Improving and Validating Data-Driven Genotypic Interpretation Systems for the Selection of Antiretroviral Therapies

Alejandro Pironti

Tag des Kolloquiums	07.12.2016
Dekan	Prof. Dr. Frank-Olaf Schreyer
Vorsitzender der Prüfungskomission	Prof. Dr. Gerhard Weikum
Erster Berichtserstatter	Prof. Dr. Thomas Lengauer
Zweiter Berichtserstatter	Prof. Dr. Hans-Peter Lenhof
Akademischer Mitarbeiter	Dr. Markus List

Thesis advisor: Prof. Dr. Dr. Thomas Lengauer   Alejandro Pironti

Improving and Validating Data-Driven Genotypic Interpretation Systems for the Selection of Antiretroviral Therapies

Short Summary

Infection with *Human immunodeficiency virus type 1* (HIV-1) requires treatment with a combination of antiretroviral drugs. This combination of drugs must be selected under consideration of its prospects for attaining sustained therapeutic success. Genotypic therapy-success interpretation systems can be used for selecting a combination of antiretroviral compounds. However, a number of shortcomings of these systems have prevented them from reaching the bedside.

In this work, I present and validate novel methods for deriving interpretable genotype interpretation systems that are trained on HIV-1 data from routine clinical practice. One method produces scores that are correlated with previous exposure of the virus to the drug and with drug resistance. A further, novel genotype interpretation system produces a prognostic score correlated with the time for which the antiretroviral therapy with a certain drug combination will remain effective.

The methods presented in this work represent an important advance in techniques for the interpretation of viral genotypes. Validation of the methods shows that their performance is comparable or, most frequently, superior to that of previously available methods. The methods are interpretable and can be retrained without the need for expert intervention. Last but not least, long-term therapeutic success is considered by the methods such that their predictions are in line with the results of clinical studies.

Kurzfassung

Eine Infektion mit dem Humanen Immunodefizienz-Virus Typ 1 (HIV-1) erfordert die Behandlung des Patienten mit einer Kombination von antiretroviralen Wirkstoffen. Die Auswahl dieser Wirkstoffkombination muss unter Berücksichtigung der Aussichten für einen lang anhaltenden Behandlungserfolg stattfinden. Bei der Auswahl von Wirkstoffkombinationen können Systeme zur Vorhersage des Behandlungserfolgs eingesetzt werden. Bisher verfügbare Systeme weisen jedoch mehrere Defizite auf, sodass sie in der klinischen Praxis kaum Verwendung finden.

In dieser Arbeit werden neuartige Methoden zur Aufstellung von Systemen zur Genotypinterpretation präsentiert und validiert. Eine dieser Methoden bewertet einen HIV-1-Genotyp bezüglich der vorhergehenden viralen Wirkstoffexposition und der Wirkstoffresistenzen. Eine weitere Genotypinterpretationsmethode errechnet eine prognostische Zahl, welche mit der Zeit korreliert, die eine antiretrovirale Therapie effektiv sein wird.

Diese Arbeit stellt eine wichtige Weiterentwicklung der Methoden zur Interpretation von viralen Genotypen dar. Zum Einen ist das Vorhersagemögen der Modelle dieser Arbeit vergleichbar oder sogar höher als diejenige von bisher verfügbaren Modellen. Zum Anderen sind die Modelle dieser Arbeit interpretierbar und können ohne Expertensupervision neu trainiert werden. Darüber hinaus berücksichtigen die Methoden den Langzeittherapieerfolg, sodass ihre Vorhersagen mit den Ergebnissen klinischer Studien übereinstimmen.

Thesis advisor: Prof. Dr. Dr. Thomas Lengauer   Alejandro Pironti

Improving and Validating Data-Driven Genotypic Interpretation Systems for the Selection of Antiretroviral Therapies

Abstract

Infection with *Human immunodeficiency virus type 1* (HIV-1) requires treatment with antiretroviral drugs. Without treatment, patients with HIV-1 infection develop symptoms referred to as acquired immunodeficiency syndrome (AIDS), ultimately leading to the death of the patient. The high productivity and variability of HIV-1 results in the continuous emergence of drug-resistant viral variants. In order to be able to suppress viral replication, several drug compounds must be used simultaneously in antiretroviral therapy. For this reason, a combination of drug compounds must be selected under consideration of the drug resistance of the virus and of the prospects that the drug combination has for attaining sustained therapeutic success.

Genotypic drug-resistance interpretation systems are frequently used for selecting combinations of antiretroviral drug compounds. These systems interpret HIV-1 genotypes in order to predict the susceptibility of the virus to each individual antiretroviral drug. However, the actual selection of an optimal drug combination needs to be carried out by the treating physician. In contrast, genotypic therapy-success interpretation systems provide their users with predictions for the success of antiretroviral drug combinations. However, a number of shortcomings of these systems have prevented them from reaching the bedside.

In this work, I present and validate novel methods for deriving genotype interpretation systems that are trained on HIV-1 data from routine clinical practice. All of these systems provide the user with an interpretation of their predictions. One system produces numbers called *drug exposure scores* (DES) for each available antiretroviral drug. DES are correlated with previous exposure of the virus to the drug and with drug resistance. I also present and validate methods for converting DES into clinically meaningful categories, such that they can readily be used by human experts for selecting optimal antiretroviral therapies. DES can be used as features for further analyses relating to antiretroviral therapy. I present a further, novel genotype interpretation system that is trained on DES to produce a prognostic score correlated with the time for which the antiretroviral therapy with a certain drug combination will remain effective.

The methods presented in this work represent an important advance in techniques for the interpretation of viral genotypes. Validation of the methods shows that their performance is comparable or, most frequently, superior to that of previously available methods. Their data-driven methodology allows for automatic retraining without the need for expert intervention. Their interpretability helps them gain the confidence of the users and delivers a rationale for predictions that could be considered surprising. Last but not least, the ability of the therapy-success interpretation system to consider cumulative, long-term therapeutic success allows it to produce predictions that are in line with the results of clinical studies.

Acknowledgments

First and foremost, I would like to thank my supervisor, Prof. Dr. Thomas Lengauer, PhD, for the following reasons. First, Thomas created a fantastic environment in which people can perform research and learn how to perform research. Second, Thomas invited to me to be part of his research environment. Third, Thomas supervised my research personally. His style of supervision provided me with an exceptional amount of freedom while imparting me all concepts needed for attaining scientific rigor. Fourth, over the course of many years, Thomas provided me with a great amount of support and encouragement while challenging me to thrive as a scientist. Last but not least, Thomas provided me with an example of how a scientist and human being can and should be.

I am thankful to Dr. Nico Pfeifer for advising me during my doctoral thesis. I would like to mention that my statistical rigor improved thanks to Nico. I owe gratitude to Dr. Rolf Kaiser who imparted me a significant amount of virological knowledge and welcomed me in his extensive network of scientists. I am truly indebted to all the people who invested a lot of energy into improving my abilities as a scientist. Dr. Hauke Walter let me profit from his impressive memory such that I could attain an intuitive understanding of the developments that occurred in HIV field during the last twenty years. Furthermore, Hauke very frequently shared his research ideas with me. Dr. Björn Jensen invested considerable energy into making me understand the way a physician thinks and into making me understand what is important when selecting an antiretroviral therapy. I attended several conferences in which Dr. Martin Obermeier was present as well. Martin was always happy to explain to me any biomedical concept of which I had not previously heard. I am indebted to Prof. Dr. Hans-Peter Lenhof for imparting me the basic principles of computational biology and for refereeing this thesis.

I am much obliged to all my collaborators without whom my research would not have been possible. I would like to mention the names of the scientists who have worked under the supervision of Dr. Rolf Kaiser: Dr. Melanie Balduin, Dr. Eva Heger, Dr. Elena Knops, Dr. Nadine Lübke, Eugen Schülter, Dr. Finja Schweizer, Dr. Saleta Sierra, and Dr. Jens Verheyen. I am grateful to Claudia Müller, Dr. Stefan Scholten, and Dr. Mark Oette. I am indebted to all members of the EuResist Network. I would like to mention the following names of members of the EuResist Network: Dr. Francesca Incardona, Dr. Ricardo Camacho, Dr. Mattia Prosperi, Dr. Anders Sönnerborg, Dr. Anne-Mieke Vandamme, Dr. Maurizio Zazzi. I am indebted to Dr. Sabina Mugusi for providing me with a strong connection to reality by showing me around in Muhimbili National Hospital. I thank Martin Däumer for sharing his knowledge on HIV with me. I am thankful to Dr. Norbert Bannert, to Dr. Andrea Hauser, and to Dr. Claudia Kücherer.

I am thankful to all past and present members of the Department of Computational Biology and Applied Algorithmics of the Max Planck Institute for Informatics. I am indebted to the following persons for proofreading my thesis: Felipe Albrecht, Dr. Nadezhda Doncheva, Mattias Döring, Peter Ebert, Anna Feldmann, Tim Kacprowski, Prabhav Kalaghatghi, Dr. Markus List, Fabian Müller, and Sarvesh Nikumbh. I thank my office mates Dr. Nadezhda Doncheva and Mattias Döring for the valuable scientific discussions we had and for making my everyday life more fun. I would like to further mention Dr. Mazen Ahmad, Dr. Mario Al-

Contents

List of Abbreviations

/c cobicistat boosting dose.

/r ritonavir boosting dose.

3TC lamivudine.

ABC abacavir.

AIDS acquired immunodeficiency syndrome.

APV amprenavir.

ARS acute retroviral syndrome.

ATV atanazavir.

AUC area under the receiver operating characteristic curve.

AZT azidothymidine, also called zidovudine (ZDV).

CA capsid.

cART combination antiretroviral therapy.

CCR5 C-C chemokine receptor type 5.

CD4 cluster of differentiation 4.

cDNA complementary deoxyribonucleic acid.

CRF circulating recombinant form.

CTL cytotoxic T lymphocyte.

CXCR4 C-X-C chemokine receptor type 4.

d4T stavudine.

ddC zalcitabine.

ddI didianosine.

DES drug-exposure score(s).

DLV delavidrine.

DNA   deoxyribonucleic acid.

DRC   Democratic Republic of Congo.

DRV   darunavir.

DTG   dolutegravir.

EFV   efavirenz.

EI   entry inhibitor.

EIDB   EuResist Integrated Database.

env   envelope.

ERM   empirical risk minimization.

ETR   etravirine.

EVG   elvitegravir.

FC   fold-change in the 50% inhibitory drug concentration.

FDA   Federal Drug Administration.

FI   fusion inhibitor.

FPR   false-positive rate.

FPV   fosamprenavir.

FTC   emtricitabine.

gag   group-specific antigen.

GPP   genotype-phenotype pair.

GSS   genetic susceptibility score.

GTHP   genotype-therapy-history pair.

HAART   highly active antiretroviral therapy.

HIV-1   *Human immunodeficiency virus type 1.*

HIV-2   *Human immunodeficiency virus type 2.*

IDV   indinavir.

IN   integrase.

INI integrase strand-transfer inhibitor.

ITT intent-to-treat.

kb kilo base pairs.

KDE kernel density estimation.

LANLSD Los Alamos National Laboratory Sequence Database.

LPV lopinavir.

MA matrix.

MISE mean integrated square error.

mRNA messenger ribonucleic acid.

MSE mean square error.

MTT 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide.

MVC maraviroc.

NAS number(s) of aviremic semesters.

NC nucleocapsid.

Nef negative regulatory factor.

NFV nelfinavir.

NNRTI non-nucleoside reverse-transcriptase inhibitor.

NRTI nucleotide or nucleoside reverse-transcriptase inhibitor.

NVP nevirapine.

OT on-treatment.

PBMC peripheral blood mononuclear cell.

PCR polymerase chain reaction.

PEP post-exposure prophylaxis.

PI protease inhibitor.

pNAS predicted number(s) of aviremic semesters.

POE probability of exposure to a drug.

pol polymerase.

PON probability of naïvety.

PR protease.

PrEP pre-exposure prophylaxis.

RAL raltegravir.

Rev regulator of expression of virion proteins.

RF resistance factor.

RKHS reproducing kernel Hilbert space.

RNA ribonucleic acid.

RPV rilpivirine.

RT reverse transcriptase.

RT-PCR reverse-transcription polymerase chain reaction.

RTI reverse-transcriptase inhibitor.

RTV ritonavir.

SD standard deviation.

SIR susceptible-intermediate-resistant.

SIV *Simian immunodeficiency virus*.

SQV saquniavir.

SVC Support Vector classifier.

SVM Support Vector machine.

T thymus.

T-20 enfuvirtide.

TAF tenofovir alafenamide fumarate.

Tat transactivator of transcription.

TCE treatment-change episode.

TDF tenofovir disproxil fumarate.

TDR transmitted drug resistance.

TE treatment episode.

TFE treatment-failure episode.

TFV tenofovir.

tMRCA time to most common recent ancestor.

TPV tipranaivr.

URF unique recombinant form.

V3 third hypervariable loop of gp120.

Vif viral infectivity factor.

VL viral load.

Vpr viral protein R.

Vpu viral protein unique.

ZDV zidovudine, also called azidothymidine (AZT).

List of Figures

List of Tables

When you give rise to that which is within you, what you have will save you. If you do not give rise to it, what you do not have will destroy you.

GTh, Nag Hammadi Library. Davies Translation.

0
Introduction

Since the first antiretroviral drug was approved by the Federal Drug Administration (FDA) in 1987, antiretroviral therapy for treating *Human immunodeficiency virus type 1* (HIV-1) infection has made remarkable progress. During its early days, the success of antiretroviral therapy was often hampered by (1) the development of drug resistance by the virus and by (2) the toxic effects of antiretroviral drugs. Over the last three decades, the frequency with which these two problems occur has declined steadily. Therefore, the rates of success of antiretroviral therapy have steadily increased, to such an extent that HIV-1 infection no longer amounts to a death sentence for those affected. Modern antiretroviral therapy has transformed HIV-1 infection into a chronic disease. Nevertheless, antiretroviral drugs cannot cure HIV-1 infection, such that life-long treatment is necessary. Improvements in antiretroviral therapy have been achieved through (1) the discovery and release of better antiretroviral drugs and (2) the continuous acquisition of knowledge on how to use antiretroviral drugs. In comparison to older antiretroviral drugs, newer antiretroviral drugs present a more favorable side-effect profile and are less prone to reductions in efficacy due to resistance development. Furthermore, novel drug formulations have augmented intake convenience for the patient by reducing both the size and number of tablets that have to be ingested each day, as well as the number of daily doses.

Successful antiretroviral treatment requires the simultaneous intake of several drug compounds. Therapeutic success (or failure) results from the interaction of the chosen drug compounds with the virus and with the body of the patient. Drug combinations act in concert, such that the effect of a given drug compound may change if the coadministered drug compounds are changed. For these reasons, *knowledge on the effective use of antiretrovirals*, i.e. knowledge on which antiretroviral drug compounds can be used in a given setting, and on how these compounds can be combined, is decisive for attaining therapeutic success. Furthermore, knowledge on the effective use of antiretrovirals is not only important for attaining therapeutic success, but also for preventing the spread of drug-resistant HIV-1 variants. It is known that inappropriate use of antimicrobial drugs can result in the development of drug resistance by microbes. HIV-1 is capable of developing resistance to antiretroviral drugs. Resistance to a certain antiretroviral drug may entail resistance to other, similar antiretroviral drugs, a phenomenon called cross resistance. At the same time, resistant viral variants can be transmitted from one host to another. Thus, the emergence and transmission of HIV-1 variants that are simultaneously resis-

tant against many (or even all) antiretroviral drugs is possible. The inappropriate use of antiretroviral drugs can increase the frequency with which this undesirable scenario takes place.

In order to effectively select an appropriate drug combination for a patient, the following information is necessary. First, the drug susceptibility of the patient's viral variant(s) must be known, since drug resistance can impede therapeutic success. The results of drug-resistance testing are intended to help the treating physician to exclude ineffective drugs. For this reason, they are frequently conveyed to the physician in a simplified manner such that they can be easily understood and applied to the drug-selection process. Specifically, the results of drug resistance testing are frequently stated in terms of resistance categories, although drug resistance is a continuum which is better expressed in a quantitative form. Additionally, drug resistance is stated for each antiretroviral drug separately, in spite of the fact that antiretroviral drug combinations act in concert. Second, patient-specific characteristics must be considered in order to exclude drugs that could be unfavorable for the patient. For example, hypersensitivity of the patient to a specific drug compound can preclude its use in therapy. Third, the overall propensities of antiretroviral drug combinations to elicit therapeutic success must be known, i.e. the probability that a certain drug combination will lead to short- or long-term therapeutic success regardless of the patient in question. This information is frequently obtained from governmental agencies regulating the prescription of antiretrovirals, from the scientific literature, from discussion with scientists and other treating clinicians, and from experience. Once the information on drug resistance, relevant patient-specific factors, and efficacy of drug combinations becomes available, the treating physician must integrate it in order to select a promising drug combination. In order to make this task feasible, simplification of the information may be necessary, which can result in a loss of accuracy when judging which drug combinations are effective and which are ineffective. Furthermore, effective drug combinations can present different degrees of efficacy. An antiretroviral therapy is considered more effective if it can control viral replication for a longer time while causing fewer and less severe side effects. Obviously, clinicians strive for selecting the most effective drug combination for a certain patient. In light of the process with which drug combinations are selected, it can be said that an improvement of the tools clinicians use for selecting antiretroviral therapies will result in the selection of more effective therapies, which translates in a direct benefit for the patient.

Since HIV-1 drug resistance results from alterations in the genetic material of the virus, the quantification of drug resistance can be accomplished through the interpretation of the genotype. For this purpose, genotypic drug-resistance interpretation systems are employed. Given an HIV-1 genotype, these systems output an assessment of the susceptibility of the viral variant to different antiretroviral drugs. Further developments of techniques for the interpretation of HIV-1 genotypes led to the emergence of genotypic *therapy-success* interpretation systems. These systems do not quantify the susceptibility of HIV-1 to individual drugs, but provide the user with an assessment of the prospects of different drug combinations. They reach this goal by integrating information on drug resistance and on the efficacy of different drug combinations. Therefore, they are a means of improving the process with which antiretroviral therapies are selected. While genotypic drug-resistance interpretation systems are routinely used for selecting antiretroviral therapies, therapy-success prediction systems are used less frequently and have not yet entered clinical routine broadly.

There are several facts that contribute to the lack of popularity of genotypic *therapy-success* interpretation systems. First, none of the available systems provide the user with interpretations for the predictions they make. Interpretability can boost the confidence of the users in a system and is of special relevance when the system makes predictions that strongly deviate from predictions obtained with established methods. In the case of such a deviation, the user is given little or no support for deciding whether the deviation can be deemed plausible or

not. Second, most available systems are constrained, by design, to provide predictions for a specific group of patients, e.g. treatment-experienced patients. Third, the systems often provide predictions that are in stark contradiction with the results of clinical studies. For example, combinations including drugs that have been deemed too toxic for use in most cases may receive an overly positive rating. This usually results from the lack of ability of the system to consider the long-term outcome of therapies. This work describes several improvements to genotypic therapy-success interpretation that contribute to their establishment as therapy selection tools in routine clinical practice.

0.1 Outline

This work is structured as follows. Chapter 1 provides the reader with a summary of important biological and medical concepts relating to HIV-1 infection. Furthermore, the history of the HIV-1 pandemic and of antiretroviral treatment are recounted in this chapter. Information on the structure, morphology and replication cycle of HIV-1 can be found in Section 1.1. The features of the transmission of HIV-1 as well as the clinical course of an untreated HIV-1 infection are reviewed in Section 1.2. Section 1.3 briefly narrates which events are thought to have led to the HIV-1 pandemic, and describes the evolution of the pandemic over time. The diversity of HIV-1 is dealt with in Section 1.4. Lastly, Section 1.5 provides the reader with an account of the history of antiretroviral therapy and with an introduction to the mechanisms of action of antiretroviral compounds, the development of drug resistance, and the future prospects of antiretroviral chemotherapy.

A summary on principles and methods for learning from datasets is given in Chapter 2. Section 2.1 provides the reader with a brief summary of the inductive principles that are commonly used by these methods. Then, Sections 2.2 and 2.3 review Support Vector machines (SVMs) and kernel density classification, which are methods used in this work.

Several methods relating to the interpretation of genotypes of HIV-1 with respect to drug resistance are presented in Chapter 3. Section 3.1 describes a popular wet-lab method with which HIV-1 drug-resistance data can be generated. Then, the interpretation of HIV-1 genotypes with respect to drug resistance is reviewed in Section 3.2. In Section 3.3, I present and validate a novel method for predicting therapeutic history and drug resistance from the genotype. The remainder of the chapter is devoted to describing and validating methods for obtaining cutoffs for continuous predictions related to drug resistance. The purpose of cutoffs is the translation of these quantities into categories that indicate their clinical meaning.

Chapter 4 is devoted to interpretation systems for predicting therapeutic success. The chapter starts by summarizing established definitions of therapeutic success (Section 4.1). After this, different therapy-success prediction systems are reviewed in Section 4.2. Subsequently, I present and validate a novel therapy-success prediction system. The method uses a measure for cumulative therapeutic success, such that it can produce long-term predictions for the course of a therapy. Furthermore, the method features interpretability. I propose a graphical depiction of the model weights in order to provide the user with an interpretation of the prediction, as well as a novel tree-based representation of therapy combinations that can make the process of comparing different therapy options more efficient.

The use of statistical models in a biomedical context dictates increased requirements for the robustness of these models. In Chapter 5, I first give an overview on the sources from which variability arises in the process of Sanger sequencing. Then, I describe a novel method for simulating the variability that is inherent to Sanger sequencing. Lastly, I present an analysis on the robustness of two genotypic interpretation systems used in the

context of antiretroviral therapy.

This work is concluded with Chapter 6. In this chapter, I recount the main achievements and provide and outlook for further improvements.

1

Biomedical Background

THIS CHAPTER AIMS AT PROVIDING essential background knowledge on HIV-1, including its origins and replication cycle, as well as its pathogenic nature and currently available therapeutic options. In the final sections of this chapter, I briefly present future prospects of antiretroviral therapy that are foreseeable. Many facts included in this chapter have been extracted from the following texts [1–3]. The summary on HIV-1 and antiretroviral therapy that is presented in this chapter is by no means exhaustive. The interested reader can consult the cited references in order to obtain more detailed information on the subject.

1.1 HIV-1: STRUCTURE AND REPLICATION CYCLE

HIV-1 is a virus that belongs to genus *Lentivirus* in the family *Retroviridae* [4]. As such, HIV-1's genome is encoded by single-stranded positive-sense ribonucleic acid (RNA), that is reversely transcribed into double-stranded deoxyribonucleic acid (DNA) by a viral RNA/DNA-dependent DNA polymerase, also called reverse transcriptase (RT). Viral DNA is then integrated into the genome of the host cell, which allows for messenger ribonucleic acid (mRNA) transcription for the production of new viral particles. Further characteristics that HIV-1 shares with other lentiviruses include [1, 2]

- Association with a disease with a long incubation period
- Impairment of the immune, hematopoetic, and central nervous systems
- Host-species specificity
- Selective cytopathogenicity
- Latent or persistent infection in certain cells
- Three main genes coding for viral proteins and appearing in the order 5'-gag-pol-env-3'
- Highly glycosylated envelope gene (Figure 1.2).

In the next sections, I provide a summary of the structure and replication cycle of HIV-1.
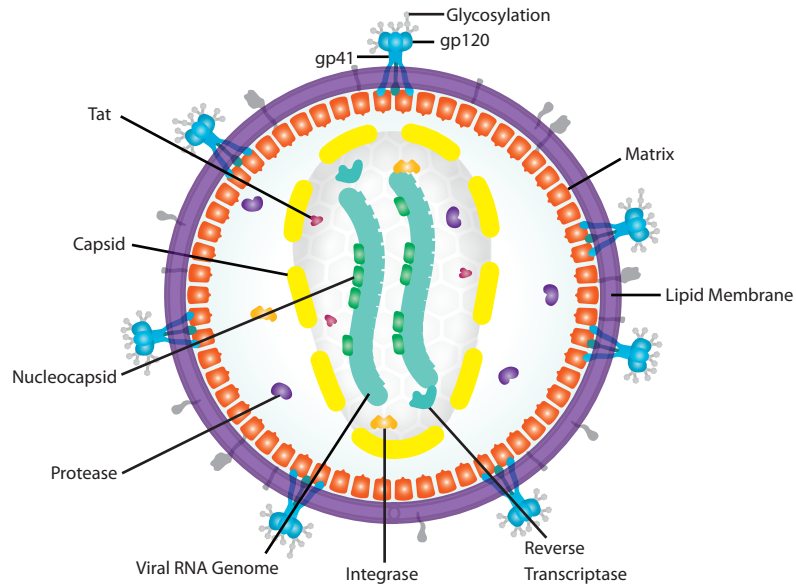
Figure 1.1: Morphology of HIV-1. A schematic representation of the morphology of HIV-1 can be found above. Source: Wikimedia Commons, with modifications.

1.1.1 STRUCTURE OF HIV-1

The morphology of HIV-1 concerns the way viral particles are constructed. The blueprint for the construction of these particles is contained in the genome of HIV-1. The genome of HIV-1 presents a distinct arrangement of genes and open reading frames. In the following, I summarize the main features of the morphology and genomic organization of HIV-1.

MORPHOLOGY OF HIV-1

Most HIV-1 proteins have a distinctive name. However, some of them are only known by their molecular weight in kilodaltons, which is appended to the letter *p* for protein or *gp* for glycoprotein. That being said, I proceed to describe the morphological structure of an HIV-1 particle. HIV-1 is a diploid virus. Specifically, the viral particles of HIV-1 enclose two copies of the single-stranded viral RNA, along with three viral enzymes that are necessary for viral replication: RT (consists of two subunits, p66+p51), integrase (IN, p31), and protease (PR, p15; Figure 1.1). In the viral particles, HIV-1 RNA is closely associated with the proteins RT and nucleocapsid (NC, p7). Viral RNA is enclosed along with other proteins in a conical structure made out of the capsid (CA, p24) protein. This conical structure is called the core or capsid. The proteins viral infectivity factor (Vif) and negative regulatory factor (Nef) are closely associated to the capsid, while viral protein R (Vpr) is likely to be located within the viral particle, but outside of the capsid. The capsid is surrounded by a phospholipid layer, whose inner side is associated to various copies of the viral matrix (MA; p17) protein. This phospholipid layer integrates trimers of two envelope (env) proteins: surface glycoprotein gp120, which remains outside the viral particle, and transmembrane glycoprotein gp41, which has a transmembrane domain. An overview of the known function of HIV-1 proteins can be seen in Table 1.1. See also *HIV-1 Replication Cycle* below.

Table 1.1: HIV-1 Protein Function. A summary of the function of most HIV-1 proteins is tabulated below. Gag *spacer proteins* SP1 and SP2 were excluded since little is known about their function. HLA: human leukocyte antigen.

Protein Name	Designation by Size (kDa)	Function
Capsid (CA)	p24	Structural protein [1]; Important for viral particle maturation [5]
Matrix (MA)	p17	Membrane targeting and assembly of viral precursor polyproteins gag and gag-pol [2]
Nucleocapsid (NC)	p7	Nucleic acid chaperone during reverse transcription [5]; Encapsulation of viral RNA [5];
gag p6	p6	Release of viral particles from cell surface [5]
Reverse transcriptase (RT)	p66+p51	Transcription of viral RNA into DNA [1]
Protease (PR)	p10	Cleavage of gag and gag-pol precursor polyproteins [2]
Integrase (IN)	p32	Integration of viral DNA into host genome [5]
Envelope surface protein	gp120	Entry into cell [1]
Envelope transmembrane protein	gp41	Entry into cell [1]
Transactivator (tat)	p14	Transcriptional activator of integrated provirus [2]
Rev	p19	Transport of viral mRNA out of cell nucleus [2, 5]; Regulation of expression of structural and regulatory HIV-1 genes [1]
Nef	p27	Downregulation of CD4 and HLA class I molecules [2]
Viral infectivity factor (Vif)	p23	Inhibits degradation of proviral DNA by the host factor APOBEC3G [2]
Vpr	p15	Component of the preintegration complex [5, 6]
Vpu	p16	Mediates virus release [1, 2]; Degradation of gp160:CD4 complexes [1, 2]
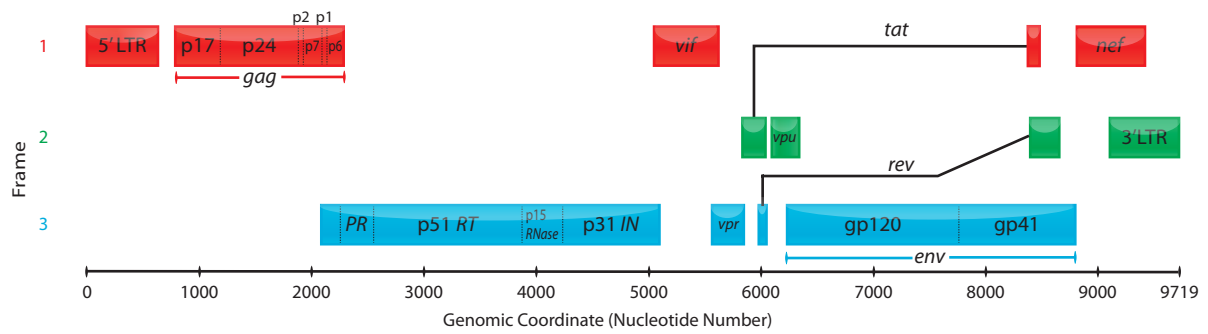
Figure 1.2: Genomic Organization of HIV-1. Above, rectangles represent open reading frames. These rectangles are colored by the number of the reading frame (first, second, or third) in which the corresponding gene is encoded. Note that the open reading frames of Tat and Rev are represented by their spliced exons. After http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

GENOMIC ORGANIZATION OF HIV-1

HIV-1 shares the viral genes group-specific antigen (*gag*), polymerase (*pol*), and envelope (*env*) with other retroviruses [7]. Additionally, the genome of HIV-1 includes the accessory genes *Vif*, viral protein unique (*Vpu*), viral protein R (*Vpr*), transactivator of transcription (*Tat*), regulator of expression of virion proteins (*Rev*), and *Nef* (Figure 1.2). In the following, I describe the genomic organization of HIV-1. The size of the genome of HIV-1 is 9.2 kilo base pairs (kb), and a schematic representation of its organization can be seen in Figure 1.2. The primary transcript of HIV-1 is a full length mRNA which is translated into the gag and pol polyproteins [1]. Through proteolytic cleavage by PR, the viral proteins CA, MA, NC, p6, p2, and p1 are produced from the gag precursor polyprotein (p55). In the same way, autocleavage of the pol polyprotein gives rise to the viral proteins PR, RT, and IN. In 95% of viral mRNA translation events, a gag polyprotein is produced. However, in 5% of viral mRNA translation events, a 1-ribosomal frame shift occurs, and a gag-pol polyprotein is translated, which is cleaved into proteins encoded by both *gag* and *pol* genes. Translation of the env polyprotein, gp160, requires a singly spliced mRNA from the full-length viral mRNA. However, proteolytic cleavage of gp160 into gp41 and gp120 is not catalyzed by PR, but by furin, a cellular endoprotease [1].

