres phénomènes sont le résultat de processus similaires: i) des virions dans un milieu physiologique entrant en contact avec des cellules cibles et où certains de ces contacts résultent en infections de cellules, ii) une molécule exogène entrant en contact avec des récepteurs cellulaires chez un hôte et génère une réponse de la cellule, ou bien encore iii) des prédateurs dans un milieu forestier rencontrant des proies et certaines de ces rencontres résulte en la mort de la proie. Bien sûr, dans la plupart des situations susmentionnées, la supposition de mélange homogène n'est pas parfaitement respectée, les éléments impliqués dans la « réaction » pouvant être regroupés. Par contre, même au niveau écologique ou épidémiologique, la loi d'action de masse permet aux modèles de prédire adéquatement le comportement de populations d'organismes.(Voit, Martens, & Omholt, 2015) Dans les chapitres qui suivent, cette loi est utilisée afin de décrire l'infection de cellules CD4 par le VIH.

Une autre loi mathématique simple permet de décrire des processus biologiques qui sont souvent issus de phénomènes décrits par la loi d'action de masse, la différence étant qu'une des variables impliquées demeure constante. Par exemple, on peut considérer le métabolisme d'une molécule par des enzymes. Si les enzymes ne sont pas saturés et que leur quantité (E) demeure approximativement constante, alors le taux de transformation de la molécule (M) peut être exprimé par $\frac{dM}{dt} = -\alpha EM = -\beta M$, où β est la constante unissant α et E. L'équation stipule que la vitesse de changement d'une quantité M est proportionnelle à cette même quantité. Les processus pouvant être décrits mathématiquement par cette équation sont dits linéaires.

La linéarité des processus saturables peut être compromise. Par exemple, lorsqu'on considère l'action d'une enzyme sur un substrat afin d'engendrer un produit, il est possible de

saturer l'enzyme de sorte que la vitesse de création du produit tende vers une valeur maximale. Un autre exemple pourrait être la saturation de la transcriptase inverse par un médicament, de sorte que l'ajout de médicament ne puisse plus nuire de surcroît à la fonction de la protéine virale. Dans ce cas, la représentation mathématique du processus implique souvent l'équation de Michaelis-Menten ainsi que sa généralisation, l'équation de Hill. (Gesztelyi et al., 2012) L'équation de Michaelis-Menten peut être dérivée d'une modélisation de la réaction suivante :



où E représente une quantité d'un enzyme libre, S, une quantité d'un substrat libre, SE, une quantité d'enzyme lié au substrat et finalement P, une quantité d'un produit, tous dans un volume fixe. La représentation par équations différentielles d'une telle réaction est :

$$\frac{d[S]}{dt} = k_{-1}[SE]$$

$$\frac{d[E]}{dt} = (k_{-1} + k_2)[SE]$$

$$\frac{d[SE]}{dt} = k_1[S][E] - k_{-1}[SE] - k_2[SE]$$

$$\frac{d[P]}{dt} = k_2[SE]$$

où k_1 , k_{-1} et k_2 représentent des taux de liaison du substrat à l'enzyme, de séparation de l'enzyme, et de transformation en produit, respectivement. (Edelstein-Keshet, 1988) Il est possible de réduire ce système d'équations. En effet, on sait que $[SE] + [E]$ doit toujours être égale à la quantité totale d'enzyme $[E]_T$. De plus, on peut supposer que $[SE]$ atteint rapidement sa valeur d'équilibre en fonction de $[S]$, i.e. la variation de $[SE]$ est entièrement déterminée par la variation de $[S]$. Ceci revient mathématiquement à supposer que $\frac{d[SE]}{dt} = 0$ est immédiatement atteint, ou de manière équivalente que

$$k_1[S][E] - k_{-1}[SE] - k_2[SE] = 0.$$

Ceci permet de déduire mathématiquement que

$$[S]([E]_T - [SE]) = \frac{k_{-1} + k_2}{k_1} [SE]$$

et d'exprimer [SE] comme

$$[SE] = \frac{[E]_T [S]}{K_M + [S]}$$

où $K_M = \frac{k_{-1} + k_2}{k_1}$. On peut donc exprimer la vitesse de production de [P] en fonction de la concentration de substrat :

$$\frac{d[P]}{dt} = k_2 [SE] = \frac{V_{max} [S]}{K_M + [S]}$$

où $V_{max} = k_2 [E]_T$. On reconnaît ici l'équation de Michaelis-Menten. L'équation de Hill peut également être dérivée de façon similaire, mais dans la situation où plusieurs substrats peuvent se lier à l'enzyme et où la liaison de substrats additionnels est altérée par les liaisons précédentes. (Gesztelyi et al., 2012; Ortiz, 2013) Ce type de liaison est fréquent en pharmacologie. En général, l'équation de Hill est exprimée par :

$$f = \frac{1}{1 + \left(\frac{K}{C}\right)^m}$$

où f représente la fraction liée d'une molécule à un récepteur, C est sa concentration dans un milieu et m est le paramètre de Hill. Ce dernier paramètre dépend du nombre de molécules pouvant se lier à la même entité possédant plusieurs récepteurs (enzyme, cellule, bactérie, etc.) et de la facilité de liaison de multiples molécules à cette même entité. (Gesztelyi et al., 2012)

1.8.3.2. Forme mathématique des processus stochastiques

Le hasard joue un rôle dans la plupart des processus susmentionnés. Par exemple, si on tentait de prédire l'emplacement d'une seule molécule dans le sang à partir de sa position précédente, il est probable que cette prédiction soit loin de la réalité: la molécule peut avoir pénétré une cellule ou avoir bifurqué dans une ramification quelconque du système sanguin. Selon la même idée, il est difficile de prédire le moment où cette même molécule sera métabolisée par un enzyme. Ainsi le temps avant la métabolisation de la molécule suit plutôt une loi de probabilité. Pourtant, les modèles déterministes arrivent à prédire avec un assez grand degré de précision l'évolution de la concentration du médicament dans le sang. La raison sous-jacente est la loi des grands nombres qui décrit le comportement d'un grand nombre

d'événements stochastiques suivant une même loi.(Tchebichef, 1846) Dans un modèle simpliste, on suppose qu'après chaque battement de cœur, toutes les molécules d'un médicament ont une probabilité p d'être métabolisée. S'il y a un très grand nombre de ces molécules, alors la loi des grands nombres stipule que la somme du nombre d'événements de métabolisation se produisant entre chaque battement de cœur peut être déterminée avec quasi-exactitude. Ceci justifie l'utilisation de lois déterministes pour modéliser un bon nombre de phénomènes. Cependant, il arrive qu'un événement soit si rare ou implique si peu d'éléments que la représentation déterministe n'est plus adéquate. Dans ce cas, le processus est modélisé par des variables aléatoires.(Jacquez, 1999) On verra un exemple de ceci au chapitre 2, où le processus de mutation qui donne lieu à l'acquisition de résistance est un événement stochastique du cycle de reproduction du virus.

Lorsque l'événement considéré se produit peu fréquemment, le nombre d'événements est modélisé par une variable suivant une loi de Poisson.(Papoulis & Pillai, 2002) Une telle variable ne peut prendre que des valeurs discrètes (0, 1, 2, etc.), et la distribution sous-jacente est entièrement déterminée par son seul paramètre λ . Ce paramètre représente une fréquence attendue exprimée en nombre d'événements par unité de temps. Dans cette thèse, on s'est limité à des variables aléatoires qui suivent cette loi.

1.8.3.3. Paramètres et variabilité interindividuelle

Ci-haut, nous avons décrit la forme des équations mathématiques permettant la représentation de processus biologiques communs. Afin d'effectuer des prédictions numériques à partir de ces équations, il est nécessaire d'attribuer des valeurs aux paramètres associés. L'estimation des paramètres du modèle correspond au « problème inverse », selon un langage de dynamique. Les types de difficultés associées à l'estimation de paramètres sont les mêmes que pour les modèles statistiques usuels, à l'exception que certaines de celles-ci sont exacerbées dans un contexte de modélisation mécanistique faisant usage de données d'origine biologique. Cette section fera mention de ces difficultés et des solutions proposées afin d'en minimiser l'impact.

Les principales difficultés émanent de la grande complexité des modèles en termes de nombre de paramètres et de structure mathématique des équations, par rapport aux modèles

statistiques paramétriques usuels (e.g. ANOVA, régression linéaire ou logistique, etc.). Estimer un grand nombre de paramètres nécessite généralement l'usage d'une grande quantité de données. De plus, ces données doivent permettre d'informer la valeur de ces paramètres. Lorsque de telles données ne sont pas disponibles, sont insuffisantes ou inadéquates, il devient difficile voire impossible d'estimer certains paramètres avec précision. Lorsque cette situation se produit, on dira qu'il y a un problème d'identifiabilité structurelle ou pratique.

Le problème d'identifiabilité structurelle fait référence à la situation où il existe plusieurs paramétrisations $\vec{\theta}$ d'un modèle qui donnent lieu à la même vraisemblance d'observation de la ou des variables prédites.(Lehmann et Casella, 1998) Par exemple, lorsque l'on ajuste un modèle à 1 compartiment aux données de concentrations plasmatiques pour un médicament administré par voie orale, il existe plusieurs valeurs différentes de l'ensemble (clairance, volume de distribution, biodisponibilité) qui sont associées à exactement la même vraisemblance d'observation des concentrations après administration orale.(Lavielle, 2014) Plusieurs approches permettent de pallier cette situation.(Chis, Banga, & Balsa-Canto, 2011) Une approche consiste à réduire la complexité du modèle ou le nombre de paramètres. Dans l'exemple précédent, il est coutume de remplacer les paramètres de clairance, de volume de distribution et de biodisponibilité par la clairance apparente (clairance/biodisponibilité) et le volume apparent (volume/biodisponibilité), ces derniers paramètres permettant aux fonctions de vraisemblance de prendre des valeurs différentes pour chaque paramétrisation. Une autre approche consiste à fixer les valeurs de paramètres avant de procéder à la modélisation. Idéalement, la valeur du paramètre devrait lui-même avoir été estimé *a priori* par une étude indépendante et avec suffisamment de précision. Ainsi, dans l'exemple ci-haut, il aurait été possible de fixer la valeur de la clairance à 80% si cette valeur avait été rapportée dans la littérature avec un degré de précision acceptable.

Le problème d'identifiabilité pratique, quant à lui, se manifeste par une grande incertitude quant aux valeurs des paramètres estimées. Cette incertitude peut être due à une faible sensibilité de la vraisemblance des observations dans les données en fonction des valeurs desdits paramètres. Elle donne généralement lieu à de très grands intervalles de confiance pour les paramètres estimés.(Raue et al., 2009) L'origine des problèmes d'identifiabilité pratique peut être le manque de précision des mesures ou un nombre insuffisant de celles-ci. La méthode de

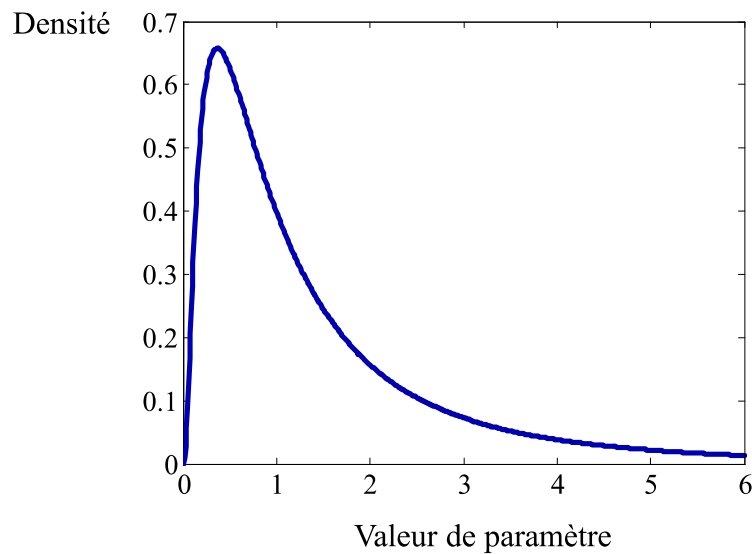
profils de vraisemblance permet de détecter à la fois les problèmes d'identifiabilité structurelle et pratique et peut être utilisée afin de reconnaître ces situations.(Raue et al., 2009) Il existe peu de méthodes permettant de pallier cette situation outre que d'augmenter la quantité et la qualité des données empiriques, ce qui s'avère un défi. En effet, du fait de la nature mécanistique de la modélisation, la plupart des paramètres représentent une caractéristique biologique ou pharmacologique, e.g. clairance d'un médicament, taux de production de cellules CD4 dans le corps, coût d'adaptation du virus, etc. Or, l'extraction de données d'ordre biologique peut être difficile, voire impossible, et très coûteuse. Il en résulte que la quantité de données disponibles afin d'effectuer l'estimation de paramètres est souvent petite ou éparse, masquant la dynamique d'un système.

À cet effet, une des tâches importantes du modélisateur est d'aller à la recherche d'information quantitative qui soit à la fois pertinente, fiable et adéquate à la modélisation. L'objectif est d'utiliser cette information afin de limiter l'estimation de paramètres par des données que le modèle tente d'expliquer. On a vu qu'il est important de pouvoir identifier l'approche qui a été utilisée dans l'obtention des valeurs de paramètres. Ainsi, on dira des paramètres qu'ils ont été estimés *a priori* (via des données indépendantes) ou à *a posteriori* (via les données à expliquer par le modèle).

Une autre difficulté surgit dans beaucoup de situations quant aux valeurs de paramètres à utiliser dans le modèle du fait que les données à expliquer ont été recueillies sur des groupes de patients. Or, si les principes biologiques et pharmacologiques dictant l'effet d'une pathologie ou d'un médicament sont souvent les mêmes, il existe une forte variabilité entre les individus quant à l'amplitude des processus sous-jacents. Par exemple, il peut y avoir des différences importantes de valeurs de clairance d'un médicament entre les individus.(Owen & Fiedler-Kelly, 2014) Cette variabilité représente une différence importante entre la modélisation mathématique de phénomènes issus de la physique et celle impliquant la biologie. Si la physique implique des constantes absolues qui peuvent être estimées avec précision (vitesse de la lumière, constante de Planck, constante gravitationnelle, masse, charge électrique, etc.), la biologie et la pharmacologie impliquent des paramètres pouvant varier significativement entre les systèmes et/ou qui sont difficilement mesurables (concentration de plaquettes, expression et efficacité des protéines, etc.).

Dans les modèles, ceci se traduit par des paramètres qui prennent des valeurs provenant de distributions. Dans la modélisation en pharmacocinétique de population, il est coutume d'utiliser la loi normale et log-normale pour représenter les distributions de paramètres biologiques et pharmacologiques tels la clairance ou le volume de distribution.(Owen & Fiedler-Kelly, 2014) Une variable aléatoire provenant d'une distribution log-normale ne peut prendre de valeurs négatives, ce qui justifie partiellement le choix de cette distribution. De plus, cette distribution est adéquate pour représenter un grand nombre de paramètres biologiques. Dans les chapitres qui suivent, on fera fréquemment appel aux variables distribuées log-normalement. Ces variables peuvent s'exprimer mathématiquement par l'expression e^X où X suit une distribution normale.(Johnson Norman, Kotz, & Balakrishnan, 1994) La Figure 1.6 illustre la fonction de densité associée, où X suit une normale centrée en 0 et d'écart-type égalant 1.

Figure 1.6 – Fonction de densité de la distribution log-normale. La distribution normale associée a une moyenne de 0 et un écart-type de 1.



1.8.3.4. Logiciels utilisés

Les simulations numériques ayant permis la production de prédictions à partir de modèles ont été effectuées via le logiciel Matlab.(The MathWorks Inc, 2014) Ce même logiciel a également permis de produire la majorité des graphiques de cette thèse. Le logiciel R de CRAN a été utilisé pour effectuer certaines analyses statistiques et produire plusieurs graphiques.(R Development Core Team, 2014) Finalement, le logiciel Mathematica a permis de vérifier les principaux résultats de calculs symboliques.(Wolfram Research Inc, 2014) Pour fin de complétude, nous pouvons également mentionner qu'il existe des logiciels permettant l'ajustement de modèles non-linéaires ou décrites par des équations différentielles avec des paramètres fixes et aléatoires (modèles mixtes, i.e. certains paramètres sont issus d'une distribution), quoique ces logiciels n'ont pas servis au cours de cette thèse. De ces logiciels, on peut noter Nonmem ou bien Monolix qui peuvent s'avérer très pratiques à cette fin.(Beal, Sheiner, & Boeckmann, 2009; Lixoft SAS, 2018)

1.9. Sujets et structure des prochains chapitres

Les trois chapitres qui suivent portent sur trois problématiques distinctes. Le chapitre 2 propose une explication à l'apparente imprévisibilité du risque de développement de résistance en fonction des traitements utilisés par les patients. Au chapitre 3, la cohérence d'une explication alternative aux phases de décroissances de la charge virale est évaluée grâce à plusieurs jeux de données empiriques. Le chapitre 4 s'intéresse au phénomène de contrôle de la virémie post-traitement, et plus particulièrement à l'interprétation d'une corrélation significative entre la virémie des patients à l'arrêt des traitements et leur capacité à maintenir cette virémie à de faibles niveaux.

Au début de chaque chapitre, un préambule décrit brièvement le contexte de l'étude, ses objectifs et l'approche adoptée. Par la suite, l'article scientifique qui a découlé de cette étude est rapporté. L'article du chapitre 2 intitulé « A Mathematical Model to Predict HIV Virological Failure and Elucidate the Role of Lymph Node Drug Penetration » a été publié dans la revue « Clinical Pharmacology and Therapeutics : Pharmacometrics and Systems Pharmacology ».(Sanche et al., 2017) L'article du chapitre 3 intitulé « Exploring an alternative explanation for the second phase of viral decay: infection of short-lived cells in a drug-limited

compartment during HAART » a été publié dans la revue « PLOS One ».(Sanche, Mesplede, Sheehan, Li, & Nekka, 2018) L'article du chapitre 4 intitulé « Chapitre 4. How is residual viremia linked to post-treatment control: analysis of treatment interruption data based on a mechanistic model » sera soumis prochainement afin d'être révisé par des pairs.

Chapitre 2. A Mathematical Model to Predict HIV Virological Failure and Elucidate the Role of Lymph Node Drug Penetration

Authors and affiliation: S Sanche^{1*}, N Sheehan^{1,2}, T Mesplède³, MA Wainberg³, J Li^{1†} and F Nekka^{1†}

¹Faculté de Pharmacie, Université de Montréal, Montréal, QC, Canada H3T 1J4, ²Chronic Viral Illness Service, McGill University Health Centre, Montréal, QC, Canada, H4A 3J1, ³McGill University AIDS Centre, Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, QC, Canada H3T 1E2. †J.L. and F.N. contributed equally as last authors.

Publié dans *CPT: pharmacometrics & systems pharmacology* 6, no. 7 (2017): 469-476

2.1. Préambule

L'étude décrite dans ce chapitre est inspirée par une apparente incohérence entre des données d'efficacité de médicaments *in vitro* et celles *in vivo*. Plus spécifiquement, certains médicaments comme l'efavirenz (EFV) promeuvent le développement de résistance lorsqu'utilisés en monothérapie. Or, lorsque les concentrations plasmatiques d'efavirenz sont transposées aux essais *in vitro*, aucune sélection de souches résistantes n'est observée. Ainsi, l'objectif principal est de formuler un modèle mécanistique qui réconcilie les données *in vitro* et *in vivo* d'efficacité des traitements et, par le fait même, d'apporter une explication physiologique au phénomène de développement de résistance *in vivo*. Le modèle est évalué via sa capacité à prédire le risque d'échec virologique de patients en fonction de l'usage de médicaments spécifiques. Une attention particulière est portée à la non-adhésion, puisque ce contexte est prompt au développement de la résistance, et à la variabilité interindividuelle, puisque les statistiques comparées ont été dérivées de groupes de patients.

2.2. Abstract and keywords

Preventing virological failure following HIV treatment remains a difficult task that is further complicated by the emergence of drug resistance. We have developed a mathematical model able to explain and predict HIV virological outcomes for various compounds and patients' drug intake patterns. Compared to current approaches, this model considers, altogether, drug penetration into lymph nodes, a refined adherence representation accounting for the propensity for long drug holidays, population pharmacokinetic and pharmacodynamic variability, drug interaction, and cross-resistance. *In silico* results are consistent with clinical observations for treatment with efavirenz, efavirenz in association with tenofovir DF and emtricitabine, or boosted darunavir. Our findings indicate that limited lymph node drug penetration can account for a large proportion of cases of virological failure and drug resistance. Since a limited amount of information is required by the model, it can be of use in the process of drug discovery and to guide clinical treatment strategies.

Keywords

mathematical model; HIV; efavirenz; tenofovir; emtricitabine; darunavir; virological failure; drug resistance; in vivo extrapolation;

2.3. Introduction

Although potent antiretroviral drugs efficaciously inhibit HIV replication, monotherapy with these drugs has been historically associated with high rates of treatment failure due to the development of drug resistance. This illustrates how the *in vitro* potency of a drug can poorly reflect its clinical efficacy. Efavirenz, a nonnucleoside reverse transcriptase inhibitor, is a typical example for which clinical outcomes differ from results obtained in tissue culture infection assays. Although viral growth is significantly inhibited *in vitro* when drug concentrations are similar to those observed in the plasma of patients, (L. Shen et al., 2008; Young et al., 1995) resistance to efavirenz is frequently observed by clinicians. Accordingly, the product

monograph advises the use of efavirenz only as part of combination therapy.(Bristol-Myers Squibb, 2015)

Unfortunately, efavirenz is not an exception, and clinical studies may be the only reliable way to provide information about the clinical efficacy of antiretroviral treatments. Such studies generally involve high costs that may discourage the initiation of clinical trials with novel treatment regimens. The resulting lack of knowledge around the efficacy of different treatments inevitably hampers clinical practice. Indeed, when facing a patient experiencing treatment failure, clinicians are limited to well-known treatment options that may be suboptimal both for the health of the patient and economically. Risk evaluation from preliminary data would help clinicians to anticipate the outcomes of various treatments, thereby guiding a better choice. In addition to clinicians, drug developers could also benefit from clinical predictions by helping them plan (or avoid) costly and lengthy clinical trials.

The potential gain resulting from the predictive ability of mathematical modeling motivated the work presented herein. Accordingly, our objective was to formulate a model that would assess the risk of virological failure and drug resistance from antiretroviral use. As the starting point, we used a mathematical model elaborated by Rosenbloom et al.(Rosenbloom et al., 2012) The model successfully shows how the relationship between drug intake and virological outcomes varies for different drugs. Unfortunately, the model falls short of explaining issues such as the lack of efficacy of efavirenz monotherapy: it foretells an absence of virological failure for patients having moderate or high drug adherence, despite clinical expectations and *in vivo* evidence.(Hofman et al., 2004) We took this discrepancy as a learning opportunity, and revisited this mathematical model.

The approach adopted herein complies with the integrative vision advocated in Quantitative Systems Pharmacology.(Craig et al., 2016; Sorger et al., 2011) In particular, the most up-to-date knowledge in immune physiology, antiretroviral pharmacology, and viral kinetics was incorporated into a model that describes the processes linking drug use to treatment effect. Since the current literature only partially informs these processes, hypotheses were used to fill the knowledge gaps that prevented the formulation of a comprehensive model. Notably, the model required a conceptualization of the *in vivo* effect of drugs. For this purpose, we hypothesized that the main cause of virological failure is insufficient drug exposure in one or

more physiological compartments hosting a large number of CD4⁺ T-cell infections, hence allowing viral replication for some of the strains. We identified lymph nodes as biological structures that are likely to exhibit both characteristics for the antiretrovirals that are considered in this study, namely, efavirenz, emtricitabine, tenofovir, and darunavir; for these drugs, concentrations in lymph node mononuclear cells were consistently lower than that of peripheral blood mononuclear cells (-66% to -99%).(Fletcher et al., 2014) The level of drug penetration in the lymph node along with results from *in vitro* experiments were used for predictions of *in vivo* drug efficacy (see Methods for details).

Once a complete model was formulated, we compared model predictions of virological failure and resistance with clinical observations. Our attention then turned towards the sensitivity of model predictions to: 1) lymph node drug penetration, 2) drug adherence behavior, and 3) the within-host viral load growth rate. The latter investigation was motivated by hypotheses of high model sensitivity to these three components.

2.4. Methods

The model inputs a specified antiretroviral treatment, potential resistance mutations, and a set of patient characteristics. These characteristics, such as drug adherence behavior, are represented by model parameters. A number of virtual patients are generated by the model, each one having specific parameter values sampled from population distributions. The model outputs, at the individual patient level, the evolution over time of the number of infected CD4⁺T cells in the lymph nodes. These cells contribute to the viral load by producing virions. It was defined that a patient experiences virological failure when this contribution is above a given threshold (the threshold value can vary). The model will be further described later in this section.

2.4.1. Antiretroviral treatment

Three daily antiretroviral treatments were considered in the modeling process: 1) monotherapy with 600 mg efavirenz, 2) combined therapy with 600 mg efavirenz, 300 mg tenofovir disoproxil fumarate, and 200 mg emtricitabine, and 3) monotherapy with darunavir (800 mg)/ritonavir (100 mg). This choice was driven by the availability of virological failure data(Antinori et al., 2015; Gordon, Gharibian, Chong, & Chun, 2015; Hofman et al., 2004) and

the required parameters for the model (lymph node drug penetration,(Fletcher et al., 2014) see below).

2.4.2. The model and parameter values

To better grasp the rationale behind the model elaboration and detail the mathematical development, see the Supplementary Material. In the following, the modeling schematic is illustrated in Figure 2.1. The model was developed for the general case of multiple patients, therapy with multiple drugs, and multiple resistant strains. However, to help understand the model structure, the simple case of monotherapy for one patient is illustrated with the assumption of only one major resistance mutation (Figure 2.1). The main model component simulates the number of infected cells in the lymph nodes of a patient. Two types of infected cells are discriminated depending on the specific strain of virus having infected them (wildtype (y_1) or resistant virus (y_2)). An initial population size for the infected cells needs to be assigned (initial viral load (t_0)). Then, for a small time step $\Delta t = t_{k+1} - t_k$, the following number of events are computed for each of the two infected cell types ($i=1, 2$): 1) the number of productively infected cells of each type that die ($N_{d,i}$, Figure 2.1a, Eq. 2.1); 2) the number of newly infected cells initiating viral production ($N_{inf,i}$, Figure 2.1a, Eq. 2.2); 3) the number of de novo mutations leading to cells infected by the resistant strain (N_{novo} , Figure 2.1a, Eq. 2.3) and 4) the number of latent cells activation ($N_{A,i}$, Figure 2.1a, Eq. 2.4). Steps 1 to 4 above are repeated until the end of the virtual trial is reached. The simulation of the number of infected cells within the host over time relies on the assignment of Poisson random variable values for each of the above-mentioned processes at each evaluated time steps. The expected values of the random variables are given by Eqs. 2.1–4, with the description of model parameters immediately following:

$$\mathbf{E}(N_{d,i}) = y_i(t)(1 - e^{\Delta t d_y}) \quad (2.1)$$

$$\mathbf{E}(N_{inf,i}) = \frac{d_y \lambda R_{0i}(t - \tau) y_i(t - \tau) \Delta t}{\lambda + d_y \sum_j R_{0j}(t - \tau) y_j(t - \tau)} \quad (2.2)$$

$$\mathbf{E}(N_{novo}) = p N_{inf,1} \quad (2.3)$$

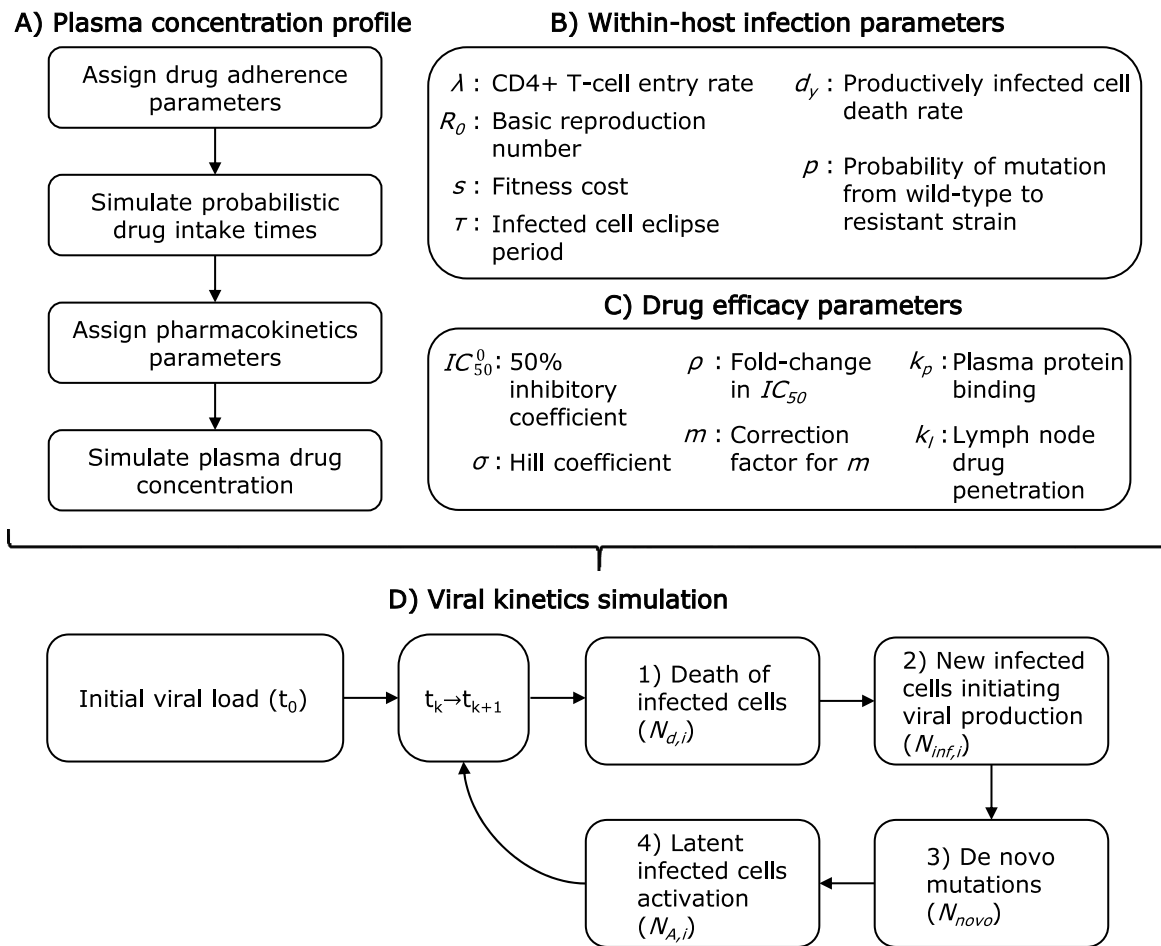
$$\mathbf{E}(N_{A,i}) = 3000 \Delta t f_i(p, s) \quad (2.4)$$

where $R_{0i}(t) = \frac{R_0(1-s_i)}{1 + \left(\frac{c_p(t)}{\rho_i IC_{50}^0 k_l k_p} \right)^{m(1+\sigma_i)}}$ and where $f_i(p, s)$ is the expected fraction of the infected cell population that would be of type i , in absence of drug.

Within-host infection parameters are required (Figure 2.1b). A viral setpoint (maximum viral load) is assigned to the patient. The setpoint value is related to the rate of entry of new CD4+ T cells (k). (Rosenbloom et al., 2012) One random value for k is generated from a distribution suggested in Rosenbloom et al. (Rosenbloom et al., 2012) A value of the basic reproduction number (R_0) is also attributed to the patient. The reproduction number characterizes the growth rate of the wildtype virus infection within the host in the absence of antiretroviral. As this can significantly vary between individuals due to immunity and quasispecies differences, the attributed R_0 is sampled from random distributions, which are reported for patients in the acute (primary infection, higher viral growth rate) and chronic (a few months after primary infection) phases. (Ioannidis et al., 2000; Ribeiro et al., 2010) For the resistant strain, growth is hampered by a nonzero fitness cost (s), which value depends on the considered mutation, and is reported for multiple strains in Sampah et al. (Sampah et al., 2011) A newly infected cell starts producing new virions after an eclipse period of 1 day (τ) and for a median period of 1 day (d_y). (Dixit, Markowitz, Ho, & Perelson, 2004; Markowitz et al., 2003) The newly produced virions can differ from the originating wildtype strains due to mutation. This occurs with probability p . (Abram, Ferris, Shao, Alvord, & Hughes, 2010; Rosenbloom et al., 2012)

Drug concentrations have an effect on the growth rate of the wildtype and resistant viruses (Figure 2.1c). Hill functions are shown to reflect the relationship between drug concentration and drug efficacy *in vitro*. (L. Shen et al., 2008) The effect of the drug is accounted in the model using the drug concentration inhibiting 50% of the infection growth *in vitro* in a medium containing no plasma protein content (IC_{50}^0) and using the Hill curve coefficient (m). The former is taken from a random distribution since significant quasispecies differences in IC_{50} are commonly reported. (Parkin et al., 2004) For the treatments investigated herein, these

Figure 2.1– Simplified model algorithm and parameters: the special case of monotherapy for one patient with the assumption of only one major resistance mutation. Index k goes from 0 to the number of evaluated time steps. Index i takes a value of 1 for the wildtype strain and 2 for the resistant strain.



parameters were obtained from the product monographs (Bristol-Myers Squibb, 2015; Gilead Sciences Inc, 2008, 2013; Janssen Inc, 2014) and from results of *in vitro* experimentations. (Sampah et al., 2011; L. Shen et al., 2008) For the resistant strain, the model uses the fold-change in IC_{50} (ρ) and the correction factor for the Hill coefficient (σ). (Rosenbloom et al., 2012; Sampah et al., 2011) If multiple drug treatment are considered, an algorithm for the combined effect is used. (Jilek et al., 2012) The *in vitro* to *in vivo*

extrapolation of drug effect needed to account for the plasma protein binding effect on the IC_{50} (k_p), and for lymph node drug penetration (k_l). (Acosta et al., 2012; Fletcher et al., 2014)

The simulation of drug concentrations is based on patient specific characteristics of adherence and drug disposition. The patient is assigned parameters representing the adherence behavior (Figure 2.1d, first two boxes). The first parameter is the general adherence level, defined as the fraction of the prescribed pills that are expected to be taken over the investigated period. The second parameter is the probability of dose intake following a missed dose (PIM), related to the tendency of the patient for long periods of nonadherence (drug holidays). For an outpatient taking medication for chronic conditions, it was shown that these two values are required to reproduce the recorded distribution of the length of time separating consecutive dose intakes. (Girard, Blaschke, Kastrissios, & Sheiner, 1998; Wong, Modi, & Ramanathan, 2003) Using the two parameters mentioned above (general adherence, PIM), we simulate a sequence of zeros (missed dose) and ones (correct dose intake) which are then transformed in time sequences according to the treatment regimen (e.g., once daily).

We then use population pharmacokinetic models to compute the expected drug concentrations in the plasma of the patient ($C_p(t)$) from drug intake (Figure 2.1d, last two boxes). This requires the random assignment of pharmacokinetic parameters whose distributions are obtained from the reported pharmacokinetic models (see Supplementary Material for a brief description of these models). For the considered treatments, pharmacokinetic drug parameters were taken from reported studies. (Baheti, Kiser, Havens, & Fletcher, 2011; Moltó et al., 2013; Sanchez et al., 2011; Valade et al., 2014) For therapy combining efavirenz, emtricitabine, and tenofovir DF, we did not consider medication interactions in absorption or elimination; efavirenz is mostly eliminated by the liver, while the other two medications are eliminated by the kidneys without noticeable interaction. (Blum, Chittick, Begley, & Zong, 2007)

2.4.3. Simulations

Using this model, three reported clinical studies were reproduced with the associated virological failure and resistance data. (Antinori et al., 2015; Gordon et al., 2015; Hofman et al., 2004) For the study of efavirenz monotherapy, a viral load of 50 copies/mL or more at the end of the trial was considered a virological failure. The mutations that were considered were SNPs

K103N, Y181C, G190S, M184V, and K65R. Virological failure could occur without resistance mutations or with resistance-associated mutations when at least 20% of the viral load has major resistance mutations. For the combination therapy, HIV patients were followed for a period of 12–18 months, with blood samples taken approximately every 3 months.(Gordon et al., 2015) Patients differed in their level of drug adherence, measured using pharmacy refill records. Virological failure was defined as two consecutive viral loads of 200 RNA copies/mL or above throughout the course of the trial. We replicated this experiment in silico. For this simulation, the same mutations (K103N, Y181C, G190S, M184V, and K65R) were considered. Finally, we used our model to predict virological failure with darunavir/ritonavir therapy for 48 weeks. No mutation was considered, since darunavir was shown to possess a high genetic barrier against resistance and because we considered a short trial in virologically suppressed patients.(Dierynck et al., 2007)

We used a stratified random sample to investigate the sensitivity of model predictions to the components mentioned above: drug adherence, and the within-host viral load growth rate. Simulations were run by groups of 300 individuals characterized by the same drug intake behavior, represented by the general level of adherence and the PIM (probability of drug intake following a missed dose), and being in the same infection phase (acute or chronic). The general level of adherence could take values from 10–100% by 10% increments. The PIM could take values from 5% to a maximum mathematically allowable value determined by the general level of adherence, by 5% increments. For each drug or drug combination, we simulated the virological outcomes of 1,200–12,000 virtual individuals per general level of adherence. The risk of virological outcome for each adherence group was computed using a weighted average. The corresponding weights were chosen based on the assumption of a stratified random sample from a patient population (see Supplementary Material for the specific values). In order to evaluate the extent of the variation in the percentage of virtual patients experiencing virological failure between trials, the minimum and maximum values were obtained for each adherence groups and are displayed in Figure 2.6 and Figure 2.7.