1.1.2 HIV-1 REPLICATION CYCLE

The replication cycle of HIV-1 is depicted in Figure 1.3. Before explaining the details of the replication cycle of HIV-1, I would like to briefly explain the concept of *viral tropism*. As with any virus, execution of the replication cycle of HIV-1 requires the infection of a cell by HIV-1. For a viral infection to occur, HIV-1 needs to encounter a suitable cell, which happens by chance as the virus roams in blood and other body compartments. The adjective *suitable* in the preceding sentence makes reference to the fact that, in general, viruses cannot infect just any cell they encounter since they have evolved to infect only certain types of cells. The process by which a virus discerns between target cells and other cells is governed by the interaction of the virion with the *receptors* on the cell surface. A *viral receptor* is a molecule on the cell surface to which the virion attaches in order to initiate infection. Sometimes, a single viral receptor is sufficient for the virus to initiate infection. In other cases, however, an additional viral receptor is necessary for infection to take place. This additional viral receptor is called the *coreceptor*. The presence of certain receptors on the cell surface determines whether a certain cell is *susceptible* to infection by a specific virus. However, intracellular factors are also essential determinants of viral infectivity. Cells with intracellular factors that allow for infection by a specific virus are called *permissive* for infection by
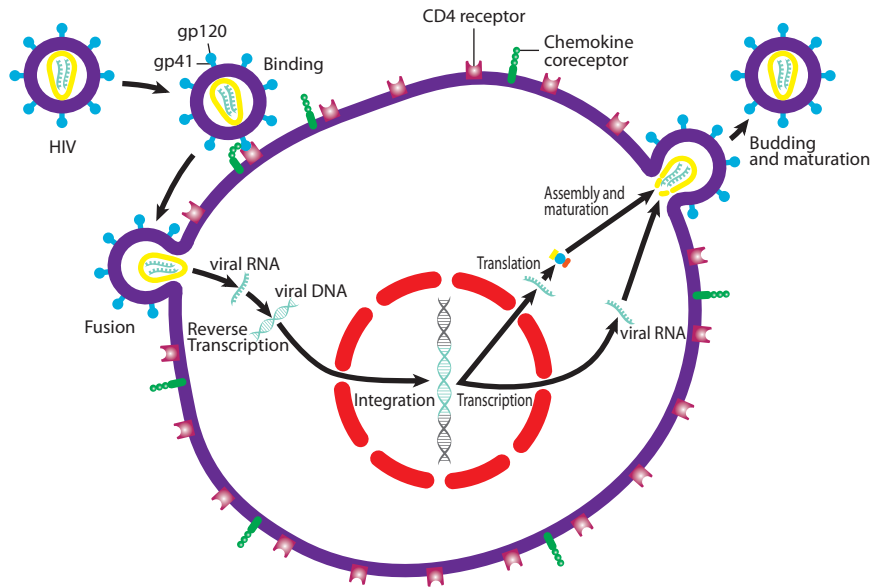
Figure 1.3: Replication Cycle of HIV-1. For entering the cell, HIV-1 binds to the CD4 cell receptor, as well as to a chemokine coreceptor. After fusion of viral and cell membranes, viral RNA is released into the cytoplasm and reverse transcribed. Following import of the preintegration complex into the cell nucleus, the viral DNA is integrated into the host's cell genome. Transcription of viral DNA ultimately results in the production of viral proteins, viral assembly and budding of new virions. Maturation of newly formed virions takes places during or after budding. Source: Wikimedia Commons, with modifications.

that virus. A concept that encapsulates both susceptibility and permissiveness to infection by a certain virus is that of *viral tropism*. Viral tropism is the preference of a virus to invade and replicate a certain type of cell. In addition to susceptibility and permissiveness, accessibility is a further determinant of viral tropism. For instance, a physical barrier can prevent the virus from making contact with its target cells, therefore rendering these cells inaccessible [7].

In the following, I provide a summary of the most important events that occur during the replication cycle of HIV-1. These can be divided into three main steps: (1) entry into the cell, (2) reverse transcription and integration, and (3) production of new viral particles.

Entry into The Cell

The primary receptor for HIV-1 is the cluster of differentiation 4 (CD4) molecule, a 58 kDa monomeric glycoprotein located on the surface of monocytes, macrophages, eosinophils, dendritic cells, microglial cells of the central nervous system, thymus (T) cell precursors within the bone marrow and thymus, as well as on approximately 60% of T lymphocytes. Interaction of HIV-1 with the CD4 receptor is necessary but not sufficient for entry into the cell. Additional interaction with a chemokine receptor is necessary as well. While many different chemokine receptors can mediate HIV-1 entry into the cell, HIV-1 most frequently uses the C-C chemokine receptor type 5 (CCR5) and the C-X-C chemokine receptor type 4 (CXCR4). R5-tropic HIV-1 strains use only the CCR5 coreceptor, X4-tropic HIV-1 strains use solely the CXCR4 coreceptor, and dual tropic HIV-1 strains can use either coreceptor. The first step in the replication cycle of HIV-1 is the attachment of the gp120 env surface protein to the CD4 molecule of the target cell. This induces conformational changes in gp120 that allow for binding to the coreceptor and subsequent exposure of a fusion peptide in gp41. This fusion peptide penetrates the plasma membrane of the cell, ultimately resulting in fusion of the viral and cellular membranes [1, 2, 8].

Reverse Transcription and Integration

After fusion of viral and host membranes, the viral core enters the cytoplasm and uncoats. This is followed by the formation of reverse-transcription complexes and the initiation of reverse transcription by RT. RT first produces an antisense complementary deoxyribonucleic acid (cDNA) copy of the viral RNA, which is subsequently duplicated in order to form a double-stranded DNA structure with long terminal repeats at each end. Double-stranded proviral DNA forms a preintegration complex with the proteins PR, RT, IN, and Vpr. After this preintegration complex is transported into the nucleus, integration into the host genome takes place. The integrated proviral DNA is called a provirus.

HIV-1 may infect CD4+ cells that exist in a resting or quiescent state. Specifically, mature T cells are released into the bloodstream after their production in the thymus. At first, these cells are immunologically naïve and exist in a resting or quiescent state. Upon recognition of an antigen, naïve T cells become active, which leads to their proliferation and differentiation into effector cells. Once the antigen is no longer present, most of the activated T cells die by apoptosis. However, a certain number of activated T cells become memory T cells, that do not die and return to a quiescent state instead [9]. Infection of quiescent cells by HIV-1 may result in the accumulation of non-integrated proviral DNA. For proviral DNA to become integrated into the host's genome, activation of the cells is necessary. *In-vivo*, activation occurs as a result of antigen contact, vaccination or opportunistic infection. Even if the proviral DNA is integrated into the genome of a quiescent cell, HIV-1 replication will be inhibited for as long as the cell does not become activated. Thus, infection of quiescent cells can build enduring cellular HIV-1 reservoirs. Antiretroviral compounds do not attack (pro-)viral DNA and are therefore inactive against viral reservoirs. Even after prolonged periods of antiretroviral therapy, previously unintegrated proviral DNA of replication-competent strains can be integrated into the host genome. For these reasons, eradication of HIV-1 through therapy with currently available antiretroviral drugs is not possible.

Production of New Viral Particles

Initial transcription of the provirus is mediated by cellular transcription factors. This results in the synthesis of the viral regulatory proteins Tat and Rev, which stimulate the transcription of further viral genes in turn. Further structural and regulatory viral proteins are produced from full-length and spliced mRNA. Specifically, the polyprotein gp160 is produced from a singly spliced full-length mRNA. The cellular protease furin subsequently cleaves this polyprotein into the viral env proteins gp41 and gp120. The gag and gag-pol polyproteins (see above) form the core of the nascent viral particle. Viral assembly occurs at the cell membrane, where viral RNA is enclosed into capsids that bud from the cell membrane, simultaneously incorporating env proteins. During or after budding, final maturation takes places. During the maturation process, gag and gag-pol polyproteins are cleaved by PR into their functional subunits.

1.2   HIV-1 Transmission and Pathogenesis

HIV-1 infection primarily occurs through sexual contact, needle sharing, pregnancy, delivery, and blood transfusion. Following infection of a patient with HIV-1, the pathogenic effects of the virus will begin to become noticeable. Left untreated, HIV-1 infection almost invariably leads to the death of the patient. In the following, I review the routes and probabilities of transmission for HIV-1. Subsequently, I describe the effects of the virus on the body, from the moment of infection until the death of the patient.

Table 1.2: Probabilities of Transmission of HIV-1. Depending on the exposure type, different probabilities of transmission of HIV-1 have been estimated.

Exposure Type	Probability of Infection per Contact (%)
Receptive penile-anal intercourse	~ 1
Receptive penile-vaginal intercourse	$0.1 - 0.32$
Insertive penile-anal intercourse	0.06
Insertive penile-vaginal intercourse	$0.01 - 0.1$
Needle sharing for drug consumption	$0.5 - 1$
Pregnancy and delivery	$12 - 50$
Consumption of breast milk	12
Reception of blood transfusion	> 90

1.2.1 HIV-1 Transmission

Infectious HIV-1 particles and cells infected with HIV-1 can be isolated from diverse body fluids. Among isolates extracted from HIV-1-infected individuals, free HIV-1 could be found in blood plasma, semen, and cerebrospinal fluid, as well as in tears, ear secretions, saliva, urine, breast milk, and cervicovaginal fluid, albeit at lower concentrations and in fewer isolates. Cells infected with HIV-1 could be isolated from blood, saliva, bronchial fluid, cervicovaginal fluid, and semen. HIV-1 spreads mainly through sexual contacts, needle sharing, pregnancy and delivery, and blood transfusions. The main factor determining whether a certain body fluid is a source of HIV-1 transmission is the infectious virus titer. Since saliva, tears, and urine (among other body fluids) contain few infectious HIV-1 particles or HIV-1-infected cells, they are not sources of HIV-1 infection. In contrast, high titers of infectious HIV-1 can be found in blood, genital fluids, and breast milk for which they *are* sources of HIV-1 infection [1]. The probability that an exposure to HIV-1 will result in an HIV-1 infection depends upon a variety of factors. Some of these factors are listed in the following:

- The amount of virus present in the inoculum[*], and thus in bodily fluids, has a positive correlation with disease transmission [10–12]
- Different HIV-1 exposure types, also called transmission routes, are associated with distinct probabilities of transmission (Table 1.2)
- Condom use, antiretroviral therapy, and male circumcision [11] can significantly reduce risk of transmission [10]
- The concomitant presence of ulcerations or inflammation in the genital mucosa can increase risk of transmission [10, 13, 14].

HIV-1 infection can occur through the transmission of cell-free viral particles or of HIV-1 infected cells that produce viral particles in the new host [10, 15]. In comparison to the viral strains that replicate in a chronically infected patient (see below), transmitted founder virus seems to be exclusively R5-tropic. Transmitted founder virus targets cells expressing a large number of CCR5 receptors, has decreased replication capacity, as well as shorter and less glycosylated env proteins [10, 16]. However, transmitted founder viruses also seem to be more infectious and to contain more env protein copies per particle than those present in a chronically infected patient [17]. Transmission of the virus across an intact mucosal barrier represents a viral population bottleneck [16, 18]. A single transmitted founder virus is responsible for 60% to 80% of productive infections [17, 19]. Nevertheless,

---

[*]An inoculum is a substance carrying an infectious or antigenic agent.
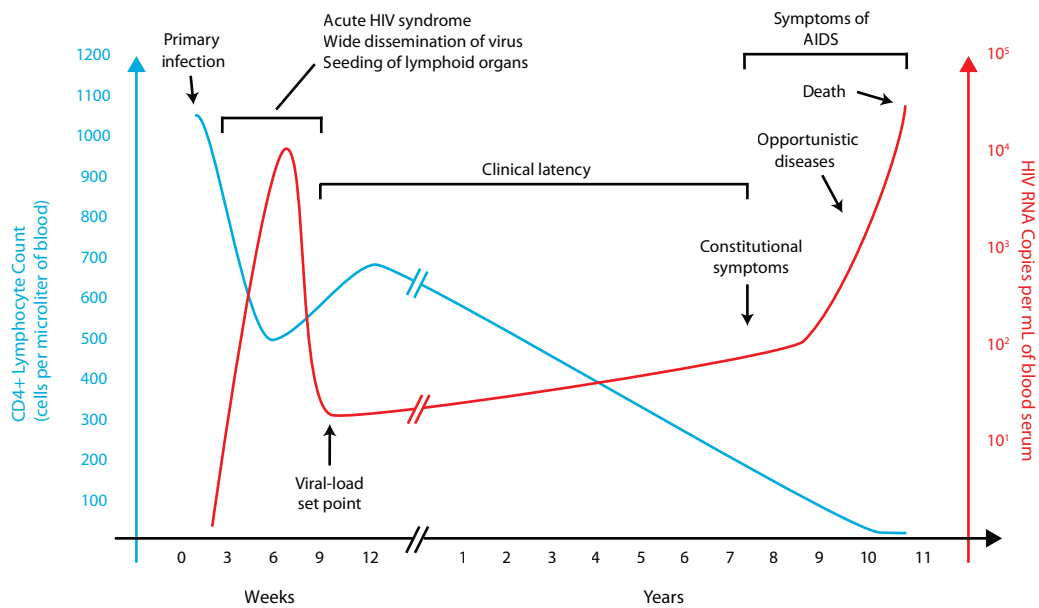
Figure 1.4: Clinical Parameters in HIV-1 Infection. The most important clinical parameters in HIV-1 infection are the CD4+-cell count and the VL. A schematic representation of the natural course of these prognostic markers during HIV-1 infection can be found above. Shortly after infection of a human host with HIV-1, the number of daily produced virions rises exponentially. As the VL rises, CD4+ cells die as a result of the infection. Some weeks after infection, antibodies appear in the blood of the host, controlling infection to a certain extent. This also results in a recovery of CD4+ cells. However, over the course of some years, the immune system gradually loses the capacity to control the infection due to the death of CD4+ cells. Thus, VL rises as CD4+-cell counts decrease. During the last years of the infection, the immune system has a substantially reduced capacity to eliminate pathogens from the body. Because of this, the patient presents constitutional symptoms and opportunistic infections, which ultimately lead to the death of the patient. Source: Wikimedia Commons, with modifications.

■ CD4+-cell count   ■ Viral load

there are factors that can decrease the bottleneck effect of mucosal barriers [14, 20, 21].

1.2.2 HIV-1 Pathogenesis

After infection with HIV-1, dendritic cells are among the first cells to come into contact with the transmitted founder virus. Infection can occur within the mucosa, where HIV-1 can infect CD4+ cells, or through the action of dendritic cells, as they transport HIV-1 to to the lymphoid tissue, where it can infect additional CD4+ cells [22, 23]. The infection of the first CD4+ cells initiates the process of HIV-1 pathogenesis. In untreated infections, this process can be divided into three phases:

1. Acute and primary HIV-1 infection
2. Chronic HIV-1 infection
3. Symptomatic HIV-1 infection

Figure 1.4 displays a schematic representation of the trajectory of the viral load (VL) and the CD4+-cell count after infection. The VL and the CD4+-cell count are the most important clinical parameters for an individual with untreated HIV-1 infection. The CD4+-cell count indicates the number of CD4+ cells that present in a microliter of blood, while the VL quantifies the number of HIV-1 RNA copies contained in one milliliter of blood *serum*. In the following, I describe the major features of the three phases of HIV-1 pathogenesis, as they occur in an untreated HIV-1 infection.

Acute and Primary HIV-1 Infection

Acute HIV-1 infection is defined as a period lasting from one to four weeks after virus transmission, with the particularity that no anti-HIV-1 antibodies can be detected. During the first weeks *post infectionem*, no immune response is present and HIV-1 can replicate and spread within the body at an increased rate. Seven to ten days after transmission, HIV-1 RNA can be detected in blood. From this point on, HIV-1 viremia begins to rise, with an average doubling time of nine to 20 hours during the first two to three weeks of infection. Within six to 21 days, virus quantities in blood reach a peak (10^6 to 10^7 HIV-1 RNA copies per milliliter of blood serum) [1, 2, 24, 25]. During this early period, a large number of cells become infected, and a predominant viral strain emerges. The viral population is very homogeneous during the acute phase. The virus is seeded in a variety of tissue reservoirs [26] and CD4+T lymphocytes are destroyed, particularly in the lymphoid tissues and in the gut. The first immunologic response against HIV-1 involves the activation of HIV-1-specific cytotoxic T lymphocytes (CTLs) [24, 27], which reduces the VL [28, 29]. The emergence of an initial immunologic response, as well as the onset of acute retroviral syndrome (ARS; described in the following) are the events that tend to be used in order to define the beginning of primary HIV-1 infection, which begins before the end of acute HIV-1 infection [26]. One to four weeks after transmission, 50% to 80% of newly infected persons may present a clinical state termed ARS. The most common symptoms present during ARS are loss of appetite, malaise, fever, rash, swollen lymph nodes, as well as pain in muscles and joints. Further symptoms can include central nervous system disorders, oral candidiasis, and ulcerations in the esophagus, anus, and vagina. These symptoms usually disappear within one to three weeks, although lymphadenopathy, malaise, and lethargy can persist for several months. Absence of ARS may predict a slow progression towards disease, while the severity and duration of ARS symptoms correlates with the speed of disease progression. The end of acute HIV-1 infection is defined by the appearance of anti-HIV antibodies in the blood of the patient. This event is termed seroconversion. ARS usually occurs before seroconversion [1, 2, 24, 26, 30, 31]. Some weeks after acute infection, viremia is substantially reduced as a result of immune response against HIV-1. However, immune response comes too late to control the infection. After seroconversion, neutralizing antibodies appear to be present only transiently [1]. After immunologic activation involving CTL response, the VL decreases until a *set point* is reached following the production of anti-HIV-1 antibodies, as well as the production of broadly neutralizing antibodies (Figure 1.4) [1, 2, 25, 26]. Primary HIV-1 infection is usually defined to end when the VL reaches its initial set point [26]. This set point is a prognostic marker of the future course of disease [1, 32].

Chronic HIV-1 Infection

Three to six months following primary infection, CD4+-cell counts rise to almost normal levels. However, from this point on, CD4+ cells will be depleted at a rate of 25 to 60 cells per microliter of blood per year. An asymptomatic period follows, lasting up to ten years. During this period, viral replication continues, at first at at a low level and especially in the lymph nodes [33, 34]. At this stage, HIV-1 RNA levels vary among individuals and range from less than 50 to more than one million HIV-1 RNA copies per milliliter of blood plasma. Even if viremia in plasma is undetectable, HIV-1 proviral DNA can still be detected in peripheral blood mononuclear cells (PBMCs) [35]. The suppression of HIV-1 replication during chronic HIV-1 infection appears to be mediated by CD8+ cells. The presence of a strong immune response results in diversification of the viral population. Specifically, as new viral variants continuously emerge, most of them are subsequently eliminated by the immune system, but some of them manage to escape immune pressure and proliferate [1]. As the infection pro-
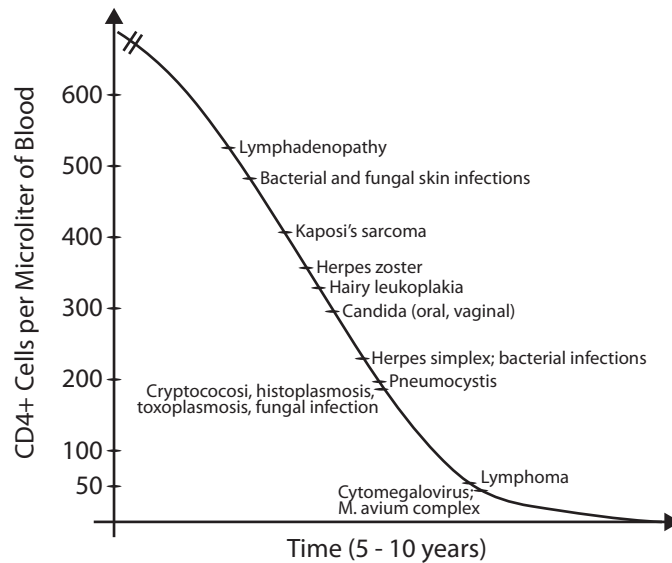
Figure 1.5: CD4-Dependent Risk of Developing Opportunistic Infections and Cancer. As CD4+-cell counts decrease, different clinical conditions may occur. The graph above depicts the CD4+-cell counts at which these conditions typically happen. Tuberculosis, Kaposi's sarcoma, and non-Hodgkin's lymphoma can occur at any time. However, tuberculosis and Kaposi's sarcoma are especially frequent when the CD4+-cell count drops below 400 cells per microliter of blood, and non-Hodgkin's lymphoma with CD4+-cell counts below 50 cells per microliter of blood. After [1].

gresses, the VL rises, while CD4+ cell counts decrease (Figure 1.4). Chronic HIV-1 infection is asymptomatic. However, during this stage, the virus is highly dynamic, replicating in high titers [35]. At the same time, the immune system is progressively damaged [29]. During chronic HIV-1 infection, the VL is the most informative prognostic marker for disease progression, followed by the CD4+ cell count [29, 32–34].

Symptomatic HIV-1 Infection

Within two to ten years after infection, the prodromes of acquired immunodeficiency syndrome (AIDS) may show up. At this stage, CD4+ cell counts usually drop below 350 cells per microliter of blood, the VL increases substantially, and antiviral CD8+ cell responses are reduced. As a result of damage and therefore reduced action of the immune system, the viral population becomes again homogeneous, as in the acute phase of the disease. As the CD4+ lymphocytes are depleted, the risk for developing opportunistic infections and cancer rises (Figure 1.5). The onset of AIDS in HIV-1-infected patients is defined as the occurrence of one or more specific diseases (see [1] for a list of these diseases) or a decrease in the CD4+ cell count below 200 cells per microliter of blood [29]. Left untreated, HIV-1 infection will almost inevitably lead to the death of the patient.

1.3 History of the HIV-1 Pandemic

In 1981, increasing numbers of deaths of previously healthy homosexual men were reported [36]. The clinical state of these men included unusual and severe opportunistic infections and malignancies. This initiated the process of recognition and definition of AIDS. Epidemiological studies at the time suggested that the disease is transmitted horizontally through sexual contact, intravenous drug administration, and blood transfusion [37]. For this reason, a pathogenic etiology of AIDS was suspected. In 1983, HIV-1 was isolated from a patient

presenting the prodromes of AIDS [37]. Further studies isolated the same pathogen from a number of patients suffering from AIDS and its prodromes [38, 39], contributing to the establishment of HIV-1 as a causative agent of AIDS. However, the history of the HIV-1 pandemic begins with the emergence of HIV-1 in Africa. Many analyses have been performed in order to clarify the etiology of HIV-1 and in order to reconstruct the pathogen's epidemiological history predating 1981. The results of these analyses are summarized in the following sections.

Several clues on the etiology of HIV-1 lead to Central and Equatorial Africa. Between 1983 and 1984, many cases of AIDS were reported. These cases concerned black patients from this African region [40–43] and white patients who had traveled to that region [41–44]. In 1984, HIV-1 was isolated from a married couple from Zaire [45]. A study published in 1985 compared the nucleotide sequence of HIV-1 with that of the visna lentivirus, providing support for the inclusion of HIV-1 into the retroviral subfamily *Lentivirinae* [46]. A pathogen related to HIV-1 was isolated from West African patients suffering from AIDS in 1986 [47]. This pathogen was later sequenced and named *Human immunodeficiency virus type 2* (HIV-2) [48], and its relationship to a variant of *Simian immunodeficiency virus* (SIV) was established [49]. The recognition of the similarities between this SIV variant and HIV-2 suggested that the etiology of HIV-1 could be elucidated by examining other SIV variants.

The etiology of HIV-1 has been reconstructed using molecular epidemiology, phylogenetic analyses, and historical information. In the following, I provide a chronological account of the history of the HIV-1 pandemic, including both reconstructed and recorded facts.

1.3.1 Zoonotic Transmission and Spread of HIV-1

Zoonoses are infectious diseases that are transmitted from animals to humans. The term *zoonosis* is also used to refer to the event of cross-species pathogen transmission [50]. The origins of HIV-1 lie in West-Central Africa, where multiple zoonotic transmissions of SIV from non-human primates to humans occurred [51]. This claim is supported by molecular phylogenetic analyses (Figure 1.7). Molecular phylogenetics studies the evolutionary relationships between and among organisms, based on comparative analyses of the nucleotide and protein sequences of their genomes. Phylogenetic analyses can deploy a molecular clock, a statistical model that describes the relationship between evolutionary distances and calendar time [52]. In the following, I first give a brief introduction to SIVs, and then describe how HIV-1 is thought to have originated via zoonosis.

Simian Immunodeficiency Virus

Serological methods have delivered evidence for SIV infection in over 40 primate species [53]. Most of these species only harbor a specific SIV strain. This has been shown with phylogenetic analyses in which the viral sequences from members of the same species cluster into the same subtree (Figure 1.7). Although almost all SIV transmissions occur within the same species, there are also several inferred cases in which SIV zoonoses have occurred [54, 55].

Depending on the primate species that an SIV lineage primarily infects, an additional three-letter abbreviation is appended to *SIV*. Thus, SIVcpz denotes SIV infecting chimpanzees. SIVcpz found in *Pan troglodytes troglodytes* has a close relationship to HIV-1 [54–56], and it is thought that this SIVcpz variant gave rise to HIV-1 type M (see below). *P. t. troglodytes* probably acquired SIVcpz from other primates [54], as the animal is known to hunt and kill other mammals, including other primates. Isolated groups of *P. t. troglodytes* living in Cameroon form natural reservoirs of SIVcpz (Figure 1.6) [54, 56].
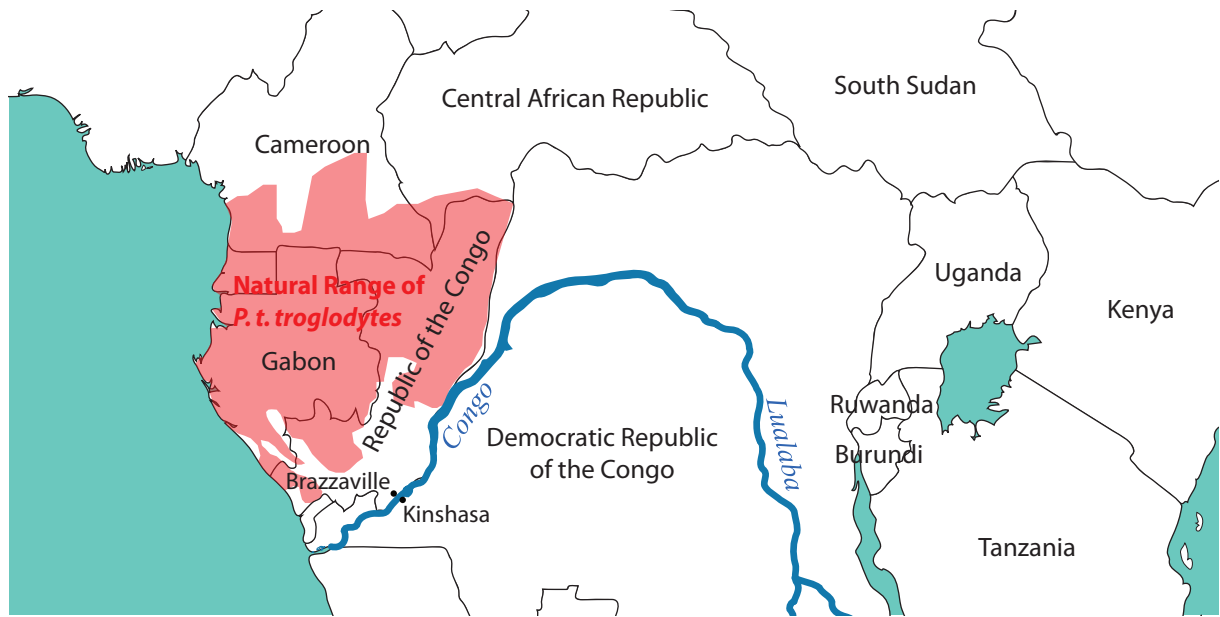
Figure 1.6: Location of Kinshasa and Brazzaville. The map above shows the locations of Kinshasa and Brazzaville as well as the natural range of *Pan troglodytes troglodytes* (marked in red). After [54], with modifications.

Human Immunodeficiency Virus Type 1

HIV-1 comprises distinct lineages classified into the groups M, N, O, and P. Each one of these groups have most likely resulted from an independent zoonosis and subsequent adaptation to the host (Figure 1.7) [50, 54, 56]. While the global pandemic is caused by HIV-1 group M, infections with other HIV-1 groups are much less prevalent and mainly confined to Cameroon, the Democratic Republic of Congo (DRC), Gabon, and other neighboring countries [50, 54]. Due to the biology of SIV, these zoonoses must have occurred through the exposure of cutaneous or mucous membranes to infected ape blood or other bodily fluids [54, 55]. It is likely that this occurred during hunting, capture, butchering or trade of nonhuman primates, or while a primate was kept as a pet by a human [50, 57].

The oldest existing HIV-1 samples were obtained in 1959 [58] and 1960 [59] in Kinshasa, DRC (formerly Leopoldville, Zaire and Leopoldville, Belgian Congo). Reconstruction of the history of HIV-1 infection can be achieved with phylogenetic analyses coupled with molecular clocks. These analyses can provide estimates of the time at which a group of related nucleotide sequences diverged from their most recent common ancestor (tMRCA). Several phylogenetic analyses have estimated that HIV-1 group M is the oldest HIV lineage (Table 1.3), with a tMRCA between 1920 and 1937 [60]. A phylogenetic study analyzing SIVcpz and HIV-1 group M nucleotide sequences simultaneously dates their most common recent ancestor to 1853 (1799-1904) [62]. With some uncertainty on the time at which it occurred, cross-species transfer resulting in HIV-1 group M is therefore inferred to have taken place between 1853 and the early 1900s [50].

Spread of HIV-1 in Central Africa

Central Africa is thought to be the epicenter of the global HIV-1 pandemic [50, 63]. In the following, I mention some of the arguments supporting this hypothesis. Strains of HIV-1 group M are further classified into subtypes and circulating recombinant forms (CRFs) (Section 1.4). CRFs are thought to have emerged at a later point in time than subtypes (Table 1.4). Subtypes and CRFs have specific distribution patterns in different regions of the
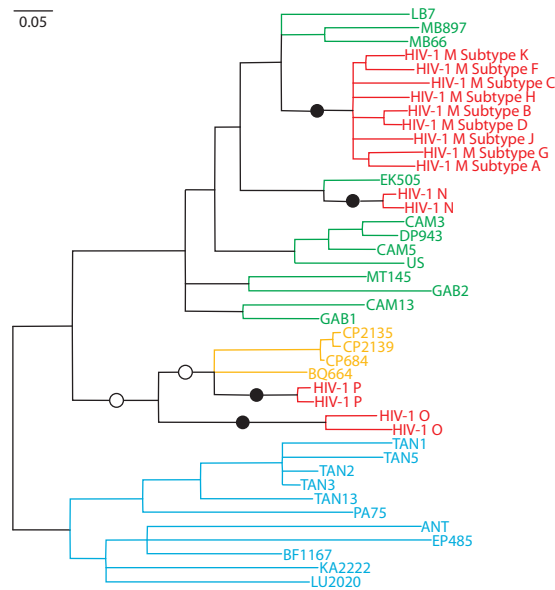
Figure 1.7: Phylogenetic Tree for HIV-1, SIVcpz, and SIVgor. This phylogenetic tree depicts the genetic relationships between the *pol* genes of representative HIV-1, SIVcpz, and SIVgor strains. Four HIV-1 groups are shown: M, N, O, and P. Each of these HIV-1 groups resulted from independent cross-species transmissions. Probably, HIV-1 Group P resulted from a gorilla-to-human cross-species transmission. The gorillas, in turn, were probably infected by chimpanzees. At the top-left corner, the scale of 0.05 substitutions per site is indicated. After [54], with modifications.

■ SIVcpz (*Pan troglodytes troglodytes*) ■ HIV-1 (*Homo sapiens sapiens*) ■ SIVgor (*Gorilla gorilla*) ■ SIVcpz (*Pan troglodytes schweinfurtii*)

● Cross-species transmission to human ○ Chimpanzee-to-gorilla cross-species transmission

Table 1.3: tMRCA of Various HIV-1 Types. Phylogenetic analyses can be used to determine the time to the most recent common ancestor of a group of nucleotide sequences. Below, a summary of several such analyses for different HIV-1 types can be found. CI: 95% confidence interval; tMRCA: time to most recent common ancestor; * 99% confidence interval. After [50].

HIV-1 group	tMRCA	CI
Group M	1920	1909 - 1930 [61]
Group M	1921	1908-1933
Group M	1920/1937	1902-1949*
Group M	1931	1915-1941
Group O	1920	1890-1940
Group N	1963	1948-1977

Table 1.4: tMRCA of Various HIV-1 Subtypes. Phylogenetic analyses can be used to determine the time to the most recent common ancestor of a group of nucleotide sequences. Below, a summary of several such analyses for different HIV-1 subtypes and circulating recombinant forms can be found. CI: 95% confidence interval, CRF: circulating recombinant form; IDU: intravenous drug users; tMRCA: time to most recent common ancestor; * 99% confidence interval. After [50].

HIV-1 group M subtype	tMRCA	CI	HIV-1 group M subtype/CRF	tMRCA	CI
Subtype A1 (Africa)	1948	1865-1969	Subtype B (USA)	1967	1960-1971
Subtype A1	1954	1939-1968	Subtype B (USA/Canada)	1969	1966-1972
Subtype B/D	1949/1954	1935-1964*	Subtype B (USA/Canada)	1969	1966-1972
Subtype B/D	1950	1938-1959	Subtype B (USA/Canada)	1973	1970-1976
Subtype B/D	1954	1946-1961	Subtype B (Trinidad and Tobago)	1976	1954-1986
Subtype B/D	1944/1946	1935-1955	Subtype B (Albania)	1984	1972-1989
Subtype D	1958/1967	1931-1984	Subtype B (South Korea)	1969	1959-1978
Subtype D (East Africa)	1966	1961-1971	Subtype G	1976	1966-1982
Subtype D	1952/1955	1933-1971	Subtype F1 (South America)	1977/1978	1963-1988
Subtype C	1958	1949-1960	Subtype F1 (Brazil)	1975	1970-1980
Subtype C	1966-1969	1959-1973	CRF01_AE (Africa)	1965/1969	1963-1973
Subtype C	1967	1962-1972	CRF01_AE	1984	1980-1986
Subtype C (Zimbabwe)	1972	1969-1974	CRF01_AE (Thailand)	1980/1981	1979-1982
Subtype C (India)	1976/1979	1972-1982	CRF01_AE (Vietnam, South, hetero)	1987/1990	1987-1990
Subtype C (China)	1981	1976-1985	CRF01_AE (Vietnam, South, IDU)	1993/1994	1992-1995
Subtype C (Ethiopia)	1982	1980-1983	CRF01_AE (Vietnam, North)	1975/1976	1969-1981
Subtype C (Ethiopia)	1983	1980-1984	CRF02_AG	1973	1972-1975
Subtype C (Brazil)	1980-1983	1966-1988	CRF02_AG	1988	1982-1992
Subtype B	1959-1966	1950-1972	CRF31_BC (Brazil)	1990	1988-1991
Subtype B (Global and Haiti)	1966	1962-1970	CRF08_BC (China)	1993	1991-1995
Subtype B	1970/1972	1963-1978*	CRF07_BC (China)	1991-1993	1987-1997
Subtype B (Brazil)	1964/1967	1946-1984	CRF33_01B (Malaysia)		

world (Figure 1.8). However, all HIV-1 groups and subtypes (not including the more recent CRFs) are found in Central Africa. Furthermore, the highest HIV-1 sequence diversity is found in this region as well. Specifically, the diversity in Kinshasa, DRC, is comparable to that among global strains [56, 65, 66] (Figure 1.8).

In a phylogenetic analysis, isolates from the DRC have a basal position in a phylogenetic tree that includes global strains with diverse subtypes. Thus, each global subtype is probably the result of the exportation of HIV-1 strains from the DRC [66]. Furthermore, the two oldest HIV-1 samples stem from Kinshasa (as mentioned above). These arguments indicate that HIV-1 originated in colonial west-central Africa [54]. Zoonosis probably took place in Cameroon, after which the disease spread along the Sangha and Congo rivers, reaching the city of Kinshasa, where the global epidemic probably initiated [50, 56, 67]. As mentioned above, phylogenetic analyses estimate the tMRCA of HIV-1 group M to be between 1920 and 1937. The necessary conditions for a zoonosis from simians to humans have existed for hundreds or thousands of years. One could assume that such transmission events have occurred in the past. In the twentieth century, however, urban centers grew rapidly and medical procedures involving injections [68, 69] and blood transfusions [70, 71] became widespread. This key difference can explain why initial HIV-1 group M infections could become a pandemic [55, 72, 73], although the sexual transmission probability of HIV-1 is at most one percent (Table 1.2).

Even though the zoonosis of an HIV-related pathogen has been possible for a long time, it was only in the twentieth century that conditions for a wide spread of the pathogen were met [67]. In the following, some of the factors that may have contributed to the adaptation to the initial HIV-1 infection, and its spread in Central Africa can be found. Central Africans have long been in contact with great apes. Key factors facilitating the spread of HIV-1 gradually came into place during the eighteenth and nineteenth centuries. Specifically, human mobility rose considerably in the region, and Central African societies began to engage in short- and long-distance trade [72]. At the beginning of the twentieth century, colonial authorities in Central Africa conscripted people as forced labor to work on railroads and other infrastructure projects. People were forced to harvest large quantities of rubber, which was coupled with a harsh quota system in the Belgian Congo. Conscription for hard forced labor led people to flee villages and settle in the forests. At the same time, those conscripted had allegedly little opportunity to gain their sustenance through agriculture, due to the hard work they were forced to perform. This probably increased people's reliance on bush meat, and hunting of large animals such as chimpanzees was facilitated through the availability of firearms. Brazzaville became the capital of the French Congo in 1910, while Leopoldville (now Kinshasa) served as the capital of Belgian Congo from 1923. Following the denomination of these cities as capitals, employment opportunities as well as the establishment of schools, hospitals, and churches attracted rural migrants, thus contributing to urban growth. As populations of Kinshasa and Brazzaville grew, their population became denser (Figures 1.9 and 1.6) [72, 73]. Male African rural migrants lived in temporary settlements in these cities. At the same time, single women were banned from these cities until the mid 1930s. In 1928, about 21,500 African men, but only 358 legally married women lived in Leopoldville. However, more than 5,000 women were smuggled into the city. Transactional sex probably offered the only means of survival for many of the women who illegally resided in Leopoldville. Transactional sexual services in Leopoldville were provided by *femmes libres* who served a few regular clients. In exchange for a regular income, these women provided their clients with sex, conversation, cooking, laundering, and hair care. Iantrogenesis represents yet another factor probably contributing to the spread of HIV. Specifically, I refer to the possible spread of HIV-1 through the use of unsterile syringes [68, 69] as well as through the transfusion of contaminated blood [70–72]. In 1929, the first of two *Dispensaires Antivénériens* was established in Leopoldville. The *Dispensaires Antivénériens* were clinics for treating sexually transmitted diseases. Male and
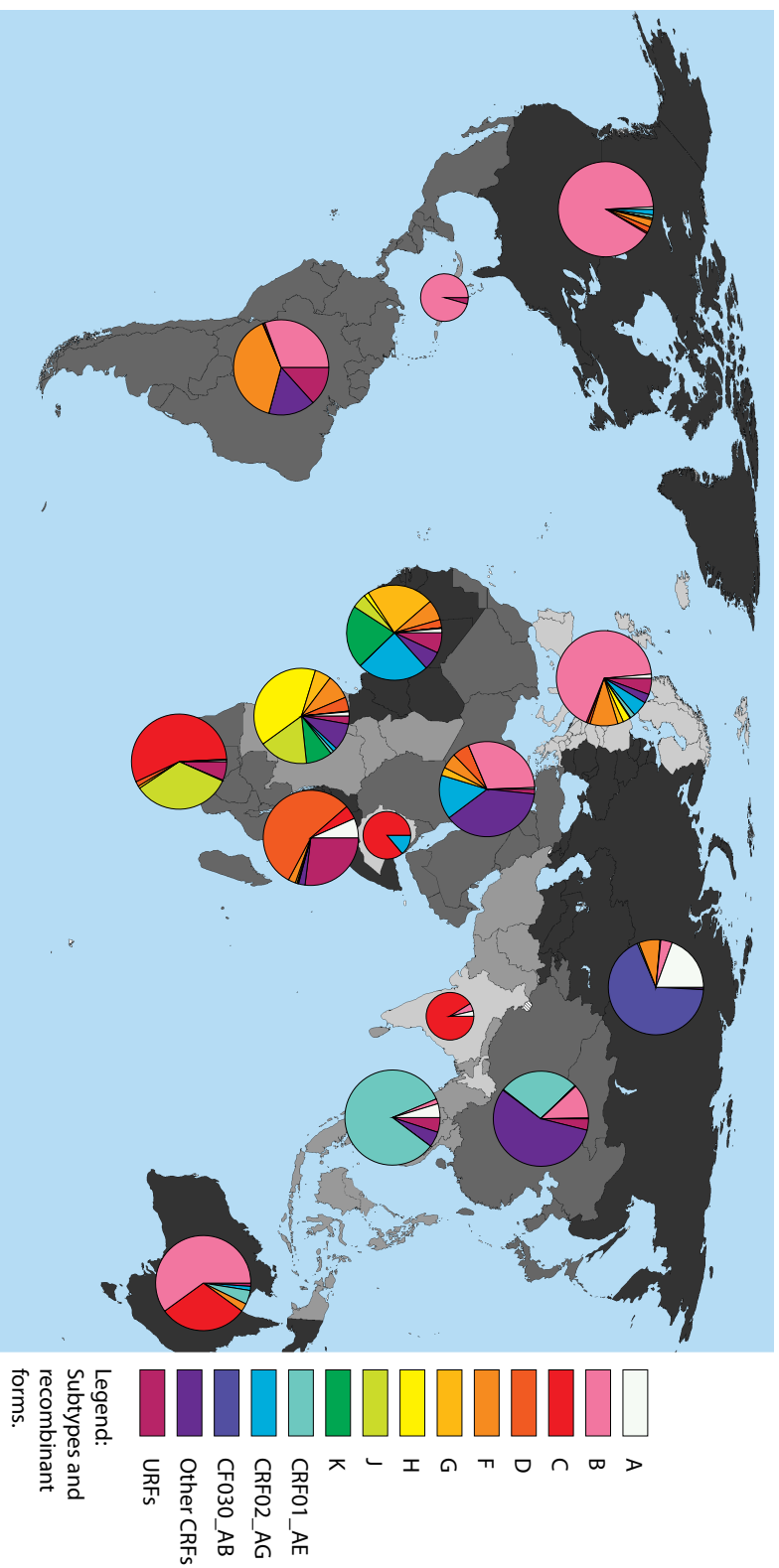
Figure 1.8: World-Wide Distribution of Subtypes and Recombinants. The pie charts above depict the distribution of subtypes and recombinants in different regions of the world from 2004 to 2007. Pie charts are superimposed on the world region for which they indicate the distribution of subtypes and recombinants. Regions are delimited by different shades of gray. The following regions were considered: North America, Caribbean, Latin America, Western and Central Europe, Eastern Europe and Central Asia, Oceania, India, South and South-East Asia (excluding India), East Asia, Middle East and North Africa, West Africa, East Africa (excluding Ethiopia), Ethiopia, Central Africa, and Southern Africa. CRF: circulating recombinant form; URF: unique recombinant form. After [50, 64], with data from [64].
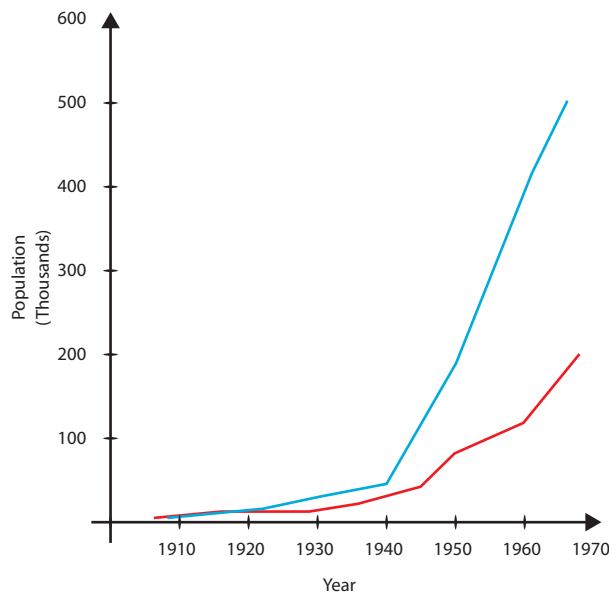
Figure 1.9: Population Growth in Kinshasa and Brazzaville. After the foundation of Kinshasa (blue line) and Brazzaville (red line) at the end of the 19th century, the populations of these cities experienced a strong and continuous growth. After [73].

■ Kinshasa   ■ Brazzaville

female patients with genital complaints who presented to these clinics were treated free of charge. Furthermore, migrants from outside Leopoldville and *femmes libres* were required to periodically attend a *Dispensaire Antivénérien* for screening. Every person who was tested serologically positive for syphilis was treated with intravenous drug injections. Unfortunately, yaws, a non-venereal cutaneous disease caused by *Treponema pallidum*, cannot be serologically distinguished from syphilis. Given the high incidence of yaws in the rural areas from which migrants originated, this resulted in many unnecessary treatments. Moreover, patients with a former syphilis infection test serologically positive for syphilis for several months after the infection has been cleared. Patients continued to receive injections for as long as they tested serologically positive for syphilis. The number of injections that *Dispensaires Antivénériens* administered annually peaked in 1953 at 154,572 injections. Evidence of iantrogenical transmission of Hepatitis through *Dispensaires Antivénériens* is strongly suggestive of inadequate sterilization of injection equipment in these clinics. Therefore, parenteral, iatrogenic transmission of HIV-1 probably took place in these clinics as well. For this reason, it has been hypothesized that the spread of HIV-1 was iantrogenically boosted by *Dispensaires Antivénériens* in the community of *femmes libres*, their customers, and further people sexually relating to them. In June 1960, the Congo became independent. This brought about many changes in sex labor, which shifted from transactional sex to high-risk prostitution, in which some women entertained up to 1,000 customers per year. Starting from the established infection in the community of (former) *femmes libres*, HIV-1 spread sexually after 1960, substantially amplifying the number of infected people [74]. The growth rate of HIV-1 group M in Kinshasa has been estimated for the period between 1920 and the early 2000s [61]. While the growth of HIV-1 group M is estimated to be exponential over the whole period, the estimates also indicate a three-fold increase in this growth rate around 1960. This is consistent with the hypothesis of iatrogenic origins of the HIV-1 pandemic I have described above.

Figure 1.10: Multiple Simultaneous Lesions of Kaposi's Sarcoma. Kaposi's sarcoma is a malignancy caused by an infection with *Human herpes virus 8* (HHV-8). The virus is predominantly transmitted by saliva, but also via sexual contacts and blood products. In some regions of the world, HHV-8 has a prevalence of up to 50% of the general population. However, the development of Kaposi's sarcoma appears to be related to immune suppression in the host, and may be present at advanced stages of HIV-1 infection [1, 79]. See also Figure 1.5. Image courtesy of AIDS Images Library `http://www.aids-images.ch`.

SPREAD OF HIV-1 TO THE REST OF THE WORLD

HIV-1 subtype B infections are most common in the United States and Canada, the Caribbean, and Western Europe (Figure 1.8). Phylogenetic analyses have shown that in phylogenetic trees, HIV-1 subtype B nucleotide sequences derived from archived blood samples of Haitian patients fall basal to other subtype B nucleotide sequences derived from patients around the world. The findings imply that Haiti experienced the oldest-known HIV-1 epidemic outside of Africa [75]. While the DRC hosts the most diverse HIV-1 group M epidemic in the world (as mentioned above), the most diverse HIV-1 subtype B epidemic in the world is hosted by Haiti [75]. Between the early 1960s and the mid-1970s, several thousand Haitian professionals temporarily migrated to Zaire, as mentioned in [76]. It is thought that they became infected with HIV-1 in Zaire, and brought back the disease to Haiti, the United States, and Europe [56, 63, 77]. The tMRCA of subtype B is estimated to be 1966 (Table 1.4), which matches the time at which Haitian emigrated to Zaire. For the epidemic in the United States, the tMRCA is estimated to be somewhere around 1968 (Table 1.4, [76, 78]), 13 years before the clinical picture of AIDS was recognized in that country (see below).

In 1981, the infection of four homosexual men with *Pneumocystis carinii* pneumonia was reported. These patients were treated in three hospitals in Los Angeles and presented symptoms of immunosuppression [36]. Further, unusual cases of patients with *Pneumocystis carinii* pneumonia and Kaposi's sarcoma (Figure 1.10) were reported. Epidemiological studies at the time suggested that the cause for the illness was a sexually-transmitted pathogen [80]. However, other etiologic agents were also proposed, such as sperm, amyl nitrate, or "chronic overexposure to foreign proteins" (as mentioned in [77]). By the beginning of 1983, the causative agent of this

immunodeficiency-related clinical picture had not been identified, however, the main transmission routes had been described, as reviewed in [81]. Later that year, the pathogen causing this clinical state was identified [37].

1.3.2   The HIV-1 Pandemic after 1983

Following the identification of HIV-1 as the etiologic agent of AIDS, key scientific accomplishments allowed for treating and preventing HIV-1 infection. In the following, I mention the most important ones. The identification of CD4 as HIV-1's main receptor in 1984 [82, 83] paved the way for an important diagnostic parameter in HIV-1 infection: the CD4 cell count. However, the development of the ability to quantify the VL in plasma required further nine years [32, 34]. (The diagnostic value of the CD4 count and of the VL determination is explained in Section 1.2.2). Preceded by the identification of antibodies in the sera of HIV-1-infected patients [84, 85], antibody testing became widely available in 1985 [86, 87]. This permitted infected patients without symptoms to become aware of their infection and prevent further transmission of the virus. Furthermore, the screening of blood supplies for HIV-1 halted the transmission of the virus by blood transfusion (at least in industrialized countries) [88]. In 1985, the nucleotide sequence of an HIV-1 variant was published, which opened the doors for phylogenetic analyses and molecular epidemiology [89]. Section 1.5.1 reviews the history of antiretroviral therapy. Here, I mention five milestones in the history of antiretroviral therapy that had a significant influence on the course of the HIV-1 pandemic. (1) The first efficacy trial for zidovudine, an antiretroviral drug, was performed in 1987 [90]. (2) This same drug was shown to be able to prevent mother-to-child transmission in 1994 [91]. (3) In 2010, the iPrex trial gave proof of the effectiveness of antiretroviral drugs as pre-exposure prophylactics [92]. (4) One year later, in 2011, it was shown that antiretroviral therapy could be effectively used by an infected person to prevent the transmission of HIV-1 [93]. (5) Last, but not least, I mention the functional cure of an HIV-1-infected patient who received a transplantation of HIV-resistant stem cells in 2009 [94].

Over the decade following 1983, knowledge on the HIV-1 infection evolved from describing a disease believed to be only prevalent in the United States, and to affect only men who have sex with men, to a global pandemic also affecting heterosexual men and women [81]. In 1989, the World Health Organization reported the existence of AIDS cases in all continents [95]. For a review of the state of the pandemic in 1990, see [96]. The world-wide incidence[†] of HIV-1 infection in 1990 was estimated to be two million, resulting in a prevalence[‡] of nine million, and 320 thousand AIDS-related deaths (Figure 1.11). More than 150 countries started to implement AIDS prevention and control programs, involving the epidemiological surveillance of HIV-1 infection, the promotion of behavioral changes in order to prevent infection by HIV-1, and the screening of blood products [96]. Nonetheless, HIV-1 incidence rose to a peak between 1996 and 1998 with an estimate of 3.4 million new infections in each of those years (Figure 1.11(a)). By 1998, HIV-1 infection was the fourth most frequent cause of death in the world, and the leading cause of death in sub-Saharan Africa ([97] and Figure 1.11(c)). This was partly a consequence of the following facts. (1) The great majority of the people infected with HIV-1 lived in developing countries (and still do; Figure 1.12), and (2) in these countries, HIV-1 is predominantly transmitted through heterosexual contacts [98]. Since most developing countries did not have sufficient resources for diagnosing and treating HIV-1 infection, the virus could continue to spread without hindrance. Furthermore,

---

[†]In epidemiology, the incidence (rate) is the number of persons acquiring a medical condition within the population at risk per given time period.

[‡]In epidemiology, the prevalence is the number of people or population proportion suffering from a medical condition at a certain point in time.
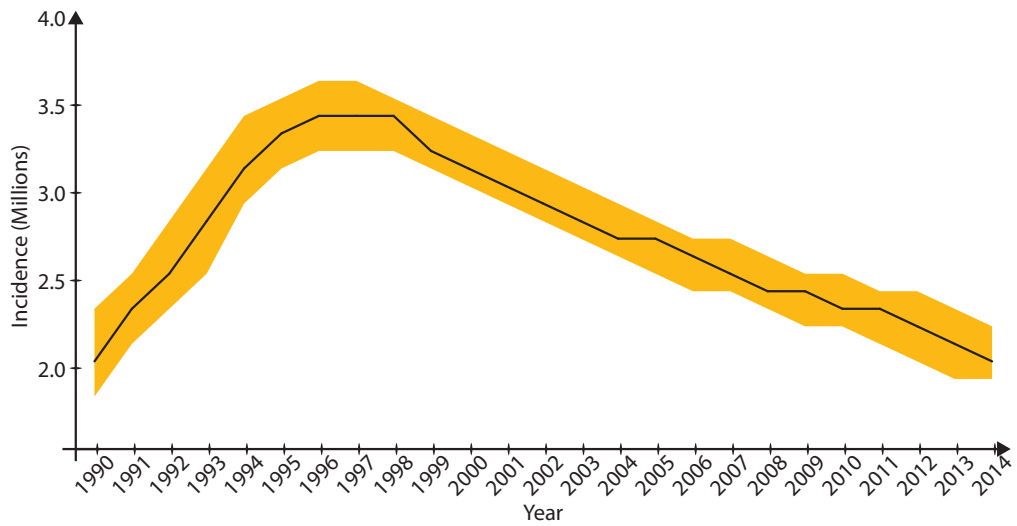
(a) Incidence of HIV
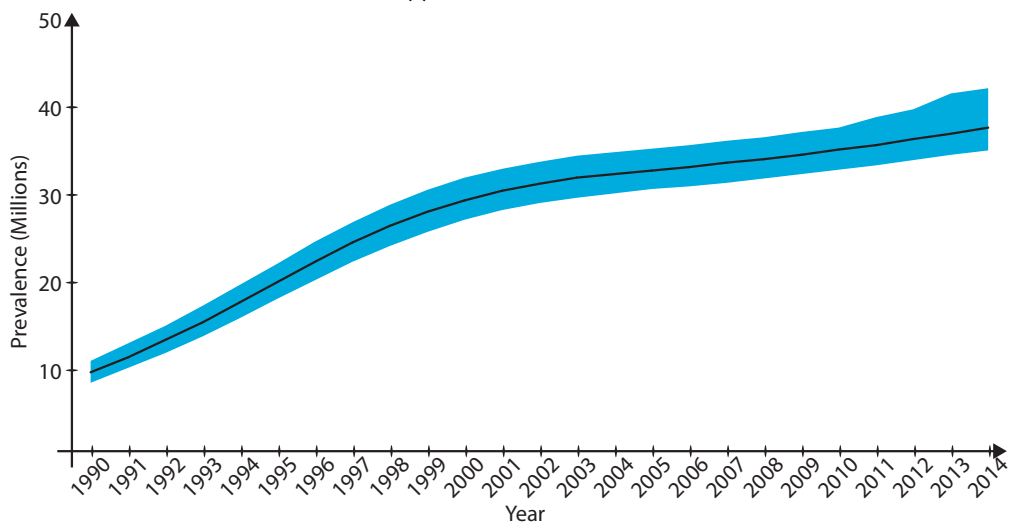
(b) Prevalence of HIV
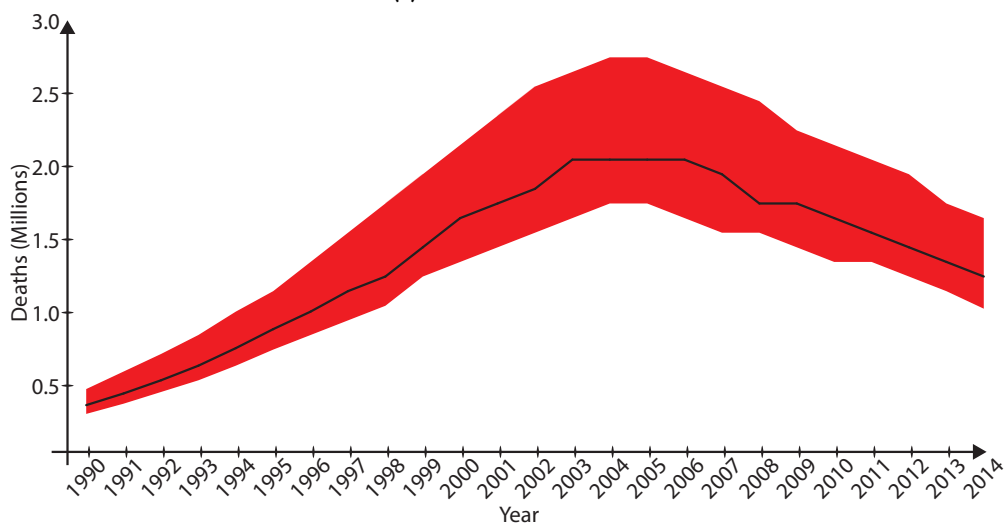
(c) AIDS-related Deaths

Figure 1.11: Prevalence, Incidence, and Mortality of HIV. Yearly estimates of the world-wide incidence (a) and prevalence (b) of HIV are plotted above, along with the numbers of AIDS-related deaths (c). The polygons surrounding the black curves indicate the range between the yearly lower and upper estimates. Source: `http://aidsinfo.unaids.org/`, accessed December 9th, 2015.

Number of people, by WHO region

Eastern Mediterranean: 280 000 [200 000–420 000]

Western Pacific: 1 300 000 [1 100 000–1 700 000]

Europe: 2 100 000 [1 900 000–2 200 000]

Americas: 3 200 000 [2 800 000–4 000 000]

South-East Asia: 3 400 000 [2 900 000–4 000 000]

Africa: 24 700 000 [23 500 000–26 100 000]

Total: 35 000 000

[33 200 000–37 200 000]

0   875   1,750   3,500 Kilometers
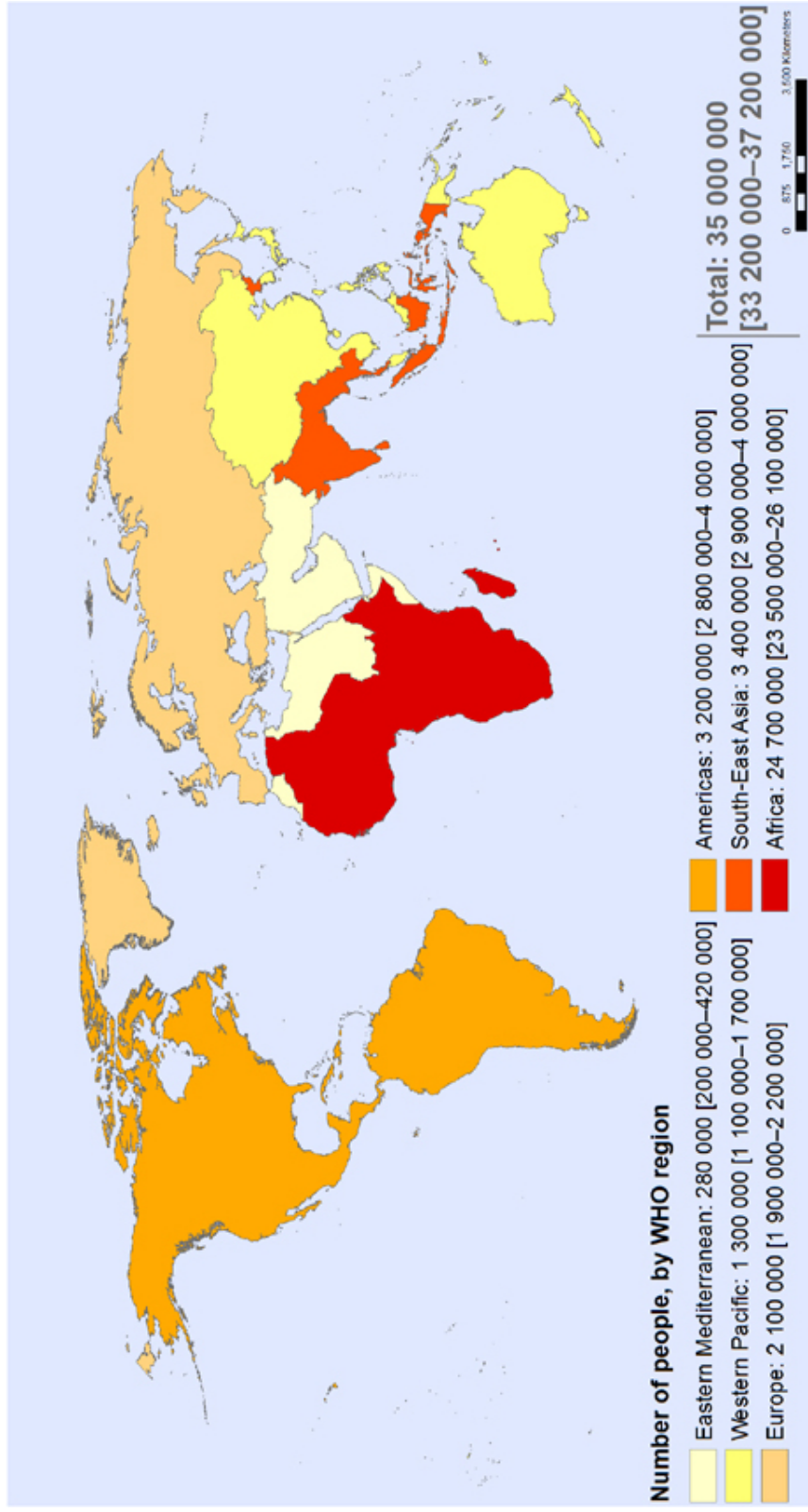
Figure 1.12: HIV-1 Prevalence by WHO Region. The map depicts the prevalence of HIV-1 in 2013 in each of the regions defined by the World Health Organization (WHO). Source: http:
//apps.who.int/gho/data/node.wrapper.HIVAIDS1?lang=en, accessed December 7th, 2015.

the fact that heterosexual contacts are more frequent than homosexual contacts increases the probability of new infections, even if the risk of transmission is higher for homosexual contacts (Table 1.2). In contrast, HIV-associated morbidity and mortality was decreasing at the time in the United States as a result of the advances in antiretroviral therapy [99] (Section 1.5.1). Fortunately, national AIDS control and prevention programs did result in a contention of the epidemic, and the world-wide incidence of HIV-1 began to decrease (Figure 1.11(a)). Nonetheless, a peak in the number of AIDS-related deaths was reached between 2004 and 2006 (Figure 1.11(c)). In 2007, there was a reduction in the number of deaths due to HIV-1. This was partly attributed to the scaling up of access to treatment [100]. The incidence of HIV-1 and the numbers of deaths attributed to HIV-1 have been continuously declining. However, since HIV-1 incidence is still considerable (Figure 1.11(a)) and people with HIV-1 are living longer, the prevalence of HIV-1 in the world continues to rise (Figure 1.11(b)).

In 2000, the United Nations Millennium Summit took place. Following that summit, eight international development goals were established, including the improvement of global health, education, environmental protection and gender equality by 2015 [101]. Goal number six includes halting and reversing the HIV-1 epidemic. This intent was complemented by two declarations adopted in 2006 and 2011 which bear the title *Political Declaration on HIV/AIDS*. These declarations recognize the urgency in providing universal access to HIV-1 treatment, prevention, support, and care. The 2011 *Political Declaration on HIV/AIDS* includes the following targets and elimination commitments, due 2015 [102]:

1. Reduce sexual transmission of HIV-1 by 50 percent
2. Reduce HIV-1 transmission among drug users by 50 percent
3. Eliminate mother-to-child HIV-1 transmission
4. Provide HIV-1 treatment for 15 million people
5. Eliminate obstacles that hinder low- and middle-income countries in providing treatment and diagnostics for HIV-1 infection as well as measures for preventing infection.

In the following, I mention some figures that can assess the extent to which these goals have been achieved, by comparing figures for the years 2011 and 2014. There has been a 13 percent reduction in the world-wide incidence of HIV-1 (Figure 1.11(a)). The percentage of pregnant, HIV-1 infected women receiving antiretroviral drugs for preventing mother-to-child transmission has increased from 58 percent in 2011 to 73 percent in 2014. The percentage of HIV-infected people receiving antiretroviral therapy has increased from 27 percent in 2011 to 40 percent in 2014 (AIDSinfo; http://aidsinfo.unaids.org; accessed November 20, 2015). Further targets regarding world-wide access to antiretroviral treatment have been defined by UNAIDS in 2014: the *90-90-90* treatment target, which I summarize in the following. By 2020, 90 percent of all people living with HIV-1 should be diagnosed. Of those infected, 90 percent should be on treatment. And of those treated, 90 percent should be virally suppressed [103] (see Section 4.1 for a definition of viral suppression).

1.4 Diversity of HIV-1

The biological characteristics of HIV-1 have resulted in substantial diversification of the virus. The variability of HIV-1 is noticeable both within a host and between different hosts. Furthermore, groups of closely related HIV-1 strains have a distinct geographical distribution. In the following, I summarize the molecular basis for HIV-1 variability and I explain the nomenclature that has been established in order to label closely related groups of HIV-1 variants that have emerged in the course of the pandemic. Furthermore, I briefly mention differences between groups of closely related HIV-1 variants with respect to their transmissibility and pathogenicity.

1.4.1 Molecular Basis for HIV-1 Variability

Compared to other pathogens, HIV-1 presents a very high degree of variability. The reasons for this large degree of variability lie in the lack of a proof-reading ability of the viral enzyme RT [104] (3.4×10^{-5} mutations per bp per cycle [105]), the mutations induced by the host DNA deaminase APOBEC3G [106], as well as the high titers in which HIV-1 replicates. Specifically, 10.3×10^9 new virions are produced each day, on average [107]. Furthermore, HIV-1 is a diploid virus with the capacity of recombination. Two viral RNA copies are packaged into assembling virions, and these need not be identical. Recombination of viral genetic material results from the ability of the viral enzyme RT to alternate between RNA templates during the process of reverse transcription (Sections 1.1.1, 1.1.2, and 1.4.2) [108]. Last but not least, selective pressure exerted by antiretroviral chemotherapy is a further factor driving HIV-1 evolution. A consequence of this selective pressure is the emergence of drug resistance, which is reviewed in Section 1.5.3.