An alternative model was used for the impact of lymph node drug penetration. The model was identical to the model described above, with the exception that the intracellular drug concentration of CD4+T cells located in the lymph node was set equal to that of peripheral blood mononuclear cells ($k_l = 1$).

All models were numerically implemented and simulations were carried using MatLab 2014a.(The MathWorks Inc, 2014)

2.5. Results

Using the model that accounts for incomplete lymph node penetration (the main model, not the alternative model), we obtained the virological outcomes for efavirenz monotherapy and efavirenz, emtricitabine, and tenofovir combination therapy (Figure 2.2, Figure 2.3). For efavirenz monotherapy, a majority of virtual patients having a general level of adherence of 20% or more had virological failure with resistance at the end of 48 weeks. In comparison, two out of three macaques having plasma drug concentrations similar to those of simulated patients had high viral loads and significant resistance within the first 42 weeks of therapy.(Hofman et al., 2004) For the treatment combination, predictions of virological failure (green bars) are displayed along a regression curve that is based on clinical data.(Gordon et al., 2015) Compared to the regression curve, more virtual patients having an adherence level between 20% and 60% experienced virological success. The opposite is true for patients with a level of drug adherence of 10% or 70% and above. For treatment with darunavir/ritonavir, simulation results are only reported for those with full adherence (100%), as no comparative data were found for the case of low adherence. The results were sensitive to small changes in the level of lymph node drug penetration k_l , as confirmed by a sensitivity analysis with k_l varying from 0.5–2.0% (the reported value was 1%).(Fletcher et al., 2014) Virological failure was obtained for 13–48% of individuals, depending on the value of k_l that was used, compared to a reported value of 14%.(Antinori et al., 2015)

Results from the alternative model (complete lymph node drug penetration from the plasma) for efavirenz monotherapy and the combination therapy are displayed in Figure 2.4 and Figure 2.5, respectively. No virological failure was predicted by this model for darunavir/ritonavir therapy. In comparison with the model accounting for incomplete lymph node drug penetration, a smaller number of cases of virological failure was obtained using the alternative model.

Figure 2.2 – Probability of virological outcomes as a function of adherence under efavirenz monotherapy. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented.

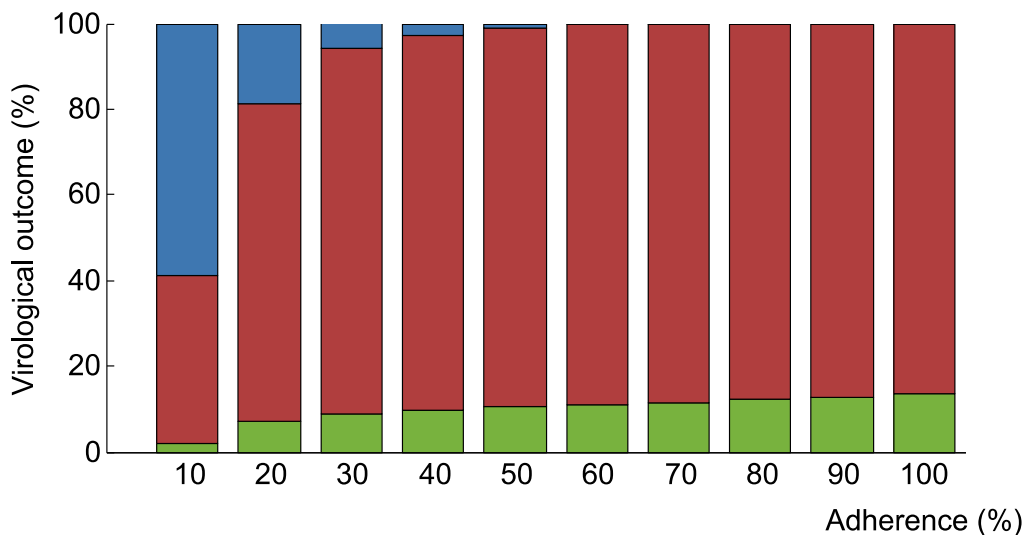


Figure 2.3 – Probability of virological success as a function of adherence under efavirenz, tenofovir disoproxil fumarate, and emtricitabine combination therapy. The model prediction for virological success is illustrated by the green bars and compared to a regression curve fitting clinical data (green line). (Dierynck et al., 2007)

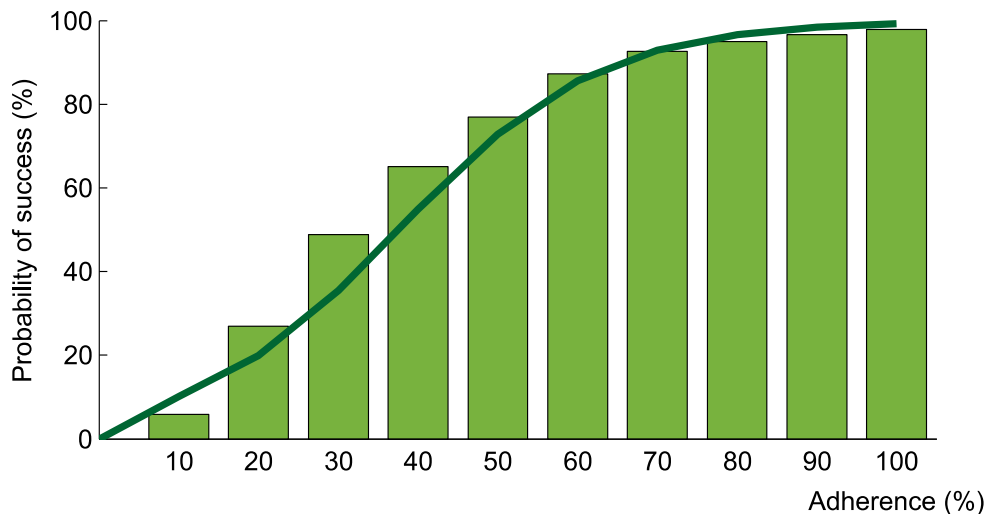


Figure 2.4 – Probability of virological outcomes as a function of adherence under efavirenz monotherapy with complete lymph node penetration. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented.

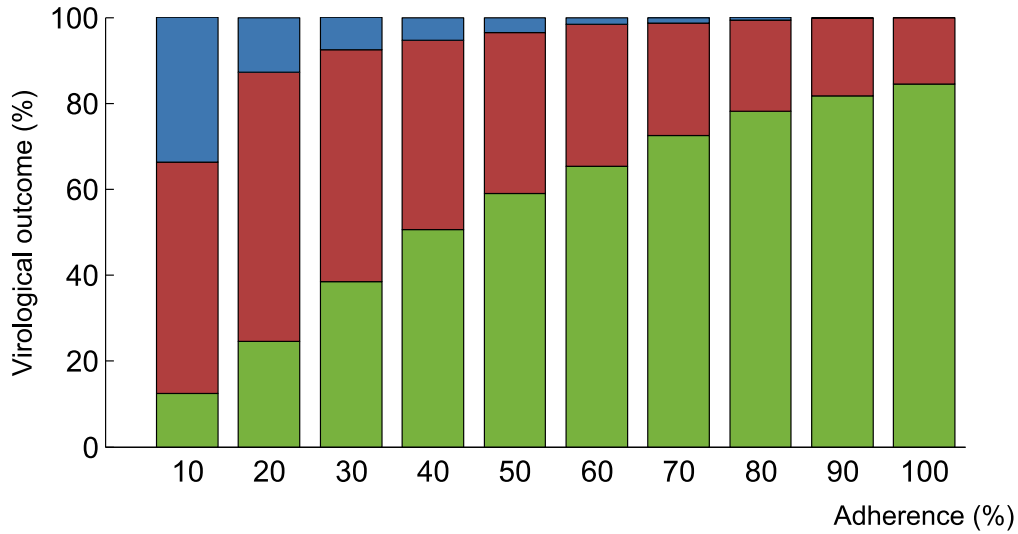
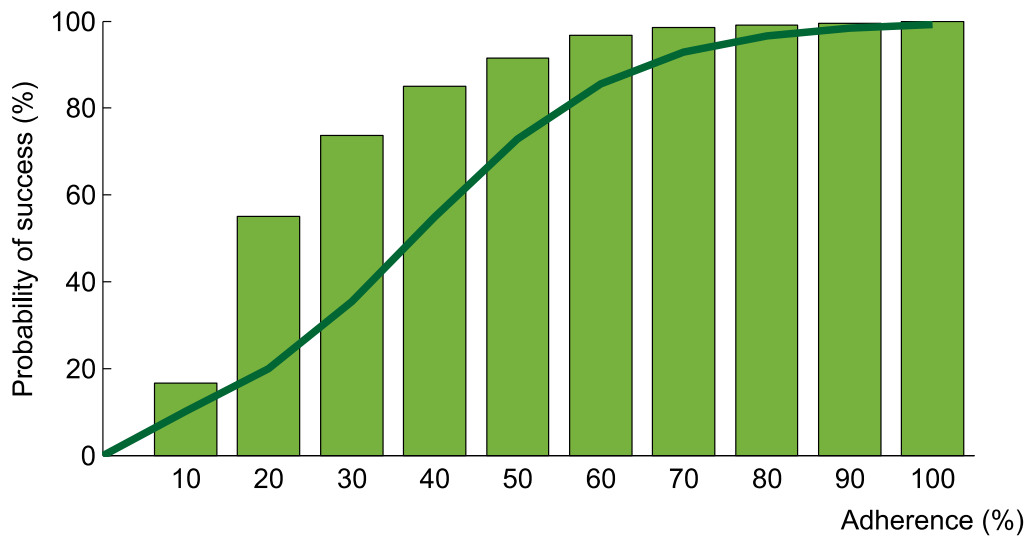


Figure 2.5 – Probability of virological success as a function of adherence under efavirenz, tenofovir disoproxil fumarate, and emtricitabine combination therapy: results from the model assuming complete lymph node penetration for all drugs. The model prediction for virological success is illustrated by the green bars and compared to a regression curve fitting clinical data (green line).(Gordon et al., 2015)



Within each adherence group, model predictions significantly differed in terms of PIM and infection phase (see Figure 2.6, Figure 2.7). The lower and higher limits for the green bar represent patients in the acute infection phase who have the lowest PIM and in the chronic infection phase who have the highest PIM, respectively.

Figure 2.6 – Interindividual variability in the probability of virological outcome under efavirenz monotherapy as a function of adherence. The impact of model variability in adherence patterns and infection phase is displayed using whiskers on the graph. The proportion of individuals without virological failure (green), with virological failure and resistance mutations (red), and with virological failure and no resistance mutations (blue) are represented.

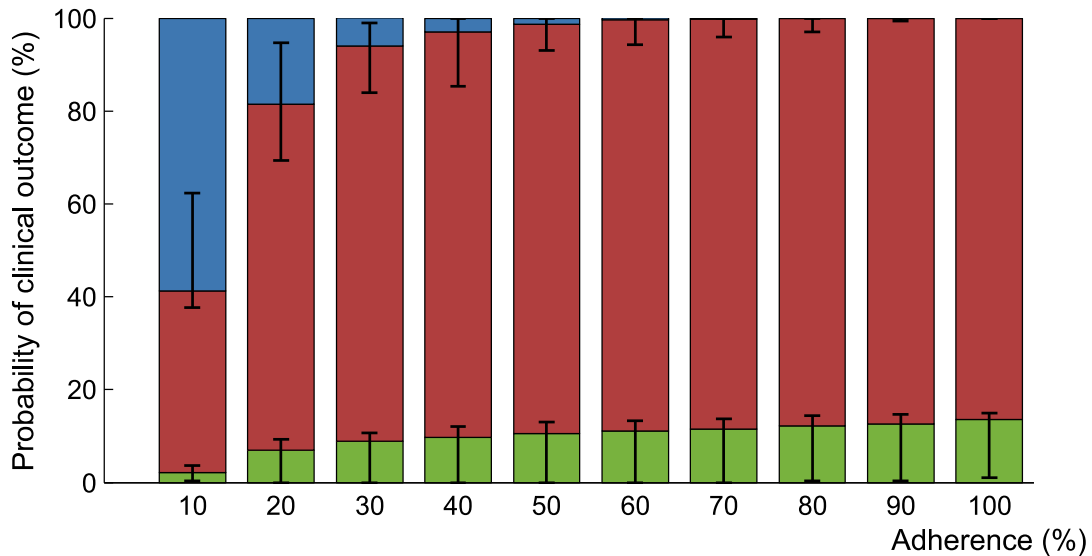
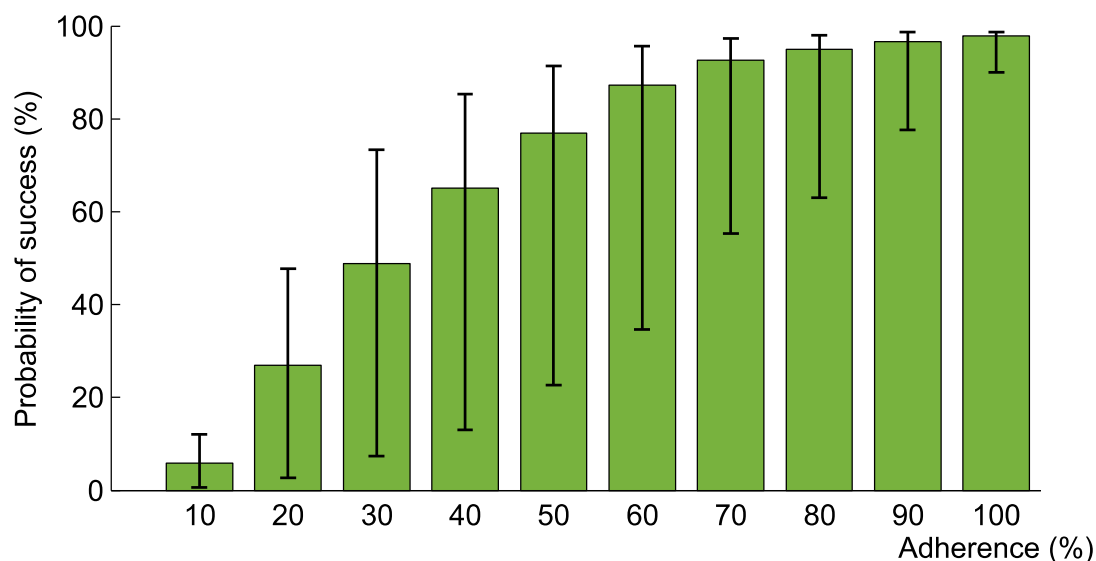


Figure 2.7 – Interindividual variability in the probability of virological success under efavirenz, tenofovir disoproxil fumarate, emtricitabine combination therapy as a function of adherence. The impact of model variability in adherence patterns and infection phase is displayed using whiskers on the graph.



2.6. Discussion

The main objective of this study was to develop a model able to reproduce the expected or observed relationships between patients’ drug use and clinical outcome for several treatments: efavirenz monotherapy, efavirenz/tenofovir disoproxil fumarate/emtricitabine combination therapy, and boosted darunavir (ritonavir) monotherapy. The effect of three model components on virological failure predictions was also investigated: lymph node drug penetration, the PIM, and the viral load growth rate through a distinction of the acute and chronic infection phases.

The approach used in this study enabled the evaluation of hypotheses regarding the pathogenesis of virological failure. Indeed, since the parameter values used in the current model are fully supported by independent studies, an observed discrepancy between predictions and observations puts into question the veracity of at least one of the model assumptions. However, a satisfactory clinical prediction under various therapeutic contexts using a specific set of hypotheses suggests both the plausibility of model assumptions and the potential utility of the

model in evaluating the risk of virological failure and resistance in patients. The model we developed has significantly increased the accuracy of predictions for efavirenz monotherapy. Indeed, a previous model predicted virological success for all adherent patients under the regimen, a result that is inconsistent with a reported animal study.(Hofman et al., 2004; Rosenbloom et al., 2012) For the combination treatment, there was a high level of agreement between the regression fit of the observed clinical data and model predictions. However, for those who were less adherent, the model underestimated the number of cases of virological failure. A possible explanation for the discrepancy could originate from the lack of precision in the obtained PIM distribution due to a smaller sample size for those with very low drug adherence (see Supplementary Material). Those with a lower PIM tended to have less virological success (Figure 2.6, Figure 2.7), hence it is possible that those with a low level of adherence are more likely to take long drug holidays than what was estimated. Furthermore, it was not possible to assess the proportion of patients from the reported trial who were in the acute infection phase. We assumed that the great majority (90%) were in the chronic infection phase, but predictions may have been impacted by deviations from this assumption. For treatment with darunavir/ritonavir, the intracellular concentration ratio between mononuclear cells from lymph nodes and peripheral blood was very small for this drug and estimated to be 1%.(Fletcher et al., 2014) When k_l is small, its variation can have a significant impact on the predicted lymph node drug concentration (compared to the reported 1% drug penetration for darunavir, a value of 0.5% leads to a decrease of 50% in drug concentration, while a value of 2% doubles the drug concentration). Model predictions coincided with the clinically observed proportion of virological failure when this ratio was assumed to be 1.9%.

The results suggest that low drug penetration in lymph nodes can account for a large proportion of cases of virological failure. It can also account for a significant number of cases that could not be solely explained by plasma drug exposure. Whether or not this means that lymph node drug exposure is responsible for these cases is debatable. However, we argue that higher drug concentrations in these biological structures are likely to have a significant impact on viral loads. Since the action of most antiretrovirals is centered around preventing the infection of healthy T cells, an underlying condition for this hypothesis would be that a large number of cell infection events occur there, as suggested by the following. Indeed, there is evidence that virus producing T cells mainly infect cells located in their vicinity,(A. T. Haase, 1999) therefore

a sufficient number of virus producing and healthy cells are required in lymph nodes in order for infection events to occur locally. Lymphatic tissues, including lymph nodes, are major sites where the viruses are produced.(A. T. Haase, 1999; Ashley T Haase et al., 1996) Whether the infection occurred in the same biological structure or in different tissues is unknown (lymphocytes exhibit complex homing and recirculation dynamics).(Wiedle, Dunon, & Imhof, 2001) Unfortunately, little information is available regarding the dynamics of healthy CD4+T cells within human organs. T cells migrate from the blood to many organs. They spend a relatively short amount of time circulating in the blood of healthy subjects (average of 2.3 h).(Savkovic et al., 2012) Furthermore, it was estimated that, at any given moment, only 5.5% of CD4+T cells capable of trafficking to peripheral blood are located therein.(Savkovic et al., 2012) In healthy mice, CD4+T cells passage in the lungs and liver has been estimated to be quite rapid (less than a minute), compared to their passage in the spleen (2 h), or in lymph nodes or gut (10 hours).(Ganusov & Auerbach, 2014) It is also known that CD4+T cells are depleted from the gut of patients upon infection.(Brenchley, 2004) In SIV-infected macaques, this phenomenon is observed at all stages of infection.(Veazey & Lackner, 2004) These arguments suggest that healthy CD4+T cells likely reside in lymph nodes for a long period compared to other secondary lymphoid organs, increasing the likelihood of infection events occurring in these structures.

From these considerations, a high drug concentration inside lymph node T cells is desirable, but not observed for many drugs.(Fletcher et al., 2014) However, we argue against changes in drug formulation or therapeutic administration that would solely favor lymph node penetration without providing therapeutic plasma concentrations. Indeed, other physiological compartments could host a large number of cell infection events (e.g., spleen). We assumed that, in such compartments, cell infection activity is either suppressed by the drugs during treatment, occurs concomitantly with lymph node activity, or occurs at a small scale without altering the predicted outcome. This consequently allowed model predictions to be solely based on the evolution of the viral population in lymph nodes. If, however, therapeutic drug concentrations for all viral strains are observed in the lymph nodes but not in a different compartment, then new cell infections in the latter may become responsible for rebounds and the selection of resistance mutations. We believe model predictions of virological failure could still be

performed in such cases by replacing the parameter value associated with lymph node drug penetration with its respective value in the new drug limited compartment.

Additional HIV compounds and combinations could be investigated using the developed model. Trial predictions can be performed by modifying model parameters according to drug pharmacokinetics, *in vitro* efficacy curves, and the intracellular concentration in peripheral blood mononuclear cells and mononuclear cells located in various secondary lymphoid tissues. Even though the model focuses on predicting the outcome of groups of patients, individualized risk evaluation can be performed. The procedure involves simulating virological outcomes for a group of virtual patients having common characteristics, such as adherence behavior, drug disposition, and the rate of viral load growth during rebounds. The utility of individualized predictions is highlighted by the impact of interindividual variability in PIM and viral load growth rate (Figure 2.6 and Figure 2.7). However, a full individualization still depends on our capacity to gather patient-specific data.

The scope of the model can be extended to predict other virological outcomes and tackle more complex therapeutic cases. Modifications to the structural model, however, could be required. Indeed, the model predicts virological failure and resistance both based on threshold values applied to a viral load. Accurate viral load predictions would likely require simulating the dynamics of viral populations and drug effect in all of the important compartments concomitantly. Furthermore, many components of the immune system (CD8+, follicular dendritic cells)(De Boer et al., 2010) as well as the processes governing latent cells activation were not explicitly modeled; this may impact the ability of the model to predict virological failure over more extended periods or in the case of patients suffering from AIDS. The model would also benefit from a better understanding of the pharmacokinetics of drugs in specific body compartments: so far, it was assumed that there is rapid equilibrium in drug concentrations between all relevant compartments, which may not be a good approximation for all drugs. Overall, the current model can be used as a backbone for the investigation of diverse situations.

In summary, the translational approach adopted here is in line with the increasingly popular Quantitative Systems Pharmacology. It suggests that antiretroviral drug outcomes, such as virological failure and the development of resistance, can be demystified by integrating current scientific knowledge. We consider that the model developed herein is a promising step

forward to clinical prediction and to the understanding of the pathogenesis of virological failure in vivo.

2.7. Supplementary Material

One specific model of HIV dynamics and antiretroviral response, hereafter referred to as the initial model, was revisited.(Rosenbloom et al., 2012) In the following, an overview of the initial model and its assumptions is presented. The section is followed by a description of the discrepancies between this model and the model we developed. This approach was chosen as it proved more natural to describe our rationale and to ensure that proper credit is attributed to previous efforts. A summary of the model we developed is later provided in the form of an algorithm.

2.7.1. The initial model

A picture of the initial model is given here. Further details on this model are provided in Rosenbloom *et al.*(Rosenbloom et al., 2012) Similarly to the model used in this study, the initial model has four components representing processes linking antiretroviral drug use to virological outcomes: patients' drug intake patterns, pharmacokinetics, the inhibition of viral replication by the drug, and within-host viral kinetics. The initial model focused on monotherapy with various drugs.

For the component mirroring patients' drug intake patterns, the moments of dose intake are fixed to those determined by the prescribed regimen (nominal times). One parameter, the general level of adherence, is used to represent the drug adherence behavior of patients. Time sequences of dose intakes are then obtained for each virtual patient from a stochastic simulation: a dichotomous random variable determines whether the dose was taken or not at each nominal time with intake probability equal to the adherence level of the patient.

From the time sequences of drug intake, the second component simulates how plasma drug concentrations vary over time. The procedure is based on three drug-specific parameters, namely the expected maximum drug concentration (C_{max}), trough concentration (C_{min}) and half-life. Concentrations are then obtained using the following algorithm. The first dose taken by a patient immediately increases plasma concentrations to $C_{max} - C_{min}$. From that point,

concentrations decay exponentially according to the half-life of the drug. Any additional dose increases the predicted concentration at the time of dose intake by $C_{max} - C_{min}$.

The third component does not directly link to the second component. Although it is used later on to link plasma drug concentrations to the inhibition of viral replication *in vivo*, we first describe it in the *in vitro* context in which it was first established. The output of the third component is the proportion of single-cycle infection events that are unaffected by the drug, as observed through *in vitro* assays ($U_i(D)$). (Sampah et al., 2011; L. Shen et al., 2008) It is a function of the antiretroviral, fixed drug concentration D in the assay and viral strain i . Hill equations, thought to fit experimental curves adequately, have been used to represent the relationship between U_i and D :

$$U_i(D) = \frac{1}{1 + \left(\frac{D}{\rho_i IC_{50}}\right)^{m(1+\sigma_i)}}, \quad (2.5)$$

where the additional parameters IC_{50} is the drug concentration inhibiting 50% of wild-type virus replication events for the drug, m is linked to the slope of the logarithm of $U_i(D)$, and parameters ρ_i and σ_i represent the mutations impact on the IC_{50} and on m , respectively. In particular, ρ_i is the IC_{50} fold-change due to resistance mutations. The index value $i = 1$ is used to represent the wild-type strain. In turn, this implies that $\rho_1 = 1$ and $\sigma_1 = 0$ for all drugs.

The relationship defined by Equation 2.5 is used in the last component of the initial model. This component represents the kinetics of the viral populations within the host. The modeled processes are identical to those described in the Methods section with one exception: in the initial model, cells being infected start producing virions immediately (no eclipse period).

The system was translated into mathematical equations. Viruses are classified in n groups. Each group represents one considered viral strain and is identified by an index taking values between 1 and n . The following quantities are allowed to vary over time: 1) the total number of virions of strain i in the body (v_i); 2) the number of CD4+ T-cells that can be infected x ; and 3) the number of CD4+ T-cells that are infected by strain i (y_i). The mathematical representation of the system takes the following form:

$$\begin{aligned}
\frac{dx}{dt} &= \lambda - \sum_{i=1}^n \beta_i x v_i - d_x x \\
\frac{dy_i}{dt} &= \beta_i x v_i + A_i - d_y y_i \\
\frac{dv_i}{dt} &= k_i y_i - d_v v_i
\end{aligned} \tag{2.6}$$

where λ is the production rate of susceptible CD4+ T-cells, d_x and d_y are the death rate of susceptible and infected T-cells respectively, A_i represents the activation rate of infected T-cells having entered a latent state after infection (the reservoir of T-cells in viral latency), k_i and d_v are the virion production rate per infected cell and virion clearance from the body, respectively, and β_i encompasses multiple concepts linked with infection. In fact, β_i takes into account the production of non-infectious virions, the probability of contact between virions and susceptible T-cells as well as the probability of infection after contact.

The model encounters scaling issues when implemented numerically since the modeled quantities vary at very different rates. One assumption enables to reduce the number of dimensions in the model and to avoid this issue. When assuming that the number of susceptible CD4+ T-cells and the number of virions are at equilibrium relative to the number of infected T-cells, the model can be described by the following set of equations:

$$\frac{dy_i}{dt} = A_i + d_y y_i \left[\frac{\lambda R_{0i}}{\lambda + \sum_{j=1}^n R_{0j} d_y y_j} - 1 \right] \tag{2.7}$$

which has the additional advantage of reducing the number of parameters and involving reproduction numbers R_{0i} , defined as the number of newly infected cells generated by one infected cell when the number of cells that can be infected is large. There is one reproduction number per viral strain i , and the quantities are allowed to vary over time.

The value taken by R_{0i} at a specific time t is based on: 1) the basic reproduction number R_0 (reproduction number associated to the wild-type virus in absence of drug); 2) the fitness cost of strain i (s_i), and 3) the plasma drug concentration at time t . The reproduction number for strain i at time t is obtained by multiplying $R_0(1 - s_i)$ by the proportion of single-cycle

replication events that are unaffected by the drug *in vivo*. This required an *in vitro* to *in vivo* extrapolation of Equation 2.5. The approach consisted in assuming that the *in vivo* proportion is equal to $U_i(C_p(t))$ where $C_p(t)$ is the plasma concentration of the drug at time t . In other words, the model assumes that the effect of the drug as observed *in vitro* will be the same as the effect of the drug *in vivo* if the tested concentration D is equal to plasma drug concentration $C_p(t)$. Because plasma drug concentration values change over time, so does the reproduction number. In absence of hysteresis, i.e. when one concentration value can be associated to only one effect without regard to past drug concentrations, this can be summarized by the following which completes the deterministic version of the model:

$$R_{0i}(t) = \frac{R_0(1 - s_i)}{1 + \left(\frac{C_p(t)}{\rho_i I C_{50}}\right)^{m(1+\sigma_i)}}. \quad (2.8)$$

The deterministic nature of the model is adequate when the number of events occurring within a short period is sufficiently high to allow applying the law of large numbers. However, the authors wished to consider rarer events: the *de novo* generation of mutant strains. A stochastic model based on the deterministic formulation and further accounting for the occurrence of new mutations was obtained. In this latter model, the number of events occurring during a short period, for each processes mentioned above (e.g. infection, cell death, T-cell activation), was modeled using random variables. The random variables were set to follow Poisson distributions having expected values equal to the deterministic prediction of the number of events of each type.

The structure of the initial model having been defined, we hereby describe how parameter values were obtained for this model. The involved parameters are those of Equations 2.7 and 2.8 addition to the probabilities of *de novo* mutations. Parameter values were obtained from quantitative knowledge produced through studies reported elsewhere. The estimates of the basic reproduction number R_0 that were reported in the literature were originally obtained by fitting viral load data collected during viral rebounds in HIV-positive patients. (Ioannidis et al., 2000; Little, McLean, Spina, Richman, & Havlir, 1999; Ribeiro et al., 2010) Accordingly, for the initial model, R_0 values of 5 and 10 were chosen to represent patients in the chronic (a few months after primary infection) and acute (primary infection, higher viral growth rate) phase of

infection, respectively. Other studies reported estimates for the other parameters of Equation 2.8. These studies investigated *in vitro* viral infection curves under conditions that varied for the antiretroviral, viral strain and drug concentrations. (Sampah et al., 2011; L. Shen et al., 2008) In these studies, the medium used for experimentation was supplemented with 50% human serum, 10% fetal bovine serum, interleukin-2 and cytokine-rich supernatant, as an attempt to reproduce plasma content.

Other assigned parameter values were: 1) the probability of mutant strain generation for each strain, which is of the order of 10^{-6} to 10^{-5} , (Abram et al., 2010) 2) the total rate of infected T-cell activation $\sum_{i=1}^n A_i$, approximately equal to 3,000 cells per day as deduced from the average residual viral load of HIV-infected patients, (Dinosa et al., 2009) 3) the specific activation rates A_i , for cells infected by strain i , corresponding to a fraction of the above, with proportions deduced from population dynamics equations under conditions of equilibrium between viral populations in absence of selection pressure (the equations are functions of the probabilities of mutant generation and fitness costs, parameters mentioned above), (Rosenbloom et al., 2012) 4) the production rate of CD4+ T-cells, derived from the viral setpoint of HIV-infected patients and allowed to vary between individuals (taking values between 5×10^6 and 1.7×10^9 cells per day), (Rosenbloom et al., 2012) and 5) the death rate of virion producing cells d_y , reported as being 1 per day (expected lifetime of 1 day). (Markowitz et al., 2003)

The formulation of the initial model involved assumptions and approximations of reality. The following lists the most important ones.

1. The host can be represented by a volume containing a well-mixed homogeneous content.
2. The fraction of reproduction cycles impacted by the drug *in vivo* on CD4+ T-cells when plasma drug concentration is C_p is similar to the equivalent *in vitro* in a medium having drug concentration $D = C_p$ and containing 50% human serum content.
3. The reservoir of latent and infected T-cells is maintained over time for relatively short periods (approximately a year).

4. There are no back-mutations: a strain having acquired a selected mutation *de novo* will produce virions with this mutation or with added mutations, without reverting back to the wild-type strain.
5. The number of new CD4+ cells entering circulation is approximately constant over time, albeit different for each individual.
6. The number of infected and virus-producing CD4+ T-cells is approximately equal to the total number of plasma virions.(Hockett et al., 1999) This important approximation was used to allow parameter estimation and for interpreting the results in terms of viral loads.
7. Relative to that of CD4+ T-cells, the impact of other virus-producing cells on the number of newly-infected cells is negligible.
8. Mutations not considered in the simulations have a small impact on the number of individuals experiencing virological failure, as compared to the impact of the considered strains.

2.7.2. Final model

In this section, we describe the modifications that were brought to the initial model in order to obtain the model used in this study, hereafter referred to as the final model.

2.7.2.1. Revisiting the drug adherence model

In the final model, we refined the drug adherence model. We thought it was important to model this phenomenon accurately since virological outcomes are very sensitive to patient adherence.(Bangsberg, Kroetz, & Deeks, 2007) The model of adherence that was used initially is based on one parameter, namely the probability of pill intake at any nominal time point. This model cannot replicate the observed distribution of the time lapse between two pill intakes, as measured from medication event monitoring systems (MEMS).(Girard et al., 1998; Wong et al., 2003) The algorithm we used differentiates two individuals with similar adherence level by their propensity for long drug holidays, i.e. extended periods of non-adherence (a lower probability of drug intake following a missed dose represents a higher propensity for drug holidays). The

inter-individual variability in adherence behavior was simulated using the algorithm described in section 2.7.4.

2.7.2.2. A model for the pharmacokinetics of drugs

We believed the effect of drugs could be very sensitive to drug concentrations, hence the need for a more precise pharmacokinetic model. Therefore, we linked drug use to plasma concentrations using population pharmacokinetics models obtained from the literature. (Baheti et al., 2011; Blum et al., 2007; Moltó et al., 2013; Sanchez et al., 2011) For the specific pharmacokinetic models that were used, see Summary of Parameter Values, below.

2.7.2.3. Adaptation for lymph node penetration

Equation 2.5 describes the *in vivo* effect of a drug on a viral strain i when plasma concentration is $C_p = D$. In the initial model, lymph node drug penetration was not considered. Mathematical expressions were thus required to represent drug efficacy in lymph nodes. Through mathematical arguments, we will show that we can use *in vitro* observations to obtain an expression of *in vivo* drug efficacy in this context. We assumed that intracellular drug concentrations completely dictate the proportion of single-cycle infection events that are affected by the drug. In other words, drug efficacy *in vivo* will be equivalent to drug efficacy *in vitro* if all considered cells have the same intracellular concentration. Since Equation 2.5 was obtained experimentally and inputs assay drug concentrations, we need to find the assay concentrations that lead to concentrations observed in lymph node T-cells of patients. We obtained the following equation, which will be explained below, describing the proportion of infection events unaffected by the drug in the lymph node, as a function of plasma drug concentrations (C_p):

$$U_i(C_p) = \frac{1}{1 + \left(\frac{C_p}{\rho_i I C_{50}^0 k_p k_l}\right)^{m(1+\sigma_i)}}. \quad (2.9)$$

The meaning of equation parameters is as described above, with the exception of: 1) IC_{50}^0 , which is the 50% inhibitory concentration for the wild-type virus obtained from an experiment performed *in vitro* using a medium containing no human plasma serum; 2) k_p , which

is a drug-dependent correction factor applied to the IC_{50}^0 in order to account for 100% plasma serum content (estimated in Acosta *et al.*, Table 2),(Acosta et al., 2012) and 3) k_l , is also a drug dependent parameter which is the ratio of drug concentrations inside a peripheral blood mononuclear cell (C_p^{in}) to that of a lymph node mononuclear cell (C_l^{in}) (as measured in Fletcher *et al.* for various drugs).(Fletcher et al., 2014) Deriving the preceding equation required the following additional assumptions:

1. the exchange of drug quantities between the different body compartments and between the intracellular and extracellular milieus are linear processes,
2. drug quantities quickly reach an equilibrium between all considered body compartments and between milieus.

These assumptions lead to a lemma that is used extensively in this section.

Lemma: Under the assumption of non-saturable linear exchange rates of the drug between two considered environments, the ratio of drug concentrations is a unique constant at equilibrium independent of total drug quantity.

Proof: Suppose Q_1 and Q_2 are the drug quantities in the two milieus of respective volumes V_1 and V_2 . We can represent the exchange of drug between the two milieus by:

$$\begin{aligned}\frac{dQ_1}{dt} &= k_{21}Q_2 - k_{12}Q_1 \\ \frac{dQ_2}{dt} &= k_{12}Q_1 - k_{21}Q_2\end{aligned}\tag{2.10}$$

where k_{12} and k_{21} represent the drug exchange rates. At equilibrium:

$$0 = k_{21}Q_2 - k_{12}Q_1.\tag{2.11}$$

Solving for $\frac{Q_1V_2}{Q_2V_1} = \frac{C_1}{C_2}$, we obtain

$$\frac{C_1}{C_2} = \frac{k_{21}V_2}{k_{12}V_1}\tag{2.12}$$

which is a constant independent of total drug quantities. \square

The lemma above is used to derive an expression of drug efficacy for a CD4+ T-cell located in a lymph node based on *in vitro* observations. We will first show that we can estimate intracellular concentrations *in vivo* from plasma concentrations. We will then use this result to obtain an expression of drug efficacy for CD4+ T-cells located in the lymph nodes based on plasma concentrations and infectivity assay results. For readability purposes, we will only consider the effect of the drug on the wild-type strain ($i = 1$). This allows to further simplify the mathematical expressions since $\rho_1 = 1$ and $\sigma_1 = 0$ in that case.