As mentioned in Section 1.2.2, HIV-1 infection begins with a nearly homogeneous viral population that starts to diversify as selective pressure from the immune system kicks in. While HIV-1's high productivity, combined with the sloppiness of its enzyme RT, leads to the production of many viral variants, mutations that allow HIV-1 to escape the immune system are fixated [108]. This leads to the emergence of a quasispecies in the host, a set of closely related yet genetically distinct viral variants [109, 110]. The distribution of variants in the quasispecies changes over time, as different forms of immune or drug pressure are exerted on the pathogen [111, 112]. Even if some viral variants cease to be present in blood, they can be archived in proviral DNA with the capacity to reemerge, depending on the selective pressure exerted at a given point in time. The mechanics of HIV-1 transmission represent a genetic bottleneck, as few transmitted founder HIV-1 variants initiate new infections (Section 1.2.1). Therefore, the transmission of the variability within a host to another host is severely limited [113].

1.4.2 HIV-1 Group M Subtypes

Over the course of epidemiological history, HIV-1 group M has formed distinct and wide-spread lineages called subtypes (Table 1.4). Initially, ten different subtypes, denoted by the letters A to J, were defined. However, the phenomenon of dual infection was not well characterized at the time. Dual infection is the result of simultaneous or sequential infection with two heterologous viral strains. Infection of a cell with more than one provirus may result in heterozygous packaging of the progeny virus. Upon infection of further cells with diploid and heterozygous progeny virus, multiple template switching events might occur during reverse transcription, giving rise to recombinant viral variants [50]. Eventually, it became evident that subtypes E and I emerged as a result of recombination from other subtypes. Because of this, subtypes E and I were removed from the lineage classification and the concept of a CRF was introduced. CRFs are viral lineages with genome regions that can be clearly mapped to other subtypes (Figure 1.13). CRFs are denoted by the abbreviation *CRF*, appended to a two-digit index designating the order of CRF discovery, an underscore, and the letters designating the subtypes from which the recombinant form emerged. For example, CRF01_AE, was the first recombinant form that was discovered, and it was determined to have emerged from subtypes A and E (see below). A recombinant form of HIV-1 is called CRF if it is detected in at least three individuals, and unique recombinant form (URF) if it could be detected in at most two individuals (http://www.hiv.lanl.gov/, accessed November 11th, 2015). An exception in the rules of nomenclature was made for Subtype E, which was determined to be a recombinant of subtype A and another subtype (Figure 1.13). Specifically, subtype E was renamed to CRF01_AE
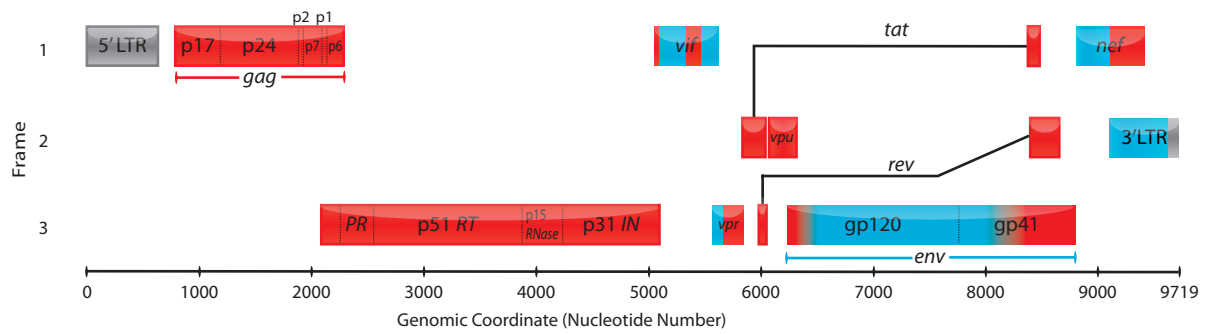
Figure 1.13: Recombination of CRF01_AE. Initially, the circulating recombinant form (CRF) CRF01_AE was though to be a *pure* subtype and therefore named subtype E. However, it was later determined that subtype E emerged through recombination of subtype A, and a further subtype for which no full-length sequence is available. Subsequently, the former subtype E was renamed to CRF01_AE. Above, the segments of the genome of CRF01_AE that correspond to subtype A and to the hypothetical subtype E, respectively, are delimited by colors. Source: `http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/breakpoints.html`, with modifications. Accessed December 7th, 2015.

■ Segment originating from subtype A ■ Segment originating from a hypothetical subtype E ■ Segment with unknown origin

although no full-length sequence of a subtype-E representative has been found. No further exceptions in the rules of nomenclature shall be made, as it was established that an unknown parent subtype of further, emerging CRFs should be denoted by *U*, for unknown [114]. In CRF nomenclature, emergence of a novel CRF as a result of recombination of a previously established CRF is denoted with the two digits indexing the parent CRF instead of the letter denominating the subtype. Furthermore, CRFs resulting from the recombination of more than two subtypes or CRFs are denoted with the letters *cpx* for complex. Thus, former subtype I was determined to have arisen from subtypes A, G, H, K, and a further, unknown lineage, and was subsequently renamed as CRF04_cpx. Up to now, 72 CRFs are known ([114] and extended version accessed November 11th, 2015 on `http://www.hiv.lanl.gov/`). Differences between subtypes have been established in terms of the speed at which the host progresses towards AIDS if the infection remains untreated. Specifically, subtype-D infection has been associated with a faster disease progression to death than subtype A [50, 115–118], and than other subtypes as well [119]. Some studies assert differences between HIV-1 subtypes in transmission efficiency, replicative capacity, and virulence. These are reviewed in [120, 121]. However, the findings are partially contradictory such that I prefer not to summarize them, in this work. Note that HIV-1 subtypes have a differential geographical distribution (Figure 1.8).

Depending on intra-subtype variability, subtypes may be further split into sub-subtypes. Sub-subtypes are labeled by appending a number to the letter identifying the subtype. In retrospect, it appears that subtypes B and D should have been labeled as sub-subtypes B1 and B2, respectively. However, relabeling has been avoided in order to retain the subtype labels that have been used in many scientific publications ([114] and extended version on `http://www.hiv.lanl.gov/` accessed November 11th, 2015). Lastly, I mention that since HIV-1 group M originated (Table 1.4), its genetic diversity has been increasing, in terms of both inter- and intra-subtype diversity, as well as the number of recombinant strains [56].

1.5 Antiretroviral Therapy

The following sections aim at providing a comprehensive overview of antiretroviral therapy. First, I review the history of antiretroviral therapy. Then, I present currently approved antiretroviral compounds along with their mechanism of action and review the consequences of the selection of drug-resistance mutations in HIV-1 for

these drug compounds. Subsequently, I summarize key features of state-of-the art antiretroviral therapy. Lastly, I provide a perspective on likely future developments of antiretroviral chemotherapy.

1.5.1 History of Antiretroviral Therapy

In this section, a historical account of antiretroviral therapy can be found. This account begins by detailing on treatment of HIV-1 with single drug compounds, describes the discovery of HIV-1 drug resistance, and the development of combination antiretroviral therapy (cART). The development of novel drug compounds and drug classes is recounted in its treatment-historical context. Furthermore, I outline the progress of antiretroviral therapy in terms of antiviral efficacy, side-effect profiles, and pill burden. Last but not least, I show how progress in antiretroviral therapy has resulted in the recommendation for immediate treatment with antiretroviral drugs for every patient with HIV-1 infection, which was not the case in the early years of antiretroviral therapy.

The Age of Monotherapy

After the identification of HIV-1 as the etiologic agent of AIDS in 1983 (Section 1.3), great efforts were undertaken in order to find a way to control the virus. By 1987, two molecular targets used by modern antiretroviral drugs had been proposed: (1) RT, through the use of nucleotide analogs as chain terminators, and (2) PR, through the use of effective inhibitors. However, other drug targets that have not led to the development of effective antiretroviral drugs were proposed as well [122, 123]. Among the compounds being tested as antiretroviral drugs at the time [122], the thymidine analog azidothymidine (AZT), a nucleotide reverse-transcriptase inhibitor (the abbreviation NRTI encompasses both nucleotide and nucleoside reverse-transcriptase inhibitors), showed a significant reduction in patient mortality at 24 weeks after treatment initiation (1/145 deaths vs. 19/137 deaths compared to placebo) [90]. (AZT was later called zidovudine (ZDV), and the abbreviations AZT and ZDV are now used interchangeably). Approval of AZT by the FDA followed in 1987 (Table 1.5). Further studies showed that the intake of four or five daily doses of AZT (up to 1,500 mg/day; the currently recommended dose is 500 mg/day) in patients infected with HIV-1 led to a significant delay in progression towards AIDS when compared to placebo [124, 125]. The benefits were evident for patients with more than 200 but less than 500 CD4+ cells per microliter of blood. However, patients who started AZT after receiving an AIDS diagnosis showed decreased mortality only in the first year of therapy, compared to placebo [126, 127]. Furthermore, patients with more than 500 CD4+ cells per microliter of blood had no benefit in starting AZT therapy immediately, when compared to patients who started therapy after their CD4+ cell count dropped below 500 cells per microliter of blood [127–129]. Since AZT therapy also resulted in neutropenia and anemia in a dose-dependent fashion [124, 129], it was clear that the risks and benefits of AZT therapy had to be weighed against each other. The benefits derived from AZT intake mainly consisted in a transient increase in CD4+-cell counts and a delay in disease progression. *In-vitro* AZT-susceptibility testing of isolates from patients who had been treated with the drug showed that HIV-1 developed resistance against AZT [130]. Comparative sequencing of AZT-resistant and wild-type HIV-1 variants showed that AZT-resistant variants accumulated certain mutations [131]. Subsequently, the development of drug resistance by HIV-1 delivered an explanation for the transient nature of the benefits of therapy with AZT. Following AZT's approval by the FDA, three further NRTIs were approved for treating HIV-1 infection: didianosine (ddI; 1991), zalcitabine (ddC; 1992), and stavudine (d4T; 1994; Table 1.5). After the HIV-1 variant infecting a patient developed resistance to AZT, the use of another antiretroviral drug resulted again in a transient benefit for the patient with a delay in disease progression, but

Table 1.5: NRTIs by Date of Approval by the FDA. Nucleoside and nucleotide reverse-transcriptase inhibitor (NRTI) formulations that have been approved by the Federal Drug Administration (FDA) are tabulated below, by date of approval. All currently approved NRTIs present risk of lactic acidosis, indicated in their respective black-box warnings. Extracted from `http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm`, accessed November 25th, 2015.

Brand Name	Generic Name	Approval Date	Comments
Retrovir	Zidovudine, azidothymidine, ZDV, AZT	March 19th, 1987	Risk of hematological toxicity, myopathy
Videx	Didanosine, dideoxyinosine, ddI	October 9th, 1991	Risk of potentially fatal adverse reactions, pancreatitis, severe hepatomegaly with steatosis
Hivid	Zalcitabine, ddC	June 19th, 1992	Risk of potentially fatal adverse reactions, severe peripheral neuropathy, pancreatitis, severe hepatomegaly with steatosis, hepatic failure. No longer marketed
Zerit	Stavudine, d4T	June 24th, 1994	Risk of hepatomegaly with steatosis, pancreatitis
Epivir	lamivudine, 3TC	November 17, 1995	Risk of lactic acidosis and severe hepatomegaly with steatosis
Combivir	3TC and AZT coformulation	September 27, 1997	See 3TC and AZT
Ziagen	Abacavir sulfate, ABC	December 17th, 1998	Risk of severe hepatomegaly. High risk of hypersensitivity reaction in patients who carry the HLA-B*5701 allele
Videx EC	Enteric coated didanosine, ddI EC	October 31st, 2000	See ddI
Trizivir	ABC, AZT, and 3TC coformulation	November 14, 2000	See ABC, AZT, and 3TC
Viread	Tenofovir disoproxil fumarate, TDF	October 26, 2001	Risk of severe hepatomegaly with steatosis
Emtriva	Emtricitabine, FTC	July 2nd, 2003	Risk of severe hepatomegaly with steatosis
Epzicom	ABC and 3TC coformulation	August 2nd, 2004	See ABC and 3TC
Truvada	FTC and TDF coformulation	August 2nd, 2004	See FTC and TDF

Table 1.6: PIs by Date of Approval by the FDA. Protease inhibitor (PI) formulations that have been approved by the Federal Drug Administration (FDA) are tabulated below, by date of approval. Extracted from http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm and https://aidsinfo.nih.gov/drugs, accessed November 25th, 2015.

Brand Name	Generic Name	Approval Date	Comments
Invirase	Saquinavir mesylate	December 6th, 1995	
Norvir	Ritonavir RTV	March 1st, 1996	Risk of life threatening drug-drug interactions
Crixivan	Indinavir, IDV	March 13th, 1996	Risk of nephrolithiasis and renal toxicity
Viracept	Nelfinavir mesylate, NFV	March 14th, 1997	
Fortovase	Saquinavir	November 7th, 1997	Soft-gel-capsule formulation which is no longer marketed
Agenerase	Amprenavir, APV	April 15th, 1999	No longer marketed
Kaletra	Lopinavir and RTV coformulation, LPV/r	September 15th, 2000	Coformulation with RTV as a pharmacokinetic enhancer
Reyataz	Atazanavir, ATV	June 20th, 2003	
Lexiva	Fosamprenavir Calcium, FPV	October 20th, 2003	
Aptivus	Tipranavir, TPV	June 22nd, 2005	Risk of hepatotoxicity and intracranial hemorrhage
Prezista	Darunavir, DRV	June 23rd, 2006	
Evotaz	ATV and cobicistat coformulation, ATV/c	January 29th, 2015	Coformulation with cobicistat as a pharmacokinetic enhancer
Prezcobix	DRV and cobicistat coformulation, DRV/c	January 29th, 2015	Coformulation with cobicistat as a pharmacokinetic enhancer

no benefit on the long run [132]. Furthermore, antiretroviral drugs approved until 1994 displayed a very unfavorable toxicity profile ([133–135]; reviewed in [136, 137]). This motivated the attempt to improve therapy by alternating monotherapy regimens [138, 139]. In 1995, a drug with a new mechanism of action was approved by the FDA: saquniavir (SQV; Table 1.6), a protease inhibitor (PI). Although the drug was substantially better tolerated than nucleoside analogs of the time [140], monotherapy with the drug failed to produce a sustained decline in plasma VL due the development of drug resistance by the virus [141].

Progress Towards Triple Therapy

In order to attempt to delay the emergence of drug-resistant HIV-1 strains, the benefits of chemotherapy with combinations of AZT and the nucleoside analogs ddC or ddI were investigated. When compared to monotherapy, dual therapy with nucleoside analogs was found to provide significant improvements in the survival, CD4+ depletion, and AIDS progression rates of patients [142, 143]. Nevertheless, the benefits were still not durable. A further NRTI, lamivudine (3TC), was approved by the FDA in 1995. *In-vitro* experiments had shown that 3TC and AZT presented synergistic effects in inhibiting HIV-1 replication. This effect was due to the fact that

Table 1.7: **NNRTIs by Date of Approval by the FDA.** Non-nucleoside reverse-transcriptase inhibitor (NNRTI) formulations that have been approved by the Federal Drug Administration (FDA) are tabulated below, by date of approval. Extracted from http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm and https://aidsinfo.nih.gov/drugs, accessed November 25th, 2015.

Brand Name	Generic Name	Approval Date	Comments
Viramune	nevirapine NVP	June 21, 1996	Immediate release formulation
Rescriptor	Delavirdine DLV	April 4th, 1997	
Sustiva	Efavirenz EFV	September 17th, 1998	
Intelence	Etravirine ETR	January 18th, 2008	
Viramune XR	Nevirapine NVP	March 25, 2011	Extended release formulation
Edurant	Rilpivirine RPV	May 20th, 2011	

viral variants that had selected for a resistance mutation against 3TC remained sensitive to AZT [144]. The synergy of the drug combination was confirmed in clinical trials [145–149].

The possible benefits of triple therapy were tested with the PI SQV. While triple therapy with SQV, AZT, and ddC was superior when compared to dual therapy, there was a progressive loss of antiviral efficacy in all study arms [150]. This was partially due to the fact that SQV presents poor oral bioavailability [151]. In 1996, a further drug with a novel mechanism of action was approved by the FDA: nevirapine (NVP; Table 1.7), a non-nucleoside reverse-transcriptase inhibitor (NNRTI). Phase I/II clinical trials demonstrated that as a monotherapy, the drug remained fully active for less than four weeks, after which HIV-1 developed resistance to the drug. In combination with AZT, however, the drug remained active for up to 12 weeks [152, 153]. Nonetheless, the antiviral activity of dual therapy waned with time. Sustained antiviral activity lasting more than one year was only seen in clinical trials testing triple therapy with AZT, ddI, and NVP. Specifically, in treatment-experienced patients, the antiviral activity of AZT + ddI + NVP decreased with time and VLs tended to return to baseline levels after 48 weeks of treatment [154]. In therapy-naïve patients, however, antiviral activity with this drug combination was sustained [155]. Two further PIs were approved by the FDA in 1996: ritonavir (RTV) and indinavir (IDV). RTV presented favorable pharmacological properties, including drug absorption over a prolonged period of time and a comparatively extended half life. This permitted twice-daily dosing of the drug [156], a major break through at the time. Used as monotherapy, RTV resulted in a reduction in the VL that tended to be less pronounced after 12 weeks of therapy. Nonetheless, after one year of monotherapy, at the latest, resistant viral variants emerged [156, 157]. In combination with two nucleoside analogs, RTV therapy lead to comparatively strong and sustained reductions in the VL, as well as increases in CD4+ cell counts. However, the drug presented a very unfavorable toxicity profile, with over 80% of patients experiencing side effects, including peri-oral paresthesia and taste perversion [158, 159]. This forced 24% of the patients in one study to discontinue medication with the drug [159]. In clinical trials testing IDV together with AZT and 3TC, an unprecedented degree of viral suppression could be demonstrated. After 52 weeks of treatment, more than 80% of the patients presented VLs less than 500 copies per milliliter of blood serum, and more than 60% of the patients presented VLs below the limit of quantification of 50 copies per milliliter of blood serum [143, 160]. After two years of therapy, viral suppression below 500 copies per milliliter of blood serum was still given for 78% of the patients [161]. These results were very encouraging and made the drug combination very popular. However, in the long run, IDV produced significant renal toxicity, with nephrolithiasis occurring in 9% of the

patients [162]. The PIs SQV, RTV, IDV, and nelfinavir (NFV; approved by the FDA in 1997; Table 1.6) are often classified as first-generation PIs. Common to all of them (except for RTV) is their poor oral bioavailability [163], resulting in a high pill burden and low concentrations in blood. In contrast, RTV is an inhibitor of the enzyme cytochrome P450 3A4. This enzyme metabolizes many drugs, including PIs. As mentioned before, RTV is tolerated poorly at virologically active doses. However, a smaller dose of RTV (100 mg) is well tolerated and, coadministrated with other PIs, greatly improves their bioavailability and half life [164]. (This is called *boosting* with ritonavir, denoted with */r*). Around 1995, important scientific progress was made with respect to the understanding of the disease. First, the fact that the quantity of HIV-1 RNA in plasma predicts the progression of the disease was proven [32]. Second, it was discovered that HIV-1 also replicates during the clinically latent stage of the disease [33, 34], and that this replication occurs in high titers, and is characterized by a rapid viral turnover [107, 165]. Both scientific achievements significantly contributed to the understanding and diagnostics of the disease. In conjunction with the benefits of triple therapy, these achievements lead to significant reductions in the morbidity and mortality caused by HIV-1 infection [99]. A further noteworthy achievement is the discovery that AZT prevents mother-to-child transmission of HIV-1 [91].

IMPROVING COMBINATION ANTIRETROVIRAL THERAPY

cART, also called highly active antiretroviral therapy (HAART), is defined as treatment of HIV-1 infection with at least three different drugs of at least two different classes. Before cART was introduced into the clinic, HIV-1 was a death sentence for those infected. The introduction of cART transformed this deadly infection into a manageable, chronic disease [99]. Although early cART rescued many lives, life quality of patients on early cART was poor. cART is associated with short-, mid-, and long-term side effects, and early cART had especially severe side effects (reviewed in [136, 137]). Less severe side effects of antiretrovirals include headache, rash, diarrhea, mood problems, insomnia, nightmares, nausea, vomiting, fatigue, and dizziness. The most detrimental adverse events occurring with all NRTIs result from the mitochondrial toxicity of these drug compounds. This toxicity arises from the fact that NRTIs not only inhibit the viral RT (Sections 1.1.2 and 1.5.2), but also the mitochondrial DNA polymerase γ [166, 167]. Inhibition of the mitochondrial DNA polymerase γ results in reductions in the amount of mitochondrial DNA present in the cell [168], and in the production of dysfunctional mitochondrial protein. This, in turn, results in severely impaired mitochondrial function [169], often manifesting as myopathy, neuropathy, lipoatrophy ([170, 171], Figure 1.14), lactic acidosis, and hepatic steatosis (reviewed in [136, 172–174]; see also Table 1.5). The more recently approved nucleoside analogs 3TC, abacavir (ABC), and emtricitabine (FTC), as well as the nucleotide analog tenofovir (TFV), are also toxic for mitochondria, but to a decreased extent [136] (Table 1.5). For improving antiviral action, TFV is formulated as the prodrug tenofovir disproxil fumarate (TDF), which is converted in the body to TFV (Section 1.5.2). Recently, a combination tablet containing a new prodrug of TFV has been approved, tenofovir alafenamide fumarate (TAF; Table 1.8). TAF promises an improved side-effect profile with regard to renal and bone toxicities [175, 176], when compared to TDF. The PIs SQV, lopinavir (LPV), IDV, tipranaivr (TPV), and RTV (Table 1.6) have been associated to dislipidemia (abnormal quantities of lipids in blood) [136, 177]. Improvement of PIs has also taken place since the more recently approved PIs fosamprenavir (FPV), atanazavir (ATV), and darunavir (DRV) (Table 1.6) do not seem to cause dislipidemia [136]. The reduced toxicity of newer compounds represents an improvement in cART.

The reduction of the pill burden is a further improvement that cART has undergone over time. The
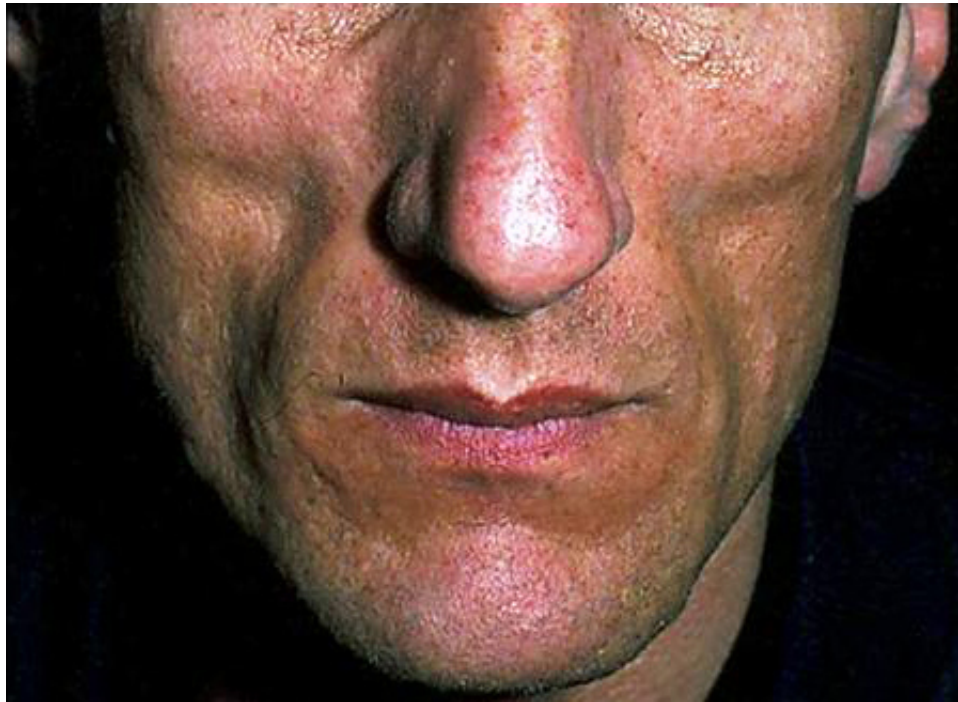
Figure 1.14: Lipoatrophy of the Face. Therapy with nucleoside reverse-transcriptase inhibitors is associated with lipoatrophy, which is part of larger clinical picture of changes in the fatty-tissue distribution in the body. Lipoatrophy is the result of asymmetric loss of body fat in the extremities and face. Lipodystrophy additionally includes the accumulation of body fat in the trunk. While lipodystrophy does not directly cause harm, it is psychologically detrimental to the patient, and may affect drug adherence when patients blame HIV-1 chemotherapy for this condition. Image courtesy of AIDS Images Library http://www.aids-images.ch.

high pill burden during early antiretroviral therapy was mainly due to three causes. First, during the time at which monotherapy was widespread, there were attempts to retain drug activity in spite of drug resistance with high and frequent doses (reviewed in [178]). Second, early PIs were characterized by low bioavailability (as mentioned above), which required frequent intake of a large number of pills [179, 180]. Third, there has been a need to counteract the side effects of antiretrovirals with further medication. Furthermore, additional drugs are needed for treating opportunistic infections [181–183]. The pill burden in a patient cohort from South Alberta, Canada, was calculated for the years 1990 to 2010 [184]. Canada boasts different private and public health insurance schemes that facilitate access to antiretroviral drugs for most residents of the country (http://www.catie.ca/en/practical-guides/managing-your-health/19, Accessed March 1st, 2016), such that the South Alberta Cohort can be thought to be exemplary for a resource-rich setting. In 1990, patients in this cohort had a total mean pill burden of 4.9 (the standard deviation is given in parenthesis, in the following; 3.1) pills per day, which is decomposed in 4.6 (2.5) pills per day, on average, for antiretrovirals and 0.2 (1.9) pills per day, on average, for other medication. Thereafter, the total mean number of pills per day peaked in 1998, with 12.1 (6.9), which is decomposed in a mean of 10 (6.1) pills per day for antiretrovirals, and 1.7 (3.2) pills per day, on average, for other medication. After this peak, the total mean number of pills per day decreased to 6.8 (5.6) in 2006, decomposed in 3.8 (3.0) pills per day, on average, for antiretrovirals and 2.4 pills per day, on average, for other medication. In the years following 2006, and until 2010, these figures did not change much. With respect to intake frequency, from 1990 to 1996, most patients in the South Alberta Cohort had to take their medication three times a day. The intake frequency of most patients could be reduced to twice daily between 1998 and 2006, with a further reduction to once-daily from 2008 on. The following factors contributed to the reduction of pill burden over time. (1) The use of drug formulations with improved phar-

Table 1.8: **Multi-Class coformulations by Date of Approval by the FDA.** Multi-class coformulations that have been approved by the Federal Drug Administration (FDA) are tabulated below, by date of approval. Extracted from http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm and https://aidsinfo.nih.gov/drugs, accessed November 25th, 2015. Note: cobicistat is a pharmacokinetic enhancer.

Brand Name	Generic Name	Approval Date	Comments
Atripla	FTC, TDF, and EFV coformulation	July 12th, 2006	See Tables 1.5 and 1.7
Complera	FTC, TDF, and RPV coformulation	August 10th, 2011	See Tables 1.5 and 1.7
Stribild	FTC, TDF, EVG, and cobicistat coformulation	August 27, 2012	Coformulation with cobicistat as a pharmacokinetic enhancer. See Table 1.5
Triumeq	3TC, ABC, and DTG coformulation	August 22, 2014	See Tables 1.5 and 1.9
Dutrebis	3TC and RAL coformulation	February 6th, 2015	See Tables 1.5 and 1.9
Genvoya	Tenofovir alafenamide fumarate (TAF), FTC, EVG, and cobicistat coformulation	November 5th, 2015	Risk of fatal hepatomegaly with steatosis. See Tables 1.5 and 1.9.

macological properties [185, 186]; (2) the use of RTV ([185]; see above), and later also cobicistat (boosting with cobicistat is denoted by /c) [187], as a pharmacokinetic enhancers; (3) the introduction of tablets containing more than one drug compound (*coformulation tablets*; Tables 1.5, 1.6, 1.8). The reduction in the pill burden of cART has improved patient adherence, and thus lowered therapy failure rates [188, 189], as well as enhanced the quality of life of HIV-1-infected patients.

After the introduction of reverse-transcriptase inhibitors (RTIs) and PIs, drugs with other mechanisms of action have followed. Specifically, enfuvirtide (T-20) [190], a fusion inhibitor (FI), was approved by the FDA on March 13th, 2003. T-20 must be administered by subcutaneous injection, and reactions localized in the injection site occur in 98 percent of patients treated with T-20. These include pain, erythema, pruritus (itching), and enduration (hardening). Due to the inconvenient administration form of the drug (parenterally with the need to reconstitute [§] the drug prior to use) and its side effects, T-20 is only used as a further active compound in patients with HIV-1 drug resistance who cannot be treated otherwise. Nonetheless, T-20 represents a further therapy option. Maraviroc (MVC) [191] is an entry inhibitor (EI) that is only active on CCR5-tropic viruses (Section 1.1.2). It was approved by the FDA on August 6th, 2007. In three clinical studies, the drug showed favorable tolerability with less (91.9%) therapy-naïve patients experiencing adverse events when treated with MVC as compared to treatment with efavirenz (EFV) (94.2 %; Table 1.7) [192, 193]. Nonetheless, the FDA has issued a black-box warning regarding the hepatotoxicity of MVC (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm, accessed November 25th, 2015). MVC has not been very successful in terms of the fraction of patients treated with the drug, for the following reasons. (1) Viral tropism determination must precede prescription of the drug [191]. (2) It is not possible to take the drug (only) once daily. (3) In therapy-naïve patients, the drug failed to show non-inferiority against efavirenz [193] and, in com-

---

[§]In the context of injectable drugs, *reconstitution* is the process of mixing a drug that is present in a powdered form with a diluent. This is done in order to be able to inject the drug in a liquid form into a patient's body.

Table 1.9: INIs by Date of Approval by the FDA. Integrase-inhibitor (INI) formulations that have been approved by the Federal Drug Administration (FDA) are tabulated below, by date of approval. Extracted from `http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm` and `https://aidsinfo.nih.gov/drugs`, accessed November 25th, 2015.

Brand Name	Generic Name	Approval Date	Comments
Isentress	Raltegravir RAL	October 12th, 2007	
Vitekta	Elvitegravir EVG	September 24, 2014	
Tivicay	Dolutegravir DTG	August 13th, 2013	

bination with DRV/r, against TDF+FTC [194]. Even so, MVC is an effective and well-tolerated therapy option for patients harboring CCR5-tropic viruses. Last, but not least, three integrase strand-transfer inhibitors (INIs) have been approved by the FDA for use in antiretroviral therapy (Table 1.9), along with coformulations of INIs with other antiretroviral drugs (Table 1.8). Raltegravir (RAL) [195, 196] was the first INI approved by the FDA, in 2007. In a phase III clinical trial on treatment-naïve patients, VL decay in patients treated with RAL was significantly faster than in patients treated with EFV [197], which was unprecedented up to that point in time. RAL is well tolerated, with less than 2.4% of the patients experiencing a severe study-drug-related clinical adverse event, and fewer therapy-naïve patients experiencing clinical adverse events while treated with RAL (90%), when compared to those treated with EFV (96.5%) [197, 198]. Despite these favorable characteristics, once-daily dosing of RAL is not possible. In part, this is due to the fact that pharmacokinetic enhancement of RAL with RTV is not possible, as the drug is not metabolized by cytochrome P450 enzymes. In contrast, the second INI approved by the FDA, elvitegravir (EVG) [199], can be boosted with a pharmacokinetic enhancer such as RTV or cobicistat, which allows for once-daily dosing [199, 200]. In one study comparing EVG to RAL, 75% of patients treated with EVG had a VL below 50 copies per milliliter of blood serum 48 weeks after treatment initiation, as compared to 73% of the patients treated with RAL [201]. With respect to toxicities, 1.1% of patients treated with EVG experienced a severe, study-drug-related adverse event, compared to 2% of patients treated with RAL. Thus, both drugs have similar efficacy. However, EVG, in contrast to RAL, permits once-daily dosing and an EVG-containing, once-daily combination tablet has been approved by the FDA (Table 1.9). Unfortunately, extensive cross-resistance between RAL and EVG can occur, with five out of six mutation patterns conferring high-level drug resistance to RAL also conferring high-level resistance to EVG, and vice versa (`http://hivdb.stanford.edu/pages/phenoSummary/Pheno.INI.Simple.html`, accessed December 2nd, 2015; see also Section 1.5.3). The third INI approved by the FDA (in 2013), dolutegravir (DTG) [202, 203], has surpassed RAL and EVG in several aspects. First, DTG remains active in several EVG- and RAL-resistant HIV-1 variants, and has a higher genetic barrier to drug resistance ¶ [204–206]. Up to now, no cases of development of resistance to DTG have been reported in therapy-naïve patients who started therapy with DTG [204]. Second, DTG can be dosed once daily without the need for a pharmacokinetic enhancer. This is advantageous, as boosting of antiretrovirals is associated with an increased risk of drug-drug interactions [202]. A DTG-containing once-daily single combination tablet has been approved by the FDA (Table 1.9). Third, in several clinical trials, DTG has been shown to be statistically superior to RAL [207], DRV/r [208], and EFV [209], 48 weeks after treatment initiation. At the same time, DTG has a favorable toxicity profile, with the odds

---

¶The genetic barrier to drug resistance is the probability that an HIV-1 variant will develop resistance to a certain drug within a certain amount of time. This probability is sometimes expressed with a correlated measure, e.g. the number of substitutions in the nucleotide sequence of an HIV-1 gene that are necessary for the virus to become resistant to the drug in question.

of experiencing an adverse event while on DTG therapy being significantly lower when compared to ATV/r, EFV, and LPV/r, and insignificantly different when compared to DRV/r, EVG/c, RAL, and rilpivirine (RPV; Table 1.5) [210]. In summary, the development of further drugs with novel mechanisms of action has improved cART in four aspects. (1) Several single-tablet once-daily regimens have become available, fostering adherence and improving intake convenience. (2) Improved and novel drug mechanisms of action allow for full antiviral activity in viral strains that could have been previously deemed untreatable. Furthermore, drugs with a higher genetic barrier to drug resistance have become available. (3) The side-effect profiles of these newer drug compounds are more favorable. (4) The efficacy of the drugs has risen, with faster, more prevalent and more durable viral suppression below the limit of quantification of 50 HIV-1 RNA copies per milliliter of blood serum.

Recommendations on when to start antiretroviral therapy have changed over time. In the time at which monotherapy was widespread, the first guidelines for the treatment of HIV-1 infection were published in 1990, recommending initiation of AZT monotherapy only after the CD4+-cell count had dropped below 500 cells per microliter of blood [129]. After the development of effective cART, theoretical reasoning led to the recommendation of initiating triple therapy early (Table 1.10). The reasons for this recommendation were the observation that treatment during primary infection had shown a clinical benefit and the understanding that the viral population in a patient becomes more diverse as the disease progresses. However, this recommendation was not backed by clinical data, as they were lacking at the time [211, 212]. Later on, however, experience with cART showed that treatment was disadvantageous for the patient, as it bore the risk of serious drug-related adverse effects and of limitation of future treatment options due to resistance development (Section 1.5.3). cART has the ability to protect and therefore allow for the restoration of CD4+ cells and general immune function, and to avoid HIV-associated complications that may occur regardless of CD4+-cell count. Furthermore, it can reduce the risk of HIV-1 transmission to nearly zero. However, it was thought that the risks of initiating treatment (serious drug-related adverse effects and limitation of future treatment options), with a CD4+ cell count above 350 cells per microliter of blood, outweighed the benefits of viral suppression, as the risk of progression is inversely proportional to CD4+-cell counts [213, 214]. As further and better antiretroviral drugs were approved (see above), the risks associated with cART decreased. Specifically, (1) single-tablet, once-daily regimens substantially increased the convenience for the patient. (2) Concerns about the preservation of future treatment options were substantially diminished as more potent drugs and drugs with novel mechanisms of action were approved. (3) Newer drugs present a better tolerability and are associated with fewer toxicities. As clinical studies showed that the benefits of new drugs and new formulations of these drugs translated into a clinical advantage for patients who start treatment earlier, recommendations on when to start treatment gradually raised the CD4+-cell-count treatment-initiation threshold [215]. The most recent guidelines recommend treatment of all persons with HIV-1 infection as early as possible, regardless of CD4+-cell counts (Table 1.10).

1.5.2 Mechanism of Action of Antiretroviral Drugs

In the following, the mechanism of action of the most important classes of antiretroviral drugs is summarized. First, NRTIs and NNRTIs are presented, which are both RT inhibitors. Then, PIs and INIs are briefly reviewed. Last, antiretroviral drugs preventing the entry of HIV-1 into the cell are presented.

Nucleoside and Nucleotide Reverse-Transcriptase Inhibitors

Phosphorylated NRTIs (Figure 1.15) are analogs of cellular 2'-deoxynucleotides (Figure 1.16). All NRTIs

Table 1.10: IAS-USA Recommendations for Treatment Initiation by Publication Year. IAS-USA guidelines (`http://www.iasusa.org/guidelines-archive/`; accessed December 2nd, 2015) on when to start antiretroviral therapy have changed over time. Below, CD4+-count-dependent criteria for treatment initiation are tabulated. The term *individualize* elliptically refers to patient-specific treatment initiation depending on CD4+-cell count, CD4+-cell decline, viral load, comorbidities, and other patient-specific factors. CD4+-cell counts are expressed in terms of cells per microliter of blood, and viral loads in terms of HIV-1 RNA copies per milliliter of blood serum. VL: viral load

Year / CD4+ cell count	< 200	$200 - 350$	$350 - 500$	> 500
1996	Treat	Treat	Treat	Treat if symptomatic or VL greater than 30 000 copies/ml or rapid decline in CD4+ cells
1998	Treat	Treat	Treat if VL greater than 5,000 copies/ml, otherwise consider treatment initiation	See left
2002	Treat	Treat if symptomatic. Otherwise, individualize	See left	See left
2004	Treat	Treat if symptomatic. Otherwise, individualize	Treat if symptomatic. Otherwise defer treatment	See left
2006	Treat	Treat if symptomatic. Otherwise, individualize	Treat if symptomatic. Otherwise defer treatment	See left
2010	Treat	Treat	Treat	Treat if symptomatic or in acute phase. Otherwise, individualize
2012	Treat	Treat	Treat	Treat
2014	Treat	Treat	Treat	Treat

(a) 3TC   **(b)** ABC   **(c)** AZT

(d) d4T   **(e)** ddI   **(f)** FTC

(g) TFV
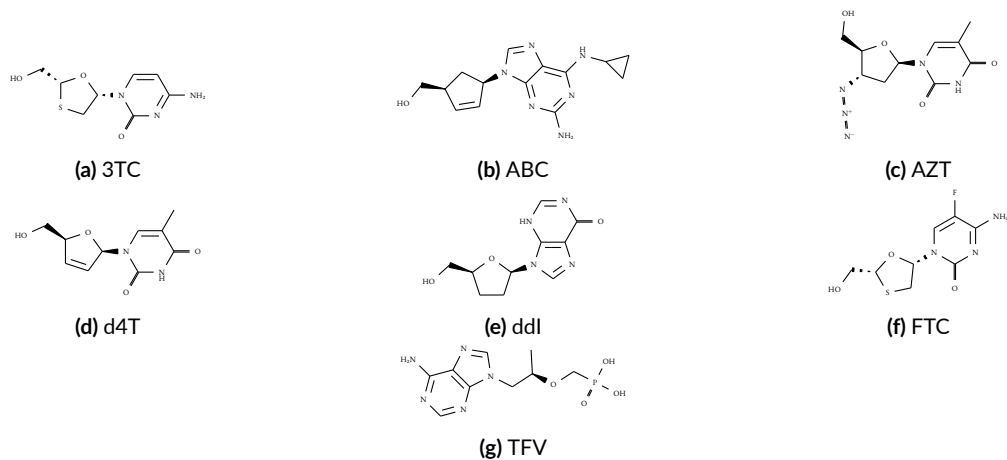
Figure 1.15: Chemical Structure of NRTIs. Nucleoside and nucleotide reverse-transcriptase inhibitors (NRTIs) are chemically similar to 2'-deoxynucleosides (Figure 1.16). All compounds above are nucleoside reverse-transcriptase inhibitors, except for TFV. TFV is a nucleotide (i.e. nucleoside monophosphate) reverse-transcriptase inhibitor formulated as a prodrug, which can be either tenofovir disoproxil fumarate or tenofovir alafenamide fumarate. NRTIs are phosphorylated in the cell, which results in their transformation into nucleoside triphosphates. The affinity of cellular polymerases for the phosphorylated forms of NRTIs is much reduced, compared to endogenous nucleotides. This results from their chemical differences to cellular 2'-deoxynucleotides. However, the affinity of HIV-1 reverse transcriptase for phosphorylated NRTIs is comparable to its affinity for endogenous nucleotides. NRTIs lack a 3'-hydroxl group, making their phosphorylated forms 2'-3'-dideoxynucleotide analogs. Therefore, their incorporation by the viral reverse transcriptase into a growing DNA molecule leads to the inability of the enzyme to incorporate further nucleotides, thus impeding reverse transcription of viral RNA into DNA. Source: `https://pubchem.ncbi.nlm.nih.gov/`. 3TC: lamivudine; ABC: abacavir; AZT: zidovudine; d4T: stavudine, ddI: didanosine, FTC: emtricitabine; TFV: tenofovir.

(a) 2'-Deoxyadenosine   **(b)** 2'-Deoxycytidine

(c) 2'-Deoxyguanosine   **(d)** 2'-Deoxythymidine

Figure 1.16: 2'-Deoxynucleosides. Living beings obtain 2'-deoxynucleosides from nutrition or by *de novo* synthesis. 5'-phosphorylated 2'-deoxynucleosides are called 2'-deoxynucleotides, and they are the building blocks of deoxyribonucleic acid (DNA). Source: `https://pubchem.ncbi.nlm.nih.gov/`.

(a) EFV

(b) ETR

(c) NVP

(d) RPV
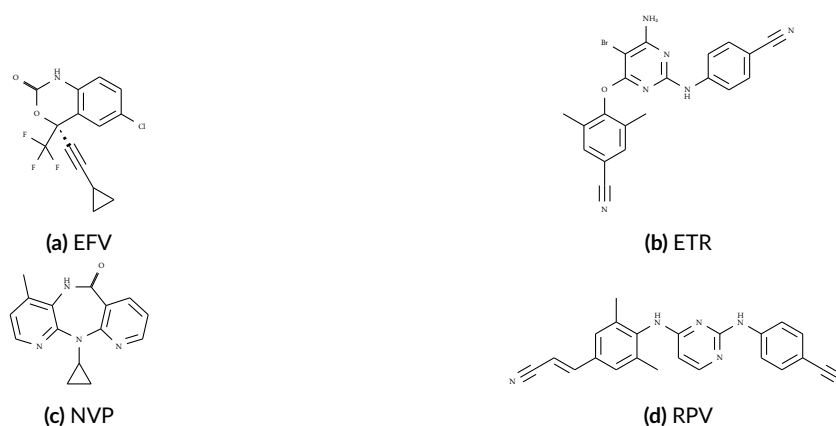
Figure 1.17: Chemical Structure of NNRTIs. Non-nucleoside revere-transcriptase inhibitors (NNRTIs) exert antiviral action by inducing and binding to a pocket in the viral enzyme reverse transcriptase. This results in conformational changes of the enzyme that reduce its catalytic activity. Source: `https://pubchem.ncbi.nlm.nih.gov/`. EFV: efavirenz; ETR: etravirine; NVP: nevirapine; RPV: rilpivirine.

are formulated as *prodrugs*. A prodrug is a drug compound that requires metabolization in the body in order to become pharmacologically active. In order to be antivirally active, NRTIs need to be converted from 2'-3'-dideoxynucleosides (except for TFV) to 2'-3'-dideoxynucleoside triphosphates through several phosphorylation reactions. TFV presents two particularities. First, TFV is formulated as either the prodrug TDF or the prodrug TAF. Second, TFV is a nucleotide analog, specifically, a 2'-3'-dideoxynucleoside monophosphate that requires further phosphorylation in order to be converted to the triphosphate, active form. Phosphorylation rates differ among NRTIs, which affects the concentrations of their active forms. Due to the chemical differences between cellular 2'-deoxynucleosides and NRTIs, most endogenous polymerases have a substantially reduced binding affinity to the triphosphate forms of NRTIs. However, the binding affinity of the viral RT to NRTI triphosphates is equal or only slightly reduced in comparison to endogenous deoxynucleotides. Thus, NRTI triphosphates compete with cellular deoxynucleoside triphosphates for incorporation by the viral RT during proviral DNA synthesis. Once incorporated into a growing DNA chain, their lack of a 3'-hydroxl prevents the incorporation of further nucleotides, thus interrupting proviral DNA synthesis [216, 217].

Non-Nucleoside Reverse-Transcriptase Inhibitors

NNRTIs (Figure 1.17) are non-competitive inhibitors of the viral RT. This inhibition occurs through the induction of the *NNRTI binding pocket* by the inhibitor and subsequent binding of the inhibitor to this pocket. The NNRTI binding pocket is not present in RT when no NNRTI is bound. The NNRTI binding pocket is hydrophobic and consists of the residues L100, K101, K103, V106, T107, V108, V179, Y181, Y188, V189, G190, F227, W229, L234, and Y318 of the RT subunit p66, and the residue E138 of the other RT subunit, p51 (Figure 1.18). The NNRTI binding pocket is located at a distance of approximately 10 Å from the catalytic site of RT. Binding of NNRTIs to RT results in conformational changes that reduce the catalytic activity of the enzyme. In contrast to NRTIs, NNRTIs present a comparatively high target selectivity which makes them HIV-1-specific RT inhibitors [217, 218].

Protease Inhibitors

PIs (Figure 1.19) prevent viral maturation by inhibiting the viral enzyme PR. PR (Figure 1.20) cleaves viral

(a) View of the Active Site   **(b)** View of the Region behind the Active Site
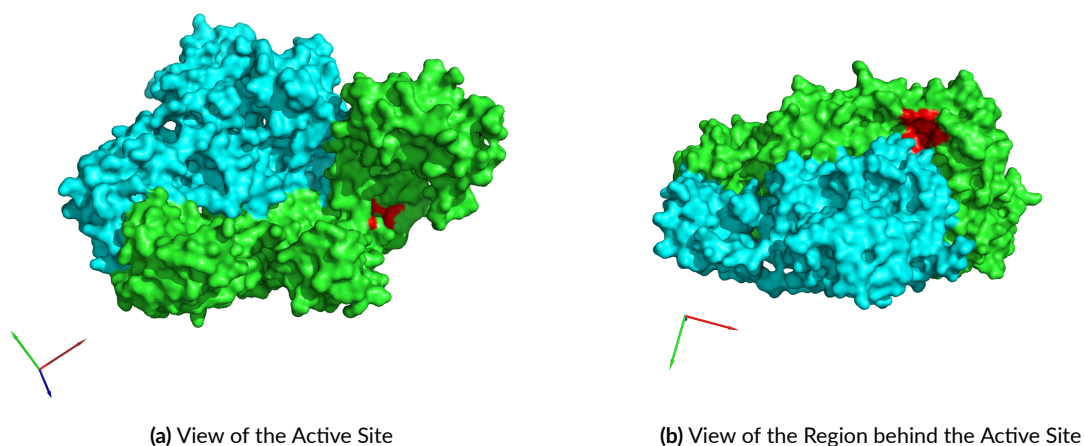
Figure 1.18: NNRTI Binding Pocket. Two views of the molecular surface of HIV-1 reverse transcriptase can be seen above. (a) shows the active site of the protein, while (b) shows the region behind the active site of the protein. The reverse transcriptase consists of two functional subunits: p51 and p66. In both figures, p51 is displayed in green, while p66 is displayed in blue. The residues L100, K101, K103, V106, T107, V108, V179, Y181, Y188, V189, G190, F227, W229, L234, and Y318 of p66, and the residue E138 of p51 constitute the NNRTI binding pocket. These residues are displayed in red. Coordinate axes have been included in the lower left corners. Source: `http://www.rcsb.org/`. ID: 3HVT. Images generated with PyMOL.

gag and gag-pol polyproteins into their functional forms. All PIs but TPV are competitive peptidomimetic inhibitors, with a structure similar to that of the natural substrate of PR. In contrast, TPV is not peptidomimetic. All PIs bind to the active site of PR. Cleavage of peptidomimetic inhibitors is prevented through a hydroxyethylene core (Figure 1.19). [185].

INTEGRASE STRAND-TRANSFER INHIBITORS

The viral enzyme IN (Figure 1.21) binds to the double-stranded proviral DNA that results from reverse transcription, and mediates the integration of proviral DNA into the host cell's genome. INIs (Figure 1.22) bind to the active site of IN following a DNA-induced conformational change. This active site is formed by the amino acids with reference-alignment positions 50 to 212. Probably, proviral DNA is part of the binding site of INIs. Their binding to IN sequesters metal ions that are essential for the strand-transfer reaction that is responsible for incorporating the proviral DNA into the cellular DNA. In this way, the integration of proviral DNA into the host's genome is inhibited [219, 220].

MVC AND T-20

EIs constitute an antiretroviral drug class of which only MVC (Figure 1.23) has been approved for treatment of HIV-1 infection. MVC binds to a hydrophobic pocket in the transmembrane helices of CCR5 (Section 1.1.2), altering the conformation of the extracellular portions of the loops formed by the transmembrane helices. It is believed that this conformation change disturbs the binding of HIV-1 to CCR5. Since the CCR5 binding site of MVC is different from the CCR5 binding site of gp120 (Sections 1.1.1 and 1.1.2), entry inhibition by MVC is allosteric. The CXCR4 coreceptor is not inhibited by MVC. Hence, MVC is only active against CCR5-tropic viruses (Section 1.1.2), and viral tropism determination must precede MVC prescription [221].

FIs block gp41-mediated membrane fusion. The drug T-20 is the only representative of this class that has

(a) ATV

(b) DRV

(c) FPV

(d) IDV

(e) LPV

(f) NFV

(g) RTV

(h) SQV

(i) TPV
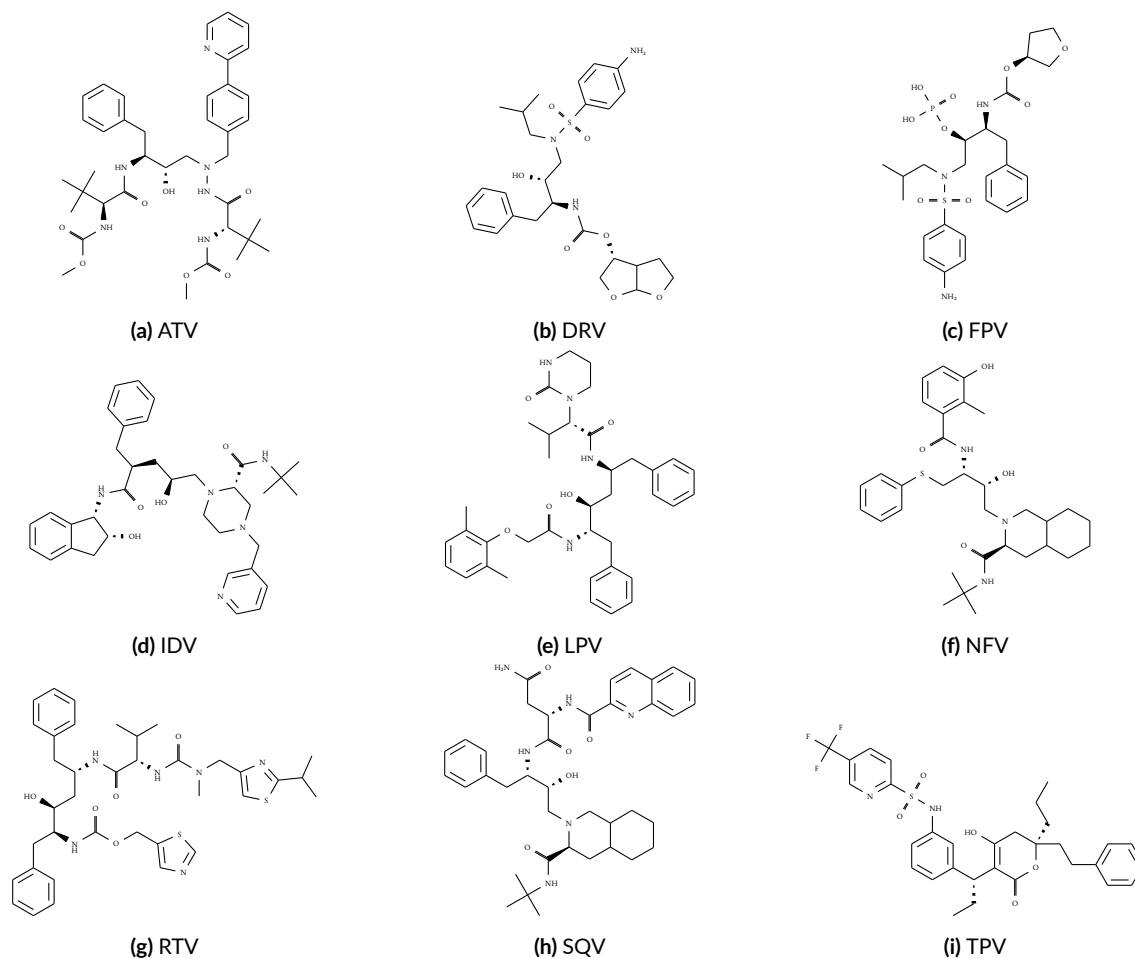
Figure 1.19: Chemical Structure of PIs. Protease inhibitors (PIs) prevent maturation of newly produced virions by competitive binding to HIV-1 protease. All protease inhibitors except for TPV are peptidomimetic and feature a hydroxylethylene core that prevents their cleavage by the viral protease. (Continued in the next figure.) Source: https://pubchem.ncbi.nlm.nih.gov/. ATV: atazanavir; DRV: darunavir; FPV: fosamprenavir; IDV: indinavir; LPV: lopinavir; NFV: nelfinavir; RTV: ritonavir; SQV: saquinavir; TPV: tipranavir.
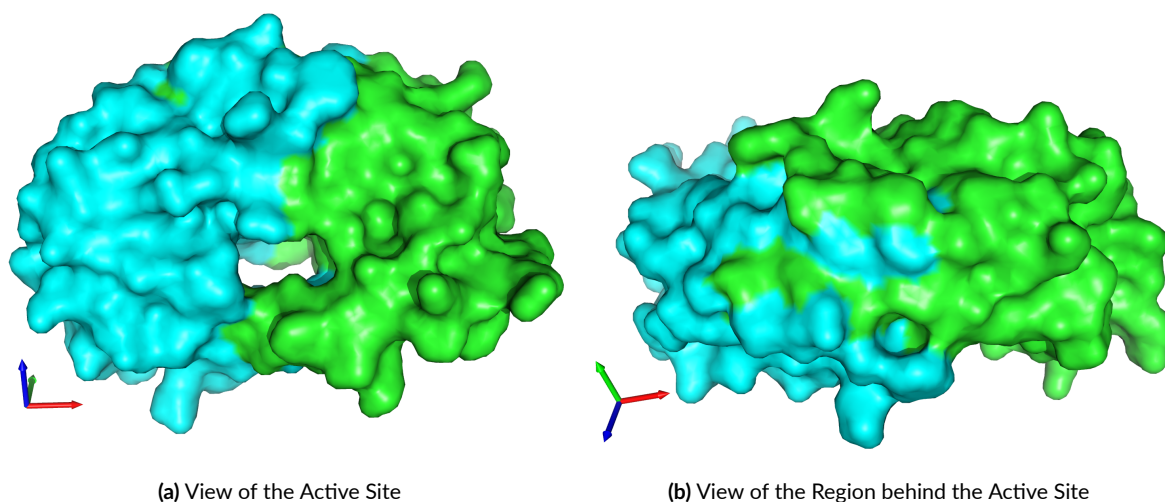
(a) View of the Active Site **(b)** View of the Region behind the Active Site

Figure 1.20: HIV-1 Protease. Two views of the molecular surface of HIV-1 protease can be seen above. (a) shows the active site of the protein, while (b) shows the region behind the active site of the protein. HIV-1 protease consists of two identical subunits, colored in green and blue above. Each one of these subunits consists of 99 amino acids. Coordinate axes have been included in the lower left corners. Source: `http://www.rcsb.org/`. ID: 3OXC. Images generated with PyMOL.

been approved for treatment of HIV-1 infection. T-20 is a linear synthetic peptide composed of 36 amino acids: Ac-Y-T-S-L-I-H-S-L-I-E-E-S-Q-N-Q-Q-E-K-N-E-Q-E-L-L-E-L-D-K-W-A-S-L-W-N-W-F-NH$_2$. The final step in the process of HIV-1 entry into the cell is the fusion of the viral membrane with the membrane of the infected cell. This last step involves the formation of a six-helix bundle structure through the interaction of two heptad-repeat domains in gp41: HR1 and HR2. The amino-acid sequence of T-20 is identical to a fragment of the sequence of HR2. Thus, T-20 is an inhibitor of the HR1-HR2 domain interaction through competitive binding to the HR1 domain. The disruption of the HR1-HR2 domain interaction precludes the fusion of the viral and cellular membranes [221].

1.5.3 The Emergence of Drug Resistance

Two characteristics of HIV-1 result in the continuous generation of many new viral variants. (1) HIV-1 replicates in high titers, with 10.3×10^9 new virions produced each day, on average [107]. (2) The viral enzyme RT lacks proof-reading ability [104], resulting in 3.4×10^{-5} mutations per bp per replication cycle [105]). Furthermore, the cellular DNA deaminase APOBEC3G contributes to the generation of new viral variants by inducing G-to-A mutations in the HIV-1 genome [222]. An antiretroviral drug exerts a selective pressure that can filter out viral variants that are susceptible to that drug, while drug-resistant variants are selected. This process is further facilitated if drug concentrations, in the body as a whole or in certain compartments of the body, sink to subinhibitory levels. Thus, the probability that drug-resistant variants will emerge is correlated with the extent of viral replication while the drug is present in the body of a patient. Usually, the simultaneous intake of at least three drugs is required in order to reduce viral replication to levels that allow for several years of antiretroviral treatment before clinically-relevant drug-resistant variants are selected [217]. The emergence of drug-resistant HIV-1 strains through treatment can occur with each of the available antiretroviral drugs. Furthermore, resistance selected by one drug can also confer resistance to other drugs that have not been previously used by the patient. This phenomenon, named cross-resistance, occurs among drugs with the same mechanism of action (Section 1.5.2).
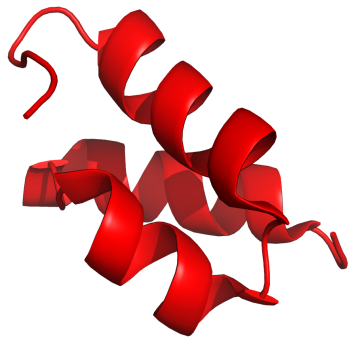
(a) Zinc-Binding Region
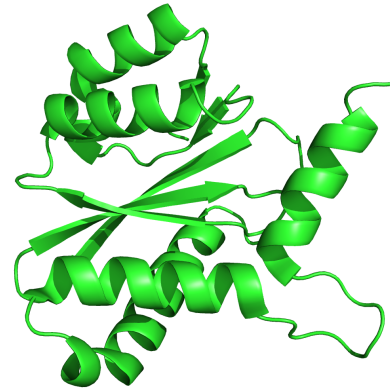
(b) Catalytic Core Domain
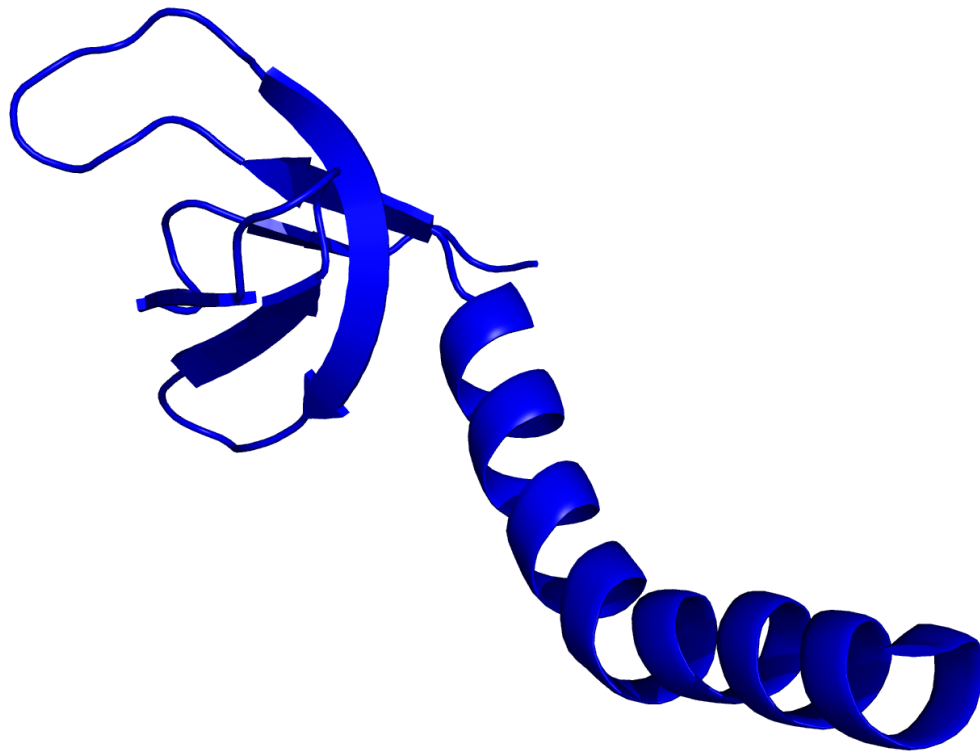
(c) DNA-Binding Region

Figure 1.21: HIV-1 Integrase. HIV-1 integrase is conformed by three domains: the zinc-binding domain (a), the catalytic core domain (b), and the DNA-Binding domain (c). To my knowledge, no crystal structure of the entire protein exists. Source: `http://www.rcsb.org/`. IDs: 1K6Y, 1QS4, and 1EX4, respectively. Images generated with PyMOL.
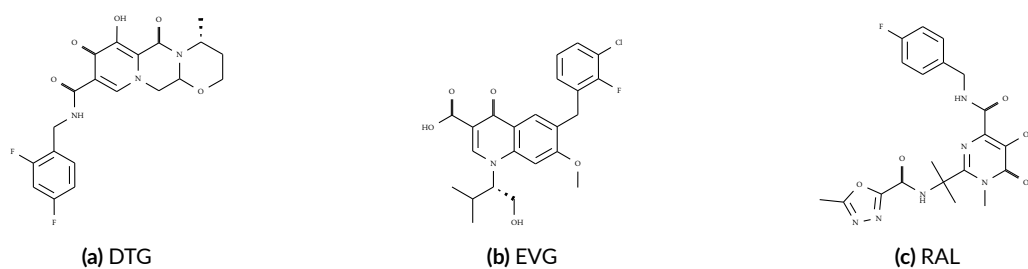
Figure 1.22: Chemical Structure of INIs. Integrase strand-transfer inhibitors (INIs) prevent the integration of proviral DNA into the host cell's genome. In this manner, they interrupt the replication cycle of HIV-1. Source: `https://pubchem.ncbi.nlm.nih.gov/`. DTG: dolutegravir; EVG: elvitegravir; RAL: raltegravir.
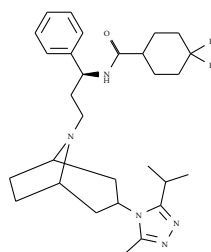
Figure 1.23: Chemical Structure of MVC. Maraviroc (MVC) binds to the CCR5 chemokine receptor on the surface of the cell and hinders the entry into the cell by CCR5-tropic HIV-1 strains. Source: `https://pubchem.ncbi.nlm.nih.gov/`.

Once a resistant variant has been selected in a patient, it can be transmitted to another patient. When drug resistance is transmitted, the therapy options of the newly infected patient can be limited. Treatment-selected drug-resistant variants often revert to the wild type when drug pressure is removed [223]. This can occur when the treatment of a patient is interrupted or when a drug-resistant viral strain is transmitted to a new host. Reversion to wild type becomes apparent through longitudinal sequencing of the virus circulating in the blood plasma (before and after reversion). However, drug-resistant variants can be archived in the body of the patient and promptly reemerge upon resumption of drug pressure [224].