Let C_l^{in} be the concentration inside a CD4+ T-cell of a lymph node in a patient. Let C_p^{in} be the concentration inside a CD+ T-cell located in peripheral blood at the same instant. Finally, let C_p be the plasma drug concentration. Under assumptions of linear exchange of drug between environments, the lemma above states that there exists unique constants k_l and K for which $k_l = \frac{C_p^{in}}{C_l^{in}}$ and $K = \frac{C_p^{in}}{C_p}$ is true when concentrations are at equilibrium. If equilibrium is reached quickly, we can conclude that these expressions lead to good approximations of drug concentrations at any time point. We can conclude from this that for any value of plasma concentration C_p , the concentration inside CD4+ T-cells of the lymph node is equal to $\frac{KC_p}{k_l}$.

This first result will be used for the *in vivo* extrapolation of infectivity assay results. For each drug and as mentioned above, the following mathematical expressions were found to correctly predict the proportion of single-cycle infection events that are unaffected by the drug *in vitro* ($U(D)$) in a medium containing 50% human plasma serum for the wild-type strain:(L. Shen et al., 2008)

$$U(D) = \frac{1}{1 + \left(\frac{D}{IC_{50}^L}\right)^m} \quad (2.13)$$

where D is the extracellular drug concentration in this medium, IC_{50}^L is the concentration required in this medium (50% of human serum content) to inhibit 50% of infection events, and m is the Hill parameter as described above (we added the identifier L for the roman numeral 50, for 50% human serum content). This formula can also be expressed as a function of the

concentration inside the cells, D^L , by acknowledging that there exists a constant k^L for which $k^L = \frac{D^L}{D}$ is always true, based on the above assumptions. Therefore we get:

$$U(D^L) = \frac{1}{1 + \left(\frac{D^L}{k^L IC_{50}^L}\right)^m}. \quad (2.14)$$

Because of our first assumption, cells having the same intracellular concentrations, whether they are located *in vitro* or *in vivo*, will experience the same drug effect. Therefore, this expression can also be used to describe drug efficacy for T-cells of the lymph node if $C_l^{in} = D^L$. Therefore :

$$U(C_l^{in}) = \frac{1}{1 + \left(\frac{C_l^{in}}{k^L IC_{50}^L}\right)^m}, \quad (2.15)$$

where C_l^{in} is as defined above. We showed that C_l^{in} can be obtained from plasma concentration. Using this result we obtain:

$$U(C_p) = \frac{1}{1 + \left(\frac{K C_p}{k_l k^L IC_{50}^L}\right)^m}. \quad (2.16)$$

The constant k_l has been estimated for mononuclear cells and for multiple drugs, but k^L and K are unknown constants.(Dixit et al., 2004) However, this issue can be avoided. Using our first assumption again (drug effect only dependent of intracellular concentrations), we can claim that our deduction remains valid when using assays of various plasma serum content. In fact, we could conclude that the efficacy of the drug on T-cells located in the lymph node would also be equal to:

$$U(C_p) = \frac{1}{1 + \left(\frac{K C_p}{k_l k^C IC_{50}^C}\right)^m}. \quad (2.17)$$

where k^C represents the ratio of intra to extracellular concentrations in a medium containing 100% plasma serum content, and IC_{50}^C is the assay concentration necessary to inhibit 50% of

infection events in this medium. In such an assay, we believe reasonable to assume that the ratios of intra to extracellular concentrations K and k^C would be approximately equal, simplifying the above equation to:

$$U(C_p) = \frac{1}{1 + \left(\frac{C_p}{k_l IC_{50}^C}\right)^m}. \quad (2.18)$$

Unfortunately, technical difficulties arise when the amount of plasma serum content rises in infectivity assays, making the value of IC_{50}^C difficult to evaluate directly. However, Acosta *et al.* noticed a linear relationship between the fold-change in the concentration value inhibiting half of the infection and the plasma serum content of the medium. They extrapolated the relationship to obtain an estimate of the IC_{50}^C as a function of the IC_{50}^0 (i.e. the inhibitory coefficient measured in a medium with 0% plasma serum content). (Acosta *et al.*, 2012) Let k_p be this fold-change, reported in Acosta *et al.*, such that $k_p = \frac{IC_{50}^C}{IC_{50}^0}$. We can conclude that:

$$U(C_p) = \frac{1}{1 + \left(\frac{C_p}{IC_{50}^0 k_l k_p}\right)^m} \quad (2.19)$$

where parameter values can be obtained for various drug and strain combinations. It can be verified that the same conclusions apply to other strains ($i \neq 1$), hence Equation 2.9. Note that parameters IC_{50}^0 , m , k_l and k_p are drug specific.

Assuming the absence of hysteresis, we can obtain an expression for the variation of the reproduction number of strain i , with time, based on Equation 2.9:

$$\boxed{R_{0i}(t) = \frac{R_0(1 - s_i)}{1 + \left(\frac{C_p(t)}{\rho_i IC_{50}^0 k_l k_p}\right)^{m(1+\sigma_i)}}}. \quad (2.20)$$

2.7.2.4. Eclipse Period

In virology, an eclipse period is the time separating the infection of a cell and the appearance of new complete viruses. An eclipse period of approximately one day is shown to occur for HIV virus.(Dixit et al., 2004) The initial model, however, neglected its impact, which resulted in unrealistic viral load curves following treatment interruptions. To replicate clinical observations, we further needed to adapt the model equations to include an eclipse period τ . The model then becomes:

$$\begin{aligned}\frac{dx}{dt} &= \lambda - \sum_{i=1}^n \beta_i x v_i - d_x x \\ \frac{dy_i}{dt} &= \beta_i x(t - \tau) v_i(t - \tau) e^{-d_x \tau} + A_i - d_y y_i \\ \frac{dv_i}{dt} &= k_i y_i - d_v v_i\end{aligned}\tag{2.21}$$

We set x and v_i to their equilibrium values as a function of y_i and derived the reproduction number under this new model (as in Rosenbloom et al. and Ribeiro et al.)(Ribeiro et al., 2010; Rosenbloom et al., 2012):

$$\begin{aligned}R_{0i}(t) &= \frac{\beta_i(t) k_i(t) \lambda}{d_x d_y d_v} e^{-d_x \tau} \\ X(t - \tau) &= \frac{\lambda^2}{d_x (\lambda + d_y \sum R_{0i}(t - \tau) y_i(t - \tau))} \\ V_i(t - \tau) &= \frac{k_i(t - \tau) y_i(t - \tau)}{d_v}\end{aligned}\tag{2.22}$$

where X and V represent the equilibrium values for x_i and v_i upon knowledge of all y_i . The model can be summarized by the following equations, similar to the set of Equations 2.7:

$$\boxed{\frac{dy_i}{dt} = A_i - d_y y_i(t) + \frac{d_y \lambda R_{0i}(t - \tau) y_i(t - \tau)}{\lambda + d_y \sum_j R_{0j}(t - \tau) y_j(t - \tau)}}\tag{2.23}$$

The eclipse period τ was set to 1 day, as used elsewhere.(Ribeiro et al., 2010) The cells going through an eclipse period were assumed to have the same lifetime expectancy as a non-infected cell (as modeled by the term $e^{-d_x \tau}$ above).

2.7.2.5. Adapting the model to study drug combinations

In order to describe the approach, a different notation will be used, introducing index j that takes values 1, 2 or 3 to identify three considered drugs. For drug j , Equation 2.9 is rewritten as:

$$U_{ij}(C_{pj}) = \frac{1}{1 + \left(\frac{C_{pj}}{\rho_{ij}IC_{50j}^0 k_{pj} k_{lj}}\right)^{m_j(1+\sigma_{ij})}}. \quad (2.24)$$

This notation has the additional advantage of highlighting those parameters which are drug-dependent and those which are dependent on both the drug and the strain. Hereafter, the notation U_i will refer to the proportion of single-cycle infection events that are unaffected by a combination of drugs (instead of a single drug treatment).

An *in vitro* study provided a strategy to estimate drug efficacy in the context of multiple drug therapies. (Jilek et al., 2012) More specifically, the study investigated the effect of pairs of drugs in infectivity assays. Empirical values of U_i were recorded for various pairs of drugs under various conditions. Using this data, the authors obtained a drug-combination-specific constant DI which is relatively independent of drug concentrations. This value allows extrapolating the value of U_i for all drug concentrations. The extrapolation relies on the expected value of U_i under two types of drug-drug interactions (Bliss independence and Loewe additivity), and on a constant, DI , which is a measure of the proximity of the empirical drug combination effect to either Loewe additivity or Bliss independence predictions. This strategy was adapted to estimate the effect of three drugs *in vivo* in the lymph nodes.

As a first step, it is possible to evaluate the proportion of single-cycle infection events that are unaffected by the combination of drugs in absence of competition for binding sites and for strain i ($U_{i,Bliss}$). In this case we have:

$$U_{i,Bliss} = \frac{1}{1 + \left(\frac{C_{p1}}{\rho_{i1}IC_{50_1}^0 k_{l1}k_{p1}}\right)^{m_1(1+\sigma_{i1})}} \times \frac{1}{1 + \left(\frac{C_{p2}}{\rho_{i2}IC_{50_2}^0 k_{l2}k_{p2}}\right)^{m_2(1+\sigma_{i2})}} \times \frac{1}{1 + \left(\frac{C_{p3}}{\rho_{i3}IC_{50_3}^0 k_{l3}k_{p3}}\right)^{m_3(1+\sigma_{i3})}} \quad (2.25)$$

Under Loewe additivity, all drugs compete for the same binding site. The associated proportion of unaffected events $U_{i,Loewe}$ satisfies the following equation:

$$1 = \frac{C_{p1}}{\rho_{i1}IC_{50_1}^0 k_{l1}k_{p1} \left(\frac{U_{i,Loewe}}{1 - U_{i,Loewe}}\right)^{-\frac{1}{m_1(1+\sigma_{i1})}}} + \frac{C_{p2}}{\rho_{i2}IC_{50_2}^0 k_{l2}k_{p2} \left(\frac{U_{i,Loewe}}{1 - U_{i,Loewe}}\right)^{-\frac{1}{m_2(1+\sigma_{i2})}}} + \frac{C_{p3}}{\rho_{i3}IC_{50_3}^0 k_{l3}k_{p3} \left(\frac{U_{i,Loewe}}{1 - U_{i,Loewe}}\right)^{-\frac{1}{m_3(1+\sigma_{i3})}}} \quad (2.26)$$

Both $U_{i,Bliss}$ and $U_{i,Loewe}$ can be computed for all drug concentrations since the parameters of these equations are equivalent to those obtained for single drug effects. (Sampah et al., 2011; L. Shen et al., 2008) To evaluate the empirical effect (U_i), we used the measure DI for drug pairs reported in Jilek *et al.*, which is assumed constant for each drug pair combined with any strain i and is defined as (Jilek et al., 2012)

$$DI = \frac{u_i - u_{i,L}}{u_{i,B} - u_{i,L}} \quad (2.27)$$

where $u_i = \log\left[\frac{1-U_i}{U_i}\right]$, $u_{i,B} = \log\left[\frac{1-U_{i,Bliss}}{U_{i,Bliss}}\right]$, and $u_{i,L} = \log\left[\frac{1-U_{i,Loewe}}{U_{i,Loewe}}\right]$. The evaluation of the effect of two drugs can be performed by computing U_i from $U_{i,Bliss}$, $U_{i,Loewe}$ and DI using Equation 2.20. Under the three drugs $j = 1$, $j = 2$ and $j = 3$, an estimate of the DI parameter (DI_{123}) is obtained from a weighted average of the DI parameter reported for each pair of drugs. More specifically:

$$DI_{123} = w_{12}DI_{12} + w_{13}DI_{13} + w_{23}DI_{23} \quad (2.28)$$

where DI_{ab} is the value of the parameter when considering the pair of drugs $j = a$ and $j = b$, as reported in Jilek *et al.*, and w_{ab} is an associated weight.[21] The value of the weight is based on the number of reproduction cycles that would be affected by each drug in absence of the other drugs (i.e. $1 - U_{ij}(C_{pj})$, for $j = 1, 2, 3$). We believe that the following is a good indication of each drug pair interaction opportunities. More precisely, we set:

$$w_{ab} := \frac{(1 - U_{ia})(1 - U_{ib})}{\sum_{A \neq B} (1 - U_{iA})(1 - U_{iB})} \quad (2.29)$$

This equation implies that, if the concentration of drug 1 is zero, for example, the value of DI_{123} is DI_{23} . Using values of $U_{i,Bliss}$, $U_{i,Loewe}$ and DI_{123} allows to evaluate the empirical value of U_i for the drug trio as a function of plasma concentrations using Equation 2.20. Therefore we have $U_i(C_{p1}(t), C_{p2}(t), C_{p3}(t))$ which enables evaluating viral replication at time t for viral strain i using the following expression, instead of Equation 2.20:

$$R_{0i}(t) = R_0(1 - s_i)U_i(C_{p1}(t), C_{p2}(t), C_{p3}(t)) \quad (2.30)$$

2.7.2.6. Adapting the model for cross-resistance

We adapted the model in order to study strains with multiple resistance mutations. The rate of emergence of new strains relies on the parent strain's replication rate and the probability that one replicated virus acquires the mutation associated with the new strain. These are taken into account in the model by two parameters: the fitness cost and the probability of *de novo* mutation. Those values could not be obtained from the literature for strains with multiple mutations. Furthermore, quantification of drug efficacy on strains with multiple mutations was performed. These parameters were approximated using the following assumptions.

We assumed that the probability of acquisition of a specific mutation is independent of the presence of other mutations on the originating strain. We also assumed that one mutation may be acquired per replication cycle. The efficacy of a specific drug on a viral strain having multiple mutations, U_{ij} from Equation 2.24, is computed by assuming that the fold-change ρ_{ij}

is equal to the product of fold-change values for single-mutation strains over the acquired mutations. More precisely, let $i = a$ be the index of a strain having acquired q resistance mutations. Let $[m_1, m_2, \dots, m_q]$ be the set identifying these mutations. Finally, let $[b_1, b_2, \dots, b_q]$ be the set of indices identifying the strains having one single resistance mutation and for which strain b_k has mutation m_k ($k \in [1, 2, \dots, q]$). Then $\rho_{aj} = \prod_{b \in B} \rho_{bj}$. This expression can be derived from assuming an independent action of each of the acquired resistance mutation on drug efficacy. The value of σ_{aj} is set equal to the average of the equivalent parameter values for single-mutation strains, over the involved mutations ($\frac{1}{q} \sum_{b \in B} \sigma_{bj}$). The fitness of a strain with multiple mutations $i = a$ is set equal to the middle point between $\min_{b \in B}$ of $(1 - s_b)$ and $\prod_{b \in B} (1 - s_b)$.

2.7.2.7. Inter-individual variability in model parameters

Different individuals can experience different outcomes despite using the same treatment. For example, it can be admitted that individuals with high drug clearance values are more likely to experience subtherapeutic concentrations. Under this context, ignoring variability in clearance between individuals would lead to underestimating the number of virological failure. For this reason, we added inter-individual variability in the most important model parameters. For most of these parameters, the distribution of attributed values mirrored observed data distributions in HIV-infected patient populations.

We identified the following factors as the main components that can impact viral loads from drug use: the specific drug intake pattern above and beyond the general adherence level, the patient's immunity toward infection (also, the impact of the infection phase on patients' immunity), drug disposition, and the quasispecies composing the patient's viral population. Each of these factors is associated with parameter values in the model.

We first researched the literature for population distributions of drug intake patterns. The two parameters defining drug intake patterns are the general adherence level p and the probability of drug intake following a missed dose q (PIM). For the majority of the clinical data used in this study, p was estimated from pill counts. We did not need to estimate population variability for this parameter since we could compare model predictions to clinical data using

groups having the same level of adherence p . However, the PIM values are not generally measurable in outpatient groups, hence the need to evaluate its variability within adherence groups. We obtained a distribution of PIM values per adherence level p using Medication Event Monitoring System data from 207 hypertensive outpatients taking medication once daily which was produced elsewhere.(Fellows et al., 2015) We assumed that there would not be major differences in pill intake behavior between this patient population and the HIV-positive individuals after accounting for p . Each patient from the sample has a level of adherence and a PIM value determined from their MEMS records. We were not interested in the distribution of patients between levels of adherence as comparative datasets reported the virological outcomes per adherence level stratum. However, we were interested in the distribution of PIM values within adherence level strata. Sample size restrictions motivated a categorization of both the level of adherence and the PIM before proceeding to estimation. The categories for the level of adherence were: 1) 0-59%, 2) 60-79% and 3) 80-100% (N=23, 29 and 137 when dismissing missing values). The categories for the PIM, within each level of adherence category, were: 1) PIM less than the value of the level of adherence minus 5%, 2) PIM greater than the value of the level of adherence plus 5% and 3) all other PIM values (close to the level of adherence). The distribution was fitted using frequency counts within adherence level groups.

It was observed that those with a general level of adherence of less than 60% have PIM values that are smaller than their general adherence level, which is associated with a higher propensity for long drug holidays. Those having general adherence levels between 60 and 79% had a wider range of patterns of drug intake. Those with a general adherence level of 80% or more had a lower tendency toward drug holidays. In fact, for this adherence group, the probability of pill intake was higher if the last dose was missed than if the last dose had been taken. Such behavior could be observed in a subject remembering missing a dose and making sure that the next one is taken. This gave the weight that was attributed to each PIM value within levels of adherence as shown in Table 2.1. Estimations of the probability of virological failure within groups relied on this distribution, as explained in the Methods section.

Table 2.1 – Distribution of weight values associated with the probability of pill intake following a missed dose (PIM), per adherence level. These are used to obtain population prevalence estimates of virological outcomes.

PIM	Adherence level								
	10	20	30	40	50	60	70	80	90
5	0.800	0.267	0.160	0.114	0.089	0.050	0.042	0.024	0.021
10	0.200	0.267	0.160	0.114	0.089	0.050	0.042	0.024	0.021
15		0.267	0.160	0.114	0.089	0.050	0.042	0.024	0.021
20		0.150	0.160	0.114	0.089	0.050	0.042	0.024	0.021
25		0.050	0.160	0.114	0.089	0.050	0.042	0.024	0.021
30			0.150	0.114	0.089	0.050	0.042	0.024	0.021
35			0.025	0.114	0.089	0.050	0.042	0.024	0.021
40			0.025	0.150	0.089	0.050	0.042	0.024	0.021
45				0.010	0.089	0.050	0.042	0.024	0.021
50				0.010	0.150	0.050	0.042	0.024	0.021
55				0.010	0.005	0.050	0.042	0.024	0.021
60				0.010	0.005	0.140	0.042	0.024	0.021
65				0.010	0.005	0.039	0.042	0.024	0.021
70					0.005	0.039	0.14	0.024	0.021
75					0.005	0.039	0.052	0.024	0.021
80					0.005	0.039	0.052	0.49	0.021
85					0.005	0.039	0.052	0.038	0.021
90					0.005	0.039	0.052	0.038	0.49
95					0.005	0.039	0.052	0.038	0.075
100					0.005	0.039	0.052	0.038	0.075

Patients' immunity was a second considered factor. There are clear differences in the therapeutic impact of the drug depending on the infection phase. In the acute infection phase, the viral population grows much more rapidly. (Ribeiro et al., 2010) We believe the main difference between the acute and chronic infection phase is the build-up of immunity towards infection. In the model, this is expressed by the basic reproduction number which becomes smaller after patients reach the chronic infection phase. Because of the tight link between the reproduction number and viral loads, we thought crucial to represent population variability in this parameter. Estimates of reproduction number values can be obtained via knowledge of the

viral load expansion rate.(Ribeiro et al., 2010) It can be shown that the relationship between this rate r and the reproduction number follows this approximate relationship:

$$R_0 = \left(1 + \frac{r}{d_y}\right)e^{r\tau} \quad (2.31)$$

Distribution values of r for individuals in the acute and chronic infection phases were obtained from publications.(Ioannidis et al., 2000; Ribeiro et al., 2010) Values of r were approximately lognormal in each case, as verified using quantile-quantile plots. The average and standard deviation of $\ln(r)$ were of -0.07 and 0.44 for individuals in the acute infection phase, -0.78 and 0.52 for individuals in the chronic infection phase, respectively, and associated with typical R_0 values 4.9 and 2.3. In the final model, individual values of $\ln(r)$ were randomly picked from the above-mentioned distributions and then transformed into R_0 using Equation 2.31.

The final model also considered inter-individual variability in drug concentrations. We included inter-individual variability in pharmacokinetic parameters using population pharmacokinetic models obtained from the literature.(Baheti et al., 2011; Blum et al., 2007; Moltó et al., 2013; Sanchez et al., 2011) Although not shown here, there can be significant differences in the observed maximum plasma drug concentration between individuals using the same regimen. The same is true for the area under the concentration time curve.

Finally, there are significant differences in estimated IC_{50} between viruses isolated from different individuals, even in the absence of known drug resistance mutations.(Parkin et al., 2004) A three-fold difference is commonly observed. Therefore, we used the range of IC_{50} values reported in product monographs. Individual IC_{50} values were obtained by randomly selecting a value from a lognormal distribution. Such distribution was shown to correctly fit observed results from population phenotypic assays.(Parkin et al., 2004) The average of the lognormal of this variable was estimated as the average between the logarithm of the minimum and maximum IC_{50} values reported in the product monograph for each drug. The standard deviation of the lognormal distribution was estimated as 1/6.5th the range of the difference between the logarithms of the maximum and of the minimum values reported in the product monograph. For example, for efavirenz, this meant a median IC_{50} value of 0.00055 mg/L

uncorrected for plasma protein binding.(Bristol-Myers Squibb, 2015) The parameter values for all drugs are listed in the section below.

2.7.3. Summary of parameter values

The above describes the modifications we brought to the initial model. For practical purposes, the parameter values we used for simulations are listed here. Most parameter values were taken from Rosenbloom *et al.*(Rosenbloom et al., 2012) The following lists the exceptions.

First, a delay τ of 1 day was added for the expected time between cell infection and production of virions. The basic reproduction number R_0 for individuals in the chronic infection phase follows a log-N(μ, σ^2) distribution with $\mu = -0.78$ and $\sigma = 0.52$.(Ioannidis et al., 2000) The basic reproduction number for individuals in the acute infection phase follows a log-N(μ, σ^2) distribution with $\mu = -0.07$ and $\sigma = 0.44$.(Ribeiro et al., 2010)

Efavirenz was simulated with a daily dosage of 600mg. The pharmacokinetics followed a 1-compartment model with absorption rate $k_a = 0.3h^{-1}$, clearance $CL = 9.61Lh^{-1}$, volume of distribution $V = 291L$. There is inter-individual variability in clearance $IIV_{CL} = 0.437$ and in volume of distribution $IIV_V = 1.01$.(Sanchez et al., 2011)

Emtricitabine was simulated with a daily dosage of 200mg. The pharmacokinetics followed a two-compartment model with absorption rate $k_a = 0.53h^{-1}$, clearance $CL = 15.1Lh^{-1}$, main compartment volume of distribution $V_1 = 42.3L$, peripheral compartment volume of distribution $V_2 = 55.4L$ and intercompartment clearance $Q = 5.75Lh^{-1}$. There was inter-individual variability of clearance $IIV_{CL} = 0.174$ and an impact of creatinine clearance on clearance of $\theta_{CLCR} = 0.278$. Creatinine clearance in liters per hour followed a log-N(μ, σ^2) distribution with $\mu = 104.5$ and $\sigma = 0.34$.(Blum et al., 2007)

Tenofovir was simulated with a daily dosage of 300mg tenofovir disoproxil fumarate. The pharmacokinetics followed a two-compartment model with absorption rate $k_a = 0.822h^{-1}$, clearance $CL = 42.1Lh^{-1}$, main compartment volume of distribution $V_1 = 224L$, peripheral compartment volume of distribution $V_2 = 512L$ and intercompartment clearance $Q = 175Lh^{-1}$. There was inter-individual variability of clearance $IIV_{CL} = 0.359$, main compartment volume $IIV_{V_1} = 0.794$ and peripheral compartment $IIV_{V_2} = 0.417$.(Baheti et al., 2011)

One combination therapy involved the intake of the following drugs: efavirenz, tenofovir disoproxil fumarate and emtricitabine. We did not consider medication interaction in absorption or elimination; efavirenz is mostly eliminated by the liver, while the other two medications are eliminated by the kidneys without noticeable interaction.(Blum et al., 2007) Consequently, drug concentrations were obtained by simulating the three drugs separately.

Another therapy involved drugs darunavir/ritonavir which was simulated for a regimen with daily dosage of 800/100mg, respectively. The pharmacokinetics followed a two-compartment model with absorption rate $k_a = 0.3h^{-1}$, clearance $CL = 10.7Lh^{-1}$, main compartment volume of distribution $V_1 = 36.9L$, peripheral compartment volume of distribution $V_2 = 1130L$ and intercompartment clearance $Q = 21.3Lh^{-1}$. There was inter-individual variability of clearance $IIV_{CL} = 0.318$, main compartment volume $IIV_{V_1} = 0.589$ and peripheral compartment $IIV_{V_2} = 0.821$.(Moltó et al., 2013)

Table 2.2 lists the parameters enabling the computation of drug efficacy on the infection of CD4+ T-cells in the lymph node. The degree of independence DI that was used for pairs of drugs are the following: 0.46 for efavirenz and tenofovir DF combination, 0.71 for efavirenz and emtricitabine combination, and 0.33 for tenofovir DF and emtricitabine combination.(Jilek et al., 2012) We supposed that ritonavir has a relatively small impact on viral growth at a dose of 100 mg daily (only impacting it through its effect on darunavir absorption and clearance). No correction was applied to account for 100% plasma serum content for tenofovir and emtricitabine as neither bind well to plasma proteins. As stated earlier, the variability in the susceptibility of wild-type virus in the population was taken from product monographs.(Bristol-Myers Squibb, 2015; Gilead Sciences Inc, 2008, 2013; Janssen Inc, 2014)

Table 2.2 – Plasma protein binding correction k_p , lymph node penetration k_l and variability in the susceptibility of wild-type virus in the population

Drugs	k_p	k_l	$IC_{50}^0(\text{mg/L})$ $\log\text{-N}(\mu, \sigma^2)$	
			μ	σ^2
Efavirenz	27	16.7	-7.49	0.41
Tenofovir	1	5	-1.79	0.82
Emtricitabine	1	2.9	-4.98	0.96
Darunavir	14	100	-6.28	0.3

2.7.4. Algorithm for the final model

We summarize the algorithm allowing to retrieve virological outcomes for an individual having a probability of pill intake p and a PIM value of q . In the following, we suppose a trial of d days. There are multiple considered viral strains, each identified with index i . Numerical experiments required time steps Δt separating two simulated time points. For all simulations used in this study, Δt was set to 0.1 hour (or about 0.004 day). All of the considered drug regimens were once daily, but the algorithm could be adapted to other dosing intervals. The following allows computing the number of infected cells of each strain, over time, for one virtual individual.

2.7.4.1. Simulating drug intake patterns

The drug intake pattern is represented by a series of 0 and 1 to represent a missed and taken dose respectively. T_j is the value associated to prescribed time j .

1. T_1 is simulated with probability p of taking a value of 1, otherwise it takes a value of 0,
2. For every subsequent dose occasion T_j : 1) if the value of T_{j-1} is 0 then we simulated a random binomial variable with probability q ; 2) if the value of T_{j-1} is 1 then we simulated a random binomial variable with probability $1 - \frac{q(1-p)}{p}$,

3. This second step is repeated until the number of trial days d is reached.

2.7.4.2. Simulation of drug concentrations

1. For each drug, the individual is attributed random pharmacokinetic parameter values using the appropriate population models,
2. Plasma concentrations are simulated using the pharmacokinetic model of the drug and individual model parameters obtained from the previous step. The drug intake pattern simulated in the step above is used for all the considered drugs (simultaneous drug intake),
3. Values of concentrations are recorded in a vector having one value for every 0.1 hour for the extent of the trial of d days. There is one vector per drug for each individual.

2.7.4.3. Simulating other individual variability

1. A random value of basic reproduction number R_0 is simulated from the adequate distribution depending on the infection phase (acute or chronic),
2. For each drug, random values of the IC_{50}^0 of the wild-type strain for the individual is simulated using the random distribution reported in Table 2.2.

2.7.4.4. Simulating drug effect

1. Using the notation proposed for Equation 2.24, each drug has a value of k_{lj} , k_{pj} and m_j , and each combination of drug and mutation has values of ρ_{ij} and σ_{ij} , enabling the computation of $U_{ij}(C_{pj}(t))$ for each of the 0.1 hours of the simulation. One vector $U_{ij}(C_{pj}(t))$ is obtained per drug being tested and per mutation,
2. If there is more than one drug, the predicted effect of the combination of drugs under Bliss independence and Loewe additivity, as per Equations 2.25 and 2.26, is computed,
3. If the treatment combines two drugs, the effect of the combination of drugs U_i is computed using Equation 2.20 and the degree of independence between drugs DI reported in Jilek *et al.*,(Jilek et al., 2012)

4. If the treatment combines three drugs, U_i is also obtained using DI , except that DI is now computed using Equation 2.28,
5. One vector is simulated per mutation (including the wild-type),
6. Using the basic reproduction number and each mutation's s_i , the result above is used to compute the $R_{0,i}(t)$ at each time point using Equation 2.23.

2.7.4.5. Simulating viral kinetics

The algorithm for the simulation of viral kinetics is similar to the one reported in the supplementary material of Rosenbloom *et al.*, (Rosenbloom et al., 2012) with the exception of the inclusion of a 1-day eclipse period. The output consists in multiple vectors y_i , one for each strain i , of $\frac{d}{\Delta t}$ elements.

The initiation of the algorithm consists in determining the number of infected cells of each type (strain) at the beginning of the simulation (or the first element of the y_i vector, $y_i[1]$). The total number of initially infected cells $\sum_i y_i [1]$ is chosen randomly from a distribution of setpoint values taken from population data (see Rosenbloom *et al.*). (Rosenbloom et al., 2012) Then the strain of each infected cell at the beginning of the simulation is simulated. The proportion of cells of each type could be obtained from solving for equilibrium a system of differential equations representing the dynamics of a population with self-limiting growth, possible mutations and fitness cost. It could also be obtained from running simulations without drugs for an elongated period from a random non-zero number of infected cells of each type. We used an approximation that keeps the order of magnitude of these proportions and assumed that the exact proportions would not make a big difference on simulation results. By default, the strain is assigned the wild-type strain $i = 1$, but these cells have a probability $P_{nov0,i}/S_i$ of being infected with strain i ($P_{nov0,i}$ is the probability of *de novo* mutation giving strain i from the wild-type strain). The value $y_i[1]$ is attributed to the first 240 elements of y_i ($\frac{24 \text{ hours}}{\Delta t \text{ hours}}$), to allow computation of the number of new infections accounting for a delay of one day. The reservoir of latent cells is composed of $\sum_i A_i = 3,000$ infected cells. The proportion of this reservoir being composed of cells infected by strain i is set to $\frac{y_i[1]}{\sum_i y_i[1]}$.

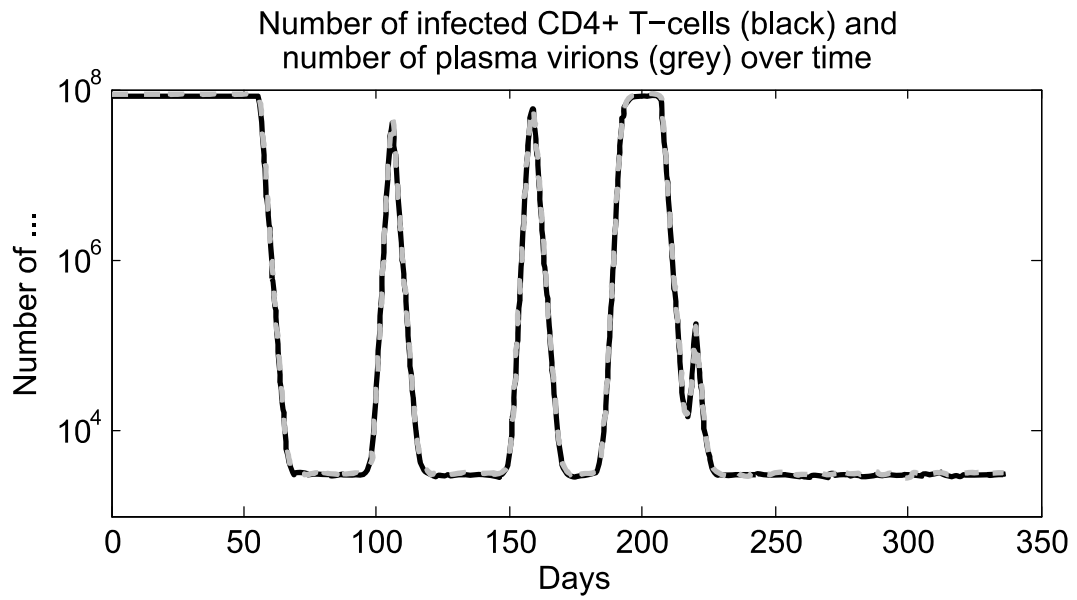
Then, iterations begin over time. For time $t + \Delta t$, the following number of events is computed for each viral strain using Poisson random variables:

1. the number of infected cell deaths (expectation: $y_i(t)(1 - e^{\Delta t \delta y})$),
2. the number of new infections for cells of each type is computed using the rightmost term of Equation 2.23 multiplied by Δt as expectation,
3. if further mutations are allowed for the simulated strain, a newly infected cell can be reassigned a mutation index from the probability of *de novo* mutation (of the order of 10^{-5} to 10^{-6} , for most),(Rosenbloom et al., 2012)
4. A certain number of infected cells from the reservoir get activated (expectation: leftmost term of Equation 2.23 times Δt).

2.7.5. Validation of one assumption

Using the initial model, there was a discrepancy in observed and predicted viral growth curves. We thought that the problem could emanate from one approximation: the number of virions in the plasma being approximately equal to the number of infected CD4+ T-cells in the body. We investigated this approximation. Clinical results convinced us of the adequacy of the assumption when an equilibrium is reached between the 2 quantities.(Hockett et al., 1999) It was less clear for us whether the assumption held true when the number of virions rapidly increases. To verify this assumption, we used a mathematical model developed in DeBoer *et al.*(De Boer *et al.*, 2010) that had for objective to explain why viral loads observed in the plasma were comparable to the number of infected CD4+ T-cells, knowing that the number of virions produced by infected cells is higher than plasma viral clearance. Their model showed that virion intake by dendritic cells within the lymph as well as high virion clearance in the tissues could explain this result. We used their dynamical system with their physiologically based parameter values to simulate the viral loads when the number of CD4 infected cells rapidly increased. The results are shown in Figure 2.8 and provided an argument toward the adequacy of the approximation.

Figure 2.8 – Model-based number of infected T-cells and number of plasma virions over time



Chapitre 3. Exploring an alternative explanation for the second phase of viral decay: infection of short-lived cells in a drug-limited compartment during HAART

Authors and affiliation: Steven Sanche^{1*}, Thibault Mesplède^{2,3}, Nancy L. Sheehan^{1,4}, Jun Li^{1,5}, Fahima Nekka^{1,5}

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

Publié dans *PloS one* 13, no. 7 (2018): e0198090.

3.1. Préambule

Afin de mieux comprendre ce qui a mené à l'objectif du présent chapitre, il est nécessaire de mentionner les pistes d'études qui ont découlé du chapitre 2 et qui ont été préalablement explorées. Le modèle décrit au chapitre précédent a mené à une première visée, celle d'évaluer le risque de développement de résistance pour un antirétroviral récemment approuvé par la FDA, soit le dolutégravir. Un obstacle a entravé l'utilisation du modèle: le taux de pénétration de cet inhibiteur de l'intégrase dans les ganglions lymphatiques, paramètre nécessaire aux prédictions, était indisponible. Deux tentatives infructueuses d'estimation de la valeur de ce paramètre ont été effectuées : une régression du taux de pénétration en fonction de propriétés physico-chimiques de médicaments à niveaux de pénétration connus, et l'élaboration d'un modèle pharmacocinétique à base physiologique. Une troisième tentative fut d'utiliser l'efficacité *in vivo* de médicament, dérivée de la vitesse de décroissance de la charge virale, afin de déduire le taux de pénétration du médicament dans les ganglions. Des lectures approfondies à ce sujet ont mené à la problématique des phases de décroissance de la charge virale (revoir la section intitulée « La persistance du virus ») et à questionner le rôle de la pénétration tissulaire des antirétroviraux dans sa genèse. Une étude ayant pour but d'explorer cette question n'a pas mené

à des résultats concluants.(Murray et al., 2007) Or, le modèle élaboré précédemment peut être utilisé afin d'évaluer la plausibilité d'une des hypothèses qui y est mentionnée. Cette hypothèse stipule que la deuxième phase de décroissance de la charge virale est due à une concentration plus faible de médicaments dans un compartiment physiologique. Le modèle que nous avons développé a été modifié afin de permettre la considération de multiples compartiments physiologiques différemment exposés à la médication. Il a ensuite été évalué par l'analyse de sa capacité à prédire deux jeux de données empiriques : i) des charges virales provenant de plusieurs patients et prélevées à court terme suite au début d'un traitement (les deux premières phases se produisent dans les premiers mois), ainsi que sa capacité à prédire la charge virale à long terme chez des groupes de patients.

3.2. Abstract and keywords

Most HIV-infected patients who initiate combination antiretroviral therapy experience a viral load decline in several phases. These phases are characterized by different rates of viral load decay that decrease when transitioning from one phase to the next. There is no consensus as to the origin of these phases. One hypothesis put forward is that short- and long-lived infected cells are responsible for the first and second phases of decay, respectively. However, significant differences in drug concentrations are observed in monocytes from various tissues, suggesting the first two phases of decay in viral loads could instead be attributed to short-lived cells being differently exposed to drugs. Compared to a well-exposed compartment, new cell infection can be expected in a compartment with limited drug exposure, thus leading to a slower viral load decay with potential virologic failure and drug resistance. In the current study, the latter hypothesis was investigated using a model of viral kinetics. Empirical datasets were involved in model elaboration and parameter estimation. In particular, susceptibility assay data was used for an *in vitro* to *in vivo* extrapolation based on the expected drug concentrations inside physiological compartments. Results from numerical experiments of the short-term evolution of viral loads can reproduce the first two phases of viral decay when allowing new short-lived cell infections in an unidentified drug-limited compartment. Model long-term predictions are however less consistent with clinical observations. For the hypothesis to hold, efavirenz, tenofovir and emtricitabine drug exposure in the drug-limited compartment would have to be

very low compared to exposure in peripheral blood. This would lead to significant long-term viral growth and the frequent development of resistant strains, a prediction not supported by clinical observations. This suggests that the existence of a drug-limited anatomical compartment is unlikely, by itself, to explain the second phase of viral load decay.

Keywords:

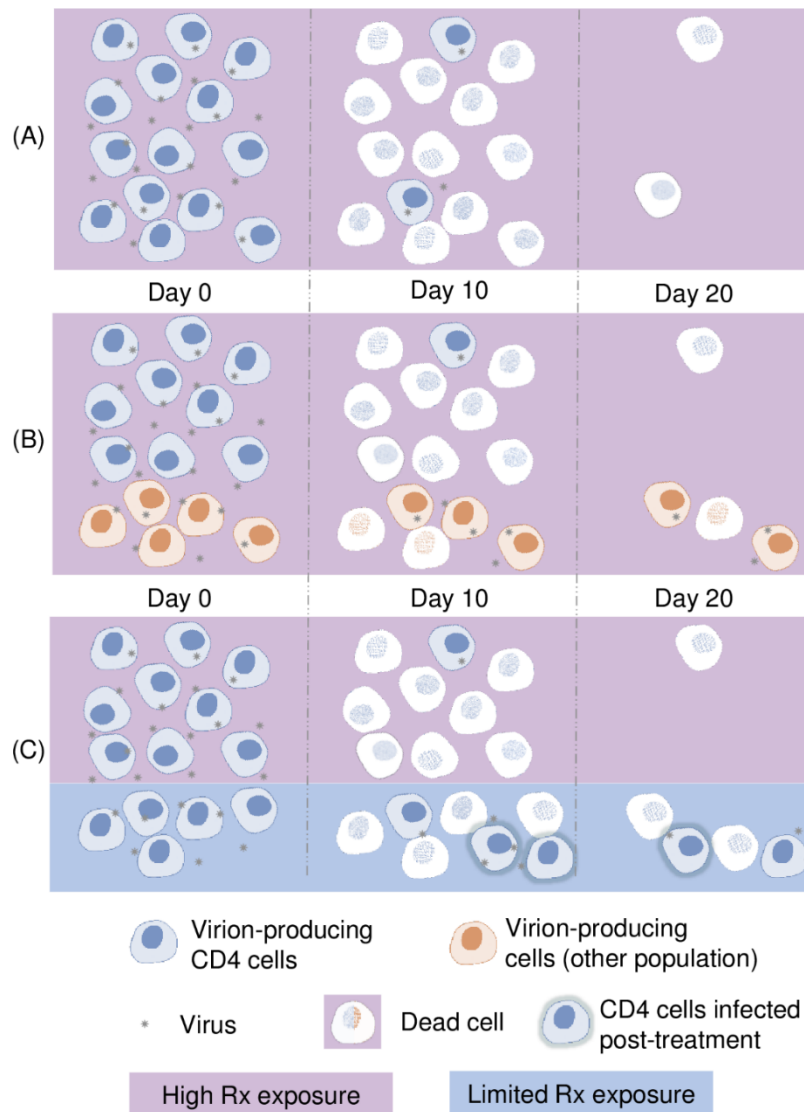
HIV; viral loads; antimicrobial resistance; highly-active antiretroviral therapy; decay rates; drug penetration

3.3. Introduction

Viral loads in the plasma of patients initiating highly active antiretroviral therapy (HAART) generally decrease very rapidly during the first days of treatment before reaching a slower second phase of decay.(Murray et al., 2007; Perelson et al., 1997) In fact, up to four phases of decreasing viral load can be observed, each new phase being slower than the previous one.(Maldarelli, 2015) These phases are the result of the complex interaction between host, drugs and virus. The existence of multiple phases of viral decay challenges our understanding of this interaction.(Blankson, Persaud, & Siliciano, 2002)

In the following, we will demonstrate that there are multiple rational explanations for the first two phases of viral load decay. First, we will infer that a set of three assumptions is inconsistent with multiple phases of viral decay. Under the first assumption, viral loads during the first and second phases of viral decay mainly come from one infected cell population: CD4 cells having a half-life of virion production of about one day (short-lived). Under the second assumption, viral loads are proportional to the number of infected cells. This assumption is partially supported by results suggesting rapid virion clearance in lymphoid tissue and plasma (no accumulation of virions).(De Boer et al., 2010; Zhang et al., 1999) Under the third assumption, HAART has the capacity to fully inhibit all new cell infections.

Figure 3.1 – Hypotheses related to the decay of virion-producing cells after treatment initiation. A) Virion-producing cells are short-lived infected CD4 cells and all new cell infections are prevented due to high drug exposure. This translates into only one phase of viral decay; B) There are two types of virion-producing cells having half-lives of around 1 (short-lived) and 14 days (long-lived), respectively, and all new cell infections are prevented due to high drug exposure. (Perelson et al., 1997) This translated into two phases of viral decay C) Virion-producing cells are mostly short-lived infected CD4 cells located in two compartments, one with high (pink) and one with low (blue) drug exposure. The compartment with low drug exposure partially allows new cell infections, effectively leading to two phases of viral decay. The compartment is not associated to a specific tissue at this point, as its existence is hypothesized.



If all these assumptions were true, there would be only one phase of viral decay, as depicted by Figure 3.1A. Indeed, under assumption one and two, the viral load (V) would equate to a proportionality constant (K) times the number of short-lived infected CD4 cells (C) which would decay at a constant rate (d). Because of assumption three, C cannot increase after treatment initiation. In other words, the viral load would be described by the following equation: $V(t) = K C(t_0) e^{-dt}$, where t_0 is the time of treatment initiation, and t represents the time since t_0 . This equation can only describe one phase of viral decay. This incoherence inevitably calls into question the above assumptions. We will now show how two of these assumptions have been revisited in order to give rise to two competing hypotheses, one of which is further studied herein.

In particular, the assumption of a single source of plasma virion (short-lived CD4 cells) has been largely disputed regarding its involvement in the slowdown of the viral load decay after the rapid first phase.(Perelson et al., 1997; Spivak et al., 2011) Other sources of virions have been suggested to explain the occurrence of this second phase decay. One possible explanation put forward is the existence of infected cells that can survive for a longer period (long-lived) (Figure 3.1 B).(Perelson et al., 1997) Each of the four decay phases would correspond to a specific virion-producing cell population.(Costiniuk & Jenabian, 2015) Recent work suggests the long-lived cells responsible for the second phase of decay could integrate viral DNA more slowly, explaining the delay before virion production.(Cardozo et al., 2017) These cells could potentially be resting CD4 cells or macrophages, which do exhibit slower integration when investigated *in vitro*.(Arfi et al., 2008; Swiggard et al., 2005) Other authors also mention dendritic cells or monocytes as a potential source of virions explaining the second phase of decay.(Hlavacek, Stilianakis, Notermans, Danner, & Perelson, 2000; Kaul, Garden, & Lipton, 2001; Lewin et al., 1998; Yukl et al., 2013) However, uncertainties prevail as to the kinetics of these infected cells and the total viral contribution of these cells *in vivo*. So far, kinetics estimates are solely based on data fitting using models that assume the involvement in viral decay of secondary sources of virions.(Cardozo et al., 2017) However, model assumptions can drastically change the values of these estimates since models solely involving cells with a uniform half-life can still fit the viral decay data quite well, as will be later seen.

Recent work challenges the notion that HAART has the capacity to fully inhibit all new cell infections in all organs. Indeed, there can be a large difference in drug concentrations across tissues within the host.(Fletcher et al., 2014) For example, one study reports very poor penetration in lymph nodes for all tested antiretrovirals, and several poorly penetrated ileum and rectum tissues.(Fletcher et al., 2014) Testicular tissue is another example of a drug sanctuary for many antiretrovirals,(Huang et al., 2016) along with the central nervous system.(Yilmaz et al., 2012) The extent of drug penetration in specific secondary lymphoid tissue is largely unknown for most antiretrovirals.(Fletcher et al., 2014) Motivated by this and by reported evidence of ongoing viral replication for patients taking HAART,(Frenkel et al., 2003; Kupfer et al., 2007; Lorenzo-Redondo et al., 2016) we were led to question the role drug penetration may play in explaining the first phases of viral decay, as proposed by Murray et al. (Murray et al., 2007) The potential involvement of new cell infections in a physiological compartment with limited drug penetration provides an alternative perspective. Under this hypothesis, cells infected before treatment initiation would produce the majority of virions lost during the first phase of decay, while cells infected under treatment within a drug-limited compartment would produce the virions lost during the second phase of decay. Long-lived infected cells would have a negligible impact on viral load, infected cells being short-lived CD4 cells. A depiction of this hypothesis is provided (see Figure 3.1 C).

Better characterizing the involvement of a drug-limited compartment is important, since a compartment harbouring a large quantity of infected cells and with little drug exposure could affect long-term treatment efficacy. Indeed, a large viral population implies a high chance for resistant mutants to be already present before treatment initiation, as the virus naturally mutates at a high frequency.(Abram et al., 2014) Low drug levels could then lead to suboptimal suppression of mutants that resist to antiretroviral drugs. However, factors such as the decreased viral fitness of resistant strains, compared to wild-type virus, could offset resistant viral growth.(Hu & Kuritzkes, 2014) Further, concomitant use of multiple drugs could prevent the exponential growth of resistant strains, as such strains may be sufficiently susceptible to one or more of the drugs.

We have developed a mathematical model accounting for these various factors. Our previous mechanistic model describes viral and short-lived infected CD4 cells dynamics in one

compartment, namely lymph nodes.(Sanche et al., 2017) The model can explain the long-term risk of virologic failure.(Sanche et al., 2017) Herein, this model was modified to include two compartments, one with low and another with high drug penetration. The compartment with low drug penetration, hereon referred as the drug-limited compartment, partially allows the infection of CD4 cells by the wild-type virus. The compartment is not associated to a specific tissue at this point, as its existence is hypothesized. The compartment with high drug penetration comprises tissues where drug concentrations are sufficient to prevent new cell infection events from wild-type virus (e.g. hypothetically the plasma and well-irrigated tissues).

Our approach was integrative and characteristic of the quantitative pharmacology of systems (QSP).(Craig et al., 2016) Compared to classical ‘top-down’ approaches that use empirical observations to deduce a mathematical relationship between measured input and output, bottom-up approaches, including QSP, integrate all the information available on the involved processes to build a mechanistic model that describes the hypothesized causal links between input and output. This mechanistic model is then used to predict the output from input values. The produced predictions are compared to empirical observations in order to evaluate if the hypothesized system can explain the phenomenon under investigation.

Our first objective is to evaluate the capacity of the new mechanistic model to predict viral loads observed at the initiation of treatment, particularly the occurrence of two phases, the time delay between phases and the rates of decay. One model was fitted to the two-month viral load data of 6 patients taking nelfinavir, zidovudine and lamivudine, using least squares.(Perelson et al., 1997) As a second objective, we examined whether results from this model are consistent with the observed risk of long-term virologic failure. Short- and long-term viral load data of patients taking efavirenz, tenofovir disoproxil fumarate (DF) and emtricitabine combination therapy were used.(Gallant et al., 2006; Karris et al., 2017) The former dataset was used to estimate model parameters using a least-square approach, while the latter was used to compare with model predictions. Depending on whether we assess the short- or long-term viral load evolution for treatment naïve patients, resistance mutations are either excluded or included as part of our model, respectively.

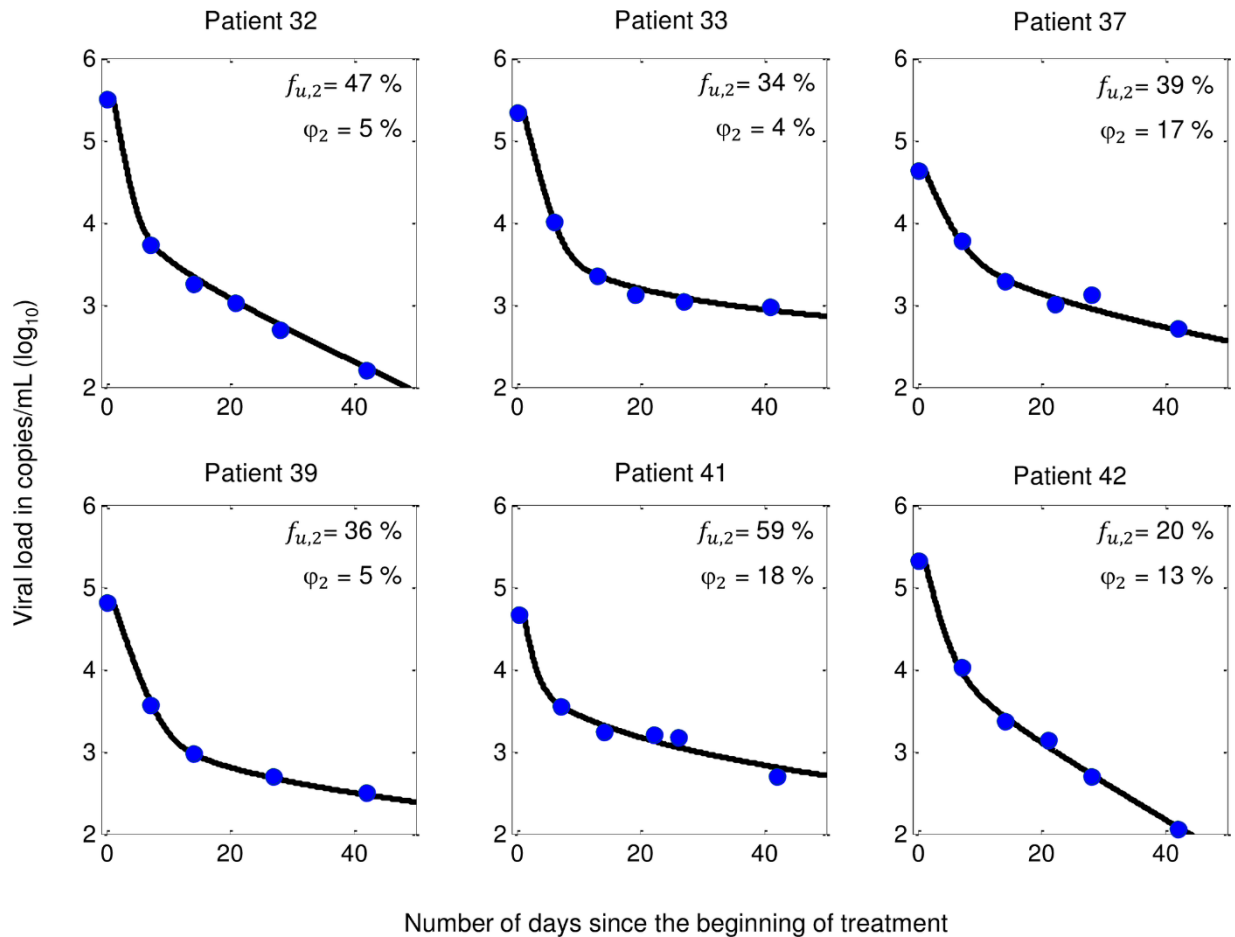
3.4. Results

The QSP approach adopted here has previously led to a model that serves as a basis for this study.(Sanche et al., 2017) This previous model, which only looked at infected CD4+ lymphocytes in active state as a source of virions, has been modified in order to simultaneously consider two physiological compartments differently exposed to antiretroviral drugs. One compartment represents the blood compartment and well-irrigated tissues where it is assumed drug penetration is sufficient to prevent new cell infection by wild-type virus. The other compartment is drug-limited. In the following, the former compartment is identified with index 1, while the drug-limited compartment is identified with index 2. The model parameters requiring *a posteriori* estimation are: 1) the average fraction of total infection events not affected by the drugs in each compartment for the wild-type virus ($f_{u,1}$, $f_{u,2}$), and 2) the contribution of each compartment to the maximum plasma viral load or setpoint, in percent (φ_1 , φ_2). The values for parameters $f_{u,1}$ and $f_{u,2}$ are mathematically associated with drug concentrations in the respective compartments, with higher drug concentrations leading to smaller f_u values (see Methods and Equation 3.1 for detail).(Rosenbloom et al., 2012) As for parameters φ_1 and φ_2 , an illustrative example would be if one compartment contributes 20% of the viral load when the patient reaches the peak viral load, then the value of φ associated with this compartment would be 20%. Since there are only two compartments, the value of φ for the other compartment would be 80%.

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

Viral load data were retrieved using WebPlotDigitizer software(Rohatgi, 2012) for the 6 patients reported in Perelson et al.(Perelson et al., 1997), displayed in Figure 3.2, blue dots. All patients were treatment naïve and initiated a therapy combining three antiretroviral drugs (nelfinavir, zidovudine and lamivudine). Patient demographics for this study can be found in Table 1 of the referenced article.(Perelson et al., 1997)

Figure 3.2 – Viral load data extracted from Perelson et al.(Perelson et al., 1997) (blue dots), model fit (black curve), and associated $f_{u,2}$ and ϕ_2 values. $f_{u,2}$ is the average fraction of total infection events not affected by the drugs and ϕ_2 is the fraction of the plasma viral load setpoint. Both parameters refer to the drug-limited compartment.



The short-term viral load decay of the 6 patients can be recovered using our mechanistic model, while keeping the model structure as well as the *a priori* determined parameter values. One of the two compartments is assumed to be sufficiently exposed to medication to prevent further CD4+ infection, i.e. $f_{u,1} = 0$. This compartment represents physiological compartments where antiretroviral drug concentrations are sufficiently high to prevent any new cell infection (e.g. hypothetically the plasma and well-irrigated tissues). Results from model fitting procedures for the remaining parameters, using a least-square approach, are shown in Figure

3.2. With these parameter values, our model predicted the two phases of viral load decay for the 6 patients. Fractional values φ_2 ranging from 4 to 18% allowed the best model fits. Also, when the drugs inhibit 41% to 80% of the infection in the drug-limited compartment (corresponding to $1-f_{u,2}$), the obtained predictions are the closest to the viral load observations based on least squares.

To assess whether the above short-term findings still hold for longer-term virologic response, the model was used to simulate viral load dynamics on a period longer than two months. For this, we first determined a $f_{u,2}$ value using short-term viral load data for patients under efavirenz, tenofovir DF and emtricitabine (600, 300 and 200 mg daily) combination therapy.(Karris et al., 2017) Secondly, values around the $f_{u,2}$ estimate were translated in terms of drug concentrations in the drug-limited compartment. Finally, these concentrations served to simulate the viral load evolution over a period of approximately one year. This time, resistant strains were allowed to emerge. Virologic failure at 48 weeks was the simulated outcome. To compare the simulation results with clinical data, the virologic failure threshold was set to 400 copies per mL of plasma.(Gallant et al., 2006)

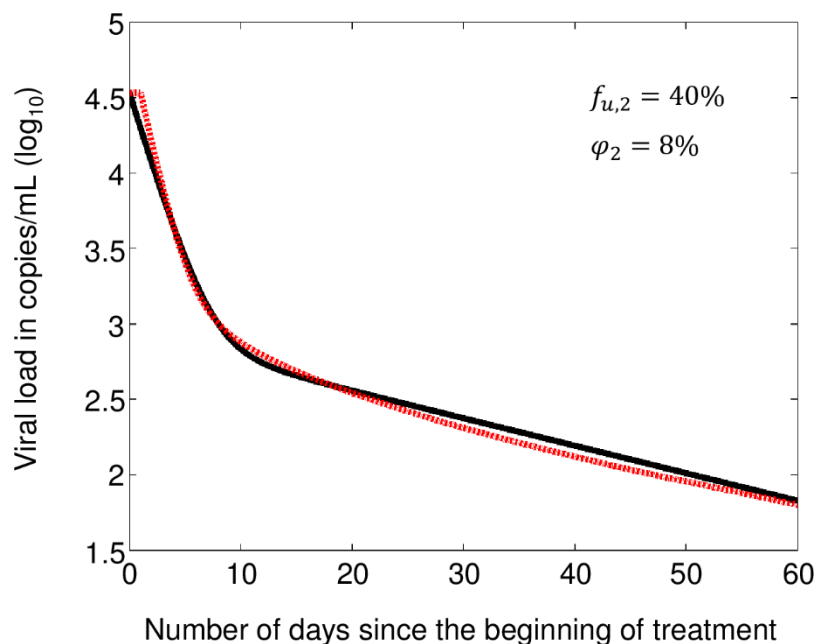
3.4.2. What values of $f_{u,2}$ and φ_2 allow reproducing the first two phases of decay in a typical patient under efavirenz, tenofovir DF and emtricitabine?

Karris et al. reports a mixed-effect biexponential regression model for the viral loads of 25 patients initiating the combination treatment containing efavirenz, tenofovir DF and emtricitabine.(Karris et al., 2017) Patient demographics for this dataset can be found in Table 2 of the referenced article.(Karris et al., 2017) To determine $f_{u,2}$ values, we focused on the typical patient's viral loads, in the sense that the patient is characterized by median values of the regression model parameters. In Figure 3.3, we superimposed the best fitting decay curve predicted by our model on the regression curve representing the viral load decay over a two-month period for the typical patient.(Karris et al., 2017) The values of $f_{u,2}$ and φ_2 used in our model were obtained in the same way as done for Figure 3.2. The values for these parameters allowing the adjustment in Figure 3.3 are 40% and 8%, respectively. A value of 40% for $f_{u,2}$ corresponds to an inhibition of 60% of the infection activity in the drug-limited compartment.

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

The parameter value $f_{u,2} = 40\%$ was translated into concentration values of efavirenz, tenofovir and emtricitabine in the drug-limited compartment. In fact, the association between $f_{u,2}$ and concentrations were mathematically described for many drugs using *in vitro* experiments. (L. Shen et al., 2008) The *in vivo* relationships can be deduced from these associations by assuming that intracellular concentrations dictate drug efficacy, independently of the medium in which the cells reside. (Sanche et al., 2017) The concentration inside the mononuclear cells of a compartment can be derived from extracellular plasma concentrations using two parameters,

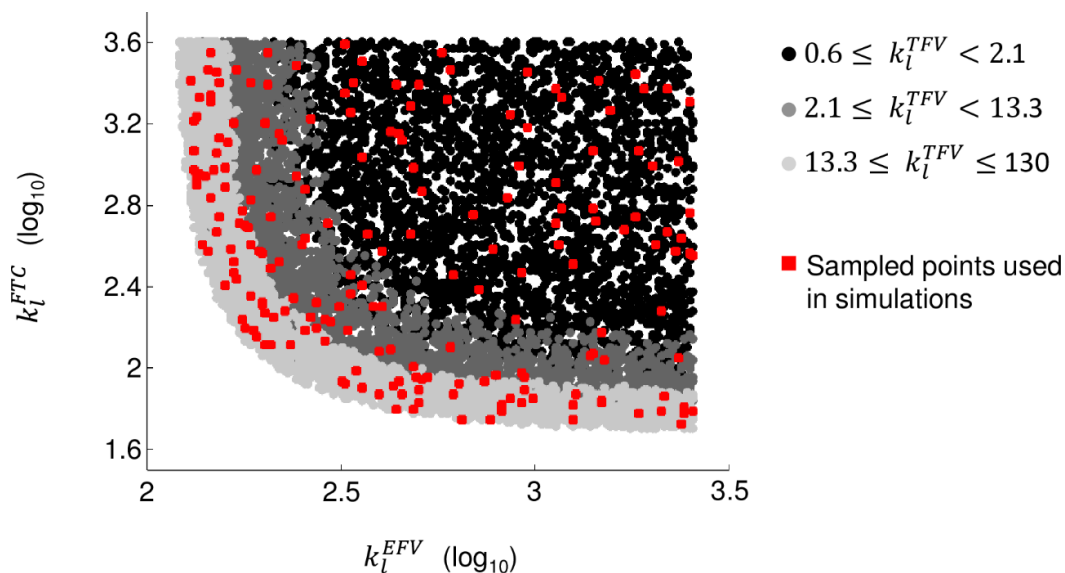
Figure 3.3 – Regression line (black) for the typical patient undertaking a treatment combining efavirenz, tenofovir DF and emtricitabine, model fit (red), and associated $f_{u,2}$ and ϕ_2 values. $f_{u,2}$ is the average fraction of total infection events not affected by the drugs and ϕ_2 is the fraction of maximum plasma viral load. Both parameters refer to the drug-limited compartment. Regression curve based on data from Karris et al. (Karris et al., 2017)



k_p and k_l (see Equation 3.1 from Material and Methods). At steady-state, the median extracellular plasma concentrations C_p were extracted from population pharmacokinetics studies.(Baheti et al., 2011; Sanchez et al., 2011; Valade et al., 2014) Parameter k_p , reflecting the effect of plasma protein binding on drug efficacy, is independent of compartmentalization and has been determined empirically elsewhere.(Acosta et al., 2012) Parameter k_l , representing the ratio of drug concentrations inside peripheral blood monocyte to the equivalent in the compartment, is the sole unknown value to determine $f_{u,2}$. When a single drug is used, the relationship between $f_{u,2}$ and k_l is one-to-one, and the latter can be uniquely determined. When multiple drugs are used concomitantly, the procedure is more complex as the values k_l for each drug need to be determined (k_l^{EFV} , k_l^{TFV} and k_l^{FTC} for efavirenz, tenofovir and emtricitabine, respectively). Further, different combinations of drug concentrations can lead to the same value of $f_{u,2}$, after accounting for interaction in drug effects.(Jilek et al., 2012) This translates into multiple combinations of k_l^{EFV} , k_l^{TFV} and k_l^{FTC} leading to the same value of $f_{u,2}$. Hence, using a computational iterative procedure, we determined the combinations k_l^{EFV} , k_l^{TFV} and k_l^{FTC} leading to an inhibition of $60 \pm 5\%$ of the infection activity.(Sanche et al., 2017)

The space of parameter values is displayed in Figure 3.4. In general, a higher k_l value indicates less drug penetration in the hypothesized compartment compared to peripheral blood, while a value of $k_l < 1$ indicates a higher drug concentration in the compartment. The k_l^{EFV} value ranges between 88 and 2593, k_l^{TFV} between 0.6 and 130 and k_l^{FTC} between 32 and 4000. Inguinal lymph nodes are known to be less exposed to efavirenz, tenofovir and emtricitabine.(Fletcher et al., 2014) For comparison purposes, we here give the reported values of k_l^{EFV} , k_l^{TFV} and k_l^{FTC} in this tissue, which are 16.7, 2.9 and 5, respectively, implying a greater drug exposure in lymph nodes as compared to the hypothesized drug-limited compartment.(Fletcher et al., 2014)

Figure 3.4 – Space of k_l parameter values associated with efavirenz (EFV), tenofovir (TFV) and emtricitabine (FTC) leading to an inhibition of 55%-65% of infection events in the drug-limited compartment. Parameter k_l represents the ratio of concentrations inside peripheral blood mononuclear cells to its equivalent in the hypothesized drug-limited compartment.



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

The values k_l^{EFV} , k_l^{TFV} and k_l^{FTC} associated with the drug-limited compartment were used to compute drug efficacy through viral load curves over 48 weeks of therapy with efavirenz, tenofovir DF and emtricitabine (600, 300 and 200mg daily). For this *in silico* experiment, we considered the possibility of the emergence of resistant strains through single nucleotide mutations responsible for the K103N, Y181C, G190S, M184V or K65R substitutions, as used in our previous study.(Sanche et al., 2017) In the simulations, these strains could be generated *de novo* or be selected. We used reported empirical curves of strain-specific drug efficacy and viral fitness.(Rosenbloom et al., 2012; L. Shen et al., 2008) Virtual patients expressed inter-individual variability in pharmacokinetics and immune response, modeled using patient population data.(Baheti et al., 2011; Ioannidis et al., 2000; Ribeiro et al., 2010; Sanchez et al., 2011; Valade et al., 2014) This inter-individual variability was added so groups of virtual

patients are representative of real patient populations, as in Sanche et al.(Sanche et al., 2017) Due to the uncertainty in drug penetration levels in the second compartment (the k_l values), one simulation was carried out for each of 200 sampled trios of parameter values k_l^{EFV} , k_l^{TDF} and k_l^{FTC} , selected using simple random sampling (points shown in red in Figure 3.4). For each trio of parameter values, we calculated the proportion of individuals expected to experience virologic failure over 100 simulated virtual patients.

To give the reader a better sense of the performed simulations, the predicted viral load of one of the simulated patients, as a function of time since treatment initiation, is shown in Figure 3.5. The transition from phase one to phase two can be observed approximately a week after treatment initiation (Figure 3.5, upper left). The long-term viral load originating from each compartment is also presented (Figure 3.5, upper right). In this example, the K103N mutation was selected in the drug-limited compartment (Figure 3.5, bottom left). There was no selection for other strains (Figure 3.5, bottom right).

We used a virologic failure threshold of 400 copies/mL to match the definition used in a previously published clinical study.(Gallant et al., 2006) Patient demographics for this study can be found in Figure 3.1 of the referenced article.(Gallant et al., 2006) A box plot for the predicted risk of virologic failure was compared to clinical data.(Gallant et al., 2006) As illustrated in Figure 3.6, we obtained a range of values for the risk of virologic failure going from 22% to 67% at the end of the trials. Most virtual cases of virologic failure were with resistance (>20% of the viral load attributed to resistant viral strains), with K103N and K65R being the most common mutations conferring resistance (data not shown). The smallest value over all 95% confidence intervals for the predicted risk of virologic failure was 17.4%. In comparison, the risk of virologic failure (>400 copies/mL) based on clinical observation is 2% of patients (4/210 patients), with a 95% confidence interval of [0.5;3.8]%.(Gallant et al., 2006) The p-value for the absence of difference of the proportion of failure between the empirical sample and the model-based sample associated with the smallest predicted risk is smaller than 1E-08. Those who left the clinical study for reasons unrelated to their viral load observations were discarded before computing the empirical risk of failure (pregnancy (n=4), adverse events (n=9), lost to follow-up (n=12), withdrew consent (n=5), other (n=4)),(Gallant et al., 2006) explaining the

difference with the value reported in the original paper (16% without a response).(Gallant et al., 2006)

Figure 3.5 – Simulation results for one random patient after 48 weeks of therapy with efavirenz (600mg), tenofovir DF (300mg) and emtricitabine (200mg) taken once daily. The model included a drug-limited compartment that allowed new short-lived cell infections. The virologic failure threshold was set at 400 copies/mL.

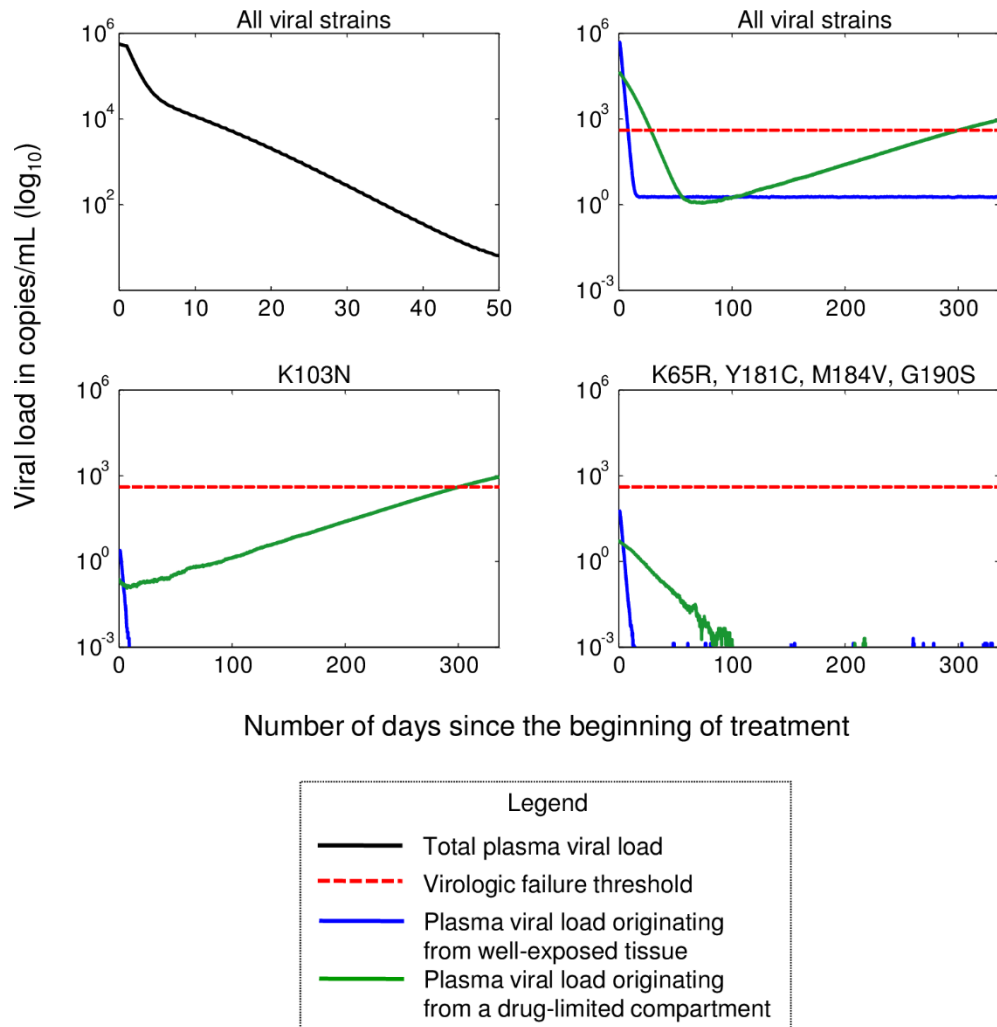
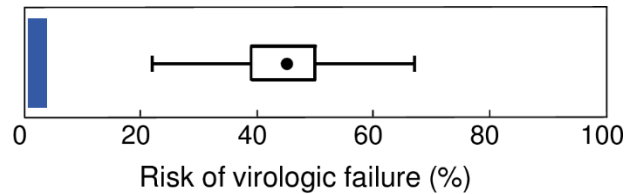


Figure 3.6 – Boxplot of the risk of virologic failure as predicted by the model and obtained from 200 simulations, one per sampling point of drug penetration values (k_l^{EFV} , k_l^{TFV} , k_l^{FTC}). For comparison purposes, the area shaded in darker blue is the confidence interval for the equivalent but observed in a real patient sample. (Gallant et al., 2006)



We then questioned whether it was possible to reproduce the empirical risk of failure using φ_2 and $f_{u,2}$ values around their estimates from the main analysis (8% and 40%, respectively). For the purpose of this sensitivity analysis, additional simulations were performed by first allowing φ_2 to take values of 6%, then 10%. For each new φ_2 value, we sampled 50 points in the space of parameter values for k_l^{EFV} , k_l^{TFV} and k_l^{FTC} associated with an unchanged $f_{u,2}$ and, for each point, computed the risk of failure over simulations of $N=100$ virtual patients. Predicted probabilities of failure fell within the range obtained in the main analysis (22% to 67%). Using a similar approach, we computed the risk of failure assuming $f_{u,2}$ values of 30%, then 50%, with an unchanged φ_2 . The smallest risk of failure over these simulations was 17% (95%CI [10.2;25.8]%). We further extended our analysis to $f_{u,2}$ values of 20% and 60%. For these parameter values, the smallest predicted risk of failure was 9% (95%CI [4.2;16.4]%) and was achieved for a $f_{u,2}$ value of 20%. Under this $f_{u,2}$ value, the slope of second phase decay for the typical individual would be more than 2.5 times greater than its empirical estimate.

3.5. Discussion

The viral load of most HIV-infected patients initiating HAART decays in distinct phases. (Blankson et al., 2002) During each phase, the viral load decays at an approximately constant rate. However, this rate decreases substantially from one phase to the next. During the final phase, the decay is so slow that low-level viremia (<50 copies/mL) persists over many years of therapy. (Palmer et al., 2008) To date, the causes behind these phases, which characterize the viral decay profiles, still draw attention and fuel a debate. One hypothesis advances that the second phase of viral decay would be caused by infected cells that could

survive longer than short-lived CD4 cells.(Perelson et al., 1997) Evidence suggests these cells could be CD4 cells exhibiting a slower integration of viral DNA.(Cardozo et al., 2017) Another hypothesis stipulates that new short-lived infected cells from a drug-limited compartment could cause the phenomenon.(Murray et al., 2007) We investigated the latter hypothesis.