RESISTANCE TO NRTIS

There are two known general mechanisms by which HIV-1 can become resistant to NRTIs. The first one involves the selection of *discriminatory* mutations at the RT reference-alignment positions 65, 74, 115, 151, and 184 (Figure 1.24), among others. Residues at these reference positions make important contacts with NRTIs, and are thus crucial for the binding affinity of the viral RT to NRTIs. Discriminatory mutations reduce the binding affinity of RT to phosphorylated NRTIs, but not to endogenous deoxynucleotides. This diminishes the incorporation of NRTIs into the growing proviral DNA, and allows for the continuation of the viral replication cycle in the presence of the drug. The second general resistance mechanism involves mutations at the reference-alignment positions 41, 67, 210, 215, and 219 (Figure 1.24), among others. Mutations at these residues can lead to an increased rate of excision of incorporated NRTIs, thus reversing chain termination and allowing for continuation of the polymerization reaction [216, 217, 220]. For tables summarizing NRTI-resistance mutations by drug, see [225].
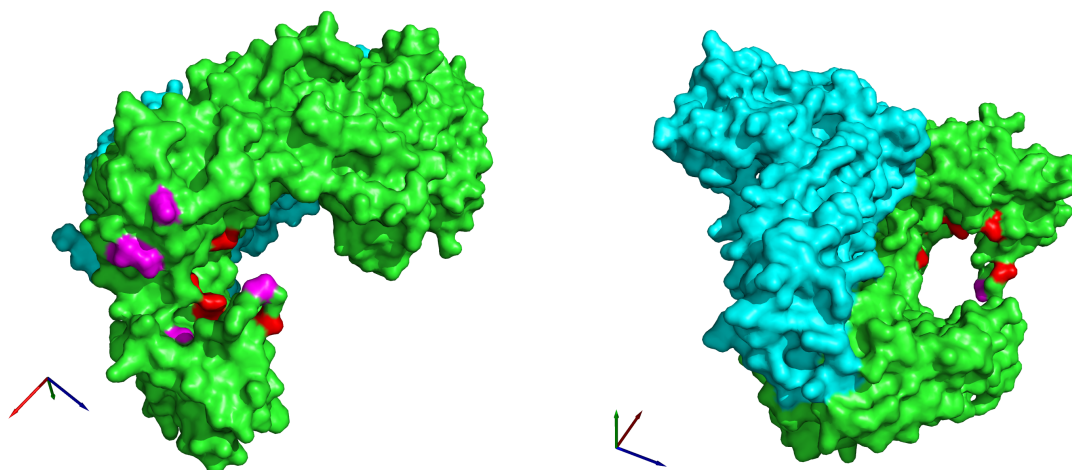
Figure 1.24: NRTI Drug-Resistance Mutations. Two views of the molecular surface of HIV-1 reverse transcriptase can be seen above. Both views depict the active site of the protein. The reverse transcriptase consists of two functional subunits: p51 and p66. In both figures, p51 is displayed in green, while p66 is displayed in blue. The residues K65, L74, Y115, Q151, and M184 of p66 (shown in red) are involved in the selection of drug-resistance mutations that confer HIV-1 the ability to discriminate between cellular nucleotides and nucleotide analogs. The residues M41, D67, L210, T215, and K219 of p66 (shown in magenta) are involved in the selection of drug-resistance mutations that increase the rate of excision of incorporated NRTIs. Coordinate axes have been included in the lower left corners. Source: http://www.rcsb.org/. ID: 3HVT. Images generated with PyMOL.

Resistance to NNRTIs

Resistance mutations at the RT reference-alignment positions 100, 101, 103, 106, 179, 181, 188, 190, and 236 can confer resistance to NNRTIs (Figure 1.18). Most NNRTI resistance mutations are found in and around the NNRTI binding pocket (Section 1.5.2). Three general mechanisms for NNRTI resistance have been described. (1) Loss or change of key hydrophobic interactions between NNRTIs and RT through mutations of the amino-acid residues Y181, Y188, and F277. (2) Steric hindrance of NNRTI binding through mutations at amino-acid residues L100 and G190. These two residues are in the central region of the NNRTI binding pocket. (3) Mutations at residues K101 and K103 which are located at the entrance of the NNRTI binding pocket. These mutations cause interference with the entry of NNRTIs into the NNRTI binding pocket [217]. For tables summarizing NNRTI-resistance mutations by drug, see [225].

Resistance to PIs

The development of resistance to PIs is a gradual process resulting from the accumulation of *major* and *minor* PI-resistance mutations. Major PI-resistance mutations (Figure 1.25) confer drug resistance by themselves, are usually selected first, and result in a widening of the active site of PR. This widening causes decreased binding of PIs to PR, but at the same time, it reduces the catalytic activity of PR, thus decreasing viral replication. Minor PI-resistance mutations do not confer drug resistance by themselves, but rather restore the decreased replication capacity caused by the selection of major resistance mutations. Minor PI-resistance mutations are usually selected after major PI-resistance mutations have been selected. However, some minor PI-resistance mutations occur naturally, i.e. in the absence of PI pressure. In addition to the selection of minor PI-resistance mutations, co-evolution of the PR cleavage sites in gag and gag-pol and their surrounding regions play a role in the restoration of replication capacity after major PI mutations have been selected [185, 220]. For tables summarizing PI-resistance mutations by drug, see [225].
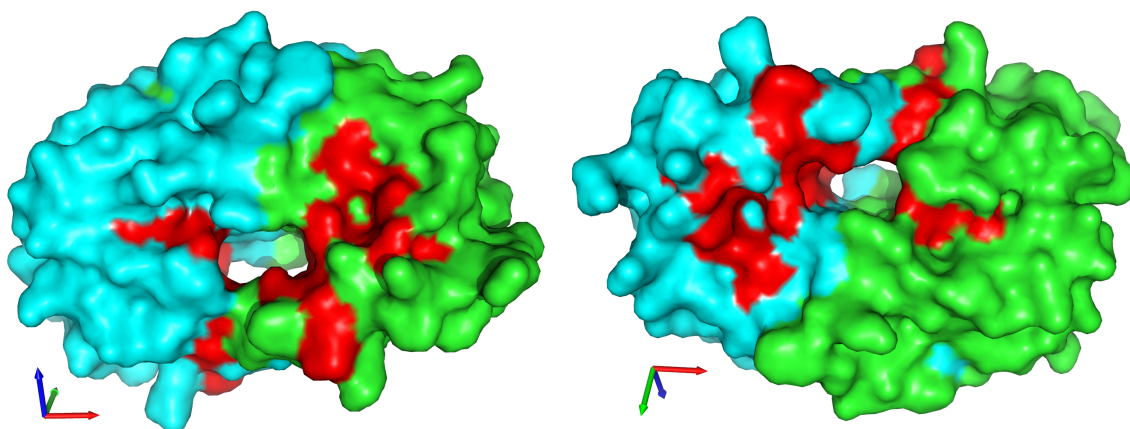
Figure 1.25: PI Major Drug-Resistance Mutations. Two views of the molecular surface of HIV-1 protease can be seen above. In both views, the active site of the protein can be seen. The protease consists of two identical subunits, colored in green and blue above. In both figures, residues D30, V32, M46, I47, G48, I50, I54, Q58, T74, L76, V82, N83, I84, N88, and L90 are marked in red, as they are involved in the selection of major drug-resistance mutations against at least one protease inhibitor. Coordinate axes have been included in the lower left corners. Source: http://www.rcsb.org/. ID: 3OXC. Images generated with PyMOL.

RESISTANCE TO INIS

The mechanism by which resistance to INIs is developed is not yet sufficiently understood. Currently, the structure of the full-length IN enzyme is not available, although structures of fragments of the enzyme containing its catalytic core (residues 50-212) do exist ([205, 226] and search query on http://www.pdb.org, December 10th, 2015). Even though the binding mode of INIs to IN fragments has been analyzed [205], resistance against INIs can only be described in terms of the mutations that are present in INI-resistant HIV-1 strains. For tables summarizing INI-resistance mutations by drug, see [225]. Resistance against RAL has been characterized with clinical data, showing that two main resistance pathways exist, as well as a third, less frequent one. The first resistance pathway involves development of the IN substitutions Q148H/K/R, often followed by *accessory* mutations at the IN residues number 74, 92, 97, 136, 138, 140, and 151 (Figure 1.26). The second resistance pathway usually starts with the substitution N155H, with occasional subsequent development of accessory mutations at the IN residues number 74, 97, 157, and 163 (Figure 1.26). The third, less frequent resistance pathway is characterized by the appearance of the substitutions Y143C/H/R, followed by accessory substitutions at IN residues number 74, 97, and 163 (Figure 1.26). The resistance-mutation profile of EVG is similar to that of RAL. Specifically, resistance to EVG occurs with pathways involving mutations at IN residues number 148 and 155, and their corresponding accessory substitutions. The resistance pathway involving IN residues number 143, however, does not affect EVG susceptibility. DTG remains effective on viruses that have selected INI drug-resistance mutations including mutations at IN residues number 148, 155, and 153 [220]. Furthermore, DTG presents the particularity that no *de novo* resistance mutations against the drug have been reported in previously treatment-naïve patients treated with the drug. However, this does not hold for treatment-experienced patients, even if they are INI naïve, as the emergence of DTG-resistance mutations in IN have been reported. Nonetheless, success rates with treatment-experienced patients treated with DTG remain high [227]. *In vitro*, the selection of DTG resistance mutations at IN residues number 121, 153, 118, 138, and 263 have been reported, with mutation R263K being the most common mutation. Although R263K confers low-level resistance to DTG, it also
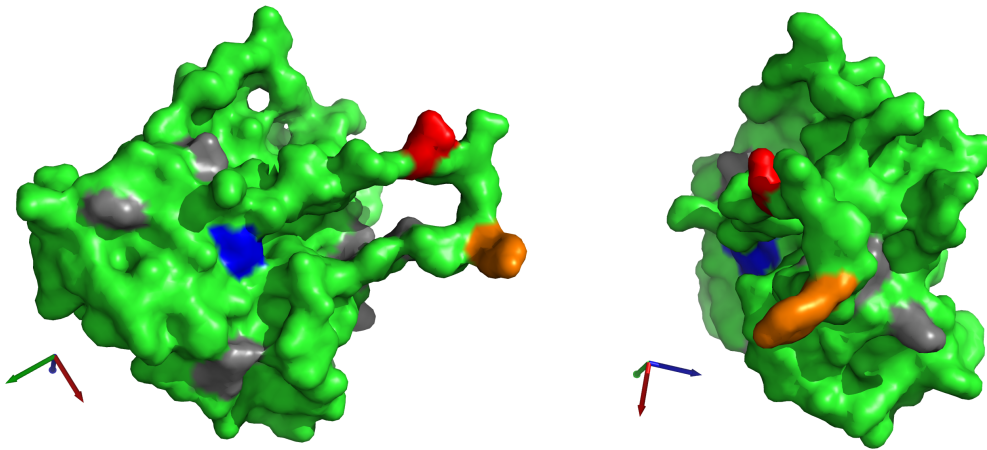
Figure 1.26: INI Drug-Resistance Mutations. Two views of the molecular surface of catalytic core of HIV-1 integrase can be seen above. In both figures, residues Q148, N155, and Y143 are colored in red, blue, and orange, respectively. Each of these residues is involved in a distinct resistance pathway. Residues L74, E92, T97, K136, E138, G140, E157, and G163 are colored in gray. These residues are involved in the selection of accessory mutations which follow the selection of mutations at residues Q148, N155, and Y143. Coordinate axes have been included in the lower left corners. Source: http://www.rcsb.org/. ID: 2ITG. Images generated with PyMOL.

impairs strand-transfer activity and viral replication capacity. Therefore, it has been suggested that during therapy with DTG, the virus is unable to select for further resistance mutations against DTG and also against other antiretroviral compounds used in the therapy [204].

Resistance to MVC and T-20

HIV-1 resistance to MVC can be given in two ways. (1) CXCR4-tropic strains exhibit a natural resistance to MVC, as the drug only inhibits the CCR5 coreceptor. Thus, the presence or emergence of CXCR4-tropic strains results in resistance to the drug compound. (2) CCR5-tropic strains can select for mutations in the *env* gene that will allow them to use the MVC-bound CCR5 coreceptor while remaining phenotypically CCR5 tropic. Key residues for MVC resistance are 316 and 326, located in the third hypervariable loop of gp120 (V3). Resistance to T-20 is mediated by substitutions at gp41 residues number 36 to 38, located at HR1, as well as other substitutions located in the vicinity of these residues. Since these mutations have a negative impact on the replicative capacity of the virus, compensatory mutations in HR2 may appear as well [226].

1.5.4 Antiretroviral Therapy Today

In this section, I summarize the goals of and recommendations for modern antiretroviral therapy, as defined by panels of American and European experts [228, 229]. Due to the high extent of infrastructure and monetary resources required to carry them out, many of the recommendations are only feasible in resource-rich settings.

Goals of Antiretroviral Therapy

The goals of antiretroviral therapy for HIV-1 infection can be summarized as follows.

- Stop replication of HIV-1

- Restore and/or preserve immune function
- Prevent transmission of HIV-1
- Alleviate symptoms of HIV-1 infection
- Stop viral diversification
- Reduce inflammation and persistent immune activation
- Preserve lymphoid tissue integrity
- Allow for pre- and post-exposure prophylaxis against HIV-1 infection.

These goals can be reached with antiretroviral chemotherapy, many times at the expense of drug toxicity and the inconvenience associated with the therapy.

When Should Antiretroviral Therapy Be Started?

As reviewed in Section 1.5.1, recommendations on when to start antiretroviral therapy are based on the presence or absence of symptoms in persons with HIV-1 infection, as well as on their CD4+ counts. Currently, initiation of antiretroviral therapy at the earliest possible time is recommended, as this is thought to bear no additional harm, entail health benefits, and be cost effective. However, the recommendation to initiate treatment is given more emphasis during primary HIV-1 infection (Section 1.2.2), if HIV-1 infection is symptomatic, or if CD4+-cells counts have sunk below 350 cells per microliter of blood. Antiretroviral treatment during primary HIV-1 infection promises benefits that can be retained throughout the life of the patient (Section 1.2.2): reduction of symptoms of ARS; reduction of VL set point and of the size of the viral reservoir; reduction of viral diversity; preservation and robust reconstitution of immune function with CD4+-cell counts above 900 cells per microliter of blood serum; preservation of lymphoid tissue integrity; reduction of inflammation and immune activation. A further, important benefit promised by treatment of primary HIV-1 infection is the prevention of HIV-1 transmission, as most transmissions of the virus occur during primary infection. Primary HIV-1 infection is thought to be a major driver of the pandemic. The recommendation to start treatment in asymptomatic patients with chronic HIV-1 infection, and CD4+-cell counts above 350 cells per microliter of blood serum is formulated in a less pressing fashion, with an emphasis on patient readiness to start treatment. Once initiated, antiretroviral therapy should not be interrupted, probably for the rest of the patient's life. Therefore, a strong commitment to therapy is required from the patient, especially in light of the possible short- and long-term drug toxicities that are to be expected as a result of treatment and the inconvenience associated to the treatment. Patient adherence is crucial to successful treatment, since lack of adherence can compromise viral suppression and foster the development of resistance.

Treatment with Antiretroviral Drugs

Genotypic resistance testing (Section 3.2) must precede initiation and, if possible, any changes in treatment. Among the possible treatment regimens for which safety and efficacy data are available, only a subset is recommended as first-line regimens (Table 1.11). The selection of these first-line regimens is based on their comparative efficacy, side-effect profiles, and convenience for the patient. The choice of the first-line regimen, among those recommended, should be based on the results of resistance testing, and on patient characteristics and preferences. Further, alternative first-line regimens have been defined for cases in which recommended regimens are not feasible [228, 229]. Following treatment change or treatment initiation, VL and CD4+-cell counts should be monitored every four weeks until suppression of the VL below the limit of quantification of 50 copies per

Table 1.11: Recommended First-Line Regimens. Recommended first-line regimens are tabulated below, along with important characteristics of these regimens. HLA-B*5701 refers to a certain human leukocyte antigen (HLA). 3TC: lamivudine; ABC: abacavir; ATV: atazanavir; CNS: central nervous system; DRV: darunavir; DTG: dolutegravir; EFV: efavirenz, EVG: elvitegravir; FTC: emtricitabine; INI: integrase strand-transfer inhibitor; NRTI: nucleoside or nucleotide reverse-transcriptase inhibitor; NNRTI: non-nucleoside reverse-transcriptase inhibitor; PI: protease inhibitor; RAL: raltegravir; RPV: rilpivirine; STR: single-tablet regimen; TDF: tenofovir disoproxil fumarate.

Regimen	Comments
2 NRTIs + INI	
3TC+ABC+DTG	Once-daily STR; do not use ABC if HLA-B*5701 positive
FTC+TDF+DTG	Once-daily regimen; long-term use of TDF is associated with increased risk of kidney injury
FTC+TDF+EVG/c	Once-daily STR; intake with food required; long-term use of TDF is associated with increased risk of kidney injury
FTC+TDF+RAL	Twice-daily regimen; long-term use of TDF is associated with increased risk of kidney injury
2 NRTIs + NNRTI	
FTC+TDF+EFV	Once-daily STR; EFV associated with risk of CNS symptoms; long-term efficacy and tolerability data available; inferior to INI-based regimens, mainly due to tolerability; long-term use of TDF is associated with increased risk of kidney injury
3TC+ABC+EFV	Once-daily regimen; risk of CNS symptoms; do not use ABC if HLA-B*5701 positive
FTC+TDF+RPV	Once-daily STR; intake with food required; RPV is not recommended for patients with VL > 100,000 copies / ml; greater risk of resistance development than other regimens; long-term use of TDF is associated with increased risk of kidney injury
2 NRTIs + PI	
FTC+TDF+DRV/r	Once-daily regimen; intake with food required; DRV causes rash in 10% of patients; long-term use of TDF is associated with increased risk of kidney injury
3TC+ABC+ATV/r	Once-daily regimen; ATV is associated with risk of jaundice, nephrolithiasis, cholelithiasis, and chronic kidney injury; do not use ABC if HLA-B*5701 positive
FTC+TDF+ATV/r	Once-daily regimen; ATV is associated with risk of jaundice, nephrolithiasis, cholelithiasis, and chronic kidney injury; long-term use of TDF is associated with increased risk of kidney injury

milliliter of blood serum has occurred. Thereafter, VL and CD4+-cell counts can be monitored every 3 months. After one year of viral suppression, if CD4+-cells are at a stable level above 350 cells per microliter of blood serum, and if the patient is adherent, VL and CD4+-cell-count monitoring frequency can be reduced to once every six months. Furthermore, toxicity monitoring of antiretroviral therapy through measurement of biomarkers is recommended as well.

The VL is the primary marker of treatment success and patient adherence. Therefore, an important goal of antiretroviral therapy is the suppression of the VL below the limit of quantification. Virological therapy failure is defined as two subsequent VL measurements above 50 copies per milliliter of blood serum, six months after the initiation of a treatment. Reasons for virological therapy failure include: the development of drug resistance by the virus, lack of patient adherence, and/or insufficient plasma drug concentrations due to drug-drug interactions and/or pharmacogenetic factors. After failure, genotypic resistance testing should be attempted. (Sequencing might not be possible if the VL is not sufficiently high.) Once virological therapy failure has been detected, the presumed reason for failure, as well as the VL measured at failure, will determine which course of action that should be taken. With VLs above 1,000 HIV-1 RNA copies per milliliter of blood serum, immediate treatment change is almost always recommended. If the VL is above 200, but below 1,000 HIV-1 RNA copies per milliliter of blood serum, a change of treatment should be considered. It is not clear if the treatment should be altered if the VL is below 200 HIV-1 RNA copies per milliliter of blood serum. In the case that resistance against the drugs in the failing regimen has been detected, it is recommended that the new therapy contain a fully active boosted PI and a drug class not previously used.

Virological failure is not the only important reason for which treatment should be switched. Further important reasons include alleviation and prevention of drug toxicities, avoidance of drug-drug interactions, planned pregnancy, aging or comorbidities, and treatment simplification. Common, nowadays avoidable drug toxicities include lipoatrophy (d4T, AZT), adverse events of the central nervous system (EFV), diarrhea (boosted PIs), and jaundice (ATV). Treatment simplification entails the possibility to get rid of treatment-related food restrictions, as well as to reduce the pill burden of the treatment in terms of the number of pills and their intake frequency.

1.5.5   Future Prospects of Antiretroviral Therapy

This last section on antiretroviral therapy is devoted to summarizing future developments of antiretroviral therapy that can be foreseen. First, I summarize attempts to reduce the toxicity of antiretroviral therapy by restricting the use of NRTIs. Then I explain the challenges related to an increasingly aging population of patients with HIV-1 infection. After briefly mentioning further antiretroviral drugs that may get approved in the near future, I lastly provide the perspective of reducing the incidence of HIV-1 with antiretroviral drug compounds.

NRTI-Sparing for Reduction of Side Effects

NRTIs have been the backbone of antiretroviral therapy since chemotherapy for HIV-1 infection exists. This has both historical and empirical reasons. Since NRTIs have been around for a long time (Section 1.5.1), there is a wealth of data available on their safety and efficacy. Furthermore, their pharmacological properties allow them to reach concentrations in the cell that cannot be easily equaled by other drugs. Several NRTIs (especially AZT, d4T, and ddI) are increasingly falling into disuse due to their association with mitochondrial toxicity (Section 1.5.1), neuropathy, and anemia. Even so, data on the long-term toxicity profile of newer, safer NRTIs (e.g.

TDF and ABC) has been accumulating. These include nephrotoxicity, increased risk of osteoporotic fracture, and increased risk of heart failure. This has motivated the investigation of NRTI-sparing regimens (also called nuke-sparing regimens) for treatment of HIV-1 infection. Deployed NRTI-sparing strategies are of two types. (1) Therapy with a drug combination including only one NRTI or no NRTIs (and further compounds), and (2) initiation of therapy with two NRTIs and further compounds (induction phase) with subsequent interruption of NRTI intake after suppression of the VL below the limit of detection (maintenance phase) [230]. Retainment of only one drug compound after viral suppression is called monotherapy maintenance. An exciting approach being tested at the moment involves nanoformulations of the drugs cabotegravir (Table 1.12) and RPV. During the induction phase, lasting 24 weeks, patients receive therapy with cabotegravir and two NRTIs. After the induction phase, the therapy is switched to nanoformulations of cabotegravir and RPV, which require intramuscular administration every one to three months. The use of the nanoformulations bears the risk that the drugs remain in the body for extended periods of time, even if treatment interruption should be indicated. The first study using this strategy, but without deploying nanoformulations, has been recently successfully completed [231]. The second part of the trials, which use the nanoformulations is underway. The use of NRTI-sparing regimens appears promising but still requires substantial research for a wide-scale roll out. The desired benefit of reduced toxicity is yet to be demonstrated in long-term clinical trials [230, 232, 233].

Aging with HIV

The success of cART is reflected by the fact that, nowadays, the life expectancy of people with HIV-1 is almost the same as that of the general population [234]. Furthermore, the mean age of people living with HIV-1 is increasing all over the world. At the same time, this development has brought up issues concerning the health of older people living with HIV-1 infection, as they are prone to acquiring the following comorbidities:

- Cardiovascular disease
- Cancer
- Declining kidney function
- Liver-associated morbidities
- Decreased bone mineral density
- Cognitive impairment
- Sexual dysfunction
- Mental disease
- Frailty

While these morbidities are frequent in persons of older age, people living with HIV-1 acquire them more frequently and at a younger age. In the general population, the age threshold at which increased monitoring and intervention is considered beneficial to health and/or required for maintaining good health lies between 60 and 75 years of age. For people living with HIV-1, however, this threshold sinks to 50 years of age. Disregarding antiretroviral treatment, older patients with HIV-1 infection are at an increased risk of progressing to AIDS than younger patients. Furthermore, old age is accompanied by *immunosenescence*, which is the age-related deterioration of the innate and the adaptive immune systems. Immunosenescence is believed to result in increased infectious and autoimmune disease, as well as to contribute to the development of osteoporosis, cognitive impairment and arteriosclerosis. Furthermore, HIV-1 infection and antiretroviral therapy contribute to the development of the morbidities mentioned above. Summarizing, the success of antiretroviral treatment is reflected

in an increased life expectancy of persons living with HIV-1. However, the effects of years of HIV-1 infection and antiretroviral therapy on the body result in further disease. For this reason, an increasing specialization of antiretroviral treatment for the elderly in the near future is inevitable [235–238].

Upcoming Antiretroviral Drugs

Several antiretroviral compounds are currently being developed. Table 1.12 gives an overview of compounds completing phase II and III clinical trials. In the following, I give some highlights on three new drugs to come. (1) Fostemsavir (BMS-663068) is the prodrug of temsavir, a gp120 inhibitor that has completed phase II clinical trials and is currently being tested in phase III clinical trials on heavily pretreated patients (clinicaltrials.gov trial ID: NCT02362503). Due to its novel mechanism of action, fostemsavir presents a new alternative for patients with multi-drug resistant HIV-1 strains. (2) Doravirine, a novel NNRTI is currently being tested in a phase III clinical trial (clinical trials.gov trial ID: NCT02275780). Doravirine presents an enhanced resistance profile when compared to other NNRTIs, as it remains active in the presence of frequent resistance mutations that render other NNRTIs inactive. 24-week results from a phase II clinical trial comparing doravirine to efavirenz indicated that the drug has similar virologic activity to efavirenz with half of the side effects (`http://www.aidsmap.com/page/2987389`). Last but not least, (3) nanoformulated cabotegravir, an INI recently completed a phase II clinical trial, the *LATTE* trial [231]. As mentioned above, in this trial, cabotegravir is given with two NRTIs in an induction phase, after which the NRTIs are replaced by RPV (maintenance phase). The second version of this clinical trial, LATTE-2, is underway, and will explore a maintenance phase using long-acting nanoformulations of cabotegravir and RPV. Apart from moving away from the imperative of including three active drugs in antiretroviral therapy, the nanoformulations used in LATTE-2 could substantially improve antiretroviral therapy by doing away with adherence-associated therapy issues, as well as enhancing the convenience of therapy.

Antiretroviral Therapy for Reducing the Incidence of HIV-1

Antiretroviral therapy can be used for preventing HIV-1 infection in three manners:

1. Treatment as Prevention: Prevention of HIV-1 transmission through antiretroviral chemotherapy for an HIV-1-infected person [93].
2. Pre-Exposure Prophylaxis (PrEP): Prevention of an HIV-1 infection through intake of antiretroviral drugs before a potential exposure to HIV-1 takes place [92, 239–241].
3. Post-Exposure Prophylaxis (PEP): Prevention of an HIV-1 infection through intake of antiretroviral drugs after a potential exposure to HIV-1 has occurred [242].

Thus, antiretroviral therapy is not only a tool for preserving the health of HIV-1-infected persons, but also for reducing HIV-1 incidence. On July 16th 2012, Truvada ®, a once-daily combination tablet that contains FTC and TDF, has been approved by the FDA for use in PrEP (`http://www.fda.gov`, accessed December 15th, 2015). Studies on the efficacy of other antiretroviral drugs in PrEP have been completed or are ongoing. However, challenges to the implementation of antiretrovirals as a means of containing the HIV-1 pandemic remain.

Several findings support the statement that treatment is effective in preventing transmission of HIV-1: reduction of HIV-1-transmission in serodiscordant couples, the negative correlation between cART programs

Table 1.12: Investigational Antiretroviral Drugs in Current Development. Investigational drug compounds currently in phases II and III of clinical development are listed below. Source: `http://aidsinfo.nih.gov`, accessed December 15th, 2015.

Name	Phase	Originator	Comments
Entry and Fusion Inhibitors			
Cenicriviroc	IIb	Takeda	CCR5 and CCR2 antagonist
Fostemsavir	III	Bristol-Myers Squibb	gp120 inhibitor
Ibalizumab	II	Biogen Idec	
PRO-140	II	Progenics	Intravenous or subcutaneous administration
UB-421	II	United Biomedical	Fusion inhibitor
Histone deacetylase inhibitors			
Vorinostat	II		
Immunostimulants			
AGS-004	II	Argos	Intradermal administration
HIV-1 DNA vaccine	II	Genetic Immunity	Topical administration
Peginterferon alpha-2a	II	Roche	Subcutaneous administration
vCP1521	III	Sanofi Pasteur	Intramuscular administration
Integrase Strand-Transfer Inhibitors			
Cabotegravir	II	ViiV	
Maturation Inhibitors			
BMS-955176	II	Bristol-Myers Squibb	
Non-Nucleoside Reverse-Transcriptase Inhibitors			
Doravirine	III	Merck (US)	
Fosdevirine	II	Idenix	
Nucleoside Reverse-Transcriptase Inhibitors			
Amdoxovir	II	RFS	
Censavudine	II	Kagoshima and Yale Universities	
Racivir	II	Gilead	
Protease Inhibitors			
TMC-310911	II	Tibotec	
Undefined Mechanism of Action			
PEHRG-214	II	Virionyx	Intravenous administration

and incidence of HIV-1, and the negative correlation between HIV-1 VL and infectiousness. It has been even suggested that, under ideal conditions, eradication of HIV-1 through treatment could be possible. However, an effective implementation of treatment as prevention on a large scale is hindered by a number of obstacles. These obstacles relate to the difficulty in fulfilling the *cascade of care*: (1) diagnosis of HIV-1 infection, (2) treatment of HIV-1 infection, and (3) suppression of the VL below the limit of quantification of 50 HIV-1 RNA copies per milliliter or blood serum. It is important that HIV-1 infection be diagnosed as early as possible, as it is the prerequisite for antiretroviral treatment. The risk of HIV-1 transmission is substantially higher in patients with primary HIV-1 infection. Although the risk of transmission is decreased after the chronic phase of the infection begins, several asymptomatic years can pass, offering many chances for transmission of the pathogen. Thus, HIV-1 testing must be performed frequently, especially in individuals at high risk of contracting HIV-1 infection. Diagnosis of HIV-1 infection must occur, at best, when patients present symptoms of ARS (Section 1.2.2), and if not, before symptomatic chronic HIV-1 infection begins. Access to antiretroviral treatment, and further, access to *effective* antiretroviral treatment is constrained by the lack of monetary resources in many countries of the world [93, 243–245]. The *90-90-90* strategy of UNAIDS aims at containing the HIV-1 pandemic through treatment with antiretroviral drugs (Section 1.3.2).

There are several challenges for the large-scale implementation of PrEP. A low awareness of PrEP has been reported among potential users, although many of them express interest in using PrEP after learning about the chemoprophylaxis. Furthermore, people at high-risk of contracting HIV-1 infection may not consider themselves being at risk, which prevents them from using PrEP. Among healthcare professionals, lack of training for prescribing PrEP, as well as the belief that condoms and other behavioral interventions should be prioritized over PrEP, deter them from prescribing chemoprophylaxis. This belief is partly based on concerns that PrEP use may encourage risky behavior. PrEP with Truvada ® has a cost of over 10,000 US dollars per year per person, which limits access to chemoprophylaxis for people without insurance or who cannot afford the insurance copayments . Even if a person has access to PrEP, a certain stigma is associated to it, which might make the person reluctant to use it. Long-term safety data for PrEP are not available, although it is known that there are toxicities associated with TDF and with FTC. Last but not least, adherence is crucial for the efficacy of orally administered PrEP. With respect to adherence, investigational compounds and formulations in the context of PrEP look very promising. Specifically, clinical phase II studies on the safety and efficacy of long-acting nanoformulations of RPV (Table 1.7) and cabotegravir (Table 1.12) for PrEP are underway. These long-acting nanoformulations could allow for monthly or even quarterly dosing of the drugs. Further routes of administration for PrEP that are currently being tested include topical administration, and administration with an intervaginal ring [241, 242].

<div style="text-align: right; font-size: 3em;">2</div>

Methods for Learning from Datasets

This chapter aims at providing a summary of key concepts and methods of the field of statistical learning that are required for understanding this work. After a brief introduction to statistical learning, I present several variants of SVMs, a family of models from the field of statistical learning. In the last sections of this chapter, I present a statistical-learning method called kernel density estimation (KDE) with an application to the task of classification. The information presented in this chapter can be regarded as a dense summary, intended to refresh previously acquired knowledge. Interested readers can consult the references cited in this chapter, should they require a deeper understanding of the concepts herein presented.

2.1 Statistical Learning

The following paragraphs are mainly based on [246]. In order to make this introduction to statistical learning more tangible, I will begin by describing a dataset which we would like to analyze. Suppose our dataset arises from the treatment of HIV-1-infected patients with a certain combination of RTIs (Section 1.5). The dataset consists of the RT amino-acid sequence (Section 1.1) of the HIV-1 variant that each patient harbors before treatment initiation, and a set of VL measurements for each patient (Section 1.2.2) that were performed while the patient was taking the drug combination. We know that the selection of resistance mutations in HIV-1's *RT* gene (Section 1.5.3) can prevent RTI-based cART from being effective in the treatment of HIV-1-infection. The presence of drug resistance to RTIs will manifest itself as VL measurements with high numeric values. Using this dataset, we can aim at extracting different types of knowledge. For example, we could choose to learn (how) to predict whether the drug combination will be successful in terms of reducing the VL below a certain threshold in a patient harboring a specific viral strain. The patient's viral strain is characterized by the amino-acid sequence of the *RT* gene. The field of statistical learning offers models and methods for selecting the parameters of these models. Common to many of the techniques offered by the field is that they use a *training set* in order to select the model parameters. The process by which model parameters are selected is called *training*.