Contrary to the role long-lived cells could play in this phenomenon, much less attention has been paid to the involvement of drug penetration in tissues. Evidence for the existence of tissues where new cell infections occur during therapy,(Frenkel et al., 2003; Kupfer et al., 2007; Lorenzo-Redondo et al., 2016) and suggesting high variability in drug penetration in terms of specific lymphoid tissue,(Fletcher et al., 2014) adds elements of controversy to the ongoing debate. In the current work, we focus on the first two phases of viral decay without the specific differentiation between the sub-phases 1a and 1b that have been recently revealed.(Cardozo et al., 2017; Gilmore, Kelleher, Cooper, & Murray, 2013) With the objective to shed additional light on the existence of multi-phasic viral load decay, we here compared predictions from a mathematical model to short-term viral load data as well as to long-term clinical outcomes.

We overlaid our model predictions on short-term viral load data of 6 patients. *De visu*, the agreement between model predictions and patient observations is substantial (Figure 3.2), and compares with the results obtained through a model assuming short- and long-lived virion-producing cells.(Perelson et al., 1997) This suggests both models are plausible mathematical explanations for the short-term viral load decay after treatment initiation. In fact, the results indicate that in order to reproduce the short-term viral loads, the infection activity in the less drug-exposed compartment has to contribute from 4% to 18% of the total plasma viral load before treatment initiation. Further, the drugs in this compartment would need to prevent 41% to 80% of all new cell infections for the model to accurately fit the data. This suggests the associated tissues need to harbour a substantial amount of the total infection activity before treatment initiation (infection hot spot). Moreover, the level of drug exposure would likely be much smaller in this compartment than what is prohibitive in infection assays.(L. Shen et al., 2008)

The existence of a drug-limited compartment allowing new short-lived CD4 infections could affect long-term drug resistance. To study this, we based ourselves on a model able to reproduce the observed risk of virologic failure as a function of the adherence level of a

population of patients initiating treatment with efavirenz, tenofovir DF and emtricitabine.(Sanche et al., 2017) This model accurately predicted the risk of virologic failure observed in patients after long-term therapy (1 to 1.5 years) under various regimen and drug adherence.(Sanche et al., 2017) Herein, this model was modified to consider two compartments differentially exposed to drugs. The principles governing viral kinetics in the compartments were kept: both models used an *in vivo* extrapolation of drug susceptibility assays along with empirical data on *in vivo* antiviral activity.(Sanche et al., 2017) Further, the extent of the immune response or whether resistant virus is present when treatment starts, which may also impact viral growth, are equivalently considered in both models.(Sanche et al., 2017)

We used the modified model to simulate the viral activity within a population of patients initiating treatment with efavirenz, tenofovir DF and emtricitabine. Here, we first deduced that even if 100% infections are prevented in well-irrigated tissues, 60% of wild-type virus infections have to be inhibited in a second compartment to explain the first two phases of decays observed in patients (Figure 3.3, $f_{u,2} = 40\%$). This value is consistent with what was obtained in Figure 3.2. Second, we determined the drug concentrations within this compartment corresponding to such level of inhibition. Finally, we used these concentrations to simulate long-term viral suppression in small groups of patients. To palliate the uncertainty in the combination of drug concentrations leading to 60% inhibition of cell infection events, we undertook 200 simulations, one per selected trio of drug penetration parameters k_l^{EFV} , k_l^{TFV} and k_l^{FTC} . We used a simple random sample, using the log scale of k_l values, to obtain the sampled values in red in Figure 3.4. The chosen points are well spread across the domain of allowable values, indicating a good representation of the entire space. Results in Figure 3.6 suggest a minimum of 22% of patients would experience virologic failure at the end of 48 weeks of treatment under the evaluated hypothesis (lower bound of 95% CI of 17.4%). The predicted risk is not consistent with the 2% risk of virologic failure reported in a clinical study we used to validate our model (95% confidence interval of empirical virologic is [0.5%;3.8%]).(Gallant et al., 2006) Sensitivity analyses also showed this inconsistency remains once uncertainty in parameter values has been accounted for. Even if the median slope of second phase viral decay was 2.5 times its empirical estimate ($f_{u,2} \sim 20\%$), it was still impossible to achieve overlap between predicted and empirical confidence intervals of the risk of failure.

The mechanistic nature of our model enables evaluating the likelihood of the hypothesis of new short-lived CD4 cell infections in a drug-limited compartment explaining the second phase of viral decay.(Craig et al., 2016) The model consists in the physiological, pharmacological and viral elements interacting to dictate the evolution of the viral population within its host. While there is an increasing body of evidence suggesting the existence of drug-limited tissues impacting the viral load dynamics, such as lymph nodes,(Fletcher et al., 2014) the results of the current study suggest persistent short-lived CD4 infections in a drug-limited compartment cannot –by itself– explain the second phase of viral decay. Here, all parameter values, with the exception of the ones that are linked to the hypothesis in question (φ_2 and $f_{u,2}$), were determined *a priori* from reported experiments. Plasma drug concentrations were simulated using models derived from population pharmacokinetic studies.(Baheti et al., 2011; Sanchez et al., 2011; Valade et al., 2014) The relationship between drug concentrations and drug effect on each separated strain was based on an *in vitro* to *in vivo* extrapolation of drug efficacy.(Acosta et al., 2012; Fletcher et al., 2014; Jilek et al., 2012; Rosenbloom et al., 2012; L. Shen et al., 2008) The within-host growth rate of viral loads was also based on reported distributions.(Ioannidis et al., 2000; Ribeiro et al., 2010) Finally, the probability of emergence of specific mutations as well as the fitness of each strain were well described in the literature.(Rosenbloom et al., 2012) However, it was not possible to find parameter values consistent with the hypothesis and allowing coherent predictions of long-term virologic responses.

The model we formulated can suggest other explanations for the second phase of viral load decay. Because mass action law dictates the dynamics of the infection in the model,(Érdi & Tóth, 1989) how quickly the number of infected cells grows or declines over time is dependent on the availability of target cells. A lower density of target cells in a compartment means that virions have a higher chance to be cleared before they can infect a cell, which leads to a decline in the infection rate while the decay rate of infected cells remains the same. When the density of target cells is too low, the model predicts a constant decline of virion production over time with a rate dependent on density. In fact, we reproduced two phases of viral decay with our model when new cell infections are completely prevented by the drug in one compartment, while in another compartment unexposed to drugs, virion production slowly

declines due to the loss of target cells (data not shown). Because of the absence of drug exposure in the second compartment, no long-term selection of resistant mutants occurs.

In reality, a reduction in the number of target cells could be initiated by immune contraction.(McKinstry et al., 2010) It is hypothesized that this phenomenon, which implies a rapid decay of effector T cells, can be triggered by a reduction in antigen availability.(McKinstry et al., 2010) We hypothesize that one explanation for the first 2 phases of decay could be that the sharp decrease in the density of virions in one compartment that is highly exposed to drugs triggers a contraction phase. This contraction phase, associated with systemic cytokine signaling,(McKinstry et al., 2010) could promote the death or inactivation of CD4 cells in another compartment unexposed to drugs.

In the current model, the effect of the immune response is represented via a single parameter that summarizes the growth rate of viral infection in absence of drug action (reproduction number).(Sanche et al., 2017) This is a consequence of a lack of quantitative knowledge about changes in patients' immunity from disease onset. In particular, little is known about the evolution of the density of activated CD4 cells in infected tissues. Although the fraction of cells expressing CD38+HLA-DR+ among CD4 cells is significantly reduced in the blood of patients after treatment initiation,(Sun et al., 2017) more precise estimates of local changes in various tissues could lead to a better understanding of the interaction between immunity and the virus.

In summary, the specific cause for the phases of viral decay is still uncertain. On the one hand, the hypothesis advancing the existence of infected cells exhibiting a decay dynamic that is different from short-lived CD4 cells has literature support.(Cardozo et al., 2017; Perelson et al., 1997) On the other hand, the hypothesis stipulating that such cells have a negligible impact on virion production and completely attributing the slower decay of virion production to new short-lived cell infections in a drug limited compartment finds long term inconsistencies. Alternative models that would measure the contribution of both poor drug exposure in tissues and the immune response remain to be studied. A better understanding of the dynamics of activated immune cells before and after treatment is initiated could shed light on the matter.

3.6. Methods

3.6.1. The model and simulations

An overview of the model is given below. It includes the description of the algorithm and main parameters, as well as the discussion of the source of parameter values. The model is adapted from Sanche et al. (Sanche et al., 2017), the main difference being the consideration of two compartments with different exposures to the drugs.

The model is implemented using an algorithm which iteratively computes the number of events involving active CD4 cells during small time intervals: i) the number of virion-producing cells that die out, ii) the number of new cell infections, and iii) the fraction of these infections that involve a newly mutated virus. Step i) is computed using a constant cell death rate (d_y). Step ii) is computed from reproduction numbers R_i^j , where i and j indicate the strain of the virus and the compartment, respectively. Details on reproduction numbers are given below. The fraction of *de novo* mutation is based on probabilities of SNP mutations estimated from empirical data and reported elsewhere. (Abram et al., 2010) Total infection activity is translated in plasma virions, by assuming one virion-producing cell supplies one plasma virion, a relationship that is consistent with independent data. (De Boer et al., 2010; Hockett et al., 1999)

The model assumes compartments are isolated from each other in terms of viral infection, i.e. a virus produced by one cell can only infect cells within the same compartment. This has been confirmed with experimental data: the infection of new cells is considered an essentially local phenomenon and the genetic makeup of the viral populations suggests a high degree of compartmentalization. (A. T. Haase, 1999; Lorenzo-Redondo et al., 2016) The modeled viral dynamics is very similar in each compartment, the main difference being the level of drug exposure influencing the reproduction numbers.

Reproduction numbers R_i^j are expressed as $R_0(1-s_i)f_{u,j}^i$, where R_0 is the mean number of CD4 cells becoming infected by viruses produced by a single infected cell when susceptible cells are abundant and when no drug is present, s_i is a fitness cost for strain i , and $f_{u,j}^i$ is the fraction of CD4 infection events unaffected by the drugs for strain i in compartment j . In this study, $f_{u,j}^i$ was either estimated for wild-type virus from viral load data (index i is omitted in

this case, since only wild-type virus was considered), or computed from drug concentrations. In the latter case, plasma drug concentrations are first simulated from reported pharmacokinetic models. Since the concentrations vary over time t , so do the $f_{u,j}^i$. In the particular case where only one drug is used, $f_{u,j}^i$ takes a relatively simple form :

$$f_{u,j}^i(t) = \frac{1}{1 + \left(\frac{C_p(t)}{\rho_i IC_{50}^0 k_p k_{l,j}} \right)^{m(1+\sigma_i)}} \quad (3.1)$$

where $C_p(t)$ is the plasma drug concentration at time t , k_p is the coefficient adjusting for plasma protein binding, $k_{l,j}$ is the coefficient adjusting for the degree of drug penetration in the compartment j , respectively, IC_{50}^0 is the concentration inhibiting 50% of infection events by the wild-type strain *in vitro* in a medium devoid of plasma proteins, m is the Hill coefficient, and finally ρ_i and σ_i are two factors adjusting the values of IC_{50}^0 and m , respectively, for the resistant viral strain i . For the wild-type strain, $\rho_i = 1$ and $\sigma_i = 0$. It should be noted that to estimate the impact of concomitant drug use, $f_{u,j}^i(t)$ no longer takes a simple analytical form (see Jilek et al. and Sanche et al.).(Jilek et al., 2012; Sanche et al., 2017)

Parameters ϕ_j are used in all simulations to limit the number of cells each compartment may contain. The overall number of cells is itself dictated by parameter λ , which is an entry rate of uninfected activated CD4 cells.

Many parameter values are patient-specific. The death rate d_y is randomly selected for each patient from a distribution of values based on empirical measures of the first phase decay.(Markowitz et al., 2003) The distribution has a median value corresponding to a half-life of about 0.7 day. The same applies for λ (directly linked to the distribution of viral set points)(Rosenbloom et al., 2012) and R_0 (based on the growth rate of viral loads during rebounds).(Ioannidis et al., 2000; Ribeiro et al., 2010) Inter-individual variability for pharmacokinetic parameters is based on the reported values of population pharmacokinetics models.(Acosta et al., 2012; Baheti et al., 2011; Blum et al., 2007; Sanchez et al., 2011) All other model parameters have *a priori* assigned values reported in Sanche et al.(Sanche et al., 2017), with the exception of ϕ_j and $f_{u,j}$.

3.6.2. Viral load data fitting

Using the developed mechanistic model, data fitting refers to finding model parameters ϕ_2 and $f_{u,2}$ for which viral load predictions are closest to empirical observations. The process was carried out in two steps to reduce numerical calculations. In the first step, we generated combinations of ϕ_2 and $f_{u,2}$ taking values between 5% and 95% in 5% increments (361 combinations). For each combination, the viral loads of 20 virtual patients were simulated. Values of patient-specific parameters d_y , λ and R_0 were randomly selected from their corresponding patient population distributions. Overall, the viral loads of a total of 7,220 virtual patients were simulated. For each set of real viral load data, we computed the sum of the squared difference between predictions and observations, using a logarithm scale for viral loads, in order to identify the virtual patient with the closest viral load predictions. This numerical experiment was repeated for values of ϕ_2 and $f_{u,2}$ around those retained for the identified virtual patient, noted ϕ'_2 and $f'_{u,2}$. More precisely, all discrete percentage values of ϕ_2 in $\phi'_2 \pm 4\%$ and $f_{u,2}$ in $f'_{u,2} \pm 4\%$ were investigated (81 combinations, 1620 virtual patients). We report the virtual patient's viral load being the closest to the observations, using the same criteria as above.

3.6.3. Identification of k_l^{EFV} , k_l^{TFV} and k_l^{FTC} in the less drug-exposed compartment

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

We used Equation 3.1 to identify the minimum and maximum values of k_l for each drug. The rationale for this procedure was the following. First, we found $k_{l,min}$ such that if $k_l < k_{l,min}$ for each of the three considered drugs, then the value of $f_{u,2}$ for the combined drug effect will necessarily be beneath $\widehat{f_{u,2}} - 5\%$. In other words, the use of a single drug would be more than sufficient to inhibit 1- $f_{u,2}$ infection events in the less exposed compartment. Further, we found

$k_{l,max}$ for which if $k_l > k_{l,max}$ implies almost no effect from the drug (too small drug concentration in the tissue to have an effect). More precisely, we computed the ratios $C_p(t)/k_l$ that correspond to $f_{u,2} = \widehat{f_{u,2}} - 10\%$ and $f_{u,2} = 99\%$, where the values for m , k_p and IC_{50}^0 are reported elsewhere.(Acosta et al., 2012; Sanche et al., 2017; L. Shen et al., 2008) We then derived $k_{l,min}$ and $k_{l,max}$ using average steady-state plasma concentrations $C_p(t)$ (2.54mg/L for efavirenz, 0.183mg/L for tenofovir and 0.55mg/L for emtricitabine).(Baheti et al., 2011; Sanchez et al., 2011; Valade et al., 2014) Finally, we discretized the range of values for k_l for each of the three drugs using 200 points between $k_{l,min}$ and $k_{l,max}$, leading to 8,000,000, or 200^3 determined trio values. For each of these trios, we could compute the expected value of $f_{u,2}$ for the drug combination, by using a procedure further accounting for the degree of independence of drug effects.(Jilek et al., 2012; Sanche et al., 2017) We retained the subset of $(k_l^{EFV}, k_l^{TFV}, k_l^{FTC})$ for which the expected value $f_{u,2} \in [\widehat{f_{u,2}} - 5\%; \widehat{f_{u,2}} + 5\%]$.

3.6.4. Statistical analysis

Parameter estimation was performed to fit our model to short term empirical data. As explained in the section entitled Data fitting for viral load data, this procedure was based on the least square difference between model predictions and empirical viral load observations in \log_{10} scale. The goodness of fit of the model was evaluated visually. We computed the 95% confidence intervals for the risk of long-term virologic using the exact method, both for the empirical and the smallest predicted risks.(Agresti, 2013) We also computed the p-value associated to a null hypothesis of equal risk of failure in both the empirical and the model-based samples and assuming the difference of proportion is normally distributed (central limit theorem).(Agresti, 2013)

Chapitre 4. How is baseline viremia linked to post-treatment control? Analysis of treatment interruption data based on a mechanistic model

Authors and affiliation: Sanche, S¹ (other authors to be determined)

¹Faculté de Pharmacie, Université de Montréal, Montréal, QC, Canada H3T 1J4

²Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA, 02115

Article en révision à l'interne

4.1. Préambule

En attente de mesures expérimentales du taux de pénétration du dolutégravir dans les tissus, nous nous sommes penchés sur la problématique d'allègement du fardeau posologique des patients. En particulier, certains cliniciens envisagent de réduire la fréquence de prise médicamenteuse de certains de leurs patients qui ont été sous suppression virologique durant de longues périodes. Une telle modification au régime posologique serait similaire à une adhésion incomplète des patients, un facteur qui est associé au développement de souches résistantes du virus.(Cadosch, Bonhoeffer, & Kouyos, 2012) Par contre, l'existence de contrôleurs post-traitement permet d'envisager un allègement du fardeau posologique chez certains patients (revoir « Adhésion et fardeau posologiques »). En effet, la tâche de maintien d'une faible charge virologique pourrait être partagée entre : i) le processus responsable du contrôle post-traitement et ii) l'action de la médication. Or, avant de pouvoir considérer un allègement posologique, il est nécessaire d'évaluer les risques de développement de résistance associés; des souches résistantes pourraient se développer si de nouvelles infections de cellules se produisent en même temps que des concentrations inappropriées de médicaments.

Afin de mieux comprendre le phénomène, il est utile d'interpréter les associations entre les caractéristiques des patients et leur capacité à maintenir de faibles niveaux de virémie à la suite d'un arrêt de traitement. Or, ces associations doivent être interprétées avec prudence,

puisque'il existe une discordance entre la durée de maintien de la virémie à de faibles niveaux et la durée de contrôle réel de la virémie. Ceci a mené au sujet du présent chapitre. Un des objectifs de l'étude de ce chapitre est d'interpréter l'association statistiquement significative entre la virémie résiduelle des patients et la mesure empirique de la suppression d'infection. Pour ce faire, on utilise un modèle de croissance de la virémie en conjonction avec des données empiriques de temps entre la cessation des traitements et l'observation de charges virologiques élevées. Par la même occasion, la durée du contrôle réel de la virémie et sa variabilité entre les patients sont estimées.

4.2. Abstract and keywords

In a small proportion of HIV+ patients, interrupting long-term therapy does not lead to viremia rebound. While several studies have been conducted to shed light on the matter, the underlying causes for the ability of these patients to control their viremia without requiring medication are still misunderstood. Results from these studies allowed uncovering many associations between patient characteristics and the duration of the period of maintained low-level viremia after therapy (low-level viremia period). However, care should be exerted before interpreting these associations in terms of actual post-treatment control. Indeed, if such control is defined as the absence of viral growth after therapy is ceased, then the period being observed empirically can also include a period in which viremia is increasing but still remain below the specified threshold (rebound period). Hence, associations between patient characteristics (including biomarkers) and the duration of the low-level viremia period could be attributed to the extent of the period of control, the extent of the period of rebound, or both. In the current study, we developed an approach that allows to quantify the duration of actual post-treatment control, its variability between patients and how much this variability can be explained by patient characteristics, using a mathematical model of viremia after therapy and data from treatment interruption trials. We show that the duration of the rebound period can entirely explain a statistically significant association between baseline viremia and the duration of the period of low-level viremia. This suggests that the rebound triggering event (RTE), hypothetically being clonal expansion, occurs at a rate that is not very sensitive to the quantity of viral RNA in the plasma. Our analysis also suggests that the expected rate of RTE differs

substantially between patients, extending from about once a day to once every several weeks. How much of this variability can be explained by other patient characteristics (including biomarkers) could be assessed using previously collected data and the approach described herein.

Key words: antiretrovirals, HIV-1, post-treatment control, residual viremia, remission induction, virus latency, adaptive immunity, mathematical modeling

4.3. Introduction

Over the years, significant progress was made in treating people with HIV. Modern antiretroviral therapy has proven very efficient in suppressing patients' viremia and in limiting the negative impact of HIV. Unfortunately, no definite cure has been found. As a result, most HIV+ patients must take their medication on a regular basis to allow continuous suppression of their viremia. Such chronic drug intake entails disadvantages that can directly impact the lives of those living with HIV, such as suffering from potentially important side effects, a financial burden, or social stigma.(Achappa et al., 2013; Al-Dakkak et al., 2013) Constant drug exposure also involves the risk of selecting drug resistant strains of the virus, which can have serious consequences on the long-term health of patients.(Hogg et al., 2006; Palella et al., 2009; Tang & Shafer, 2012)

It is however interesting to notice that some patients are able to maintain low-level viremia for a significant period of time without medication.(Deeks & Walker, 2007; Okulicz & Lambotte, 2011; B. D. Walker & Yu, 2013) Post-treatment controllers are patients for whom this viremia control is observed after ceasing therapy.(Hocqueloux et al., 2010; Saez-Cirion et al., 2013) This situation raises hopes that, if the right conditions are achieved, patients could temporarily or even permanently interrupt treatment without significantly impacting their health. Some of these conditions may be unveiled by identifying characteristics (including biomarkers) that are more common in those patients able to maintain low-level viremia after therapy.

Several treatment interruption studies have been conducted with this purpose in mind, and interesting associations have been reported. Early post-treatment controllers are those who

had initiated treatment soon after being infected, specifically during primary infection, a situation that was only observed much later in patients treated during chronic infection.(Assoumou et al., 2015; Calin et al., 2016; Hocqueloux et al., 2010) Many biomarkers have also been associated with the duration of the period in which patients maintained low-level viremia after treatment interruption (low-level viremia period). For example, those having a smaller reservoir of cell-associated DNA were more likely to maintain low-level viremia after therapy.(Saez-Cirion et al., 2013; Williams et al., 2014) In terms of viral replication biomarkers, a longer period of low-level viremia was associated with lower viremia set points(Treasure et al., 2016), baseline viremia(Li et al., 2016; Martin et al., 2017), baseline cell-associated RNA and, in macaques, a lower area-under-the-curve of the viral load since infection.(Whitney et al., 2014) Further, T-cell biomarkers such as a lower frequency of exhaustion markers,(Hurst et al., 2015) a higher CD4 count(Martin et al., 2017) and a higher CD4:CD8 ratio(Martin et al., 2017), all predicted longer periods of low-level viremia after therapy.

Before these associations can be interpreted, it is important to understand the nature of these studies' endpoint, which is often defined as the duration of the period where patients' viremia is maintained below a specific threshold. These studies usually involve patient characteristics and biomarker measurements obtained a near treatment interruption (baseline measures). Then, viremia measurements are obtained at various time points following interruption. Data suggests the existence of two distinct phases during this period: one phase during which viremia is being maintained at its baseline value, i.e. actual post-treatment control (hereafter called period of control), followed by a second phase where viremia grows exponentially until reaching the low-level threshold (the period of rebound).(Bar et al., 2016; Colby et al., 2018) The first phase (period of control) is an indication that some patients can naturally limit the number of new cell infections after therapy. Indeed, if the extent of other processes leading to virion production remains constant after treatment cessation (e.g. activation of latently infected cells), so should remain the number of newly infected cells. The second phase (period of rebound), on the contrary, implies that virion production is increasing and will reach the low-level threshold within some time. Improving patients' ability to extend the period of control would be beneficial, as it could alleviate or even negate their need for drug intake.

This goal arouses a great scientific interest, showcased by the numerous publications on the subject.

However, the second phase (period of rebound) implies that the associations unveiled in the above-mentioned studies cannot be directly interpreted in terms of actual viremia control. Difficulties arise when the involved patient biomarkers are also associated with the extent of the second phase. For example, we mentioned earlier one association involving baseline viremia and the period of low-level viremia (which includes both phases).(Li et al., 2016) However, it is also reasonable to expect that the time it takes for viremia to reach the threshold would be longer for those with lower baseline viremia. The extent by which this association with the duration of the second phase affects the strength of the association between baseline viremia and the time of low-level viremia is uncertain.

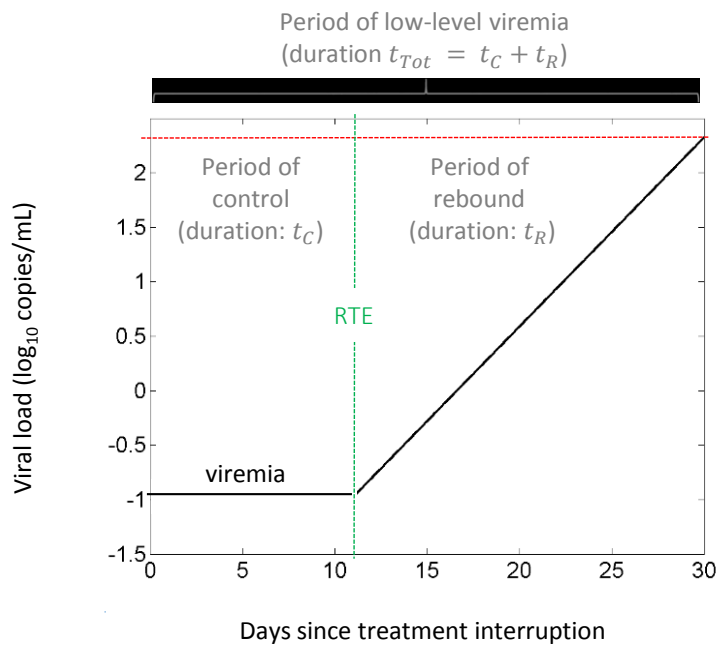
The objective of the current study is to shed more light on this last issue. Specifically, we will first evaluate the strength of the association between baseline viremia and the duration of the actual post-treatment control period (first objective). For this purpose, we mathematically modelled viremia after treatment cessation in consideration of the two phases mentioned above (see Figure 4.1). Model parameters were estimated using pooled data from studies on patients having ceased prolonged antiretroviral therapy.(Li et al., 2016) We used data from those patients who had started treatment during the chronic infection phase (asymptomatic phase following acute infection). This choice was driven by the small number of participants treated during acute infection and by studies suggesting the latter patients may substantially differ in terms of post-treatment control duration. Then, the mathematical model allowed estimating expected durations of post-treatment control and their variability between patients (second objective). These findings may prove useful when designing future studies that involve treatment interruptions, such as those further investigating post-treatment control or exploring alternative dosing regimen after long-term therapy.

4.4. Methods

In this section, the empirical data used to analyze the link between baseline viremia and the duration of the period of control is first described. We will then elaborate on the

mathematical model of the duration of this period of post-treatment control. The section closes with details on the data fitting procedure.

Figure 4.1 – Modeling the distinct viremia phases during the delay between treatment cessation and virologic failure observation (>200 copy/mL). The black line represents viremia. The dotted green line represents the occurrence of a rebound triggering event (RTE) separating the first phase (left) from the second phase (right). The dotted red line represents the virologic failure threshold. Baseline viremia was set at around 0.1 copy/mL in this example.



4.4.1. Empirical data

We used data collected in six clinical trials and reported in a study that aimed to analyze viral rebounds following treatment interruption. (Li et al., 2016) The pooled sample comprised 155 patients who had been on suppressive ART for a median of 5.1 years (3.2, 6.5) before ceasing therapy. Among these patients, we selected those who had initiated treatment during the chronic infection phase and who had all required measurements for the present study (n=82). Plasma HIV-1 RNA had been measured at multiple time points after treatment cessation (week 1, 2, 3, 4, 6, 8, 10, 12 and 16 for 29 patients, and week 4, 8, 12 and 16 for 53 patients). Baseline

viremia had been defined as the most recent measure preceding the date of ART discontinuation, using single copy assays when viremia was below the threshold from usual assays.(Palmer et al., 2003) Because of the long period of viral suppression, baseline viremia was considered equivalent to residual viremia. The low-level viremia threshold that was used in our analysis was 200 copies/mL, considered as a virologic failure. The distribution of baseline viremia in the sample of patients is shown in Table 4.1. Table 4.2 shows the number of participants by time of observed virologic failure and by baseline viremia, along with percentages.

Table 4.1 – Distribution of baseline viremia in the sample of patients

Baseline viremia (v, copies/mL)	$v < 0.1$	$0.1 \leq v < 1$	$1 \leq v < 2$	$2 \leq v < 5$	$5 \leq v < 60$
Percentage	52.4%	13.4%	12.2%	9.8%	11.0%

Table 4.2 – Number of participants by duration of the period of low-level viremia and by baseline viremia. (Li et al., 2016) Percentages were computed over patients having the same level of baseline viremia. A threshold of 200 copies per mL of plasma was used

Baseline viremia	Duration of low-level viremia period [in weeks after treatment interruption]			
	[0;4]]4;8]]8;	All
< 1 cp/mL	29 (53.7%)	17 (31.5%)	8 (14.8%)	54
> 1 cp/mL	21 (75.0%)	6 (21.4%)	1 (3.6%)	28
All	50	23	9	

4.4.2. The model

For all patients, the viremia profile from treatment cessation to virologic failure is separated into two phases (see Figure 4.1): one phase of duration t_C , in which viremia is maintained at baseline levels, and another one of duration t_R in which viremia grows exponentially from baseline levels to the virologic failure threshold. The union of the two phases covers the entire period of maintained low-level viremia. The duration of this entire period t_{TOT} , is hence the sum of the two time durations ($t_{TOT} := t_C + t_R$). This formulation was also adopted in a previous study.(Pinkevych et al., 2015) Duration t_C was modeled as the time until the occurrence of a random event, the rebound triggering event (RTE). The RTE is depicted by the dotted green line on Figure 4.1. How quickly the RTE occurs is the main subject of interest.

We expressed differences in time of occurrence of a RTE between patients in terms of two variability components. First, we allowed each patient to have its own propensity for viremia control (interindividual variability). Further, patient-specific propensity only dictated an expected time of occurrence of RTE, the actual time of RTE varying stochastically around this expected value (intraindividual variability). Using mathematical terms, this is equivalent to considering the actual occurrence time $t_{C,i}$ of a RTE as a realization of a random variable $T_{C,i}$. Specifically, we used Poisson random variables, deemed adequate for rare events, with parameter λ being the expected occurrence rate of RTE in number of events per unit time. (Cameron & Trivedi, 2013) The parameter λ represents patients' propensity for viremia control. Since this propensity varies between individuals, so do λ values. To represent parameter values specific to the i^{th} patient in the sample, parameters were additionally indexed with the symbol i , in the following. Hence, we used the notation λ_i to represent the expected rate of RTE for patient i while $t_{C,i}$ was used to represent this patient's duration of the period of control after treatment interruption. It immediately follows that $T_{C,i}$ is an exponential random variable ($T_{C,i} \sim \text{Exp}(\lambda_i^{-1})$). We allowed some of the variability in the rate of RTE (λ_i) to be explained by patients' baseline viremia. Specifically, we used the following relationship to describe λ_i :

$$\log(\lambda_i) := \alpha_i + \beta f(v_i), \quad (4.1)$$

where $f(v_i)$ is a non-parameterized function of baseline viremia v_i , and α_i and β are parameters. The logarithm scale was used to avoid negative values for λ_i and is often used to describe interindividual variability in biological measures. The part of interindividual variability in λ that cannot be explained by baseline viremia is expressed by allowing patient-specific α_i values. In the population, we assumed α_i to be normally distributed ($\alpha_i \sim \text{N}(\mu_\alpha, \sigma_\alpha^2)$). Equivalently, λ_i is lognormally distributed with three free parameters: μ_α , σ_α and β ($\lambda_i \text{ iid } \sim \text{logN}(\mu_\alpha + \beta f(v_i), \sigma_\alpha^2)$).

This formulation is sufficiently flexible to describe different types of associations between patient characteristics and their propensity to control viremia after therapy. An absence of interindividual variability is represented by one constant parameter $\lambda = \mu_\alpha$ attributed to every patient ($\sigma_\alpha = 0$ and $\beta = 0$), i.e. differences in observed times of RTE between patients

are due to chance alone, not baseline patient characteristics. A parameter σ_α being further different from 0 indicates that patient characteristics other than baseline viremia explain interindividual variability. If baseline viremia is linked to the propensity to control viremia, then β is expected to take a non-zero value for one or more functions $f(v)$. We explored four formulations for the function $f(v)$: i) $f(v) = 1$, ii) $f(v) = \ln(1 + v)$, iii) $f(v) = \sqrt{v}$, and iv) $f(v) = v$. Note that function $f(v) = 1$ is the same as considering that baseline viremia is not affecting λ and that $\beta = 0$.

4.4.3. Data fitting

It was necessary to assess whether the above formulation allows an adequate description of the empirical data. This was performed *de visu* by comparing empirical and model-based survival profiles, where survival was defined as having viremia below the virologic failure threshold (<200 copies/mL). Empirical survival was computed using a non-parametric maximum likelihood estimator (Turnbull) implemented in the *icenReg* package of R software.(Anderson-Bergman, 2016) Four model-based survival profiles were obtained, one for each function $f(v)$. For those models deemed acceptable, parameter estimates maximizing the likelihood of the observations were further inspected to characterize the link between actual post-treatment control and baseline viremia. These parameters were also used to quantify the average time of RTE occurrence (equivalent to the average time of control t_C) and its interindividual variability.

The three free parameters μ_α , σ_α and β in the model were estimated by maximizing a likelihood function for the data. The database did not contain exact $t_{TOT,i}$ values for the sample; it rather reports an open or a closed time interval during which the viremia of patient i has reached the 200 copies/mL threshold. In technical terms, dependent variables $t_{TOT,i}$ are left-right- or interval-censored.(E. T. Lee & Wang, 2003) In such case, the likelihood function LF can be expressed by:

$$LF \propto \prod_{i=1}^{n=82} [P(T_{TOT,i} < t_{i,sup} | v_i, \mu_\alpha, \sigma_\alpha, \beta) - P(T_{TOT,i} < t_{i,inf} | v_i, \mu_\alpha, \sigma_\alpha, \beta)] \quad (4.2)$$

where $]t_{i,inf}; t_{i,sup}]$ or $]t_{i,inf}; t_{i,sup} = \infty[$ describes the time interval during which the threshold was attained for patient i (i.e. $t_{TOT,i}$ is inside that interval)(E. T. Lee & Wang, 2003)

Maximizing LF requires the function to take unique values when all free parameters μ_α , σ_α and β , and patient baseline viremia v_i are known. To obtain such formulation for LF , we can first assume that the viremia growth rate after the occurrence of the RTE is also known (hereafter expressed as ω_i , for patient i). Recalling that viremia grows exponentially following rebound in the model, the time duration of the period of rebound for patient i can then be expressed by:

$$t_{R,i} = \frac{\log(200) - \log(v_i)}{\omega_i}, \quad (4.3)$$

The time duration of actual viremia control $T_{C,i}$, as described in the section above (The model), is an exponential random variable with probabilities defined by:

$$P(T_{C,i} \leq t_c | v_i, \omega_i) = 1 - e^{-\lambda_i t_c}. \quad (4.4)$$

Based on Equation 4.1, Equation 4.4 is dependent on α_i ; it is not uniquely determined when β , μ_α and σ_α are known. Using probability calculus, we can integrate Equation 4.4 over α_i to obtain an expression that only depends on the parameters of interest, when v_i and ω_i are known:

$$P(T_{C,i} \leq t_c | v_i, \omega_i) = \int_{\alpha=0}^{\infty} (1 - e^{-\lambda_i(\alpha)t_c}) g_A(\alpha) d\alpha \quad (4.5)$$

where $g_A(\alpha)$ is the lognormal density function expressed in terms of μ_α and σ_α . It follows that $T_{TOT,i}$ is a random variable having cumulative probability distribution:

$$\begin{aligned}
& P(T_{TOT,i} \leq t_{TOT} | v_i, \omega_i) \\
& = \begin{cases} 0 & \text{if } t_{TOT} \leq t_{R,i} \\ \int_{\alpha=0}^{\infty} [1 - e^{-\lambda_i(\alpha)(t_{TOT}-t_{R,i})}] g_A(\alpha) d\alpha & \text{otherwise.} \end{cases} \quad (4.6)
\end{aligned}$$

The upper case of the equation represents situations where baseline viremia and the growth rate are so small that even if immediate rebound follows treatment cessation, viremia could not have reached the threshold by time t_{TOT} .

Computing model-based probabilities described by Equation 4.2, from Equations 4.3 and 4.6 encounters challenges originating from the fact that v_i and ω_i are unknown for many or all patients in the sample, respectively (some of the patients had baseline viremia below the single-copy assay threshold (0.1 copy/mL)). We accounted for uncertainty in v_i and ω_i by integrating Equation 4.6 over possible values using density functions for these entities, similarly to what was done in 4.5 to discard this equation's dependency on α_i . The density function for v_i was inferred from another dataset, which consisted of baseline viremia measurements collected from patients having been on suppressive therapy for 4.8 years on average (3.1, 6.4). (Zheng et al., 2013) We used a data fitting procedure assuming a lognormal distribution in the population. For detail on the data fitting procedure and on goodness-of-fit, please refer to the Supplementary Material section. The derived viremia density function for patients having values below 0.1 copy/mL is hereon referred to as $g_V(v)$. We used a similar procedure to infer the viral growth rate ω density function, as detailed in the Supplementary Material section. The data used for this procedure is provided in Ioannidis et al. (Ioannidis et al., 2000) The derived viral growth rate density function is hereon referred to as $g_{\Omega}(\omega)$.

Using these density functions and probability calculus, we first obtain:

$$\begin{aligned}
& P(T_{TOT,i} \leq t_{TOT} | V_i \leq 0.1) \\
& = \int_{\omega=0}^{\infty} \int_{v=0}^{0.1} P(T_{TOT,i} \leq t_{TOT} | v_i = v, \omega_i = \omega) g_V(v) g_{\Omega}(\omega) dv d\omega, \quad (4.7)
\end{aligned}$$

which can be used to compute components of the likelihood function associated with patients having baseline viremia below the 0.1 copy/mL threshold, and

$$\begin{aligned}
& P(T_{TOT,i} \leq t_{TOT} | V_i = v) \\
&= \int_{\omega=0}^{\infty} P(T_{TOT,i} \leq t_{TOT} | v_i = v, \omega_i = \omega) g_{\Omega}(\omega) d\omega
\end{aligned} \tag{4.8}$$

for the rest of the patients. Equations 4.3, 4.6, 4.7 and 4.8 allow expressing the probabilities inside Equation 4.2 (LF) only in terms of the parameters of interest.

Maximizing the likelihood function LF was achieved numerically by minimizing the negative log-likelihood and by using an unconstrained nonlinear optimization algorithm (Nelder-Mead Simplex Method) implemented in Matlab 2014a. (Lagarias, Reeds, Wright, & Wright, 1998; The MathWorks Inc, 2014) Exact numerical integration methods implied a computer intensive minimization of the likelihood function LF . Monte Carlo integration was instead performed using randomly drawn values for α , ω (and v , for those with missing baseline viremia values) from their respective distribution. (Yakowitz, Krimmel, & Szidarovszky, 1978) We opted for 200,000 draws, a number chosen to ensure 0.1% accuracy compared to Matlab's numerical integration output over a sample of integrations. The output of the algorithm is one set of $\hat{\mu}_{\alpha}$, $\hat{\sigma}_{\alpha}$ and $\hat{\beta}$ values per function $f(v)$, with β being fixed to 0 in the case where $f(v) = 1$. We used a bootstrap approach to obtain 95% confidence intervals for all parameters. (Efron & Tibshirani, 1994) Specifically, we randomly drew 500 samples of 82 patients within our sample (with replacement), and repeated the above-mentioned parameter estimation procedures to obtain empirical-based distributions for $\hat{\mu}_{\alpha}$, $\hat{\sigma}_{\alpha}$ and $\hat{\beta}$.

4.5. Results

All minimization algorithms converged with tolerance for likelihood function and parameter values of $1e-4$. On Figure 4.2, the empirical (grey line) and model-based (red line) survival curves, when survival is defined as the maintenance of viremia below 200 copies/mL, are displayed. All the model-based survival curves superimpose almost perfectly on the graph (4 models: $f(v) = 1$, $\log(1 + v)$, \sqrt{v} or v). Parameters maximum likelihood estimates along with 95% confidence intervals obtained using bootstrap samples are displayed in Table 4.3. None of the estimates for parameters β linking the rate of occurrence of a RTE and baseline viremia were significantly different from zero at 5% significance, with p-values for models

$f(v) = \log(1 + v)$, \sqrt{v} and v being $p=0.20$, 0.22 and 0.38 , respectively. Consequently, only inferences based on the model where $f(v) = 1$ are detailed below.

4.5.1. The base model ($f(v) = 1$)

The meaning of estimates shown in Table 4.3 is not straightforward. Recall that, in the model, the expected frequency of RTE is lognormally distributed in the patient population, as described by $\lambda_i \sim \log N(\mu_\alpha + \beta f(v), \sigma_\alpha^2)$. We simulated the time of occurrence of a RTE (or,

Figure 4.2 – Non-parametric empirical (grey line) and model-based (red line) survival curves for the duration of the period of low-level viremia (<200 copies/mL). The four model-based survival curves superimpose almost perfectly.

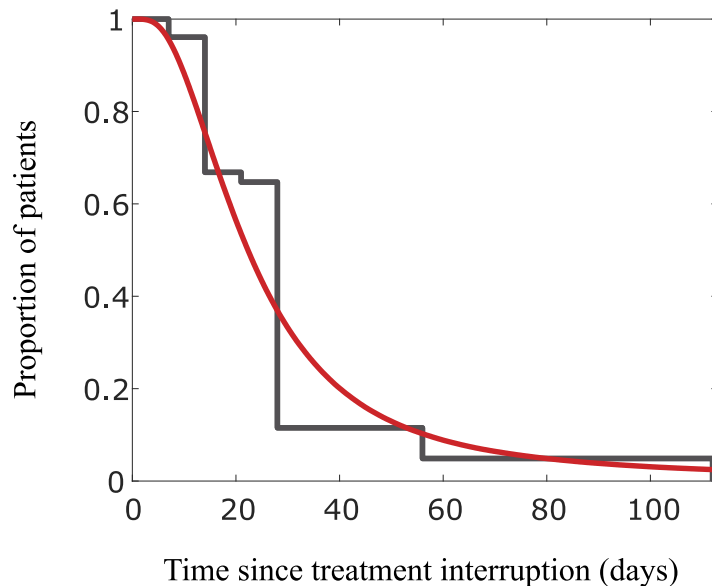


Table 4.3 – Parameter maximum likelihood estimates for the four considered models.

Model ($f(v) \dots$)	Parameters (time unit: days)		
	μ_α	σ_α	β
1	-1.89 [-2.5;-1.05]	1.23 [0.46;1.90]	NA
$\log(1 + v)$	-1.60 [-2.4;-0.69]	1.27 [\sim 0.0;1.90]	-0.34 [-0.84;0.14]
\sqrt{v}	-1.65 [-2.45;-0.61]	1.26 [0.41;1.91]	-0.21 [-0.69;0.12]
v	-1.86 [-2.52;-0.83]	1.24 [0.49;1.91]	-0.01 [-0.22;0.04]

equivalently, the time of control $t_{C,i}$ in a population of 100,000 patients having expected rates λ_i randomly drawn from $\log N(-1.89, 1.51)$. The cumulative distribution of the time of RTE over the population of patients is shown in Figure 4.3 Under the base model, more than 10% of individuals in the patient population are expected to experience a RTE within a day of treatment cessation. The median time of occurrence of RTE is 4.1 days (95% CI [1.7;7.8] days). The distribution is highly skewed, with an average time of RTE of 14.1 days (95% CI of [8.2;25.3] days). Overall, the model-predicted that 1.4 patients out of 1,000 patients would not experience a RTE during the first year following treatment cessation (95% CI of [0;9.9] patients). The interindividual variability in the expected rate of RTE is represented by parameter σ_α . This variability, after translating in expected time of RTE (λ_i^{-1}), is illustrated in Figure 4.4. As stated above, the actual time of occurrence of a RTE ($t_{C,i}$) could differ from the patient’s expected time. Hence, in Figure 4.5, we display the model-predicted cumulative probability of the time of RTE for patients represented by colored dots in Figure 4.4.

Figure 4.3 – Cumulative distribution of the time of occurrence of rebound triggering event (or $t_{C,i}$) in a patient population, as predicted by model where $f(v) = 1$. Intervals with level of confidence of 95% are displayed for 0.1, 1, 10 and 100 days after treatment cessation. The x-axis is under a \log_{10} scale.

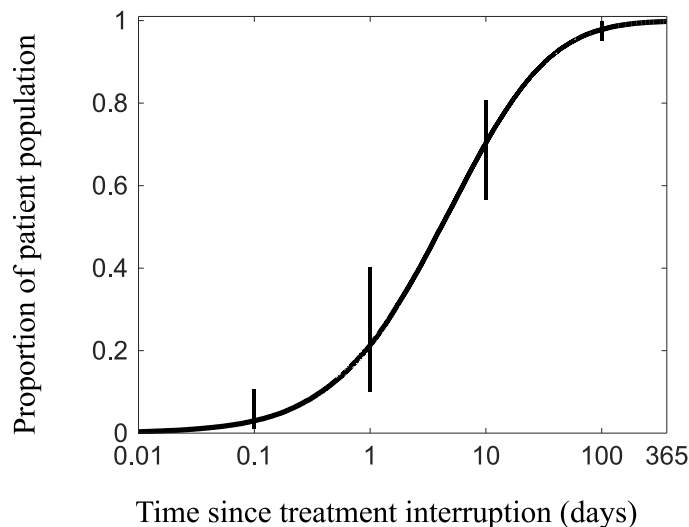


Figure 4.4 – Proportion of the patient population having an expected time of RTE (λ_i^{-1}) below the values displayed on the x-axis, as predicted by the base model ($f(v) = 1$). From left to right, the colored dots represent the 10th, 25th, 50th, 75th, 90th and 97.5th percentiles.

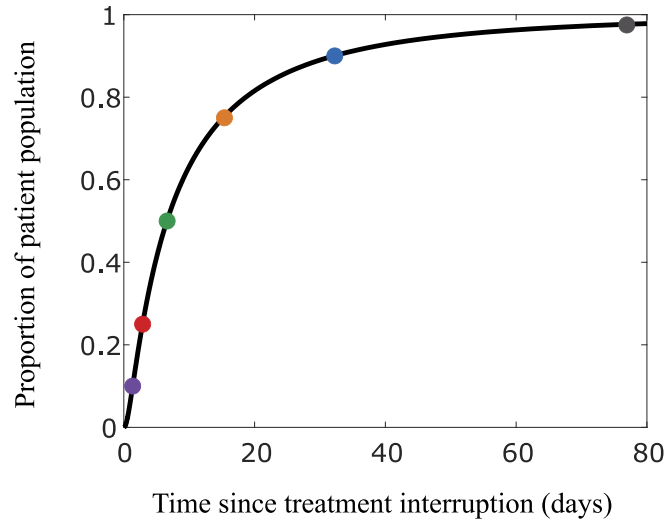
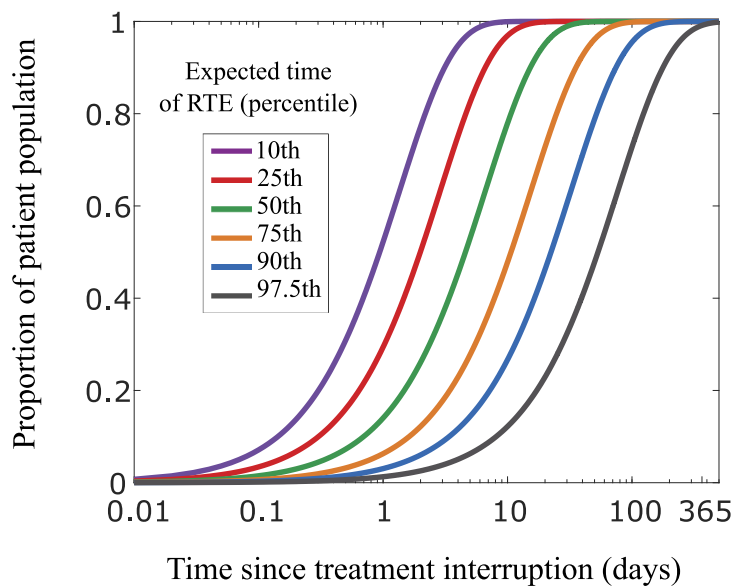


Figure 4.5 – Probability that a specific patient experiences the rebound triggering event (RTE) before the time indicated on the x axis, as predicted by the base model. Each curve represents a patient-specific parameter value λ_i^{-1} . The colors match those of Figure 4.4.



4.6. Discussion

The objective of the current study was to measure the strength of the association between baseline viremia and the duration of the period of actual control, once the period of rebound is accounted for (see Figure 4.1). A secondary objective was to provide estimates of the duration of actual post-treatment control as well as its variability among patients who initiated treatment in the chronic infection phase. For these purposes, we modeled the delay between treatment cessation and the occurrence of virologic failure (>200 copies/mL) as the sum of two periods. The first period represents the time of control or equivalently, the time from treatment cessation to the occurrence of a rebound triggering event (RTE). The second period, which we called the rebound period, represents the time for viral loads to reach 200 copies/mL from the occurrence of a RTE. The expected rate λ of RTE was formulated as dependent on baseline viremia v . To investigate the level of involvement of v in explaining the rate of RTE, four different empirical functions (expressions of v) were considered. We also considered the case where inter-individual variability is not explained by baseline viremia ($f(v) = 1$). To model the extent of the second period, we used empirical-based distributions of viral growth. (Ioannidis et al., 2000)

The four considered functions of viremia $f(v)$ led to very similar results in terms of the likelihood functions. This is reflected in Figure 4.2, where all models display very similar survival curves. The model that includes inter-individual variability in expected rate of RTE, but which variability cannot be explained by baseline viremia, gave the best model fit after accounting for the number of parameters that were estimated. Further, estimates of the effect of baseline viremia on the rebound triggering rate (λ) was not significant over all considered functions $f(v)$. This suggests that, in the current sample, the association between baseline viremia and the time of virologic failure found in Li et. al can be entirely explained by the second phase, i.e. the time it takes to viremia to grow to the virologic failure threshold from the beginning of a rebound, i.e. t_R . Inferring at population level that baseline viremia has no effect on the rate of occurrence of the RTE would require a proper power analysis. However, we argue that if such an effect exists, it is likely small. Indeed, if we assume that smaller baseline viremia should lead to less frequent RTE, the fact we obtained negative point estimates of β (indicating the inverse) suggests the actual value may be close to zero.

Using the model where the rate of RTE is independent of viremia ($f(v) = 1$), we obtained an average of $t_c = 14.1$ (8.2-25.3) days separating treatment interruption and the occurrence of a RTE in the studied population (patients treated during chronic infection for multiple years). This estimate is greater than an equivalent estimate based on different cohorts of patients, which was reported to be of around 6 days.(Pinkevych et al., 2015) The discrepancy could result from sampling variability within a highly skewed distribution of actual control times (median of around 4.1 days, see Figure 4.3). We further provide estimates of the variability in the expected rate of RTE between patients, which has no comparable value in the literature to our knowledge (coefficient of variation of around 123%). The ensuing high interindividual variability is illustrated in Figure 4.4, where the model predicts that some individuals can expect the RTE to occur within a day, while others, only after several weeks. We argue that a significant proportion of this inter-individual variability could be explained by the size of the reservoir of latently infected cells. Indeed, in a study using only 10 patients treated during chronic infection and having an extremely small reservoir, 1 patient was found to control viremia for at least 48 weeks, a result that would be unlikely if reservoir size was independent of patients' propensity for rebound (less than 1.4%, based on our model).(Calin et al., 2016)

A common hypothesis for the rebound triggering event is that it would correspond to the activation and expansion of latently infected cells.(Conway & Perelson, 2015; Pinkevych et al., 2015) During expansion, a latently infected cell can give rise to a great number of activated and virion-producing clones.(Maldarelli, 2015) This could break the equilibrium between the emergence of virion-producing cells and the capacity of cytotoxic T-cells to destroy these cells, resulting in viremia rebound.(Conway & Perelson, 2015) A greater number of latently infected cells is expected to lead to an increased probability of such cells undergoing activation and expansion, which could explain the association between the loss of post-treatment control and the size of the reservoir.(Saez-Cirion et al., 2013; Williams et al., 2014) Early treatment is also associated with having a smaller reservoir,(Gianella et al., 2011; Hocqueloux et al., 2013) which can explain, to some extent, why it would lead to more durable post-treatment control. Additional hypotheses are that antigen presentation is responsible for activation of T-cells and their clonal expansion and that viremia is a good proxy measure of the amount of antigen available for presentation by the immune system. Under this hypothesis, it is expected that

higher baseline viremia would lead to a greater frequency of activation or clonal expansion and, in turn, higher rates of rebound. The current data suggests that either it is untrue, in which case independent processes like homeostasis may be triggering rebound after treatment interruption or, if true, this process is unlikely very sensitive to the quantity of antigen in circulation in the host. In either case, the analysis herein suggests that choosing whether a patient may temporarily interrupt therapy should not be based on baseline viremia if one wishes to maintain patients' viremia and T-cell activation levels.

The above also shows that the duration of the period of rebound (see Figure 4.1) can generate difficulties in interpreting associations between patient characteristics (and biomarkers) and measures of post-treatment control. As observed in this study, the issue originates from the nature of the observations, which consist of time intervals during which the viremia of a patient has reached a threshold, not during which post-treatment control was lost. A rule-of-thumb could be that an association between characteristics and post-treatment control may need to be revisited if: i) the characteristics is thought to correlate with baseline viremia or the viral load growth rate (hence t_R), and ii) a shift in survival curves can explain the differences in survival profiles between groups of patients that differ based on these characteristics. Under this eventuality, the approach used herein may prove useful.

In conclusion, we modeled the empirical measure of post-treatment control as viremia being maintained at the baseline level until a rebound triggering event occurs, after which viremia grows exponentially to a threshold value. The rebound triggering event is expected to occur within a week of treatment interruption for most patients, although the expected rebound rate may vary significantly between patients. Our analysis suggests this variability remains largely unexplained by patients' baseline viremia. It would be interesting to use the approach elaborated herein to determine how much interindividual variability could be explained by viral reservoir or immune activation biomarkers.

4.7. Supplementary Material

Density function of patients' viremia when below 0.1 copy/mL ($g_V(v)$)

We fitted baseline viremia data obtained from Zheng et al. (Zheng et al., 2013) For this purpose, we assumed a lognormal distribution in the patient population. Estimates of the mean and variance of the underlying Gaussian distribution were obtained by maximizing the likelihood of the empirical data displayed in Table 4.4 below. We numerically minimized the log-likelihood function using an unconstrained nonlinear optimization algorithm (Nelder-Mead Simplex Method) implemented in Matlab 2014a. (Lagarias et al., 1998; The MathWorks Inc, 2014). We obtained point estimates for the mean and standard deviation of $\mu = -0.76$ and $\sigma = 2.30$. Expected frequencies based on lognormal density are also shown in Table 4.4. Model-based and empirical percentages agree substantially. The density function for baseline viremia $g_V(v)$ is conditional on viremia being below 0.1 copy/mL. Hence, we get:

$$g_V(v) = \frac{g_{LN}(v | \mu = -0.76, \sigma = 2.30)}{\int_0^{0.1} g_{LN}(v | \mu = -0.76, \sigma = 2.30)} \quad (4.9)$$

where g_{LN} represents the lognormal density function.

Table 4.4 – Empirical percentages (Zheng et al.) and model-based percentage of patients by baseline viremia categories.

Baseline viremia category	Empirical	Model-based
[0;0.2[37.6%	37.6%
[0.2;1[27.3%	28.8%
[1;2[11.7%	11.2%
[2;5[12.7%	11.8%
[5;10[5.9%	6.3%
[10;15[3.9%	2.7%
[15;20[1.0%	1.5%

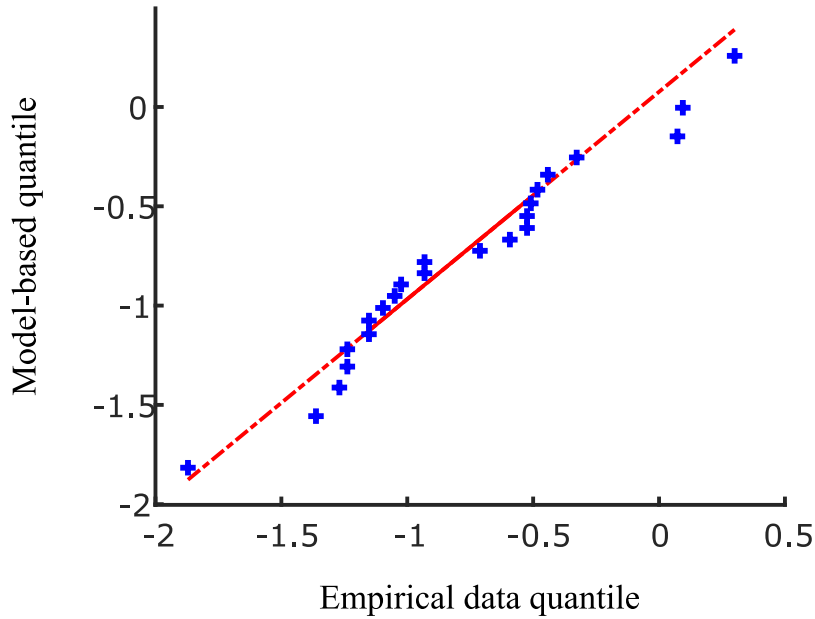
Density function for patients’ viremia growth rate after treatment interruption $g_{\Omega}(\omega)$

We fitted viremia growth rates obtained from Ioannidis et al.(Ioannidis et al., 2000) For this purpose, we used a procedure equivalent to the one described above. We obtained point estimates for the mean and standard deviation of $\mu = -0.78$ and $\sigma = 0.51$. The empirical and model-based quantile-quantile plot is displayed in Figure 4.6 and demonstrate reasonable agreement between the two distributions. The density function we used for viremia growth rates was therefore:

$$g_{\Omega}(\omega) = g_{LN}(\omega|\mu = -0.78, \sigma = 0.51) \tag{4.10}$$

where g_{LN} represents the lognormal density function.

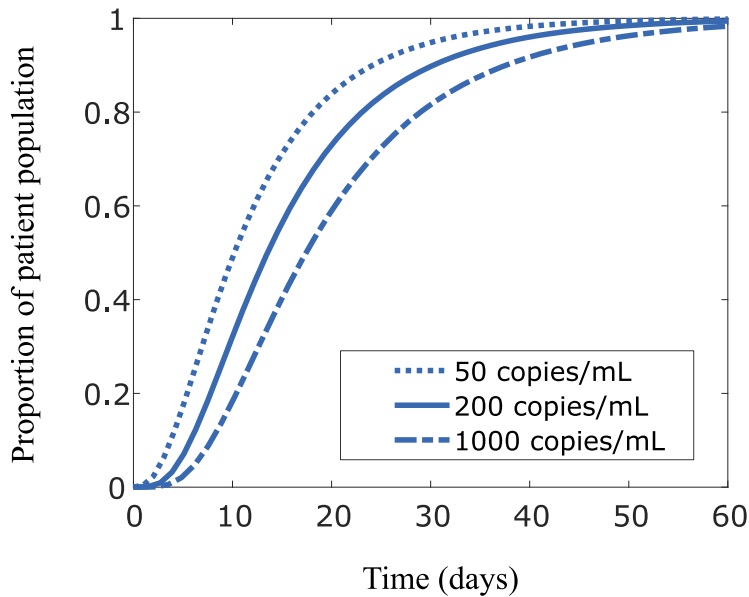
Figure 4.6 – Quantile-quantile plot of log-transformed empirical (Ioannidis et al., 2000) and model-based data.



Time to reach the virologic threshold from the rebound triggering event

We also estimated the distribution of time to reach a viremia of 50, 200 or 1000 copies/mL from rebound among patients treated in the chronic infection phase. For this purpose, we used: i) Equation 4.3, ii) a distribution of baseline viremia obtained elsewhere and using single-copy assays in patients who had been treated for around 4.8 years (3.1 to 6.4) (Zheng et al., 2013), and iii) the distribution of growth rates mentioned above and illustrated on Figure 4.6.(Ioannidis et al., 2000) Assuming independence between viremia and growth rate, the ensuing distribution for the duration of the second phase is shown in Figure 4.7.

Figure 4.7 – Cumulative distribution of the time to reach viremia threshold from the start of rebound in population of patients treated during the chronic infection phase



Chapitre 5. Discussion

La recherche dans le domaine du VIH a permis de mieux comprendre et traiter la maladie. Ces recherches ont également révélé la grande complexité des interactions entre le virus, l'hôte et les molécules exogènes. Cette complexité s'avère être la source de nombreux obstacles à la compréhension qui nuisent aux efforts menés afin de minimiser l'impact de la maladie sur ceux et celles qui en souffrent. Dans cette thèse, voulant reconstituer du moins partiellement le puzzle du VIH où de nombreuses pièces restent encore manquantes, nous nous sommes intéressées à trois questions en particulier. D'abord (Chapitre 2), comment est-il possible que le virus puisse développer de la résistance aux antirétroviraux malgré des niveaux élevés de concentration plasmatique de médicaments? Ensuite (Chapitre 3), qu'est ce qui empêche le virus de décroître de façon constante un patient traité aux antirétroviraux? Finalement (Chapitre 4), qu'est-ce qui corrèle avec le temps de maintien d'un niveau très bas de virémie après avoir cessé un traitement antirétroviral prolongé?

Ces mystères sont difficiles à élucider, entre autres puisque seules des bribes d'information peuvent être soutirées *in vivo*. En effet, malgré les avancées techniques, il demeure très difficile d'obtenir les mesures nécessaires à une compréhension globale des dites-interactions sans nuire considérablement à l'être vivant étudié. De façon illustrative, le système à l'étude est similaire à une boîte noire électrique dont on tente de deviner le circuit, mais avec seulement quelques points de mesures. Comme pour cet exemple, il en revient aux scientifiques de pallier le manque d'information en élaborant des modèles qui permettent d'expliquer les données disponibles. C'est cette approche qui a été adoptée dans cette thèse afin d'apporter des réponses aux questionnements susmentionnés. Plus précisément, une modélisation mathématique des phénomènes étudiés a été effectuée. Par l'outil informatique et l'usage de techniques d'analyse statistique, nous avons ensuite évalué ces modèles par leur capacité à expliquer un ou plusieurs jeux de données empiriques.

5.1. Principaux messages

Résumons tout d'abord les modèles, les jeux de données reproduits et les principaux messages issus des trois articles composant le corps de cette thèse. Au chapitre 2, nous avons

élaboré un modèle de cinétique virale *in vivo* qui tient compte de plusieurs facteurs associés au développement et la sélection de souches résistantes du virus, soient : i) la composition de la population virale de l'hôte et les propriétés de survie et de résistance des diverses souches qui la composent, ii) l'immunité de l'hôte, iii) son comportement d'adhésion au médicament, iv) la pharmacocinétique ainsi que l'interaction entre les antirétroviraux utilisés. Se basant sur une modélisation antérieure, (Rosenbloom et al., 2012) notre modèle se distingue par la considération des ganglions lymphatiques au centre du développement de souches résistantes. Cette modélisation décrit plus fidèlement le comportement d'adhésion des patients, et est une représentation plus réaliste de la triple interaction entre la concentration des médicaments, leur pharmacodynamie et la résistance croisée, tout en tenant compte de la variabilité interindividuelle représentant les populations de patients. Les valeurs des paramètres ont toutes été déterminées *a priori*, ce qui a l'avantage de permettre la comparaison des données prédites par le modèle à des données empiriques indépendantes. En effet, les prédictions de notre modèle sont en accord avec les risques rapportés pour les traitements étudiés, contrairement au modèle antérieur (Rosenbloom et al., 2012) qui sous-estimait de manière importante le risque de développement de résistance à ces traitements. Plus particulièrement, trois jeux de données rapportant des proportions de patients ayant subi un échec virologique principalement dû au développement de résistance ont pu être expliqués par le modèle. (Antinori et al., 2015; Gordon et al., 2015; Hofman et al., 2004) Ces résultats suggèrent que la faible quantité de médicaments qui arrive à pénétrer les ganglions peut être la raison pour laquelle un grand nombre de patients développent une résistance aux antirétroviraux, malgré des concentrations plasmatiques élevées. Par le fait même, les résultats de cette étude motivent le développement d'antirétroviraux à forte pénétration dans ces organes lymphatiques, en particulier les ganglions. Finalement, la capacité du modèle à expliquer divers jeux de données suggère que ce modèle pourrait être utilisé afin de guider le choix de combinaisons de traitements en fonction des souches virales des patients ou autres caractéristiques individuelles (vitesse de rebond, pharmacocinétique, adhésion, etc.).

Au chapitre 3, le modèle mentionné ci-haut a été adapté pour expliquer les différentes phases de plus en plus lentes de déclin de charges virales. L'adaptation considère deux compartiments physiologiques : un compartiment représentant les tissus dont l'exposition aux médicaments est largement suffisante pour inhiber toutes nouvelles infections de cellules, et un

compartiment insuffisamment exposé, où ont lieu de nouvelles infections de cellules. Nous avons tout d'abord démontré qu'il existe des valeurs de paramètres de ce modèle qui permettent de reproduire des données de charges virales à court terme chez des patients nouvellement traités.(Perelson et al., 1997) Par contre, nous avons également démontré que pour un traitement combinant efavirenz, tenofovir et emtricitabine, il n'y a vraisemblablement pas de valeurs de paramètres qui permettent aux prédictions d'être à la fois cohérentes avec les données empiriques à court ainsi qu'à long terme.(Gallant et al., 2006) Plus particulièrement, les résultats suggèrent que si un compartiment peu exposé aux médicaments est responsable de la deuxième phase de décroissance de la charge virale, alors ce même compartiment serait trop propice au développement de souches résistantes contrairement à la réalité observée. Or, une même modélisation de la cinétique virale a permis de reproduire le risque de développement de résistance pour plusieurs traitements, suggérant que cette partie du modèle est adéquate. Nous en avons conclu que la faible pénétration des médicaments dans les tissus ne peut, à elle seule, expliquer les phases de décroissances de la virémie.

Au chapitre 4, nous avons modélisé une mesure expérimentale du contrôle de la virémie post-traitement, c'est-à-dire le temps séparant l'interruption d'un traitement et l'observation de charges virologiques élevées. Avec ce modèle, nous avons représenté la progression de la virémie durant cette période en deux phases, tel qu'il est observé chez les patients : une première phase où la virémie est maintenue à sa valeur résiduelle, suivie d'une deuxième phase où la virémie croît de manière exponentielle. La fréquence de l'événement séparant les deux phases, soit l'événement causant le rebond virologique, est d'un grand intérêt et est de nature qui n'est que partiellement connue. Nous avons supposé que l'événement en question est de nature stochastique et que chaque patient a une propension spécifique au rebond. Cette propension au rebond est représentée dans le modèle par la fréquence attendue de rebond virologique par unité de temps. Quatre expressions de cette fréquence ont été formulées et chacune était une fonction de la virémie de référence des patients : i) absence de relation entre virémie et fréquence de rebond, ii) relation logarithmique, iii) racine carrée ou iv) linéaire. Ces fonctions ont été explorées afin d'évaluer si la virémie peut expliquer une partie de la variabilité interindividuelle pour la propension au rebond. Nous avons ensuite évalué laquelle de ces fonctions permet un meilleur ajustement de données de virémies post-traitement recueillies chez des patients.(Li et

al., 2016) L'analyse a démontré qu'une association rapportée entre la virémie résiduelle et le temps de maintien de la virémie sous un seuil de 200 copies/mL (Li et al., 2016) peut être entièrement expliquée par le temps nécessaire pour atteindre le seuil de virémie à partir du rebond, et non par une variabilité de fréquence de rebond. Ceci suggère que ce qui cause le rebond n'est vraisemblablement pas associé à la virémie et que, si cette dernière est étroitement liée à la concentration d'antigènes, que ces derniers ont conséquemment aussi peu d'influence sur l'occurrence dudit rebond. L'analyse effectuée dans l'article du chapitre 4 a également démontré que les associations précédemment rapportées entre les caractéristiques de patients et les mesures de contrôle de virémie post-traitements pourraient être trompeuses. (Li et al., 2016)

Les problématiques étudiées au cours de cette thèse ont grandement été influencées par les travaux de Fletcher et al. (Fletcher et al., 2014) Ceux-ci ont démontré qu'il existe de grandes différences de niveau d'exposition aux antirétroviraux entre les tissus. Cette découverte a été d'une grande inspiration pour cette thèse. Nous avons donc étudié si ce phénomène pouvait expliquer le risque de développement de résistance observé chez les patients (Chapitre 2), s'il pouvait être responsable des phases de décroissance de la charge virale (Chapitre 3), et si ce phénomène pouvait influencer l'activation des cellules infectées latentes dans le cas où la pénétration des antirétroviraux affecte la virémie résiduelle des patients (Chapitre 4).

Les résultats de cette thèse suggèrent qu'une meilleure pénétration des antirétroviraux dans certains tissus lymphatiques comme les ganglions pourrait être bénéfique. Plus précisément, une amélioration de la pénétration des antirétroviraux pourrait limiter le nombre de nouveaux cas d'échec virologiques et le développement de résistance. Il est également possible qu'elle puisse affecter la persistance du virus, bien que les résultats de cette thèse et la littérature scientifique n'appuient pas une telle possibilité. Au chapitre 3, il n'est pas exclu qu'il puisse exister un tissu si peu exposé à l'ensemble des antirétroviraux que la réplication virale par la souche sauvage du virus s'y produirait, pratiquement, librement. Il est également possible que ces nouvelles infections nourrissent le réservoir en nouvelles cellules infectées latentes. Si on suppose que l'homéostasie ne peut pas à elle seule maintenir le réservoir de telles cellules, il serait alors possible que le tissu très peu exposé aux médicaments empêche une diminution de la taille du réservoir au cours du temps et ainsi l'atteinte d'une rémission complète. Cependant, si tel était le cas, on s'attendrait à ce que la diversité génétique du réservoir varie avec le temps,

puisque les nouvelles cellules infectées latentes seraient issues de nouvelles infections qui nécessitent l'action de la transcriptase inverse, enzyme génératrice de nombreuses erreurs de transcription. Or, des données indiquent plutôt que la diversité génétique du réservoir est relativement stable.(Evering et al., 2012; Hong & Mellors, 2015; Josefsson et al., 2013), suggérant que le réservoir est formé tôt après l'initiation des traitements. Par ailleurs, les résultats du chapitre 4 suggèrent que si la seule conséquence d'une meilleure pénétration est le niveau des plateaux de virémies atteints pour les dernières phases de déclin, sans affecter le réservoir, alors il est invraisemblable qu'elle favorise le contrôle post-traitement. Ainsi, si une solution au problème de persistance au VIH existe, elle nécessiterait vraisemblablement la découverte de nouvelles classes de médicaments aptes à diminuer la taille du réservoir de cellules latentes.

5.2. Limites et perspectives

Les modèles que nous avons développés au cours de cette thèse ont fait appel à des suppositions. Le recours à des suppositions est synonyme d'un manque d'information qui, dans certains cas, aurait avantage à être comblée. Cette section fera mention des connaissances manquantes, des suppositions qui ont été formulées afin de pallier à ce manque, et de la possibilité d'études futures qui permettraient de les informer davantage. Par la suite, on discutera des limites que ces suppositions ainsi que celles des données utilisées imposent à l'applicabilité des modèles de prédiction de la charge virale. On élaborera enfin sur la façon d'étendre ce champ d'applicabilité.

5.2.1. Connaissances manquantes

La plupart des mesures de charges virales sont obtenues à partir d'échantillons de sang. Il en va de même pour les mesures de concentration d'antirétroviraux. Or, la relation entre ces valeurs et les mesures équivalentes (charges virales et concentrations) dans les organes lymphatiques secondaires demeure incertaine. Pourtant, ces organes abritent la majorité des infections de nouvelles cellules en absence de traitement. Ainsi, il a été nécessaire d'avoir recours à des approximations pour compléter les modèles, sous forme des suppositions suivantes: i) une cellule infectée et productrice de virions est représentée par environ un virion

dans le plasma et ce, qu'importe la localisation de cette cellule dans le corps, et ii) le profil de concentrations des antirétroviraux dans les tissus lymphatiques secondaires diffère du profil plasmatique par une constante de proportionnalité, unique par médicament et par tissu.

La première de ces deux suppositions est appuyée par des données suggérant une clairance très rapide des virions dans les tissus lymphatiques et le sang, ce qui en empêcherait l'accumulation de virions.(De Boer et al., 2010; Ramratnam et al., 1999) Si on suppose également que la contribution des cellules productrices de virions est indépendante de leur tissu de résidence, alors la virémie plasmatique et le nombre de cellules infectées diffèreraient par une constante de proportionnalité. Or, cette dernière supposition n'est pas vérifiée. Ainsi, il est possible que la virémie reflète majoritairement l'activité de production de virions dans un ensemble restreint de tissus. De plus, puisque l'infection de cellules est un phénomène principalement local,(A. T. Haase, 1999) une absence d'activité virale dans ces tissus pourrait mener à une virémie très basse, et ce malgré une activité virale potentiellement importante dans d'autres tissus. Ceci pourrait causer un accroissement de fatigue immunitaire, de dommage tissulaire, d'inflammation chronique ou même de résistance aux traitements, si les concentrations antirétrovirales y sont inappropriées. Une divergence entre la supposition sous-jacente aux modèles des chapitres 2 et 3 et la réalité aurait comme conséquence une surestimation de l'efficacité des régimes thérapeutiques prédits. Afin d'éviter une telle éventualité, il serait intéressant de distinguer l'activité virale dans les différents tissus et d'observer leur représentation dans le plasma. Une étude intéressante menée sur des macaques suggère qu'il est possible d'obtenir les mesures nécessaires pour faire une telle investigation chez cette espèce.(Santangelo et al., 2015) En effet, la méthode proposée par les auteurs utilise une technique d'imagerie afin de quantifier l'activité virale dans les divers tissus de l'ensemble du corps de l'animal sous traitement antirétroviral. Si des échantillons de sang sont également collectés au moment de l'imagerie, de telles mesures pourraient être mises en relation avec la charge virale plasmatique et informer la contribution des divers tissus.