For predicting the success of cART with a statistical-learning model, it would be appropriate to define variables describing the amino-acid sequence of the virus as *input* variables, since we desire to input these variables

into a model with which we can produce a prediction for therapeutic success. *Therapeutic success* would be best captured with a binary variable stating whether after treatment initiation, the VL declined below the threshold that we selected or not. Thus, it would be appropriate to define *therapeutic success* as an *output* variable, since it is the variable that we wish to predict and that should be output by our model. In the context of statistical learning, input variables are also called *features*, *predictors*, or *independent variables*. Output variables, in turn, are also called *responses* or *dependent variables*. We may choose not to use all of the variables that are available to us as input variables, disregarding some of them instead. We would disregard some variables if we thought that rather than helping us in predicting the value of the output variable, these variables could lead us astray, because they are non-informative (in which case they could cause confusion) or because they are highly correlated to other variables (in which case they would be redundant). This is known as the problem of *feature selection*. The field of statistical learning offers methods for tackling the problem of feature selection as well. If we choose to encode the information contained in the amino-acid sequence of the RT gene with p input variables, then our vector of input variables $x = (x_1, \ldots, x_p)$ has dimensionality p. We may choose to model the dependency between x and our output variable y as

$$y = f(x) + \epsilon, \tag{2.1}$$

where f is a function that takes an instance of x as an input, in order to produce an output. ϵ is an error term thought to assume values according to a certain probability distribution with mean equal to zero. We include this error term in (2.1) in order to accommodate for factors that influence the value of y although they are not captured by x (due to randomness or because we do not have measurements for these factors). Such factors are often referred to as *noise*. In (2.1), noise is modeled as being additive, since it is added to the value assumed by f. Note that noise can be modeled in many ways, depending on what is appropriate, e.g. as being multiplicative. In order to accomplish our goal of learning (how) to predict whether the drug combination will be successful, we need to estimate f. When selecting a method for estimating f, we need to decide whether our goal is *prediction* or *inference*. If we solely want to get the most accurate estimate for y, then our goal is prediction. However, if we additionally want to understand the relationship between the input and the output variables, then our goal is inference. The reason why we need to decide whether we want to perform prediction or inference has to do with the characteristics of the models used in statistical learning. When comparing the characteristics of statistical-learning models, we can see that there is a tradeoff between the models' *interpretability* and their accuracy. Interpretable models are those that can help us in understanding the relationship between the input and the output variables. Models that can easily provide quantitative summaries of the relationship between input and output variables cannot capture complex relationships between the input and the output variables. In contrast, models that *are* able to accurately capture intricate relationships between the input and the output variables tend to encode these relationships in an equally intricate way, such that they provide little help in understanding these relationships, but produce accurate predictions. Only if the relationship between input and output variables is sufficiently simple will we be able to select a method that simultaneously features prediction accuracy and interpretability. The degree of complexity of the relationship between input and output variables is tantamount to the complexity of the function that we want to estimate. Examples of questions that can be answered with interpretable models are the following. Which input features have the largest influence on the output features? What is the correlation of the input features with the output features? Is the relationship between the input and the output features linear or non-linear?

Statistical learning can be regarded as a problem of function estimation (e.g. estimation of f in (2.1)). *Func-*

tion estimation subsumes three problem types. (1) *Classification*, (2) *regression*, and (3) *density estimation*. The problem of predicting therapeutic success, as described above, is a classification problem, since we create a binary output variable (*therapeutic success*) which can assume discrete values that correspond to *categories* (in this case, *yes* or *no*). Function-estimation problems in which the output variables are *quantitative* are called regression problems. If we choose to predict the difference between the VL at treatment initiation and the lowest VL measured while the patient was on therapy, we deal with a regression problem, since we try to predict a continuous quantity. In density estimation, we assume that the data in the training set was drawn from a certain joint probability distribution, and we want to use the training set in order to estimate this probability distribution. All methods in statistical learning implement an inductive principle in order to perform prediction or inference using a dataset. In the following, I briefly characterize the two most wide-spread inductive principles.

2.1.1 Frequentist Induction

Frequentist inference has the goal of making inference about the (fixed) parameters of a distribution. Data are viewed as a random sample from a probability distribution. Thus, propositions are generated in terms of repeated sampling of that probability distribution. The basis of the propositions is the frequency or proportion of the data, linked to the frequentist interpretation of probability:

$$P(x) = \lim_{n \to \infty} \frac{n_x}{n}, \tag{2.2}$$

where x denotes an event, n_x the number of samples in which event x occurred, and n the total number of samples. A particular experiment is considered to be one of an infinite number of possible repetitions of that experiment. The events in each experiment occur with a certain relative frequency, and the probability of a given event is considered to be the limit of the relative frequency as the number of experiments goes to infinity. In frequentist inference, distribution parameters are estimated with unbiased estimators and the estimates converge in probability to the *true* parameters. The value of parameters is given as a point estimate with a confidence interval. The uncertainty in the value of the parameters is considered to be due to randomness [247–250].

In the context of the problem of function estimation (see above), frequentist induction principles revolve around the idea of minimizing the risk functional

$$R(\theta) = \int L(z, \theta) \, df(z), \tag{2.3}$$

where $z = (x, y)$ are the input and output variables, $\theta \in \Theta$ is a choice for the parameters of a certain model $m_\theta(x)$, $f(z)$ is an unknown probability measure, and L is a loss function quantifying the discrepancy between the output of the model (using the selection of parameters) and the output variable. In other words, the risk functional quantifies the cumulative discrepancy between the (measured) output variable and the output of the model. Each individual discrepancy is quantified by a loss function. Popular loss functions include:

- Squared loss: $L(z, \theta) = (y - m_\theta(x))^2$, used in regression
- Zero-one loss: $L(z, \theta) = \begin{cases} 0 & \text{if } y = m_\theta(x) \\ 1 & \text{if } y \neq m_\theta(x) \end{cases}$, used in classification
- Log likelihood: $L(z, \theta) = -\log(m_\theta(x))$, used in density estimation.

A frequentist inductive principle is that of empirical risk minimization (ERM). Specifically, this inductive

principle selects model parameters using a set of observations z_i, $i \in \{1, \ldots, n\}$. This set of observations is the training set. Model parameters are selected such that they minimize the empirical risk functional:

$$\underset{\theta \in \Theta}{\arg\min}\, \mathrm{R_{emp}}(\theta) = \underset{\theta \in \Theta}{\arg\min}\, \frac{1}{n} \sum_{i=1}^{n} \mathrm{L}(z_i, \theta). \tag{2.4}$$

Note that while the risk functional $\mathrm{R}(\theta)$ (2.3) quantifies the risk for all possible observations, the empirical risk functional $\mathrm{R_{emp}}(\theta)$

$$\mathrm{R_{emp}}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{L}(z_i, \theta) \tag{2.5}$$

does so for a certain set of observations (the training set). In a rough manner, the principle of ERM is said to be consistent for a certain model and a certain training set if the value of both $\mathrm{R}(\theta)$ and $\mathrm{R_{emp}}(\theta)$ decreases as the size of the training set increases [251], until they reach the lowest possible value. Thus, if ERM is consistent, the predictions of the model should gain accuracy as the size of the training set is incremented.

Selection of a parameter $\theta \in \Theta$ such that the empirical risk functional is minimized by a model is called *training* the model. As mentioned in Section 2.1, models differ in the maximum complexity of the functions that they are able to estimate. If the function that we want to estimate is more complex than the maximum complexity of the model, ERM will not be consistent for the model, i.e. the model will not become more accurate if the size of its training set is incremented. Models for which the principle of ERM is consistent can differ with respect to the size of the training set that they require in order to reach the highest possible accuracy. Typically, ERM-consistent learning methods that minimize the empirical risk functional to a very low number are at risk of reproducing the noise in the limited number of observations that is available for their training. This phenomenon is called *overtraining* or *overfitting*. Thus, for a small number of observations, the empirical risk might be very low, but the expected risk (for observations outside the training set) high. The concept of the *generalization ability* of the learning methods refers to the capacity of the learning methods to minimize the (unknown) expected risk based on a limited number of observations.

2.1.2 Bayesian Induction

The goal of Bayesian inference is the assignment of a probability distribution to the parameters of the distribution of the data. Bayes' theorem builds the core of Bayesian inference:

$$\begin{aligned} P(\theta \mid X) &= \frac{P(\theta \cap X)}{P(X)} \\ &= \frac{P(X \mid \theta)P(\theta)}{P(X)}, \end{aligned} \tag{2.6}$$

where θ is the parameter of the distribution of the data and X is the data. The following terminology is used for the components of (2.6):

- $P(\theta)$: prior distribution of the parameters
- $P(X \mid \theta)$: likelihood or sampling distribution (given the model parameters θ)
- $P(X)$: marginal likelihood (independent of the model parameters θ)
- $P(\theta \mid X)$: posterior distribution

Bayesian inference requires the specification of the prior distribution. The prior distribution represents the state of knowledge before the experiment was performed with which the data was generated. In the Bayesian inference process, the prior distribution $P(\theta)$ is updated by multiplying it with the likelihood $P(X \mid \theta)$. Normalization of $P(X \mid \theta)P(\theta)$ with the marginal likelihood $P(X)$ yields the posterior distribution. The marginal likelihood is the distribution of the observed data marginalized over the parameters,

$$P(X) = \int_\theta P(X \mid \theta) \, d\theta. \tag{2.7}$$

Bayesian inference allows for the combination of data with prior beliefs. The prior distribution encodes and quantifies which hypotheses are believed to be likely and which hypotheses are believed to be unlikely, according to one's prior beliefs. These prior beliefs exist before examining any data. The likelihood of the data under a certain hypothesis is quantified by the sampling distribution. Thus, Bayesian inference states that a hypothesis (regarding the value of the parameters) should be rejected if it is not compatible with the data (as quantified by the sampling distribution), but also if it contradicts prior beliefs (as quantified by the prior distribution), even if it matches the data. The prior distribution requires specification and an arbitrary distribution can be specified as a prior. However, Bayesian inference allows for sequential application of Bayes' theorem as further experiments are carried out: the posterior of an older experiment becomes the prior of a newer one. Through sequential application of Bayes' theorem, the effects of the initially selected prior distribution are diminished. Thus, Bayesian inference assumes that initial prior beliefs exist (these are represented with the prior distribution), and provides a method for updating these prior beliefs according to the data (by multiplication of the prior distribution with the sampling distribution). The posterior is the result of updating prior beliefs with the data (and normalizing). In contrast to frequentist inference, where variability is regarded to be caused solely by randomness, Bayesian inference accepts the notion of variability as the result of lack of knowledge [250, 252–255].

2.2 Support Vector Machines

SVMs are a family of methods for solving problems in machine learning. In this section, I review several variants of SVMs. Section 2.2.1 reviews the idea of using a hyperplane for separating two classes of points. Section 2.2.3 shows how functions computing the inner product in *enlarged feature spaces* are an efficient means for increasing classification performance. SVMs for classification are reviewed in Section 2.2.2, followed by SVMs for regression (Section 2.2.4). Finally, SVMs for the regression of right-censored data are presented in Section 2.2.5. This review on SVMs is mostly based on [256, 257].

2.2.1 The Concept of a Separating Hyperplane

Let us consider the concept of a hyperplane which is used for separating points $x_i \in \mathbb{R}^p$, $i \in \{1, \ldots, n\}$, of two classes, labeled with -1 and $+1$. Let the class of each of these points be denoted by $y_i \in \{-1, +1\}$. Furthermore, let these two classes of points be linearly separable in $\mathbb{R}^p$. This means that a separating hyperplane exists, $h : \{x \mid \beta \cdot x + \beta_0 = 0\}$, $\beta \in \mathbb{R}^p$, $\beta_0 \in \mathbb{R}$, such that all points of one class lie on one side of the hyperplane, and all points of the other class lie on the other side of the hyperplane. Let β be normal to the hyperplane and let $\|\beta\|$ be its Euclidean norm. Thus, $\frac{\beta_0}{\|\beta\|}$ is the signed Euclidean distance of h to the origin,

and the signed distance of h to some point x_i is given by

$$d(h, x) = \frac{\beta^{\mathrm{T}} \cdot x + \beta_0}{\|\beta\|}. \tag{2.8}$$

A classification rule for each x_i is then given by

$$y_i = \mathrm{sign}(d(h, x_i)), \tag{2.9}$$

which is equivalent to

$$y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq 0. \tag{2.10}$$

Note that there are an infinite number of different hyperplanes h for which (2.10) holds. However, each hyperplane will differ with respect to its *classification margin*, $2C$, with

$$C = \min_{x_i} d(h, x_i), \ i \in \{1, \ldots, n\}. \tag{2.11}$$

Among all separating hyperplanes h, some of them separate the classes -1 and +1 with a maximum margin, their parameters given by

$$\begin{aligned}
&\underset{\beta, \beta_0}{\arg\max} \ C \\
&\text{subject to} \quad \frac{1}{\|\beta\|} y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq C, \ \forall i \in \{1, \ldots, n\}.
\end{aligned} \tag{2.12}$$

Normalization of β leads to a unique solution of (2.12) and therefore, we (arbitrarily) normalize with $\|\beta\| = \frac{1}{C}$, yielding

$$\begin{aligned}
&\underset{\beta, \beta_0, \|\beta\| = \frac{1}{C}}{\arg\max} \ C \\
&\text{subject to} \quad C y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq C, \ \forall i \in \{1, \ldots, n\}.
\end{aligned} \tag{2.13}$$

In the inequality in (2.13), C cancels out, and we can see that maximizing C is equivalent to minimizing $\|\beta\|$. Thus, we can restate (2.13) as

$$\begin{aligned}
&\underset{\beta, \beta_0}{\arg\min} \ \frac{1}{2}\|\beta\|^2 \\
&\text{subject to} \quad y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq 1, \ \forall i \in \{1, \ldots, n\}.
\end{aligned} \tag{2.14}$$

The constraints in (2.14) require that there is an empty space of thickness $\frac{1}{\|\beta\|}$ to each side of the linear classification boundary, the classification margin. This problem is quadratic with linear inequality constraints, and can be solved with Lagrange multipliers, as follows. The Lagrange primal function is

$$L_{\mathrm{p}} = \frac{1}{2}\|\beta\|^2 - \sum_{i=1}^{n} \alpha_i[y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) - 1]. \tag{2.15}$$

76

In order to minimize L_p, and in partial fulfillment of the Karush-Kuhn-Tucker conditions [258], we set its derivatives to zero:

$$\frac{\partial L_p}{\partial \beta} = \beta - \sum_{i=1}^{n} \alpha_i y_i x_i \overset{!}{=} 0 \Leftrightarrow$$

$$\beta = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{2.16}$$

$$\frac{\partial L_p}{\partial \beta_0} = \sum_{i=1}^{n} \alpha_i y_i \overset{!}{=} 0 \Leftrightarrow$$

$$0 = \sum_{i=1}^{n} \alpha_i y_i \tag{2.17}$$

Substituting (2.16) and (2.17) into L_p (2.15), yields the Wolfe dual:

$$L_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_i \alpha_k y_i y_k x_i^T \cdot x_k \tag{2.18}$$

$$\text{subject to } \alpha_i \geq 0, \forall i \in \{1 \ldots, n\}.$$

Maximization of L_D leads to the solution of (2.14). In order to fully fulfill the Karush-Kuhn-Tucker conditions [258], the following condition must be satisfied as well:

$$\alpha_i[y_i(\beta^T \cdot x_i + \beta_0) - 1] = 0, \forall i \in \{i, \ldots, n\}. \tag{2.19}$$

The vector β, satisfying the constraints above, will be a linear combination of the points x_i (2.16). Furthermore, only a subset of points x_i will be used in this linear combination (2.19). Specifically, if α_i is not equal to zero, then $y_i(\beta^T \cdot x_i + \beta_0) = 1$, and therefore x_i is on the classification margin. Such a point x_i is called a *support vector*. Since the points are separable, α_i equal to zero implies that x_i is on the correct side of the hyperplane, further away from the hyperplane than the margin.

Summarizing, a hyperplane can be used for determining the class labels of two classes of points that are linearly separable in $\mathbb{R}^p$. Furthermore, among all possible separating hyperplanes, some of them separate the classes with a maximum margin. In addition to being unique for a given set of points x_i with respective class labels y_i, $i \in \{1, \ldots, n\}$, it can be shown that the normalized maximum-margin separating hyperplane has the best prospects or generalization among all separating hyperplanes[251, 259–261] . Figure 2.1 was created for facilitating intuitive understanding of this fact.

2.2.2 SUPPORT VECTOR MACHINES FOR CLASSIFICATION

Let us expand the concept of a maximum-margin separating hyperplane for the case in which the points $x \in \mathbb{R}^p$ are not separable. For attaining this goal, we introduce slack variables $\xi_i \geq 0$ into (2.14):

$$\underset{\beta, \beta_0}{\arg\min} \ \frac{1}{2}\|\beta\|^2 + \gamma \sum_{i=1}^{n} \xi_i \tag{2.20}$$

$$\text{subject to } \quad y_i(\beta^T \cdot x_i + \beta_0) \geq 1 - \xi_i, \ \xi_i \geq 0, \ \forall i \in \{1, \ldots, n\},$$

(a) Classification Boundary

(b) X_{10}

(c) X_{100}

(d) X_{1000}
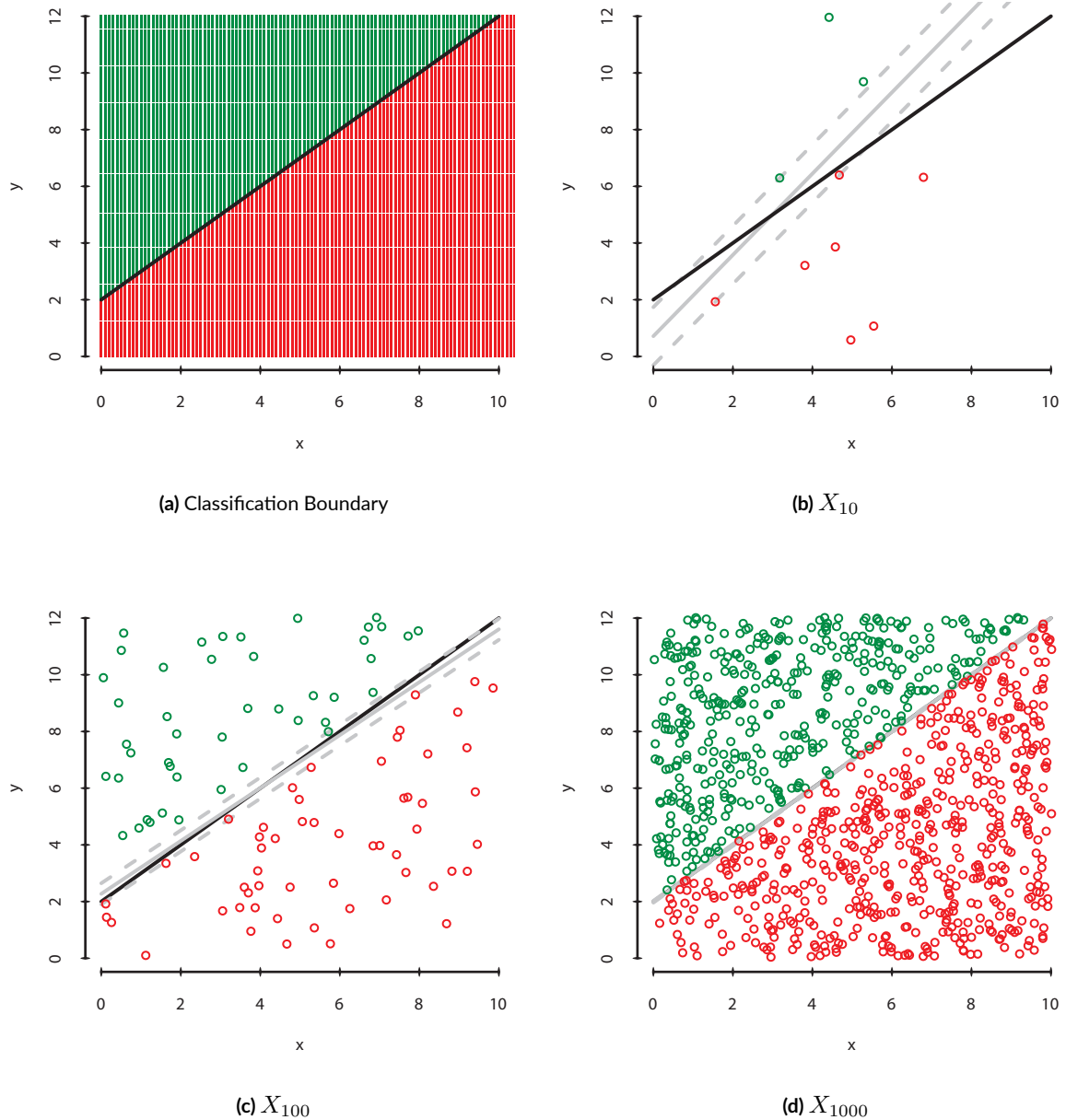
Figure 2.1: Maximum-Margin Separating Hyperplane. The hyperplane $h = \{x \mid (1, -1) \cdot (x - (1, 3))^{\mathrm{T}} = 0\}$ was defined to be the classification boundary of linearly separable points x_i, $i \in \{1, \ldots, n\}$, of two classes $y_i \in \{-1, +1\}$ (a). These two classes of points, -1 and $+1$, are represented by green and red colors, respectively. Three sets of points X_n with cardinality n, $n \in \{10, 100, 1000\}$, were sampled from $[0, 10] x [0, 12]$ according to the uniform distribution. Class labels were generated for each point with the function $\mathrm{sign}((1, -1) \cdot (x - (1, 3))^{\mathrm{T}})$. A maximum-margin separating hyperplane was fitted for separating the two classes in each X_n. Figures (b), (c), and (d) show the true classification boundary along with the estimated maximum-margin separating hyperplanes for X_{10}, X_{100}, and X_{1000}, respectively. As n increases, the maximum-margin separating hyperplane converges to the true classification boundary. Other separating hyperplanes could also be fit to the data. However, their convergence to the true classification boundary would require more sample points than the maximum-margin separating hyperplane.

■ Class -1 ■ Class $+1$ — True Classification Boundary   — Estimated Classification Boundary
— Margin of Estimated Classification Boundary

with $\gamma \geq 0$. In (2.20), a sum with two terms is minimized. The first term, $\frac{1}{2}\|\beta\|^2$, ensures that β's norm is as small as possible while still satisfying the classification constraint $y_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq 1 - \xi_i$, $\forall i \in \{1, \ldots, n\}$. This classification constraint is *relaxed* by including the slack variables ξ_i which allow a certain point x_i to be on the *wrong* side of the corresponding classification margin by the amount ξ_i. The second term of the minimization in (2.20), $\gamma \sum_{i=1}^{n} \xi_i$, enforces that the slack variables ξ_i assume values that are as small as possible. γ is a variable that controls the tradeoff between empirical risk minimization and the sparseness of β. Specifically, setting γ to a small value will reduce the value of $\gamma \sum_{i=1}^{n} \xi_i$, therefore reducing its influence in the selection of the minimizing β. At the same time, the influence of $\frac{1}{2}\|\beta\|^2$ in selecting the minimizing β will be increased. Hence, small values of γ favor values of the minimizing β that have a small norm and are therefore sparser. In comparison to larger values of γ, smaller values of γ may allow a larger number of points to be located on the wrong side of the corresponding margin. In contrast, larger values of γ will result in less points being located on the wrong side of the margin, but will also translate into more complex classifiers. γ is also called the *cost* parameter of the SVM. The separable case (2.14) corresponds to $\gamma = \infty$. In the same way as in the maximum-margin classifier presented in Section 2.2.1, the problem can be solved using quadratic programming. The maximum-margin classifier for non-separable classes is called an *SVM for classification* or a *Support Vector classifier (SVC)*.

2.2.3 The Kernel Trick

Let us review the concept of a reproducing kernel Hilbert space (RKHS). Let $\mathcal{X}$ be an non-empty set. An RKHS is a Hilbert space $\mathcal{H}$ of functions $f : \mathcal{X} \to \mathbb{C}$, in which the evaluation functionals, $\mathrm{F}_x : (\mathcal{X} \to \mathbb{C}) \to \mathbb{C}; f \mapsto f(x)$, are continuous for all $f \in \mathcal{H}$. It can be proven that there is a one-to-one correspondence between each RKHS $\mathcal{H}$ and a kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ with the reproducing property:

$$f(x) = \langle f(y), k(\cdot, x)\rangle_{\mathcal{H}}, \forall f \in \mathcal{H}, k(\cdot, x) \in \mathcal{H}. \tag{2.21}$$

Furthermore, it can be shown that all positive definite functions satisfy (2.21) in some RKHS, and therefore, defining a positive definite function is equivalent to defining a RKHS, and vice-versa. From (2.21), it follows that the kernel function is identical with the inner product of $\mathcal{H}$ [262]:

$$\langle k(\cdot, x), k(\cdot, y)\rangle_{\mathcal{H}} = k(x, y). \tag{2.22}$$

Let $\phi : \mathbb{R}^p \to \mathbb{R}^q$, with $q > p$, be a function that maps a point $x \in \mathbb{R}^p$ to another point $x' \in \mathbb{R}^q$ using q orthogonal functions $\phi_i, i \in \{1, \ldots, n\}$. Furthermore, let $\mathcal{H}_\phi = \{f \mid \exists a \in \mathbb{R}^q \text{ such that } f(x) = a^{\mathrm{T}}\phi(x)\}$ be a RKHS space. The kernel function for $\mathcal{H}_\phi$ is $k_\phi(x, y) = \phi(x)^{\mathrm{T}}\phi(y)$. Such a map ϕ can be useful in cases in which two classes of points $x_i \in \mathbb{R}^p, i \in \{1, \ldots, n\}$ are not separable in $\mathbb{R}^p$, but separable in the higher-dimensional $\mathbb{R}^q$, after application of ϕ. A naïve approach for attaining separability would be the substitution of x_i by $\phi(x_i)$ in (2.20). However, there is another approach that is computationally less expensive: the substitution of the scalar product in (2.20) with the kernel function k_ϕ. This is known as the kernel trick. Application of a kernel function does not require explicit knowledge of the RKHS to which it belongs. Thus, every positive definite function is adequate for this goal. In the following, I list three popular kernel functions:

- Polynomial kernel of degree p: $k(x, y) = (x \cdot y + 1)^p$
- Radial basis kernel: $k(x, y) = \exp(-\|x - y\|^2 / c)$

- Neural network kernel: $k(x, y) = \tanh(\kappa_1 x \cdot y + \kappa_2)$.

Scalar products can be interpreted to be similarity functions. Let us recall the geometric definition of the scalar product of two Euclidean vectors $a, b \in \mathbb{R}^p$:

$$a \cdot b = \|a\| \|b\| \cos(\theta_{a,b}), \tag{2.23}$$

where $\theta_{a,b}$ is the angle between a and b. Let us consider what happens with the value of $a \cdot b$ if a and b retain their lengths, but $\theta_{a,b}$ changes. If a and b are pointing in the same direction, $\theta_{a,b}$ is zero and the value of the cosine function in (2.23) will be equal to one. Hence, $a \cdot b$ will assume its highest possible value (the exact value depends on the lengths of a and b). If the value of $\theta_{a,b}$ is increased, the value of the cosine function in (2.23) will decrease, reaching its minimum when $\theta_{a,b}$ is equal to 180 degrees. At this value of $\theta_{a,b}$, the cosine function is equal to minus one and $a \cdot b$ assumes the smallest possible value. Hence, the scalar product quantifies the similarity of two vectors in terms of the similarity of the directions at which they are pointing. If we examine (2.8) and (2.16), we can see that the (not normalized) signed distance of a point to the hyperplane of an SVC is given by

$$\hat{f}(x) = \hat{\beta}^{\mathrm{T}} \cdot x + \hat{\beta}_0 = \sum_{i=1}^{n} \alpha_i y_i x_i \cdot x + \hat{\beta}_0, \tag{2.24}$$

where $\hat{\beta}$ and $\hat{\beta}_0$ parametrize a maximum-margin classifier obtained with the training set (x_i, y_i), $i \in \{1, .., n\}$, as described above. $\hat{\beta}$ is a linear combination of the vectors x_i with the scalars $\alpha_i y_i$. The value of $\hat{f}(x)$ can be interpreted to be a weighted sum of similarity values, the similarity values being $x_i \cdot x$ and the weights being $\alpha_i y_i$.

As I explain in this section, kernel functions compute the scalar product of two vectors in spaces of a larger dimension than the input space. In the following, I propose two approaches for gaining intuitive understanding of the kernel functions I list above. The first approach is based on consideration of the computations performed when applying the kernel to two specific vectors. Let us consider two two-dimensional vectors $x = (x_1, x_2)^{\mathrm{T}}$ and $y = (y_1, y_2)^{\mathrm{T}}$, $x_1, x_2, y_1, y_2 \in \mathbb{R}$, and the polynomial kernel of degree two, $k(x, y) = (x \cdot y + 1)^2$. Then,

$$
\begin{aligned}
k(x, y) &= (x \cdot y + 1)^2 \\
&= (1 + x_1 y_1 + x_2 y_2)^2 \\
&= 1 + 2x_1 y_1 + 2x_2 y_2 + (x_1 y_1)^2 + (x_2 y_2)^2 + 2x_1 y_1 x_2 y_2.
\end{aligned}
\tag{2.25}
$$

In comparison to the scalar product in Euclidean space, $x \cdot y = x_1 y_1 + x_2 y_2$, (2.25) includes terms that consider the square of the components of the vectors, $x_1^2 y_1^2$ and $x_2^2 y_2^2$, as well as a term that *mixes* the coordinates multipicatively, $2x_1 y_1 x_2 y_2$. Thus we can see that in this example, the polynomial kernel of degree two virtually expands the dimensionality of the feature space by incorporating the squared value of the features as well as the multiplication of the two coordinates (see also *Training of Final Models* in Section 4.2.1). The second approach that I propose for gaining intuitive understanding of a kernel function is based on the consideration of the values produced by the kernel, which can be interpreted to be similarity values. The radial basis kernel $k(x, y) = \exp(-\|x - y\|^2 / c)$ computes the squared distance between two vectors and applies the exponential function to this distances after division by $-c$. Figure 2.2 shows the values that the kernel assumes as the squared distance between two points is increased from zero to three for three different values of c. As can be seen, the
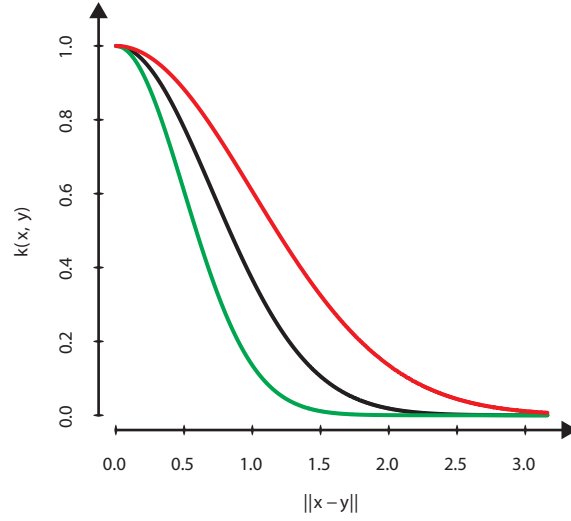
Figure 2.2: Radial Basis Kernel. The radial basis kernel $k(x,y) = \exp(-\|x-y\|^2/c)$ applies the exponential function to the squared distance between x and y divided by $-c$. Above, the value of the kernel is plotted as a function of the distance between x and y for different values of c. If the distance between x and y is equal to zero, the kernel will assume its maximal value, 1. As the distance between x and y increases, the value that the kernel assumes decreases exponentially. The minimum value that the kernel can assume is zero (asymptotically). The parameter c can be used in order to control the slope of the decrease in the values of the kernel as the distance between x and y increases.

🟩 $c = 0.5$ ⬛ $c = 1$ 🟥 $c = 2$

radial basis kernel assumes its highest value, one, if the distance between x and y is zero. As the distance between x and y increases, the values produced by the kernel decrease exponentially. Thus, when an SVC is trained with the radial basis kernel, it will classify a given point on the basis of its proximity to training points of one class or the other class and of the Lagrange multipliers α_i in (2.24). Note that $\alpha_i \neq 0$ if and only if x_i is a support vector (Section 2.2.1).