Les modèles utilisés aux chapitres 2 et 3 ont également nécessité des prédictions de profils de concentrations d'antirétroviraux dans les divers tissus. L'approximation effectuée estime que ces profils imitent celui du plasma à une constante de proportionnalité près. Or, il existe un manque de connaissance à ce sujet qui empêche la vérification de l'approximation,

manque qui peut être expliqué par les difficultés éprouvées à obtenir des mesures directes de concentrations dans les tissus. Il existe une technique de modélisation nommée pharmacocinétique à base physiologique qui pourrait permettre d'effectuer des prédictions de profils de concentrations dans divers tissus.(Sager, Yu, Ragueneau-Majlessi, & Isoherranen, 2015) Cette technique utilise des paramètres de débit sanguin, de perméabilité de membranes, de liaisons aux protéines et membranes afin d'estimer les concentrations. Si ces paramètres sont connus pour beaucoup d'organes du corps et de médicaments, ils ne sont pas définis pour la plupart des organes lymphatiques. Or, ces derniers jouent un rôle majeur dans la progression des maladies comme le VIH et le cancer.(Pereira, Jones, Jung, & Padera, 2015; Sanche et al., 2017) Une estimation des paramètres présentement manquants dans la littérature, particulièrement ceux déterminant la vitesse de transit des médicaments entre système lymphatique et sanguin, pourrait être très profitable pour la prédiction de l'effet des médicaments. Cela pourrait permettre d'estimer le risque de développement de résistance dans l'ensemble des tissus lymphatiques.

Les résultats de l'étude décrite au chapitre 3 dépendent également des suppositions mentionnées ci-haut. En effet, le modèle qui a adéquatement décrit la cinétique virale à long terme dans les ganglions lymphatiques a été utilisé afin d'effectuer des prédictions dans un autre compartiment physiologique. Ce compartiment, hypothétique, serait responsable de la deuxième phase de décroissance de la charge virale. Une analyse de sensibilité a été effectuée afin de pallier partiellement l'incertitude liée aux suppositions du modèle. Il serait possible de renforcer les conclusions de l'article et d'enquêter sur les mécanismes responsables des phases de décroissance à partir de données qui ont déjà été rassemblées. En effet, une différence de phases de décroissance a déjà été démontrée entre traitements faisant intervenir ou non des inhibiteurs de l'intégrase.(Murray et al., 2007) Ceci suggère que l'intégration du matériel génétique viral pourrait être impliquée dans la cause des phases de décroissance.(Cardozo et al., 2017) Il serait intéressant de vérifier s'il existe d'autres différences entre traitements. L'étude de Fletcher et al. démontre qu'il existe des différences importantes de niveau de pénétration de médicaments dans les tissus entre les antirétroviraux.(Fletcher et al., 2014) Dans le cas où cette pénétration est impliquée dans la formation des phases de décroissance (malgré les résultats du chapitre 3), il serait attendu que des différences de phases soient observables entre traitements, en comparant

les traitements avec ou sans inhibiteurs de l'intégrase entre eux. Une absence de différence irait dans le sens des résultats du chapitre 3. La présence de différence, quant à elle, pourrait suggérer l'implication du degré de pénétration dans les tissus ou, plus vraisemblablement, l'implication d'autres processus (transcription inverse, clivage de protéines, etc.) dans la formation des phases.

Tel que mentionné brièvement ci-haut, une hypothèse trouvant appui dans la littérature est qu'un plus grand temps d'intégration d'ADN viral pourrait être responsable de la deuxième phase.(Cardozo et al., 2017) Un ralentissement de l'activité cellulaire associée à la contraction immunitaire pourrait être impliqué. Ceci nous amène à un autre sujet : les données qui permettraient d'évaluer quantitativement la dynamique des populations de lymphocytes et de cellules dendritiques en fonction de leur niveau activation ne sont pas disponibles. Qualitativement, plusieurs études suggèrent que ces populations peuvent croître et décroître rapidement en phase d'expansion ou de contraction, respectivement.(McKinstry et al., 2010; Pennock et al., 2013) Or, la vitesse à laquelle ceci se produit, l'implication des antigènes, de cellules mémoires et la variabilité à même les populations de cellules quant à cette dynamique demeurent inconnues. Pourtant, ces paramètres sont essentiels pour déterminer la taille de population de lymphocytes T auxiliaires actifs, dont la densité influe sur la capacité des virions à les infecter. Ces paramètres sont également nécessaires pour mieux caractériser l'efficacité des lymphocytes T cytotoxiques à limiter l'infection.(Dimmock et al., 2016) Des estimations quantitatives permettraient d'étudier plusieurs phénomènes, dont le contrôle de l'infection post-traitement (attribuable pour le moment à l'efficacité des cellules T cytotoxiques)(Conway & Perelson, 2015) ainsi que l'activation et la destruction de cellules latentes (surtout des cellules mémoires), des processus vraisemblablement essentiels à l'atteinte d'une guérison complète au VIH. Ils pourraient également permettre de mieux comprendre, par modélisation mécanistique, les raisons derrière l'efficacité limitée des vaccins actuels. En fait, de telles estimations pourraient permettre une modélisation de plusieurs maladies faisant intervenir les mêmes populations de cellules. Ceci motive le design d'expériences qui permettraient de telles estimations quantitatives.

5.2.2. Limites et amélioration des modèles visant la prédiction

Comme ce fut mentionné au chapitre 2, le modèle développé pourrait être utile afin d'effectuer des prédictions de charges virales dans divers contextes. Or, pour le moment, le champ d'applicabilité, i.e. l'ensemble des situations se prêtant à de telles prédictions, est limité par les données qui ont été utilisées pour la validation et par la disponibilité des mesures associées à un des paramètres du modèle (pénétration dans les ganglions lymphatiques).

Les données externes ayant servi à la validation du modèle étaient ponctuelles : à un moment situé à environ un an du début de la prise d'antirétroviraux, les prédictions et observations empiriques d'une mesure basée sur la charge virale de patients ont été comparées. La mesure en question pouvait prendre deux valeurs qui indiquaient si la charge virale était en dessous ou au-dessus d'un seuil d'échec virologique. Les antirétroviraux considérés étaient de la classe des inhibiteurs de la transcriptase inverse et des inhibiteurs de la protéase. Ainsi, le champ d'applicabilité se restreint à des prédictions pour ces classes de médicaments, à moyen ou à long terme (autour de 1 an), et pour l'échec virologique seulement. Il est également restreint par la disponibilité de la mesure de pénétration du médicament dans les ganglions lymphatiques, inconnue pour plusieurs médicaments.

Or, une adaptation du modèle est en cours. Cette adaptation pourrait permettre d'effectuer des prédictions de charges virales en mesures continues, à plus courts termes et pour des inhibiteurs de l'intégrase. En effet, le modèle décrit au chapitre 2 ne peut pas reproduire les phases de décroissance de la charge virale. Il prédit plutôt une seule phase (chapitre 3) où la charge virale décroît avec une demi-vie de près d'une journée jusqu'à l'atteinte d'un seuil situé à 3000 cellules infectées, ou 3000 virions dans le sang. Or, ce modèle a été adapté de manière à considérer 4 populations de cellules productrices de virions, chacune des populations possédant une taille et un taux de décroissance par patient, et dont ces dernières valeurs sont basées sur des observations de la virémie chez des populations de patients. (Palmer et al., 2008; Perelson et al., 1997; Zheng et al., 2013) Les inhibiteurs de l'intégrases réduisent considérablement la taille de la deuxième population de cellule (celle associée à la deuxième phase de décroissance). (Murray et al., 2007) Un facteur de réduction de taille de population a donc été ajouté dans l'adaptation du modèle. Ce dernier modèle a produit des prédictions de charges virales en accord avec les phases de décroissance et leur variabilité dans la population, lorsque

le traitement antirétroviral efficace mène à un succès virologique (données non montrées, traitements avec ou sans inhibiteur de l'intégrase). De plus, pour le traitement en combinaison d'antirétroviral considéré au chapitre 2 (efavirenz, tenofovir disoproxil fumarate et emtricitabine), les prédictions de risques d'échec virologique environ 1 an après le début du traitement sont encore plus près des observations empiriques qu'il ne le fut pour le modèle du chapitre 2 (données non montrées). Ceci suggère que les phases de décroissance, gardant un nombre de virions plus élevé pour plus longtemps, pourraient accroître les risques de développement de la résistance. De plus, ces résultats préliminaires pourront être évalués pour plusieurs autres traitements. En effet, depuis la publication de l'article du chapitre 2, des données de pénétration des médicaments dans les ganglions lymphatiques ont été obtenues pour le dolutégravir, le raltégravir et l'elvitégravir (données en voie de publication, Fletcher et al.). Il sera très intéressant de constater si l'adaptation du modèle, muni de valeurs de paramètres pour ces médicaments, prédira des risques d'échec virologique semblables à la réalité observée pour des traitements en combinaison avec ces antirétroviraux.

Le modèle pourrait également être amélioré pour permettre des prédictions à très long terme, soit sur plusieurs années. Par contre, il devra vraisemblablement considérer des phénomènes qui ont été ignorés jusqu'à maintenant, comme la possibilité que des mutants résistants à des traitements diminuent leur coût d'adaptation, (Irwin, Renzette, Kowalik, & Jensen, 2016) ou bien que le virus développe des moyens d'échapper davantage à l'action du système immunitaire de l'hôte (Simon, Bloch, & Landau, 2015).

Éventuellement, l'outil développé pourrait être très utile dans un contexte clinique ou de recherche de nouveaux antirétroviraux. Dans ce premier cas, le développement d'une application qui permettrait au clinicien de montrer au patient les risques qu'il encoure à ne pas prendre ces médicaments de façon régulière, par exemple, pourrait être bénéfique. Cette application pourrait aussi aider les cliniciens à choisir de nouveaux traitements lorsqu'un changement de régime est nécessaire. Éventuellement, l'outil pourrait servir afin de personnaliser les régimes thérapeutiques en fonction de mesures obtenues chez les patients (caractéristiques pharmacocinétiques et autres biomarqueurs). Toutefois, il est important de souligner que le modèle devra auparavant démontrer ses capacités à prédire les réponses individuelles, en plus de sa capacité de prédiction au niveau populationnel.

5.3. Retour sur les études par modèles mécanistiques

Les modèles mécanistiques sont amenés à prendre une plus grande place dans l'étude de la pharmacologie des systèmes. Si, auparavant, les capacités informatiques imposaient des contraintes aux modèles pour rendre possible leur analyse numérique, celles-ci permettent dorénavant l'étude de représentations plus fidèles à la réalité de la complexité biologique. Les capacités techniques grandissantes facilitent également l'acquisition d'information sur un plus grand nombre de paramètres, par des mesures *in vitro* ou *in vivo*, augmentant par le fait même la pertinence des modèles et des conclusions quantitatives découlant de leur analyse.

Or, comparativement aux modèles et techniques statistiques, la communauté scientifique n'est pas encore suffisamment formée ni informée de l'utilité des modèles mécanistiques, de leurs limites et de la portée des messages issus de leur analyse. Ceci peut limiter les bienfaits et retombées de ces modèles dans le domaine expérimental. Par exemple, il est possible qu'un chercheur expérimental ne puisse pas reconnaître les situations où la modélisation mécanistique lui serait d'un apport additionnel. En retour, l'opportunité d'une collaboration synergique avec des chercheurs œuvrant en modélisation pourra être manquée. Il est également envisageable que la non-maîtrise des limites des modèles mécanistiques puisse causer des attentes démesurées, voire miraculeuses, des portées des modèles ainsi que la difficulté à juger de leur pertinence ou de la qualité des études qui se basent dessus. Sans pouvoir en évaluer la juste valeur, les chercheurs concernés doivent souvent accorder toute leur confiance aux conclusions de ces études. Il en va donc de la responsabilité des auteurs des modèles de bien documenter leur portée ainsi que leurs limites.

Une attention particulière doit être portée aux suppositions non établies, leur impact potentiel sur les résultats, ainsi que sur la force des messages rapportés. Par ailleurs, pour le moment, peu de chercheurs effectuent des études de sensibilité des résultats en termes des différentes sources d'incertitudes du modèle (valeurs de paramètres incertaines, suppositions et approximations). Ce type d'études n'est toutefois pas toujours réalisables, puisqu'elles impliquent de répéter les expériences numériques ayant déjà nécessité une grande quantité de ressources informatiques, sous de nombreuses conditions. Néanmoins, des règles de bonnes

pratiques pourraient être émises afin d'ajouter à la crédibilité des résultats d'analyses de modèles mécanistiques.

En introduction, des lignes directrices ont été mentionnées quant à l'utilité des modèles mécanistiques ainsi qu'à l'étendue des messages, et ce, en tenant compte de la façon dont les données empiriques ont été utilisées. J'ai fait de mon mieux pour que ces lignes directrices soient respectées tout le long des études réalisées dans cette thèse. Aux chapitres 2 et 3, un mécanisme hypothétique (pénétration dans certains tissus) a été évalué afin de déterminer si celui-ci peut vraisemblablement ou non expliquer un phénomène (risque de développement de résistance ou phases de décroissance de la charge virale, par exemple). Pour ce faire, une comparaison a été effectuée entre les prédictions des modèles et des données empiriques indépendantes de celles ayant servi à l'estimation des paramètres du modèle. Au chapitre 4, on a donné un sens plus mécanistique à une corrélation rapportée entre virémie résiduelle et la durée de contrôle post-traitement, en développant un modèle mathématique pour la dérivation du temps de rebond virologique. On y discute de la dépendance de cette corrélation vis-à-vis la modélisation de la virémie durant la période de contrôle, celle-ci appuyée par des observations empiriques de charges virologiques durant cette période.

Il existe en revanche dans la littérature, des situations où les résultats d'études par modélisation mécanistique se trouvent à être sujets à une interprétation inappropriée. Par exemple, il est difficilement justifiable qu'un modèle non validé sur des données recueillies chez l'humain, puisse servir pour des recommandations dans un contexte clinique. Ceci est particulièrement problématique lorsque ces recommandations mènent à des conséquences graves pour les patients. Par exemple, les prédictions du modèle développé par Rosenbloom et al. (Rosenbloom et al., 2012), et qui avaient inspiré l'article du chapitre 2, n'avaient pas été vérifiées *in vivo*. Or, comme il est mentionné dans le présent manuscrit, des données cliniques viennent contredire certaines prédictions de ce modèle, notamment celle annonçant qu'une monothérapie avec efavirenz mènerait à un succès thérapeutique sans développement de résistance. Malheureusement, le titre de l'article en question ainsi que son résumé peuvent laisser croire que ce modèle pourrait servir à des fins cliniques. Il faut cependant mentionner que les auteurs eux-mêmes rectifient le tir à la fin de leur manuscrit. Une lecture inattentive ou incomplète pourrait alors mener à un usage non justifié de ce modèle pour fins de détermination

des doses et régimes thérapeutiques de patients, avec le risque potentiel de causer le développement de résistance à long terme.

En résumé, la modélisation mécanistique jouera un rôle de plus en plus dominant en pharmacologie, en biologie et en médecine. Des lignes directrices sur les bonnes pratiques en pharmacologie quantitative des systèmes seraient souhaitables pour mieux guider la façon de rapporter les résultats des études basées sur la modélisation. Également, un travail de vulgarisation visant à outiller les chercheurs non issus des domaines quantitatifs serait très profitable, favoriserait les collaborations et améliorerait nettement la transmission des connaissances. Le clivage parfois trop grand entre chercheurs expérimentaux et mathématiciens représente un obstacle à cette collaboration et aurait avantage à être diminué. Ainsi, les uns et les autres pourraient bénéficier d'une formation hybride afin de développer un langage commun. On observe, heureusement, un rapprochement de plus en plus grand entre chercheurs qui transcendent les disciplines et font fi des difficultés qui peuvent émaner de travailler en dehors de leur cadre habituel. Les retombées de ces efforts valent largement les obstacles rencontrés.

Bibliographie

- Abram, M. E., Ferris, A. L., Das, K., Quinones, O., Shao, W., Tuske, S., . . . Hughes, S. H. (2014). Mutations in HIV-1 reverse transcriptase affect the errors made in a single cycle of viral replication. *J Virol*, *88*(13), 7589-7601. doi:10.1128/JVI.00302-14
- Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G., & Hughes, S. H. (2010). Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. *J Virol*, *84*(19), 9864-9878. doi:10.1128/JVI.00915-10
- Achappa, B., Madi, D., Bhaskaran, U., Ramapuram, J. T., Rao, S., & Mahalingam, S. (2013). Adherence to Antiretroviral Therapy Among People Living with HIV. *N Am J Med Sci*, *5*(3), 220-223. doi:10.4103/1947-2714.109196
- Acosta, E. P., Limoli, K. L., Trinh, L., Parkin, N. T., King, J. R., Weidler, J. M., . . . Petropoulos, C. J. (2012). Novel method to assess antiretroviral target trough concentrations using in vitro susceptibility data. *Antimicrob Agents Chemother*, *56*(11), 5938-5945. doi:10.1128/AAC.00691-12
- Agresti, A. (2013). *Categorical data analysis*: John Wiley & Sons.
- Al-Dakkak, I., Patel, S., McCann, E., Gadkari, A., Prajapati, G., & Maiese, E. M. (2013). The impact of specific HIV treatment-related adverse events on adherence to antiretroviral therapy: a systematic review and meta-analysis. *AIDS Care*, *25*(4), 400-414. doi:10.1080/09540121.2012.712667
- Alexander, T. S. (2016). Human Immunodeficiency Virus Diagnostic Testing: 30 Years of Evolution. *Clin Vaccine Immunol*, *23*(4), 249-253. doi:10.1128/CVI.00053-16
- Altfeld, M., & Gale, M., Jr. (2015). Innate immunity against HIV-1 infection. *Nat Immunol*, *16*(6), 554-562. doi:10.1038/ni.3157
- Anderson-Bergman, C. (2016). icenReg: Regression Models for Interval Censored Data (Version R package version 1.3.4). Retrieved from <http://CRAN.R-project.org/package=icenReg>
- Andrake, M. D., & Skalka, A. M. (2015). Retroviral Integrase: Then and Now. *Annu Rev Virol*, *2*(1), 241-264. doi:10.1146/annurev-virology-100114-055043
- Antinori, A., Clarke, A., Svedhem-Johansson, V., Arribas, J. R., Arenas-Pinto, A., Fehr, J., . . . Moecklinghoff, C. (2015). Week 48 efficacy and central nervous system analysis of darunavir/ritonavir monotherapy versus darunavir/ritonavir with two nucleoside analogues. *AIDS*, *29*, 1811-1820. doi:10.1097/QAD.0000000000000778
- Archin, N. M., Liberty, A. L., Kashuba, A. D., Choudhary, S. K., Kuruc, J. D., Crooks, A. M., . . . Margolis, D. M. (2012). Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature*, *487*(7408), 482-485. doi:10.1038/nature11286
- Arfi, V., Riviere, L., Jarrosson-Wuilleme, L., Goujon, C., Rigal, D., Darlix, J. L., & Cimarelli, A. (2008). Characterization of the early steps of infection of primary blood monocytes by human immunodeficiency virus type 1. *J Virol*, *82*(13), 6557-6565. doi:10.1128/JVI.02321-07
- Asquith, B., Edwards, C. T., Lipsitch, M., & McLean, A. R. (2006). Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. *PLoS Biol*, *4*(4), e90. doi:10.1371/journal.pbio.0040090

- Assoumou, L., Weiss, L., Piketty, C., Burgard, M., Melard, A., Girard, P. M., . . . group, A. S. s. (2015). A low HIV-DNA level in peripheral blood mononuclear cells at antiretroviral treatment interruption predicts a higher probability of maintaining viral control. *AIDS*, *29*(15), 2003-2007. doi:10.1097/QAD.0000000000000734
- Baheti, G., Kiser, J. J., Havens, P. L., & Fletcher, C. V. (2011). Plasma and intracellular population pharmacokinetic analysis of tenofovir in HIV-1-infected patients. *Antimicrob Agents Chemother*, *55*(11), 5294-5299. doi:10.1128/AAC.05317-11
- Baldauf, H. M., Pan, X., Erikson, E., Schmidt, S., Daddacha, W., Burggraf, M., . . . Keppler, O. T. (2012). SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nat Med*, *18*(11), 1682-1687. doi:10.1038/nm.2964
- Bangsberg, D. R., Kroetz, D. L., & Deeks, S. G. (2007). Adherence-resistance relationships to combination HIV antiretroviral therapy. *Current HIV/AIDS Reports*, *4*, 65-72. doi:10.1007/s11904-007-0010-0
- Bar, K. J., Sneller, M. C., Harrison, L. J., Justement, J. S., Overton, E. T., Petrone, M. E., . . . Chun, T. W. (2016). Effect of HIV Antibody VRC01 on Viral Rebound after Treatment Interruption. *N Engl J Med*, *375*(21), 2037-2050. doi:10.1056/NEJMoa1608243
- Bar, K. J., Tsao, C. Y., Iyer, S. S., Decker, J. M., Yang, Y., Bonsignori, M., . . . Shaw, G. M. (2012). Early low-titer neutralizing antibodies impede HIV-1 replication and select for virus escape. *PLoS Pathog*, *8*(5), e1002721. doi:10.1371/journal.ppat.1002721
- Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., . . . Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, *220*(4599), 868-871.
- Beal, S., Sheiner, L., & Boeckmann, A. (2009). *Bauer R.J. NONMEM User's Guides*. Ellicott City, MD. Retrieved from
- Berger, J. O., Pericchi, L. R., Ghosh, J., Samanta, T., De Santis, F., Berger, J., & Pericchi, L. (2001). Objective Bayesian methods for model selection: introduction and comparison. *Lecture Notes-Monograph Series*, 135-207.
- Biswas, P., Tambussi, G., & Lazzarin, A. (2007). Access denied? The status of co-receptor inhibition to counter HIV entry. *Expert Opin Pharmacother*, *8*(7), 923-933. doi:10.1517/14656566.8.7.923
- Blankson, J. N., Persaud, D., & Siliciano, R. F. (2002). The challenge of viral reservoirs in HIV-1 infection. *Annu Rev Med*, *53*, 557-593. doi:10.1146/annurev.med.53.082901.104024
- Blum, M. R., Chittick, G. E., Begley, J. A., & Zong, J. (2007). Steady-state pharmacokinetics of emtricitabine and tenofovir disoproxil fumarate administered alone and in combination in healthy volunteers. *J Clin Pharmacol*, *47*(6), 751-759. doi:10.1177/0091270007300951
- Bowman, R. R., Hu, W. S., & Pathak, V. K. (1998). Relative rates of retroviral reverse transcriptase template switching during RNA- and DNA-dependent DNA synthesis. *J Virol*, *72*(6), 5198-5206.
- Brenchley, J. M. (2004). CD4+ T Cell Depletion during all Stages of HIV Disease Occurs Predominantly in the Gastrointestinal Tract. *Journal of Experimental Medicine*, *200*, 749-759. doi:10.1084/jem.20040874
- Bristol-Myers Squibb. (2015). *Sustiva (efavirenz) : Highlight of Prescribing Information*.
- Buchholz, V. R., Schumacher, T. N., & Busch, D. H. (2016). T Cell Fate at the Single-Cell Level. *Annu Rev Immunol*, *34*, 65-92. doi:10.1146/annurev-immunol-032414-112014

- Cadosch, D., Bonhoeffer, S., & Kouyos, R. (2012). Assessing the impact of adherence to anti-retroviral therapy on treatment failure and resistance evolution in HIV. *J R Soc Interface*, 9(74), 2309-2320. doi:10.1098/rsif.2012.0127
- Calin, R., Hamimi, C., Lambert-Niclot, S., Carcelain, G., Bellet, J., Assoumou, L., . . . Group, U. S. (2016). Treatment interruption in chronically HIV-infected patients with an ultralow HIV reservoir. *AIDS*, 30(5), 761-769. doi:10.1097/QAD.0000000000000987
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53): Cambridge university press.
- Caplan, M. R., Daar, E. S., & Corado, K. C. (2018). Next generation fixed dose combination pharmacotherapies for treating HIV. *Expert Opin Pharmacother*, 19(6), 589-596. doi:10.1080/14656566.2018.1450866
- Cardozo, E. F., Andrade, A., Mellors, J. W., Kuritzkes, D. R., Perelson, A. S., & Ribeiro, R. M. (2017). Treatment with integrase inhibitor suggests a new interpretation of HIV RNA decay curves that reveals a subset of cells with slow integration. *PLoS Pathog*, 13(7), e1006478. doi:10.1371/journal.ppat.1006478
- Catie.ca. (2017). Un historique du VIH et du sida. Retrieved from <http://www.catie.ca/fr/journee-mondiale-contre-le-sida/historique>
- Centers for Disease Control and Prevention. (January 7, 1983). Epidemiologic Notes and Reports Immunodeficiency among Female Sexual Partners of Males with Acquired Immune Deficiency Syndrome (AIDS) - New York. *Morbidity and Mortality Weekly Report*. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001163.htm>
- Centers for Disease Control and Prevention. (June 1, 2001). HIV and AIDS - United States, 1981-2000. *Morbidity and Mortality Weekly Report*. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5021a2.htm>
- Centers for Disease Control and Prevention. (June 3, 2011). HIV Surveillance - United States, 1981-2008. *Morbidity and Mortality Weekly Report*. Retrieved from https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6021a2.htm?s_cid=mm6021a2_w
- Centers for Disease Control and Prevention. (June 5, 1981). Pneumocystis Pneumonia - Los Angeles. *Morbidity and Mortality Weekly Report*. Retrieved from https://www.cdc.gov/mmwr/preview/mmwrhtml/june_5.htm
- Centers for Disease Control and Prevention. (September 24, 1982). Current Trends Update on Acquired Immune Deficiency Syndrome (AIDS) - United States. *Morbidity and Mortality Weekly Report*. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001163.htm>
- Chambers, L. A., Rueda, S., Baker, D. N., Wilson, M. G., Deutsch, R., Raeifar, E., . . . Stigma Review, T. (2015). Stigma, HIV and health: a qualitative synthesis. *BMC Public Health*, 15, 848. doi:10.1186/s12889-015-2197-0
- Chappel, R. J., Wilson, K. M., & Dax, E. M. (2009). Immunoassays for the diagnosis of HIV: meeting future needs by enhancing the quality of testing. *Future Microbiol*, 4(8), 963-982. doi:10.2217/fmb.09.77
- Chatenoud, L. (2008). Immunologie.
- Chesney, M. A., Morin, M., & Sherr, L. (2000). Adherence to HIV combination therapy. *Soc Sci Med*, 50(11), 1599-1605.
- Chis, O.-T., Banga, J. R., & Balsa-Canto, E. (2011). Structural identifiability of systems biology models: a critical comparison of methods. *PLoS One*, 6(11), e27755.

- Chomont, N., DaFonseca, S., Vandergeeten, C., Ancuta, P., & Sekaly, R. P. (2011). Maintenance of CD4⁺ T-cell memory and HIV persistence: keeping memory, keeping HIV. *Curr Opin HIV AIDS*, 6(1), 30-36. doi:10.1097/COH.0b013e3283413775
- Chun, T. W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D., & Siliciano, R. F. (1995). In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nat Med*, 1(12), 1284-1290.
- Cohen, M. S., Shaw, G. M., McMichael, A. J., & Haynes, B. F. (2011). Acute HIV-1 Infection. *N Engl J Med*, 364(20), 1943-1954. doi:10.1056/NEJMra1011874
- Colby, D. J., Trautmann, L., Pinyakorn, S., Leyre, L., Pagliuzza, A., Kroon, E., . . . group, R. V. s. (2018). Rapid HIV RNA rebound after antiretroviral treatment interruption in persons durably suppressed in Fiebig I acute HIV infection. *Nat Med*. doi:10.1038/s41591-018-0026-6
- Collier, A. C., Coombs, R. W., Fischl, M. A., Skolnik, P. R., Northfelt, D., Boutin, P., . . . Corey, L. (1993). Combination therapy with zidovudine and didanosine compared with zidovudine alone in HIV-1 infection. *Ann Intern Med*, 119(8), 786-793.
- Collins, G. (1990). Fundamental numerical methods and data analysis. *Fundamental Numerical Methods and Data Analysis, by George Collins, II*.
- Conway, J. M., & Perelson, A. S. (2015). Post-treatment control of HIV infection. *Proc Natl Acad Sci U S A*, 112(17), 5467-5472. doi:10.1073/pnas.1419162112
- Costiniuk, C. T., & Jenabian, M. A. (2015). HIV reservoir dynamics in the face of highly active antiretroviral therapy. *AIDS Patient Care STDS*, 29(2), 55-68. doi:10.1089/apc.2014.0173
- Craig, M., Gonzalez-Sales, M., Li, J., & Nekka, F. (2016). Approaching Pharmacometrics as a Paleontologist Would: Recovering the Links Between Drugs and the Body Through Reconstruction. *CPT Pharmacometrics Syst Pharmacol*, 5(3), 158-160. doi:10.1002/psp4.12069
- Cuevas, J. M., Geller, R., Garijo, R., Lopez-Aldeguer, J., & Sanjuan, R. (2015). Extremely High Mutation Rate of HIV-1 In Vivo. *PLoS Biol*, 13(9), e1002251. doi:10.1371/journal.pbio.1002251
- De Boer, R. J., Ribeiro, R. M., & Perelson, A. S. (2010). Current estimates for HIV-1 production imply rapid viral clearance in lymphoid tissues. *PLoS Comput Biol*, 6(9), e1000906. doi:10.1371/journal.pcbi.1000906
- Deeks, S. G., Tracy, R., & Douek, D. C. (2013). Systemic effects of inflammation on health during chronic HIV infection. *Immunity*, 39(4), 633-645. doi:10.1016/j.immuni.2013.10.001
- Deeks, S. G., & Walker, B. D. (2007). Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity*, 27(3), 406-416. doi:10.1016/j.immuni.2007.08.010
- Delves, P. J., Martin, S. J., Burton, D. R., & Roitt, I. M. (2017). *Essential immunology*: John Wiley & Sons.
- Delviks-Frankenberry, K., Galli, A., Nikolaitchik, O., Mens, H., Pathak, V. K., & Hu, W. S. (2011). Mechanisms and factors that influence high frequency retroviral recombination. *Viruses*, 3(9), 1650-1680. doi:10.3390/v3091650
- Demers, K. R., Reuter, M. A., & Betts, M. R. (2013). CD8(+) T-cell effector function and transcriptional regulation during HIV pathogenesis. *Immunol Rev*, 254(1), 190-206. doi:10.1111/imr.12069

- Dierynck, I., De Wit, M., Gustin, E., Keuleers, I., Vandersmissen, J., Hallenberger, S., & Hertogs, K. (2007). Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *Journal of virology*, *81*, 13845-13851.
- Dimmock, N. J., Easton, A. J., & Leppard, K. N. (2016). *Introduction to modern virology*. John Wiley & Sons.
- Dinoso, J. B., Kim, S. Y., Wiegand, A. M., Palmer, S. E., Gange, S. J., Cranmer, L., . . . Maldarelli, F. (2009). Treatment intensification does not reduce residual HIV-1 viremia in patients on highly active antiretroviral therapy. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 9403-9408. doi:10.1073/pnas.0903107106
- Dixit, N. M., Markowitz, M., Ho, D. D., & Perelson, A. S. (2004). Estimates of intracellular delay and average drug efficacy from viral load data of HIV-infected individuals under antiretroviral therapy. *Antivir Ther*, *9*(2), 237-246.
- Doitsh, G., Galloway, N. L., Geng, X., Yang, Z., Monroe, K. M., Zepeda, O., . . . Greene, W. C. (2014). Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*, *505*(7484), 509-514. doi:10.1038/nature12940
- Dornsife, R. E., St Clair, M. H., Huang, A. T., Panella, T. J., Koszalka, G. W., Burns, C. L., & Averett, D. R. (1991). Anti-human immunodeficiency virus synergism by zidovudine (3'-azidothymidine) and didanosine (dideoxyinosine) contrasts with their additive inhibition of normal human marrow progenitor cells. *Antimicrob Agents Chemother*, *35*(2), 322-328.
- Durack, D. T. (1981). Opportunistic infections and Kaposi's sarcoma in homosexual men. *N Engl J Med*, *305*(24), 1465-1467. doi:10.1056/NEJM198112103052408
- Edelstein-Keshet, L. (1988). *Mathematical models in biology* (Vol. 46): Siam.
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*: CRC press.
- Eisele, E., & Siliciano, R. F. (2012). Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity*, *37*(3), 377-388. doi:10.1016/j.immuni.2012.08.010
- Érdi, P., & Tóth, J. (1989). *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*: Manchester University Press.
- Evering, T. H., Mehandru, S., Racz, P., Tenner-Racz, K., Poles, M. A., Figueroa, A., . . . Markowitz, M. (2012). Absence of HIV-1 evolution in the gut-associated lymphoid tissue from patients on combination antiviral therapy initiated during primary infection. *PLoS Pathog*, *8*(2), e1002506. doi:10.1371/journal.ppat.1002506
- FDA. (2018). Antiretroviral drugs used in the treatment of HIV infection. Retrieved from <https://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm>
- FDA. (2018,-a). HIV/AIDS Historical Time Line 1981-1990. Retrieved from <https://www.fda.gov/forpatients/illness/hivaids/history/ucm151074.htm>
- FDA. (2018,-b). HIV/AIDS Historical Time Line 1995-1999. Retrieved from <https://www.fda.gov/ForPatients/Illness/HIVAIDS/History/ucm151079.htm>
- Fellows, K., Rodriguez-Cruz, V., Covelli, J., Droopad, A., Alexander, S., & Ramanathan, M. (2015). A Hybrid Markov Chain–von Mises Density Model for the Drug-Dosing Interval and Drug Holiday Distributions. *The AAPS Journal*, *17*, 427-437. doi:10.1208/s12248-014-9713-5

- Fettig, J., Swaminathan, M., Murrill, C. S., & Kaplan, J. E. (2014). Global epidemiology of HIV. *Infect Dis Clin North Am*, 28(3), 323-337. doi:10.1016/j.idc.2014.05.001
- Fink, S. L., & Cookson, B. T. (2005). Apoptosis, pyroptosis, and necrosis: mechanistic description of dead and dying eukaryotic cells. *Infect Immun*, 73(4), 1907-1916. doi:10.1128/IAI.73.4.1907-1916.2005
- Fletcher, C. V., Staskus, K., Wietgreffe, S. W., Rothenberger, M., Reilly, C., Chipman, J. G., . . . Schacker, T. W. (2014). Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc Natl Acad Sci USA*, 111(6), 2307-2312. doi:10.1073/pnas.1318249111
- Ford, N., Meintjes, G., Vitoria, M., Greene, G., & Chiller, T. (2017). The evolving role of CD4 cell counts in HIV care. *Curr Opin HIV AIDS*, 12(2), 123-128. doi:10.1097/COH.0000000000000348
- Frenkel, L. M., Wang, Y., Learn, G. H., McKernan, J. L., Ellis, G. M., Mohan, K. M., . . . Mullins, J. I. (2003). Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy. *J Virol*, 77(10), 5721-5730.
- Gallant, J. E., DeJesus, E., Arribas, J. R., Pozniak, A. L., Gazzard, B., Campo, R. E., . . . Study, G. (2006). Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirenz for HIV. *N Engl J Med*, 354(3), 251-260. doi:10.1056/NEJMoa051871
- Ganusov, V. V., & Auerbach, J. (2014). Mathematical modeling reveals kinetics of lymphocyte recirculation in the whole organism. *PLoS computational biology*, 10, e1003586. doi:10.1371/journal.pcbi.1003586
- Geginat, J., Paroni, M., Facciotti, F., Gruarin, P., Kastirr, I., Caprioli, F., . . . Abrignani, S. (2013). The CD4-centered universe of human T cell subsets. *Semin Immunol*, 25(4), 252-262. doi:10.1016/j.smim.2013.10.012
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*: Chapman and Hall/CRC.
- Gesztelyi, R., Zsuga, J., Kemeny-Beke, A., Varga, B., Juhasz, B., & Tosaki, A. (2012). The Hill equation and the origin of quantitative pharmacology. *Archive for history of exact sciences*, 66(4), 427-438.
- Gianella, S., von Wyl, V., Fischer, M., Niederoest, B., Battegay, M., Bernasconi, E., . . . Swiss, H. I. V. C. S. (2011). Effect of early antiretroviral therapy during primary HIV-1 infection on cell-associated HIV-1 DNA and plasma HIV-1 RNA. *Antivir Ther*, 16(4), 535-545. doi:10.3851/IMP1776
- Gilead Sciences Inc. (2008). *Product Monograph : Emtriva (Emtricitabine Capsules)*.
- Gilead Sciences Inc. (2013). *Product Monograph : Viread (Tenofovir Disoproxil Fumarate)*.
- Gilmore, J. B., Kelleher, A. D., Cooper, D. A., & Murray, J. M. (2013). Explaining the determinants of first phase HIV decay dynamics through the effects of stage-dependent drug action. *PLoS Comput Biol*, 9(3), e1002971. doi:10.1371/journal.pcbi.1002971
- Girard, P., Blaschke, T. F., Kastrissios, H., & Sheiner, L. B. (1998). A Markov mixed effect regression model for drug compliance. *Statistics in Medicine*, 17, 2313-2333. doi:10.1002/(SICI)1097-0258(19981030)17:20<2313::AID-SIM935>3.0.CO;2-V
- Goicoechea, M., & Best, B. (2007). Efavirenz/emtricitabine/tenofovir disoproxil fumarate fixed-dose combination: first-line therapy for all? *Expert Opin Pharmacother*, 8(3), 371-382. doi:10.1517/14656566.8.3.371

- Gonzalo-Gil, E., Ikediobi, U., & Sutton, R. E. (2017). Mechanisms of Virologic Control and Clinical Characteristics of HIV+ Elite/Viremic Controllers. *Yale J Biol Med*, *90*(2), 245-259.
- Gordon, L. L., Gharibian, D., Chong, K., & Chun, H. (2015). Comparison of HIV Virologic Failure Rates Between Patients with Variable Adherence to Three Antiretroviral Regimen Types. *AIDS Patient Care and STDs*, *29*, 384-388. doi:10.1089/apc.2014.0165
- Gottlieb, M. S. (1998). Discovering AIDS. *Epidemiology*, *9*(4), 365-367.
- Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A., & Saxon, A. (1981). Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N Engl J Med*, *305*(24), 1425-1431. doi:10.1056/NEJM198112103052401
- Goudsmit, J., de Ronde, A., de Rooij, E., & de Boer, R. (1997). Broad spectrum of in vivo fitness of human immunodeficiency virus type 1 subpopulations differing at reverse transcriptase codons 41 and 215. *J Virol*, *71*(6), 4479-4484.
- Greenberg, M. L., & Cammack, N. (2004). Resistance to enfuvirtide, the first HIV fusion inhibitor. *J Antimicrob Chemother*, *54*(2), 333-340. doi:10.1093/jac/dkh330
- Griffiths, A. J., Wessler, S. R., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., & Miller, J. H. (2005). *An introduction to genetic analysis*: Macmillan.
- Haase, A. T. (1999). Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues. *Annu Rev Immunol*, *17*, 625-656. doi:10.1146/annurev.immunol.17.1.625
- Haase, A. T., Henry, K., Zupancic, M., Sedgewick, G., Faust, R. a., Melroe, H., . . . Perelson, A. S. (1996). Quantitative Image Analysis of HIV-1 Infection in Lymphoid Tissue. *Science*, *274*, 985-989.
- Hall, H. I., Song, R., Tang, T., An, Q., Prejean, J., Dietz, P., . . . Mermin, J. (2017). HIV Trends in the United States: Diagnoses and Estimated Incidence. *JMIR Public Health Surveill*, *3*(1), e8. doi:10.2196/publichealth.7051
- Harrell, F. E. (2014). Regression modeling strategies. *as implemented in R package 'rms' version*, *3*(3).
- Hawkins, T. (2010). Understanding and managing the adverse effects of antiretroviral therapy. *Antiviral Res*, *85*(1), 201-209. doi:10.1016/j.antiviral.2009.10.016
- Haynes, B. F., Moody, M. A., Alam, M., Bonsignori, M., Verkoczy, L., Ferrari, G., . . . Kelsoe, G. (2014). Progress in HIV-1 vaccine development. *J Allergy Clin Immunol*, *134*(1), 3-10; quiz 11. doi:10.1016/j.jaci.2014.04.025
- Hill, A. L., Rosenbloom, D. I., & Nowak, M. A. (2012). Evolutionary dynamics of HIV at multiple spatial and temporal scales. *J Mol Med (Berl)*, *90*(5), 543-561. doi:10.1007/s00109-012-0892-1
- HIV.gov. (2016). A Timeline of HIV and AIDS. Retrieved from <https://www.hiv.gov/hiv-basics/overview/history/hiv-and-aids-timeline>
- Hlavacek, W. S., Stilianakis, N. I., Notermans, D. W., Danner, S. A., & Perelson, A. S. (2000). Influence of follicular dendritic cells on decay of HIV during antiretroviral therapy. *Proc Natl Acad Sci U S A*, *97*(20), 10966-10971. doi:10.1073/pnas.190065897
- Hockett, R. D., Kilby, J. M., Derdeyn, C. A., Saag, M. S., Sillers, M., Squires, K., . . . Bucy, R. P. (1999). Constant mean viral copy number per infected cell in tissues regardless of high, low, or undetectable plasma HIV RNA. *J Exp Med*, *189*(10), 1545-1554.

- Hocqueloux, L., Avettand-Fenoel, V., Jacquot, S., Prazuck, T., Legac, E., Melard, A., . . . Virales, A. C. o. t. A. N. d. R. s. l. S. e. l. H. (2013). Long-term antiretroviral therapy initiated during primary HIV-1 infection is key to achieving both low HIV reservoirs and normal T cell counts. *J Antimicrob Chemother*, *68*(5), 1169-1178. doi:10.1093/jac/dks533
- Hocqueloux, L., Prazuck, T., Avettand-Fenoel, V., Lafeuillade, A., Cardon, B., Viard, J. P., & Rouzioux, C. (2010). Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection. *AIDS*, *24*(10), 1598-1601.
- Hofman, M. J., Higgins, J., Matthews, T. B., Pedersen, N. C., Tan, C., Schinazi, R. F., . . . Hemother, A. N. A. G. C. (2004). Efavirenz Therapy in Rhesus Macaques Infected with a Chimera of Simian Immunodeficiency Virus Containing Reverse Transcriptase from Human Immunodeficiency Virus Type 1. *Antimicrobial agents and chemotherapy*, *48*, 3483-3490. doi:10.1128/AAC.48.9.3483
- Hogg, R. S., Bangsberg, D. R., Lima, V. D., Alexander, C., Bonner, S., Yip, B., . . . Harrigan, P. R. (2006). Emergence of drug resistance is associated with an increased risk of death among patients first starting HAART. *PLoS Med*, *3*(9), e356. doi:10.1371/journal.pmed.0030356
- Holmes, M., Zhang, F., & Bieniasz, P. D. (2015). Single-Cell and Single-Cycle Analysis of HIV-1 Replication. *PLoS Pathog*, *11*(6), e1004961. doi:10.1371/journal.ppat.1004961
- Hong, F. F., & Mellors, J. W. (2015). Changes in HIV reservoirs during long-term antiretroviral therapy. *Curr Opin HIV AIDS*, *10*(1), 43-48. doi:10.1097/COH.0000000000000119
- Hu, Z., & Kuritzkes, D. R. (2014). Altered viral fitness and drug susceptibility in HIV-1 carrying mutations that confer resistance to nonnucleoside reverse transcriptase and integrase strand transfer inhibitors. *J Virol*, *88*(16), 9268-9276. doi:10.1128/JVI.00695-14
- Huang, Y., Hoque, M. T., Jenabian, M. A., Vyboh, K., Whyte, S. K., Sheehan, N. L., . . . Bendayan, R. (2016). Antiretroviral drug transporters and metabolic enzymes in human testicular tissue: potential contribution to HIV-1 sanctuary site. *J Antimicrob Chemother*, *71*(7), 1954-1965. doi:10.1093/jac/dkw046
- Hurst, J., Hoffmann, M., Pace, M., Williams, J. P., Thornhill, J., Hamlyn, E., . . . Frater, J. (2015). Immunological biomarkers predict HIV-1 viral rebound after treatment interruption. *Nat Commun*, *6*, 8495. doi:10.1038/ncomms9495
- Ioannidis, J. P., Havlir, D. V., Tebas, P., Hirsch, M. S., Collier, A. C., & Richman, D. D. (2000). Dynamics of HIV-1 viral load rebound among patients with previous suppression of viral replication. *AIDS*, *14*(11), 1481-1488.
- Irwin, K. K., Renzette, N., Kowalik, T. F., & Jensen, J. D. (2016). Antiviral drug resistance as an adaptive process. *Virus Evol*, *2*(1), vew014. doi:10.1093/ve/vew014
- Jacquez, J. A. (1999). *Modeling with compartments*: BioMedware.
- Jakobsen, M. R., Olganier, D., & Hiscott, J. (2015). Innate immune sensing of HIV-1 infection. *Curr Opin HIV AIDS*, *10*(2), 96-102. doi:10.1097/COH.0000000000000129
- Janssen Inc. (2014). *Monographie de produit: Prezista*.
- Jilek, B. L., Zarr, M., Sampah, M. E., Rabi, S. A., Bullen, C. K., Lai, J., . . . Siliciano, R. F. (2012). A quantitative basis for antiretroviral therapy for HIV-1 infection. *Nat Med*, *18*(3), 446-451. doi:10.1038/nm.2649

- Johnson Norman, L., Kotz, S., & Balakrishnan, N. (1994). Lognormal distributions, continuous univariate distributions (Vol. 1). In: Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics.
- Josefsson, L., von Stockenstrom, S., Faria, N. R., Sinclair, E., Bacchetti, P., Killian, M., . . . Palmer, S. (2013). The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A*, *110*(51), E4987-4996. doi:10.1073/pnas.1308313110
- Karris, M. Y., Jain, S., Day, T. R., Perez-Santiago, J., Goicoechea, M., Dube, M. P., . . . California Collaborative Treatment Group 589 Study, T. (2017). HIV viral kinetics and T cell dynamics in antiretroviral naive persons starting an integrase strand transfer inhibitor and protease inhibitor regimen. *HIV Clin Trials*, *18*(2), 67-74. doi:10.1080/15284336.2017.1282578
- Kaul, M., Garden, G. A., & Lipton, S. A. (2001). Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature*, *410*(6831), 988-994. doi:10.1038/35073667
- Kirk, P., Thorne, T., & Stumpf, M. P. (2013). Model selection in systems and synthetic biology. *Current opinion in biotechnology*, *24*(4), 767-774.
- Krakower, D. S., & Mayer, K. H. (2015). Pre-exposure prophylaxis to prevent HIV infection: current status, future opportunities and challenges. *Drugs*, *75*(3), 243-251. doi:10.1007/s40265-015-0355-4
- Kupfer, B., Matz, B., Daumer, M. P., Roden, F., Rockstroh, J. K., Qurishi, N., . . . Kaiser, R. (2007). Frequent detection of cell-associated HIV-1 RNA in patients with plasma viral load <50 copies/ml. *J Med Virol*, *79*(10), 1440-1445. doi:10.1002/jmv.20993
- Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder--Mead simplex method in low dimensions. *SIAM Journal on optimization*, *9*(1), 112-147.
- Lavielle, M. (2014). *Mixed effects models for the population approach: models, tasks, methods and tools*: Chapman and Hall/CRC.
- Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (Vol. 476): John Wiley & Sons.
- Lee, G. Q., & Lichterfeld, M. (2016). Diversity of HIV-1 reservoirs in CD4+ T-cell subpopulations. *Curr Opin HIV AIDS*, *11*(4), 383-387. doi:10.1097/COH.0000000000000281
- Lewin, S. R., Kirihara, J., Sonza, S., Irving, L., Mills, J., & Crowe, S. M. (1998). HIV-1 DNA and mRNA concentrations are similar in peripheral blood monocytes and alveolar macrophages in HIV-1-infected individuals. *AIDS*, *12*(7), 719-727.
- Li, J. Z., Etemad, B., Ahmed, H., Aga, E., Bosch, R. J., Mellors, J. W., . . . Gandhi, R. T. (2016). The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *AIDS*, *30*(3), 343-353. doi:10.1097/QAD.0000000000000953
- Little, S. J., McLean, a. R., Spina, C. a., Richman, D. D., & Havlir, D. V. (1999). Viral dynamics of acute HIV-1 infection. *The Journal of experimental medicine*, *190*, 841-850. doi:10.1084/jem.190.6.841
- Lixoft SAS. (2018). Monolix (Version 2018R1). Antony, France.
- Lorenzo-Redondo, R., Fryer, H. R., Bedford, T., Kim, E. Y., Archer, J., Pond, S. L. K., . . . Wolinsky, S. M. (2016). Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature*, *530*(7588), 51-56. doi:10.1038/nature16933

- Lund, E. (1965). Guldberg and Waage and the law of mass action. *Journal of Chemical Education*, 42(10), 548.
- Maldarelli, F. (2015). HIV-infected cells are frequently clonally expanded after prolonged antiretroviral therapy: implications for HIV persistence. *J Virus Erad*, 1(4), 237-244.
- Mansky, L. M., & Temin, H. M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol*, 69(8), 5087-5094.
- Margolis, A. M., Heverling, H., Pham, P. A., & Stolbach, A. (2014). A review of the toxicity of HIV medications. *J Med Toxicol*, 10(1), 26-39. doi:10.1007/s13181-013-0325-8
- Markowitz, M., Louie, M., Hurley, A., Sun, E., Di Mascio, M., Perelson, A. S., & Ho, D. D. (2003). A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *J Virol*, 77(8), 5037-5038.
- marquis de Laplace, P. S. (1847). *Oeuvres de Laplace: Théorie analytique des probabilités* (Vol. 7): Imprimerie royale.
- Martin, G. E., Gossez, M., Williams, J. P., Stohr, W., Meyerowitz, J., Leitman, E. M., . . . the S. T. I. (2017). Post-treatment control or treated controllers? Viral remission in treated and untreated primary HIV infection. *AIDS*, 31(4), 477-484. doi:10.1097/QAD.0000000000001382
- Massanella, M., & Richman, D. D. (2016). Measuring the latent reservoir in vivo. *J Clin Invest*, 126(2), 464-472. doi:10.1172/JCI80567
- McKinstry, K. K., Strutt, T. M., & Swain, S. L. (2010). Regulation of CD4+ T-cell contraction during pathogen challenge. *Immunol Rev*, 236, 110-124. doi:10.1111/j.1600-065X.2010.00921.x
- McLeod, G. X., & Hammer, S. M. (1992). Treatment of HIV infection: the antiretroviral nucleoside analogues. Nucleoside analogues: combination therapy. *Hosp Pract (Off Ed)*, 27 Suppl 2, 14-25.
- Melikyan, G. B. (2014). HIV entry: a game of hide-and-fuse? *Curr Opin Virol*, 4, 1-7. doi:10.1016/j.coviro.2013.09.004
- Menendez-Arias, L. (2002). Molecular basis of fidelity of DNA synthesis and nucleotide specificity of retroviral reverse transcriptases. *Prog Nucleic Acid Res Mol Biol*, 71, 91-147.
- Menendez-Arias, L., Sebastian-Martin, A., & Alvarez, M. (2017). Viral reverse transcriptases. *Virus Res*, 234, 153-176. doi:10.1016/j.virusres.2016.12.019
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- Mitsuya, H., & Broder, S. (1987). Strategies for antiviral therapy in AIDS. *Nature*, 325(6107), 773-778. doi:10.1038/325773a0
- Mittrucker, H. W., Visekruna, A., & Huber, M. (2014). Heterogeneity in the differentiation and function of CD8(+) T cells. *Arch Immunol Ther Exp (Warsz)*, 62(6), 449-458. doi:10.1007/s00005-014-0293-y
- Moir, S., Chun, T. W., & Fauci, A. S. (2011). Pathogenic mechanisms of HIV disease. *Annu Rev Pathol*, 6, 223-248. doi:10.1146/annurev-pathol-011110-130254
- Moltó, J., Xinarianos, G., Miranda, C., Pushpakom, S., Cedeño, S., Clotet, B., . . . Valle, M. (2013). Simultaneous pharmacogenetics-based population pharmacokinetic analysis of

- darunavir and ritonavir in HIV-infected patients. *Clinical Pharmacokinetics*, 52, 543-553. doi:10.1007/s40262-013-0057-6
- Moore, R. D. (2011). Epidemiology of HIV infection in the United States: implications for linkage to care. *Clin Infect Dis*, 52 Suppl 2, S208-213. doi:10.1093/cid/ciq044
- Muppidi, J., Porter, M., & Siegel, R. M. (2004). Measurement of apoptosis and other forms of cell death. *Curr Protoc Immunol*, Chapter 3, Unit 3 17. doi:10.1002/0471142735.im0317s59
- Murray, J. M., Emery, S., Kelleher, A. D., Law, M., Chen, J., Hazuda, D. J., . . . Cooper, D. A. (2007). Antiretroviral therapy with the integrase inhibitor raltegravir alters decay kinetics of HIV, significantly reducing the second phase. *AIDS*, 21(17), 2315-2321. doi:10.1097/QAD.0b013e3282f12377
- Mylvaganam, G. H., Silvestri, G., & Amara, R. R. (2015). HIV therapeutic vaccines: moving towards a functional cure. *Curr Opin Immunol*, 35, 1-8. doi:10.1016/j.coi.2015.05.001
- Ng, C. T., Mendoza, J. L., Garcia, K. C., & Oldstone, M. B. (2016). Alpha and Beta Type 1 Interferon Signaling: Passage for Diverse Biologic Outcomes. *Cell*, 164(3), 349-352. doi:10.1016/j.cell.2015.12.027
- Noble, S., & Faulds, D. (1996). Saquinavir. A review of its pharmacology and clinical potential in the management of HIV infection. *Drugs*, 52(1), 93-112.
- Okulicz, J. F., & Lambotte, O. (2011). Epidemiology and clinical characteristics of elite controllers. *Curr Opin HIV AIDS*, 6(3), 163-168. doi:10.1097/COH.0b013e328344f35e
- Okwundu, C. I., Uthman, O. A., & Okoromah, C. A. (2012). Antiretroviral pre-exposure prophylaxis (PrEP) for preventing HIV in high-risk individuals. *Cochrane Database Syst Rev*(7), CD007189. doi:10.1002/14651858.CD007189.pub3
- Ortiz, A. (2013). Derivation of Hill's Equation from Scale Invariance. *Journal of Uncertain Systems*, 7(3), 198-202.
- Osterberg, L., & Blaschke, T. (2005). Adherence to medication. *N Engl J Med*, 353(5), 487-497. doi:10.1056/NEJMra050100
- Owen, J. S., & Fiedler-Kelly, J. (2014). *Introduction to population pharmacokinetic/pharmacodynamic analysis with nonlinear mixed effects models*: John Wiley & Sons.
- Palella, F. J., Jr., Armon, C., Buchacz, K., Cole, S. R., Chmiel, J. S., Novak, R. M., . . . Investigators, H. (2009). The association of HIV susceptibility testing with survival among HIV-infected patients receiving antiretroviral therapy: a cohort study. *Ann Intern Med*, 151(2), 73-84.
- Palmer, S., Maldarelli, F., Wiegand, A., Bernstein, B., Hanna, G. J., Brun, S. C., . . . King, M. S. (2008). Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proc Natl Acad Sci U S A*, 105(10), 3879-3884. doi:10.1073/pnas.0800050105
- Palmer, S., Wiegand, A. P., Maldarelli, F., Bazmi, H., Mican, J. M., Polis, M., . . . Coffin, J. M. (2003). New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol*, 41(10), 4531-4536.
- Palmisano, L., & Vella, S. (2011). A brief history of antiretroviral therapy of HIV infection: success and challenges. *Ann Ist Super Sanita*, 47(1), 44-48. doi:10.4415/ANN_11_01_10

- Papoulis, A., & Pillai, S. U. (2002). *Probability, random variables, and stochastic processes*: Tata McGraw-Hill Education.
- Parkin, N. T., Hellmann, N. S., Whitcomb, J. M., Kiss, L., Chappay, C., & Petropoulos, C. J. (2004). Natural variation of drug susceptibility in wild-type human immunodeficiency virus type 1. *Antimicrobial agents and chemotherapy*, *48*, 437-443. doi:10.1128/AAC.48.2.437
- Pasternak, A. O., Lukashov, V. V., & Berkhout, B. (2013). Cell-associated HIV RNA: a dynamic biomarker of viral persistence. *Retrovirology*, *10*, 41. doi:10.1186/1742-4690-10-41
- Pennings, P. S. (2013). HIV Drug Resistance: Problems and Perspectives. *Infect Dis Rep*, *5*(Suppl 1), e5. doi:10.4081/idr.2013.s1.e5
- Pennock, N. D., White, J. T., Cross, E. W., Cheney, E. E., Tamburini, B. A., & Kedl, R. M. (2013). T cell responses: naive to memory and everything in between. *Adv Physiol Educ*, *37*(4), 273-283. doi:10.1152/advan.00066.2013
- Pereira, E. R., Jones, D., Jung, K., & Padera, T. P. (2015). The lymph node microenvironment and its role in the progression of metastatic cancer. *Semin Cell Dev Biol*, *38*, 98-105. doi:10.1016/j.semcdb.2015.01.008
- Perelson, A. S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., . . . Ho, D. D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*, *387*(6629), 188-191. doi:10.1038/387188a0
- Perko, L. (2013). *Differential equations and dynamical systems* (Vol. 7): Springer Science & Business Media.
- Perry, C. M. (2014). Elvitegravir/cobicistat/emtricitabine/tenofovir disoproxil fumarate single-tablet regimen (Stribild(R)): a review of its use in the management of HIV-1 infection in adults. *Drugs*, *74*(1), 75-97. doi:10.1007/s40265-013-0158-4
- Phair, R. D. (2014). Mechanistic modeling confronts the complexity of molecular cell biology. *Molecular biology of the cell*, *25*(22), 3494-3496.
- Pinkevych, M., Cromer, D., Tolstrup, M., Grimm, A. J., Cooper, D. A., Lewin, S. R., . . . Davenport, M. P. (2015). HIV Reactivation from Latency after Treatment Interruption Occurs on Average Every 5-8 Days--Implications for HIV Remission. *PLoS Pathog*, *11*(7), e1005000. doi:10.1371/journal.ppat.1005000
- Polonsky, M., Chain, B., & Friedman, N. (2016). Clonal expansion under the microscope: studying lymphocyte activation and differentiation using live-cell imaging. *Immunol Cell Biol*, *94*(3), 242-249. doi:10.1038/icb.2015.104
- Popovic, M., Sarngadharan, M. G., Read, E., & Gallo, R. C. (1984). Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*, *224*(4648), 497-500.
- R Development Core Team. (2014). R: A language and environment for statistical testing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://cran.r-project.org/>
- Ramratnam, B., Bonhoeffer, S., Binley, J., Hurley, A., Zhang, L., Mittler, J. E., . . . Ho, D. D. (1999). Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet*, *354*(9192), 1782-1785. doi:10.1016/S0140-6736(99)02035-8

- Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., & Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, *25*(15), 1923-1929.
- Reust, C. E. (2011). Common adverse effects of antiretroviral therapy for HIV disease. *Am Fam Physician*, *83*(12), 1443-1451.
- Ribeiro, R. M., Qin, L., Chavez, L. L., Li, D., Self, S. G., & Perelson, A. S. (2010). Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. *J Virol*, *84*(12), 6096-6102. doi:10.1128/JVI.00127-10
- Riedhammer, C., Halbritter, D., & Weissert, R. (2016). Peripheral Blood Mononuclear Cells: Isolation, Freezing, Thawing, and Culture. *Methods Mol Biol*, *1304*, 53-61. doi:10.1007/7651_2014_99
- Roberts, J. D., Bebenek, K., & Kunkel, T. A. (1988). The accuracy of reverse transcriptase from HIV-1. *Science*, *242*(4882), 1171-1173.
- Rohatgi, A. (2012). WebPlotDigitizer v4.0. Retrieved from <https://automeris.io/WebPlotDigitizer>
- Rosa, M. C., Lobato, R. C., Goncalves, C. V., Silva, N. M., Barral, M. F., Martinez, A. M., & Hora, V. P. (2015). Evaluation of factors associated with vertical HIV-1 transmission. *J Pediatr (Rio J)*, *91*(6), 523-528. doi:10.1016/j.jped.2014.12.005
- Rosenberg, N. E., Pilcher, C. D., Busch, M. P., & Cohen, M. S. (2015). How can we better identify early HIV infections? *Curr Opin HIV AIDS*, *10*(1), 61-68. doi:10.1097/COH.0000000000000121
- Rosenbloom, D. I., Hill, A. L., Rabi, S. A., Siliciano, R. F., & Nowak, M. A. (2012). Antiretroviral dynamics determines HIV evolution and predicts therapy outcome. *Nat Med*, *18*(9), 1378-1385. doi:10.1038/nm.2892
- Rouzioux, C., & Avettand-Fenoel, V. (2018). Total HIV DNA: a global marker of HIV persistence. *Retrovirology*, *15*(1), 30. doi:10.1186/s12977-018-0412-7
- Saez-Cirion, A., Bacchus, C., Hocqueloux, L., Avettand-Fenoel, V., Girault, I., Lecuroux, C., . . . Group, A. V. S. (2013). Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathog*, *9*(3), e1003211. doi:10.1371/journal.ppat.1003211
- Sager, J. E., Yu, J., Ragueneau-Majlessi, I., & Isoherranen, N. (2015). Physiologically Based Pharmacokinetic (PBPK) Modeling and Simulation Approaches: A Systematic Review of Published Models, Applications, and Model Verification. *Drug Metab Dispos*, *43*(11), 1823-1837. doi:10.1124/dmd.115.065920
- Sahu, G. K. (2015). Potential implication of residual viremia in patients on effective antiretroviral therapy. *AIDS Res Hum Retroviruses*, *31*(1), 25-35. doi:10.1089/AID.2014.0194
- Sampah, M. E., Shen, L., Jilek, B. L., & Siliciano, R. F. (2011). Dose-response curve slope is a missing dimension in the analysis of HIV-1 drug resistance. *Proc Natl Acad Sci U S A*, *108*(18), 7613-7618. doi:10.1073/pnas.1018360108
- Sanche, S., Mesplede, T., Sheehan, N. L., Li, J., & Nekka, F. (2018). Exploring an alternative explanation for the second phase of viral decay: Infection of short-lived cells in a drug-limited compartment during HAART. *PLoS One*, *13*(7), e0198090. doi:10.1371/journal.pone.0198090
- Sanche, S., Sheehan, N., Mesplede, T., Wainberg, M. A., Li, J., & Nekka, F. (2017). A Mathematical Model to Predict HIV Virological Failure and Elucidate the Role of

- Lymph Node Drug Penetration. *CPT Pharmacometrics Syst Pharmacol*, 6(7), 469-476. doi:10.1002/psp4.12200
- Sanchez, A., Cabrera, S., Santos, D., Valverde, M. P., Fuertes, A., Dominguez-Gil, A., . . . Tormes, G. (2011). Population pharmacokinetic/pharmacogenetic model for optimization of efavirenz therapy in Caucasian HIV-infected patients. *Antimicrob Agents Chemother*, 55(11), 5314-5324. doi:10.1128/AAC.00194-11
- Santangelo, P. J., Rogers, K. A., Zurla, C., Blanchard, E. L., Gumber, S., Strait, K., . . . Villinger, F. (2015). Whole-body immunoPET reveals active SIV dynamics in viremic and antiretroviral therapy-treated macaques. *Nat Methods*, 12(5), 427-432. doi:10.1038/nmeth.3320
- Savkovic, B., Macpherson, J. L., Zaunders, J., Kelleher, A. D., Knop, A. E., Pond, S., . . . Murray, J. M. (2012). T-lymphocyte perturbation following large-scale apheresis and hematopoietic stem cell transplantation in HIV-infected individuals. *Clinical immunology (Orlando, Fla.)*, 144, 159-171. doi:10.1016/j.clim.2012.06.004
- Schlub, T. E., Grimm, A. J., Smyth, R. P., Cromer, D., Chopra, A., Mallal, S., . . . Davenport, M. P. (2014). Fifteen to twenty percent of HIV substitution mutations are associated with recombination. *J Virol*, 88(7), 3837-3849. doi:10.1128/JVI.03136-13
- Schmeisser, H., Bekisz, J., & Zoon, K. C. (2014). New function of type I IFN: induction of autophagy. *J Interferon Cytokine Res*, 34(2), 71-78. doi:10.1089/jir.2013.0128
- Schneider, H. C., & Klabunde, T. (2013). Understanding drugs and diseases by systems biology? *Bioorg Med Chem Lett*, 23(5), 1168-1176. doi:10.1016/j.bmcl.2012.12.031
- Schraml, B. U., & Reis e Sousa, C. (2015). Defining dendritic cells. *Curr Opin Immunol*, 32, 13-20. doi:10.1016/j.coi.2014.11.001
- Sengupta, S., & Siliciano, R. F. (2018). Targeting the Latent Reservoir for HIV-1. *Immunity*, 48(5), 872-895. doi:10.1016/j.immuni.2018.04.030
- Serrano-Villar, S., Sainz, T., Lee, S. A., Hunt, P. W., Sinclair, E., Shacklett, B. L., . . . Deeks, S. G. (2014). HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and increased risk of non-AIDS morbidity and mortality. *PLoS Pathog*, 10(5), e1004078. doi:10.1371/journal.ppat.1004078
- Shapiro, S. Z. (2014). Clinical development of candidate HIV vaccines: different problems for different vaccines. *AIDS Res Hum Retroviruses*, 30(4), 325-329. doi:10.1089/AID.2013.0114
- Shen, L., Peterson, S., Sedaghat, A. R., McMahon, M. A., Callender, M., Zhang, H., . . . Siliciano, R. F. (2008). Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nat Med*, 14(7), 762-766. doi:10.1038/nm1777
- Shen, L., & Siliciano, R. F. (2008). Viral reservoirs, residual viremia, and the potential of highly active antiretroviral therapy to eradicate HIV infection. *J Allergy Clin Immunol*, 122(1), 22-28. doi:10.1016/j.jaci.2008.05.033
- Shen, P., & Fillatreau, S. (2015). Suppressive functions of B cells in infectious diseases. *Int Immunol*, 27(10), 513-519. doi:10.1093/intimm/dxv037
- Shibutani, S. T., Saitoh, T., Nowag, H., Munz, C., & Yoshimori, T. (2015). Autophagy and autophagy-related proteins in the immune system. *Nat Immunol*, 16(10), 1014-1024. doi:10.1038/ni.3273
- Simon, V., Bloch, N., & Landau, N. R. (2015). Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nat Immunol*, 16(6), 546-553. doi:10.1038/ni.3156

- Sogaard, O. S., Graversen, M. E., Leth, S., Olesen, R., Brinkmann, C. R., Nissen, S. K., . . . Tolstrup, M. (2015). The Dipeptide Romidepsin Reverses HIV-1 Latency In Vivo. *PLoS Pathog*, *11*(9), e1005142. doi:10.1371/journal.ppat.1005142
- Sorger, P. K., Allerheiligen, S. R. B., Abernethy, D. R., Altman, R. B., Brouwer, K. L. R., Califano, A., . . . others. (2011). Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms. *An NIH white paper by the QSP workshop group*.
- Spivak, A. M., Rabi, S. A., McMahon, M. A., Shan, L., Sedaghat, A. R., Wilke, C. O., & Siliciano, R. F. (2011). Short communication: dynamic constraints on the second phase compartment of HIV-infected cells. *AIDS Res Hum Retroviruses*, *27*(7), 759-761. doi:10.1089/AID.2010.0199
- Sun, Y., Fu, Y., Zhang, Z., Tang, T., Liu, J., Ding, H., . . . Jiang, Y. (2017). The investigation of CD4+T-cell functions in primary HIV infection with antiretroviral therapy. *Medicine (Baltimore)*, *96*(28), e7430. doi:10.1097/MD.00000000000007430
- Sundquist, W. I., & Krausslich, H. G. (2012). HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med*, *2*(7), a006924. doi:10.1101/cshperspect.a006924
- Swiggard, W. J., Baytop, C., Yu, J. J., Dai, J., Li, C., Schretzenmair, R., . . . O'Doherty, U. (2005). Human immunodeficiency virus type 1 can establish latent infection in resting CD4+ T cells in the absence of activating stimuli. *J Virol*, *79*(22), 14179-14188. doi:10.1128/JVI.79.22.14179-14188.2005
- Sze, A., Olganier, D., Lin, R., van Grevenynghe, J., & Hiscott, J. (2013). SAMHD1 host restriction factor: a link with innate immune sensing of retrovirus infection. *J Mol Biol*, *425*(24), 4981-4994. doi:10.1016/j.jmb.2013.10.022
- Tang, M. W., & Shafer, R. W. (2012). HIV-1 antiretroviral resistance: scientific principles and clinical applications. *Drugs*, *72*(9), e1-25. doi:10.2165/11633630-000000000-00000
- Tchebichef, P. (1846). Démonstration élémentaire d'une proposition générale de la théorie des probabilités. *Journal für die reine und angewandte Mathematik*, *33*, 259-267.
- The MathWorks Inc. (2014). MATLAB (Version 8.3.0.532 (R2014a)): The MathWorks Inc.
- Treasure, G. C., Aga, E., Bosch, R. J., Mellors, J. W., Kuritzkes, D. R., Para, M., . . . Li, J. Z. (2016). Brief Report: Relationship Among Viral Load Outcomes in HIV Treatment Interruption Trials. *J Acquir Immune Defic Syndr*, *72*(3), 310-313. doi:10.1097/QAI.0000000000000964
- UNAIDS. (2018a). AIDS-related deaths. Retrieved from <http://aidsinfo.unaids.org/>
- UNAIDS. (2018b). UNAIDS Data. Retrieved from <http://www.unaids.org/en/topic/data>
- Valade, E., Treluyer, J. M., Bouazza, N., Ghosn, J., Foissac, F., Benaboud, S., . . . Hirt, D. (2014). Population pharmacokinetics of emtricitabine in HIV-1-infected adult patients. *Antimicrob Agents Chemother*, *58*(4), 2256-2261. doi:10.1128/AAC.02058-13
- Varol, C., Mildner, A., & Jung, S. (2015). Macrophages: development and tissue specialization. *Annu Rev Immunol*, *33*, 643-675. doi:10.1146/annurev-immunol-032414-112220
- Veazey, R. S., & Lackner, A. a. (2004). Getting to the guts of HIV pathogenesis. *The Journal of experimental medicine*, *200*, 697-700. doi:10.1084/jem.20041464
- Voit, E. O., Martens, H. A., & Omholt, S. W. (2015). 150 years of the mass action law. *PLoS computational biology*, *11*(1), e1004012.
- Vrijens, B., De Geest, S., Hughes, D. A., Przemyslaw, K., Demonceau, J., Ruppar, T., . . . Team, A. B. C. P. (2012). A new taxonomy for describing and defining adherence to

- medications. *Br J Clin Pharmacol*, 73(5), 691-705. doi:10.1111/j.1365-2125.2012.04167.x
- Wainberg, M. A., Rooke, R., Tremblay, M., Li, X., Parniak, M. A., Gao, Q., . . . Ruedy, J. (1991). Clinical significance and characterization of AZT-resistant strains of HIV-1. *Can J Infect Dis*, 2(1), 5-11.
- Walker, B., & McMichael, A. (2012). The T-cell response to HIV. *Cold Spring Harb Perspect Med*, 2(11). doi:10.1101/cshperspect.a007054
- Walker, B. D., & Yu, X. G. (2013). Unravelling the mechanisms of durable control of HIV-1. *Nat Rev Immunol*, 13(7), 487-498. doi:10.1038/nri3478
- Walker, R. E., Parker, R. I., Kovacs, J. A., Masur, H., Lane, H. C., Carleton, S., . . . Fauci, A. S. (1988). Anemia and erythropoiesis in patients with the acquired immunodeficiency syndrome (AIDS) and Kaposi sarcoma treated with zidovudine. *Ann Intern Med*, 108(3), 372-376.
- Whaley, K. J., & Mayer, K. H. (2014). Strategies for preventing mucosal cell-associated HIV transmission. *J Infect Dis*, 210 Suppl 3, S674-680. doi:10.1093/infdis/jiu398
- Whitney, J. B., Hill, A. L., Sanisetty, S., Penaloza-MacMaster, P., Liu, J., Shetty, M., . . . Barouch, D. H. (2014). Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature*, 512(7512), 74-77. doi:10.1038/nature13594
- Wiedle, G., Dunon, D., & Imhof, B. a. (2001). Current concepts in lymphocyte homing and recirculation. *Critical reviews in clinical laboratory sciences*, 38, 1-31. doi:10.1080/20014091084164
- Williams, J. P., Hurst, J., Stohr, W., Robinson, N., Brown, H., Fisher, M., . . . Investigators, S. P. (2014). HIV-1 DNA predicts disease progression and post-treatment virological control. *Elife*, 3, e03821. doi:10.7554/eLife.03821
- Wolfram Research Inc. (2014). Mathematica (Version 10.1). Champaign, Illinois: Wolfram Research Inc.
- Wong, D., Modi, R., & Ramanathan, M. (2003). Assessment of Markov-dependent stochastic models for drug administration compliance. *Clinical Pharmacokinetics*, 42, 193-204. doi:10.2165/00003088-200342020-00006
- Yakowitz, S., Krimmel, J., & Szidarovszky, F. (1978). Weighted monte carlo integration. *SIAM Journal on Numerical Analysis*, 15(6), 1289-1300.
- Yilmaz, A., Price, R. W., & Gisslen, M. (2012). Antiretroviral drug treatment of CNS HIV-1 infection. *J Antimicrob Chemother*, 67(2), 299-311. doi:10.1093/jac/dkr492
- Young, S. D., Britcher, S. F., Tran, L. O., Payne, L. S., Lumma, W. C., Lyle, T. A., . . . et al. (1995). L-743, 726 (DMP-266): a novel, highly potent nonnucleoside inhibitor of the human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents Chemother*, 39(12), 2602-2605.
- Yukl, S. A., Shergill, A. K., Ho, T., Killian, M., Girling, V., Epling, L., . . . Wong, J. K. (2013). The distribution of HIV DNA and RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral persistence. *J Infect Dis*, 208(8), 1212-1220. doi:10.1093/infdis/jit308
- Zhang, L., Dailey, P. J., He, T., Gettie, A., Bonhoeffer, S., Perelson, A. S., & Ho, D. D. (1999). Rapid clearance of simian immunodeficiency virus particles from plasma of rhesus macaques. *J Virol*, 73(1), 855-860.

- Zheng, L., Bosch, R. J., Chan, E. S., Read, S., Kearney, M., Margolis, D. M., . . . Team, D. S. C. T. G. A. (2013). Predictors of residual viraemia in patients on long-term suppressive antiretroviral therapy. *Antivir Ther*, *18*(1), 39-43. doi:10.3851/IMP2323
- Zhou, S., Martin, K., Corbett, A., Napravnik, S., Eron, J., Zhu, Y., . . . Wohl, D. A. (2014). Total daily pill burden in HIV-infected patients in the southern United States. *AIDS Patient Care STDS*, *28*(6), 311-317. doi:10.1089/apc.2014.0010