2.2.4   SUPPORT VECTOR MACHINES FOR REGRESSION

The concept of an SVM can be expanded for performing regression. In this case, points $x_i \in \mathbb{R}^p$ are associated with values $y_i \in \mathbb{R}$, $i \in \{1, \ldots, n\}$, and we require $\beta \in \mathbb{R}^p$ and $\beta_0 \in \mathbb{R}$ that minimize a loss function $\mathrm{L}(y_i, \beta^{\mathrm{T}} x_i + \beta_0)$ for all $i \in \{1, \ldots, n\}$. This can be attained by solving

$$\underset{\beta, \beta_0}{\arg\min} \ \frac{1}{2}\|\beta\|^2 + \gamma \sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$\text{subject to} \begin{cases} \beta^{\mathrm{T}} \cdot x_i + \beta_0 \geq y_i - \xi_i - \epsilon \\ -(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq -y_i - \xi_i^* - \epsilon \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases}$$

$$\forall i \in \{1, \ldots, n\},$$

(2.26)

81
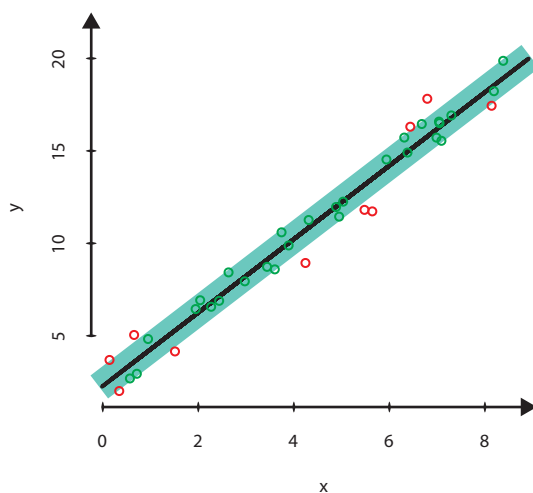
Figure 2.3: Support Vector Machines For Regression. Fifty real numbers $x_i \in [0, 10]$, $i \in \{1, \dots, 50\}$ were drawn from the uniform distribution. These were used in order to obtain $y_i = 2x_i + 2 + \mathcal{E}$, $\mathcal{E} \sim \mathcal{N}(0, 1)$. A Support Vector machine for regression was trained with (x_i, y_i), $i \in \{1, \dots, 50\}$, selecting $\epsilon = 1$. Above, (x_i, y_i), $i \in \{1, \dots, 50\}$ are plotted as circles, and the insensitivity margin is plotted in light blue. Circles inside the insensitivity margin are plotted in green while those outside or on the insensitivity margin are plotted in red. These red circles are the support vectors.

○ Training points inside margin   ○ Training points outside margin (support vectors)
— Estimated linear function   ■ Error margin

for $\gamma \geq 0$ and $\epsilon > 0$. Let us compare (2.20) with (2.26). Both optimization problems strive for minimization of the absolute values of the coefficients in β, as well as of the values of the slack variables ξ_i. The slack variables ξ_i^* are additionally included in (2.26) and are subject to minimization as well. ξ_i (and ξ_i^*) also appear in the minimization constraints of (2.20) and (2.26). On the one hand, minimization constraints that do not include β and β_0 require that the slack variables be non-negative. On the other hand, the slack variables appear in the constraints related to β and β_0, counteracting these constraints whenever the slack variables are greater than zero. This is done in order to make the optimization problem feasible for all possible training sets. While the minimization constraints in (2.20) require β and β_0 to parametrize a maximum-margin separating hyperplane, the minimization constraints in (2.26) strive for obtaining the coefficients of a linear function, β and β_0, such that $\beta^{\mathrm{T}} \cdot x_i + \beta_0 - \epsilon \leq y_i \leq \beta^{\mathrm{T}} \cdot x_i + \beta_0 + \epsilon$. Thus, the value of $\beta^{\mathrm{T}} \cdot x_i + \beta_0$ is allowed to deviate from the value of y_i by ϵ, thereby defining a margin of error tolerance or insensitivity (Figure 2.3). The insensitivity margin is sometimes called ϵ-*tube* [263], however, I do not like this term, as it conveys the idea of roundness. However, the insensitivity margin is not round. In SVMs for regression, mathematical optimization strives for finding β and β_0 such that the values y_i associated to the vectors x_i lie *inside* the insensitivity margin, for all $i \in \{1, \dots, n\}$. Examination of the Karush-Kuhn-Tucker conditions [258] allows us to spot the support vectors: those x_i whose y_i lie on or *outside* the regression margin. As in Section 2.2.2, γ is a parameter controlling the tradeoff between empirical risk minimization and the sparseness of β. For more information on Support Vector regression, the interested reader can consult [263].

2.2.5 Support Vector Machines for the Regression of Right-Censored Data

When analyzing datasets originating from clinical studies, it is not uncommon to find data points that are *censored*. A typical scenario in which this can be observed is a clinical study in which a new drug is tested, e.g. in order to find out whether the drug can prevent or delay an undesired event. Typically, four different outcomes are possible in such a setting: (1) the patient experiences the undesired event during the study, (2) the patient is lost to followup (and possibly experiences the undesired event thereafter), (3) the patient experiences the undesired event after the study has finished, or (4) the patient does not experience the undesired event during his or her lifetime. Since, in this example, the clinical study is the only setting in which information from the patients can be gained, the undesired event can only be observed while the patient is participating in the clinical study. Typically, we know how much time passes between the initiation of the study and the occurrence of the undesired event in a certain patient. In the same way, we know when was the last time that a patient presented if he or she was subsequently lost to followup. Let us count time in days after initiation of the study. If t_i denotes the time at which the undesired event occurred in patient i and the undesired event occurs in patient i at time d_1, then $t_i = d_1$. However, if the patient is lost to followup at time d_2, before the undesired event occurs, then all we know is $t_i \geq d_2$. In the same way, if the study ends at time d_3 and the undesired event is not observed before this time, then $t_i \geq d_3$. Measurements whose exact value is not known, but for which a lower bound to their value is known, are called *right-censored*. If only an upper bound to their value is known instead, then we call them *left-censored*. If we do not know the exact value of a measurement, but we know both an upper and a lower bound to its value, then we call the measurement *interval-censored* [264].

In Section 3.3.1, I define a measure for the effectiveness of an antiretroviral therapy, the *number(s) of aviremic semesters (NAS)*. In the following, I briefly explain how the NAS works, as it can be right-censored and is thus related to this section. Typically, before an antiretroviral treatment is initiated, the patient is *viremic*, which means that at least a certain quantity of viral particles (or viral RNA copies) can be detected in the blood of the patient. This quantity is called the VL. If the therapy is (initially) effective, we should observe a reduction in the VL to such a low quantity that we cannot quantify it or detect it any longer. We call this quantity the lower limit of quantification or the (lower) limit of detection, respectively. If this is the case, viremia has been suppressed and the patient has transitioned from being viremic to being aviremic. If, in contrast, the therapy is not effective, the VL will remain above the limit of detection. Among the therapies that are *initially* effective, some therapies remain effective for longer than other therapies. The NAS aims at quantifying the effectiveness of a therapy by counting the number of semesters during which the therapy suppressed viremia. However, there are three situations in which the NAS can produce a right-censored value (Figure 2.4). (1) If we decide to compute the NAS for a therapy while the therapy is still ongoing. This might cause right censoring since the NAS for that therapy will increase if the therapy continues to be successful and if we consider later VL measurements. (2) If we do not know whether the patient is viremic or aviremic during some therapy semester. This can happen if, for example, the patient does not present to the doctor when he or she should and no VL measurement is performed for some therapy semester. However, the fact that VL measurements are missing does not entail therapy failure. We just do not know whether the therapy was being effective during this period of time or not. Thus, if we compute the NAS even though VL measurements for at least one semester are missing, the resulting value will be right censored. (3) If the treating physician decides to interrupt the treatment while the patient is aviremic. The treating physician might take such a decision if, for example, the treatment causes side effects in the patient. If we compute the NAS in such a case, we could choose to think of this value as being right-
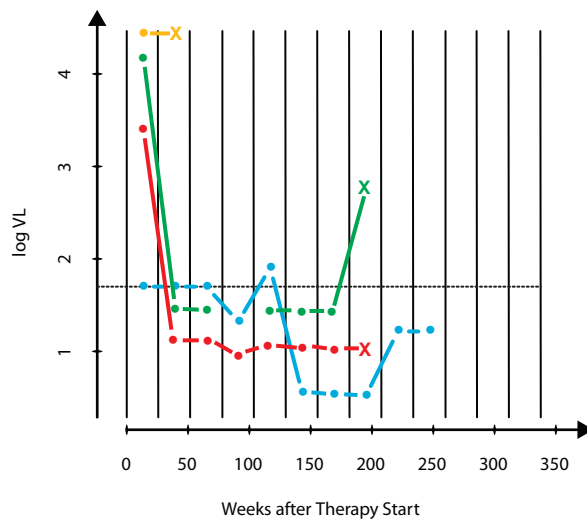
Figure 2.4: Censoring of the Number of Aviremic Semesters. A schematic representation of the viral-load trajectory of four patients receiving antiretroviral therapy can be found above. The viral-load measurements for each patient are represented with a distinct color. Contiguous viral-load measurements are joined with a line. The horizontal, black, dotted line represents the lower limit of quantification of the viral-load assay. Each therapy semester is delimited with a vertical black line. The therapy of Patient 1 was not effective, as it could not suppress the viral load (below the limit of quantification). Thus, it attained zero *uncensored* aviremic semesters. After initiation of the therapy of Patient 2, 5 semestral viral-load measurements below the limit of quantification were measured. In the eighth therapy semester, a strong rebound in the viral load led to treatment interruption. Note that a viral-load measurement for the fourth semester after treatment initiation is missing. For this reason, the number of aviremic semesters for this therapy is *censored*. If the missing viral-load measurement were available, and if the patient's viral load had been below the limit of quantification during that semester, then the number of aviremic semesters would had been six, and it would had been uncensored. The therapy of Patient 3 is still ongoing. For this reason, the last viral-load measurement for this patient is not marked with an *X*. This therapy could attain nine censored aviremic semesters, up to the last measurement. If the therapy continues to be successful, the number of aviremic semesters that it attains will increase. The therapy for Patient 4 could attain 7 aviremic semesters. In the eighth therapy semester, the treating physician decided to interrupt the treatment due to severe side effects. In this case, we could choose to regard this measurement as censored, since continuation of therapy could have led to further aviremic semesters. However, we could also choose to regard this measurement as uncensored, since severe side effects are as undesired as the rebound of the viral load.

●Viral-load measurement   **X** Viral-load measurement with subsequent therapy interruption

censored, for the following reason. If the treating physician had not decided to interrupt the treatment, then it could have resulted in further aviremic semesters. However, we could also choose to think of this value as being uncensored, if we assume that the reason for treatment interruption was as undesirable as a rebound in the VL.

The fact that the NAS can be censored may pose a problem if we wish to train statistical models in order to predict them using other, uncensored input variables. Some statistical models may be able to cope with data in which some data points have been censored, if the difference between the censored values and the *true* values is not too large. However, we typically do not know whether this difference is large or not. For this reason, it is necessary to adapt the models such that they can be trained with censored data. In the following, I introduce a modification of the SVM for regression (2.26) that is able to handle right-censored data (i.e. it is able to handle the NAS). This modification was extracted from [265]. Subsequently, I use an artificial dataset in order to explore the differences in performance between the regular SVMs for regression and SVMs for the regression of right-censored data. Let $x_i \in \mathbb{R}^p$ be vectors with associated, possibly right-censored responses $y_i \in \mathbb{R}$, with associated uncensored responses $r_i \in \mathbb{R}$, and with associated censoring labels $\delta_i \in \{0, 1\}$, for $i \in \{1, \ldots, n\}$. If a response y_i is censored, then $\delta_i = 0$ and $y_i \leq r_i$. If a response y_i is uncensored, then $r_i = y_i$ and $\delta_i = 1$. We do not allow cases where $r_i \leq y_i$. Typically, r_i is latent [*] if $d_i = 0$. We wish to find $\beta \in \mathbb{R}^p$ and $\beta_0 \in \mathbb{R}$ that minimize a loss function $L(r_i, \beta^{\mathrm{T}} \cdot x_i + \beta_0)$, $\forall i \in \{1, \ldots, n\}$, in a setting in which we know y_i, x_i, and δ_i, but we do not know r_i. We produce the estimate $\hat{r}_i = \beta^{\mathrm{T}} \cdot x_i + \beta_0$ by solving the following minimization problem:

$$
\underset{\beta, \beta_0}{\arg\min} \ \frac{1}{2} \|\beta\|^2 + \gamma \sum_{i=1}^{n} (\xi_i + \xi_i^*)
$$

$$
\text{subject to} \begin{cases} \beta^{\mathrm{T}} \cdot x_i + \beta_0 \geq y_i - \xi_i - \epsilon \\ -\delta_i (\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq -\delta_i y_i - \xi_i^* - \delta_i \epsilon \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \tag{2.27}
$$

$$
\forall i \in \{1, \ldots, n\},
$$

where $\gamma \geq 0$ is a regularization parameter controlling the tradeoff between empirical risk minimization and the sparseness of β, and $\epsilon \geq 0$ is an error-insensitivity parameter (as in Section 2.2.4). In (2.27), we use the censoring labels δ_i in order to annul the constraint $\delta_i(\beta^{\mathrm{T}} \cdot x_i + \beta_0) \geq -\delta_i y_i - \xi_i^* - \delta_i \epsilon$ for right-censored data points, since $\delta_i = 0$ for these points. In this way, $\beta^{\mathrm{T}} \cdot x_i + \beta_0$ can be larger than y_i without the need for $\xi_i^* > 0$. Note that in the minimization in (2.27), the value of ξ_i^* (and of ξ_i) trades off with the value of $\|\beta\|$.

In order to demonstrate the differences between SVMs for regression and SVMs for the regression of right-censored data, I present an analysis on synthetic data, in the following. The data consists of the independent and dependent variables, which are initially uncensored. The set X of independent variables, x_i, was created by randomly sampling 1,000 points from $[0, 100]^4$, according to the uniform distribution. Thus, $i \in \{1, \ldots, 1000\}$. From X, the dependent, uncensored variables r_i were created with the following function,

$$
r_i = 4x_{i,1} - 3x_{i,2} + 2x_{i,3} - x_{i,4}, \tag{2.28}
$$

---

[*]A latent variable is a variable that is not directly observed, but inferred from other variables by means of a mathematical model.

(a) $\epsilon_i \sim \mathcal{N}(0,1)$ **(b)** $\epsilon_i \sim \mathcal{N}(0,16)$
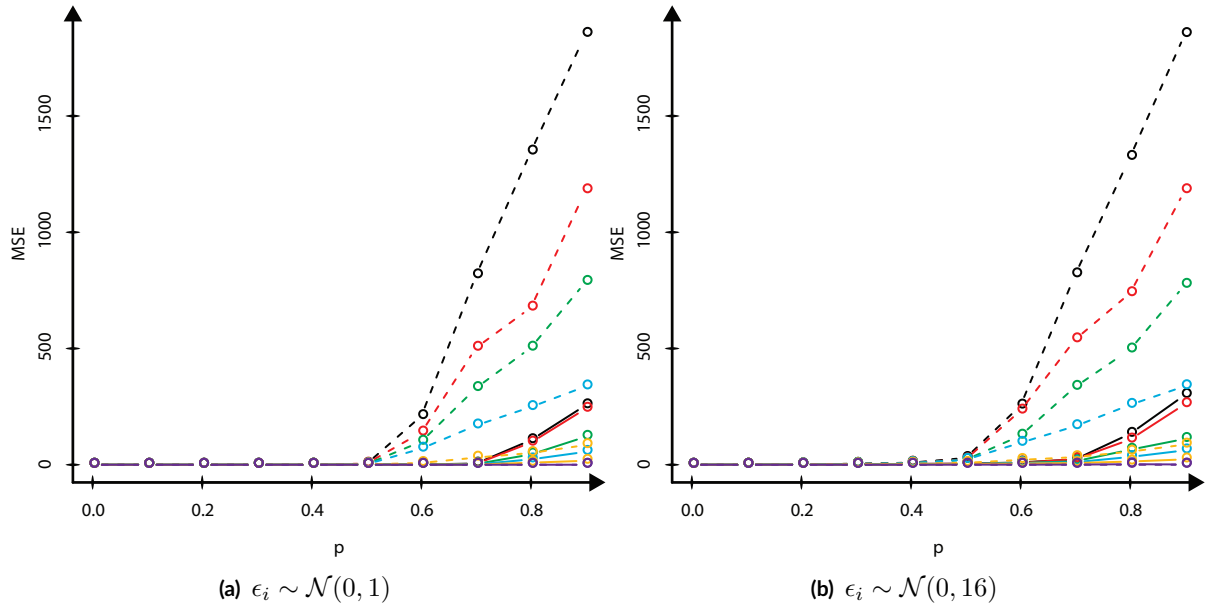
Figure 2.5: Regression of Artificial, Right-Censored Data with SVMs. Artificial, independent data points x_i were created by sampling 1,000 data points from $[0,100]^4$, according to the uniform distribution. These were used for creating the dependent data points $y_{i,\sigma,p,m} = (4x_{i,1} - 3x_{i,2} + 2x_{i,3} - x_{i,4})c_i + \epsilon_i$, where $x_{i,j}$ denotes the j-th component of x_i, $c_i \sim \mathcal{U}(m,1)$ with probability p or $c_i = 1$ with probability $1-p$, and $\epsilon_i \sim \mathcal{N}(0,\sigma^2)$. For each tested value of σ, p, and m, 900 data points were used for training an SVM for regression (2.26) and an SVM for regression of right-censored data (2.27). The remaining 100 data points were used for testing. Above, the censoring probability (p) is plotted against the resulting test mean squared error (MSE) of the trained SVMs for regression and for the regression of right-censored data, for different values of σ, the standard deviation of the additive Gaussian noise, and m, the minimum censoring factor. Note that MSE were calculated with the predicted $\hat{y}_{i,\sigma,p,m}$ and their uncensored, denoised counterparts. MSE for $\sigma = 1$ is shown in (a), while MSE for $\sigma = 4$ is shown in (b).

■ $m = 0.5$ ■ $m = 0.6$ ■ $m = 0.7$ ■ $m = 0.8$ ■ $m = 0.9$ ■ $m = 1$
- - SVM for Regression   — SVM for the Regression of Right-Censored Data

where $x_{i,j}$ denotes the j-th component of x_i. Using r_i, $i \in \{1,\ldots,1000\}$, several sets $Y_{\sigma,p,m}$ containing censored, noisy versions of r_i were created, using the following function,

$$y_{i,\sigma,p,m} = r_i c_i + \epsilon_i, \tag{2.29}$$

where $c_i \sim \mathcal{U}(m,1)$ with probability p or $c_i = 1$ with probability $1-p$ and $\epsilon_i \sim \mathcal{N}(0,\sigma^2)$. In (2.29), r_i is censored if $c_i < 1$. The number of r_i that are censored is controlled by the censoring probability p, while the extent to which r_i are shrunk, if they are shrunk, is controlled by the lower bound of the censoring factor m. After potential censoring of r_i, additive Gaussian noise with standard deviation σ is added to it. For each tested value of σ, p, and m, $(x_i, y_{i,\sigma,p,m})$ with $i \in \{1,\ldots,900\}$ were used for training an SVM for regression (2.26) and an SVM for the regression of right-censored data (2.27). SVMs were trained with all possible combinations of the SVM parameters $\gamma \in \{2^{-5},\ldots,2^4\}$ and $\epsilon \in \{2^{-5},\ldots,2^4\}$ (2.26 and 2.27). The remaining 100 points were used for testing the SVMs, using the mean square error (MSE) as a loss function. MSE was calculated with the predicted $\hat{y}_{i,\sigma,p,m}$ and their uncensored, denoised counterparts, $r_i = 4x_{i,1} - 3x_{i,2} + 2x_{i,3} - x_{i,4}$. For each tested σ, p, and m, the MSE for the best-performing set of SVM parameters is displayed in Figure 2.5. As can be seen, SVMs for regression and SVMs for the regression of right-censored data do not differ much in performance (MSE), as long as the censoring probability p is less than 0.5. However, for censoring probabilities greater than

0.5, the MSE of the SVM for regression is higher than that of the SVM for the regression of right-censored data. For both SVM formulations, MSE increases with p, but SVMs for the regression of right-censored data perform better, for all tested values of p. The minimum value of the censoring factor m also influences performance, with higher values of m leading to higher MSE. Again, SVMs for the regression of right censored data perform better than SVMs for regression for all tested values of m. The tested values for the standard deviation of the additive Gaussian noise, however, do not influence prediction performance much. In conclusion, this artificial example supports the use of SVMs for the regression of right-censored when the data presents variable degrees of right censoring.

When training a statistical model with right-censored data, we typically know whether a variable has been censored or not. However, if a variable has been censored, we do not know what its value would have been if it had not been censored. The SVM for the regression of right-censored data deals with right censoring of the data by disabling the constraint that the predicted value $\hat{y}_j = \beta^{\mathrm{T}} \cdot x_j + \beta_0$ of right-censored responses y_j be less than $y_j + \epsilon$, $j \in \{j : \delta_j = 0\}$). It is correct that $\hat{y}_j$ be larger than y_j if y_j is right-censored, since the *true* value of y_j is larger than has been empirically observed. However, it is challenging to decide how much larger the value of $\hat{y}_j$ should be with respect to the (right-censored) value of y_j if uncensored versions of y_j are not available. The SVM for the regression of right-censored data strives to capture the functional relationship between x_i and y_i, $i \in \{1, \cdots, n\}$ by choosing β and β_0 such that they reproduce what is considered to be certain. Specifically, $\hat{y}_k$ strives not to deviate from y_k by more than an *acceptable* or *inevitable* extent ϵ, $\forall k \in \{k : d_k = 1\}$, since we deem certain that the empirical measurement of value of y_k was sufficiently accurate. However, $\hat{y}_j$ only strives to be larger than $y_j - \epsilon$, $\forall j \in \{j : d_j = 0\}$, since we deem certain that the empirical measurement of the value of y_j resulted in a value that is too small (but not too large). The increase in the value of $\hat{y}_j$, when compared to y_j, results from the consistency of $\hat{y}_j$ with the functional relationship that was estimated with those characteristics of the dataset which we consider to be certain.

2.2.6 Harrel's Concordance Index

Popular measures for evaluating the predictive performance of regression models (for uncensored data), such as the mean squared error [256] and the Pearson correlation coefficient [266], are not adequate for evaluating the prediction for right-censored data points, for the following reason. Predictions for right-censored data points that are larger than their empirically observed counterparts will result in a less favorable value of the performance measure, although the larger, predicted value might be closer to the unknown, uncensored value. Therefore, the assessment of the predictive performance of statistical models for the regression of right-censored data requires a performance measure that accounts for the fact that the data consist of both right-censored and uncensored points. In the following, I present *Harrel's concordance index* [267, 268], a performance measure designed for evaluating the predictive performance of regression models that are tested on right-censored data.

Let $y_i \in \mathbb{R}$ be possibly right-censored measurements with uncensored counterparts $r_i \in \mathbb{R}$ and with associated censoring labels $\delta_i \in \{0, 1\}$, for $i \in \{1, \ldots, n\}$. If a measurement y_i is censored, then $\delta_i = 0$ and $y_i \leq r_i$. If a measurement y_i is uncensored, then $r_i = y_i$ and $\delta_i = 1$. Let $\hat{r}_i$ be estimates for the values of r_i. We would like to assess the accuracy of the estimates $\hat{r}_i$ in a setting where we know y_i and δ_i, but r_i is latent. Consider all pairs (j, k) of the indices $i \in \{1, \ldots, n\}$, with $j \in \{1, \ldots, n-1\}$, $k \in \{2, \ldots, n\}$, and $j < k$ such that the sets $\{j, k\}$ do not repeat and (i, i) is excluded. In the following, I state the rules for assigning each pair (j, k) one of three labels.

- *Concordant* pairs (j, k) are those for which $y_j < y_k$ and $\hat{r}_j < \hat{r}_k$. Furthermore, $\delta_j = 1$ for all concordant pairs.
- *Discordant* pairs (j, k) are those for which $y_j < y_k$ and $\hat{r}_j > \hat{r}_k$. Furthermore $\delta_j = 1$ for all discordant pairs.
- All pairs that are neither discordant nor concordant are *unusable*.

Additionally, we label pairs that are either concordant or discordant as *usable*. Harrel's concordance index (C) is defined as the fraction of usable pairs that are concordant. C can be interpreted to be the probability of concordance given usability. Let I be the indicator function. We can express Harrel's concordance index in symbols as follows

$$C = \frac{\sum\limits_{\{i|i<j\}} \sum\limits_{\{j|i<j\}} \{ \mathrm{I}(\delta_i = 1)\mathrm{I}(y_i < y_j)\mathrm{I}(\hat{r}_i < \hat{r}_j) + \mathrm{I}(\delta_j = 1)\mathrm{I}(y_j < y_i)\mathrm{I}(\hat{r}_j < \hat{r}_i) \}}{\sum\limits_{\{i|i<j\}} \sum\limits_{\{j|i<j\}} \{ \mathrm{I}(\delta_i = 1)\mathrm{I}(y_i < y_j) + \mathrm{I}(\delta_j = 1)\mathrm{I}(y_j < y_i) \}}. \tag{2.30}$$

Harrel's concordance index assesses the accuracy of the estimates $\hat{r}_i$ based on the order relationships between two censored measurements $y_i < y_j$ and the corresponding uncensored predictions $\hat{r}_i < \hat{r}_j$, without consideration for the differences between the values of y_i and $\hat{r}_i$ or y_j and $\hat{r}_j$, for $i < j$. When calculating C, only such $y_i < y_j$ and $\hat{r}_i < \hat{r}_j$ are considered for which $\delta_i = 1$. This means that y_i, which have been empirically measured, are smaller than y_j and uncensored, i.e. $y_i = r_i$. This requirement ensures that the true value r_i is never greater than the true value r_j. Thus, C quantifies the number of *usable* pairs of predictions $(\hat{r}_i, \hat{r}_j)$ which respect the order relationship of their measured counterparts $(y_i = r_i, y_j \leq r_j)$. This number is normalized by the total number of *usable* pairs and can thus assume a value between zero and one. Note that C is a measure of correlation that does not account for the deviation between predicted and true (or measured) values.

2.3 Kernel Density Classification

Kernel density classification relies on KDE for nonparametric classification. Therefore, in this section, I first review KDE, after which I review kernel density classification. KDE is a method for estimating a probability density function from observations drawn from that function. Although KDE can be used for estimating multivariate densities, I only review the univariate KDE. This section is based on [256, 269] to a large extent. Furthermore, I extensively quote from [269].

2.3.1 Kernel Density Estimation

KDE is a local method for estimating the probability density of a continuous distribution with points sampled from that distribution. In comparison to histogram density estimation, KDE presents the advantage of yielding smooth estimates while inherently providing the possibility of interpolation. Suppose we draw n points $x_i \in \mathbb{R}$ independently from the probability density $f(x)$ of a continuous distribution. The kernel density estimate of $f(x)$ at x_0 is given by

$$\hat{f}(x_0) = \frac{1}{n} \sum_{i=1}^{n} k_\lambda(x_0 - x_i), \tag{2.31}$$

where $k_\lambda(x)$ is a symmetric kernel function with bandwidth $\lambda > 0$. Popular choices for the kernel function include:

- Epanechnikov kernel: $k(x) = \begin{cases} |\frac{3}{4}\frac{1-x^2}{\lambda}| & \text{if } |x| \le 1; \\ 0 & \text{otherwise} \end{cases}$

- Gaussian kernel: $k(t) = \phi_\lambda(t)$, the probability density of a normal distribution with mean equal to zero and standard deviation equal to λ

- Tricube kernel: $k(x) = \begin{cases} |\frac{(1-|x|^3)}{\lambda}| & \text{if } |x| \le 1; \\ 0 & \text{otherwise.} \end{cases}$

λ is a smoothing parameter that determines the width of a local neighborhood around x_0. The points in this neighborhood are considered for producing the estimate $\hat{f}(x_0)$. Larger values of λ amount to averages over more observations. While the Epanechnikov and the tricube kernels are compact, the Gaussian kernel has infinite support. Thus, when using the Gaussian kernel, all observations are always considered for producing a density estimate, albeit to different degrees. While the choice of the probability density function does not have a large influence on the quality of the estimates, an adequate choice of λ is crucial for obtaining good results [270]. In the following, I briefly review a bandwidth estimation method that is used in this work.

The performance of a kernel density estimator as a function of its bandwidth λ can be quantified using the mean integrated square error (MISE),

$$\text{MISE}(\lambda) = E \int_{\mathbb{R}} |\hat{f}_\lambda(x) - f(x)|^2 \, dx. \tag{2.32}$$

Thus, the optimal bandwidth λ can be determined by minimizing the MISE. However, in general, the MISE does not have a closed form, which is computationally disadvantageous. For this reason, the MISE is often replaced by an asymptotic estimate

$$\text{AMISE}(\lambda) = \frac{1}{2\sqrt{\lambda \pi}n} + \frac{1}{4}\lambda^2 \psi_2, \tag{2.33}$$

where $\psi_2 = \int_{\mathbb{R}} f^{(2)}(x)f(x) \, dx$ is an integrated density derivative functional of order 2. As can be seen, AMISE (2.33) depends on the target density f through ψ_2. Plug-in bandwidth estimation methods estimate ψ_2 from the data in order to select an adequate bandwidth. This estimate is given by

$$\hat{\psi}_r(g) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}^{(r)}(x_i; g) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} k_g^{(r)}(x_i - x_j), \tag{2.34}$$

where $k_g(x)$ is the selected kernel function with bandwidth g which, in turn, must be estimated. Since the selection of g is not as important as the selection of λ for the quality of the density estimates, rough estimates suffice. Further details on bandwidth estimation go beyond the scope of this work. However, the interested reader can consult [271–273].

KDE can also be performed in a weighted fashion. This is useful when the frequencies of available observations sampled from a probability distribution do not match that probability distribution, and weights for each observation are available. Let $x_i \in \mathbb{R}, i \in \{1, \ldots, n\}$ be drawn from a probability density function $f(x)$, and $w_i \in \mathbb{R}^+$ be the corresponding weight for x_i. The weighted kernel density estimate for $f(x)$ at x_0 is

$$\hat{f}_w(x_0) = \frac{1}{\sum_{i=1}^{n} w_i}\sum_{i=1}^{n} w_i k_\lambda(x_0 - x_i), \tag{2.35}$$

where $k_\lambda(x)$ is a symmetric kernel function with bandwidth $\lambda > 0$ [274].

2.3.2 Use of Kernel Density Estimation for Classification

KDE in conjunction with Bayes' theorem (Section 2.1.2) can be used for classification, as follows. Let $\hat{f}_j(x)$, $j \in \{1, \ldots, J\}$, be the kernel density estimates for J different classes, and let π_j be the corresponding class priors. An estimate of the posterior probability of class j given an observation x_0 is

$$P(C = j \mid X = x_0) = \frac{\pi_j \hat{f}_j(x_0)}{\sum_{k=1}^{J} \pi_k \hat{f}_k(x_0)}. \tag{2.36}$$

Using (2.36), we can calculate the probability estimates for each of the classes $j \in \{1, \ldots, J\}$ given an observation x_0, and classify x_0 to the class of highest probability:

$$\mathrm{C}(x_0) = \underset{j \in \{1, \ldots, J\}}{\arg\max} \; P(C = j | X = x_0). \tag{2.37}$$

This is known as kernel density classification. Kernel density classification presents a number of advantages, some of which are listed in the following. (1) Due to its non-parametric nature, kernel density classification can be used for solving linear and non-linear classification problems. (2) For the purpose of interpretation, the kernel density estimates for each of the classes can be easily plotted along with their posterior probability estimates. This yields very intuitive plots (Section 3.5). (3) Multivariate KDE allows for straight-forward integration of heterogeneous measurements which can be jointly used for performing classification.

3

Inferring Drug Resistance and Drug Exposure from the Genotype

Prolonged chemotherapy against HIV-1 bears the risk of selection of resistant viral strains, ultimately leading to therapy failure (Section 1.5.3) [199, 228, 275–278] . Once a drug-resistant HIV-1 variant has been selected in a host, it can be transmitted to another host [278, 279]. In order to prevent premature therapy failure, the susceptibility of an HIV-1 variant to available antiretroviral drugs can be measured phenotypically or genotypically [228, 280–283]. Due to the high cost, limited accessibility and high turnaround time of phenotypic resistance tests, genotypic resistance testing has become the standard of treatment [228, 280]. This chapter is mainly concerned with genotypic resistance testing. The first two sections of this chapter give a brief review on phenotypic and genotypic determination of resistance. Section 3.3 describes a method for training a system for genotypic drug-resistance interpretation without the need of expert intervention. This system outputs numbers that are correlated with drug exposure, drug resistance, and therapeutic success. However, the use of this system by experts requires the translation of these numbers into clinically meaningful categories. For this reason, I devoted Section 3.4 to methods for the determination of cutoffs for the system. Using these cutoffs, the numerical output of the system can be translated into clinically meaningful categories regarding drug exposure, drug resistance, and therapeutic success, which are partially discriminative of these three matters. Section 3.5 is also devoted to cutoff determination for the translation of a numerical quantity related to resistance into clinically meaningful categories. However, the methods presented in Section 3.5 are designed to be applied to an established genotypic drug-resistance interpretation system, geno2pheno$_{[resistance]}$. Since geno2pheno$_{[resistance]}$ provides genotypic drug-resistance interpretation, the methods presented in Section 3.5 only aim at translating the numerical output of geno2pheno$_{[resistance]}$ into categories that describe the extent of drug resistance.

3.1 Overview of Phenotypic Drug-Resistance Testing

Phenotypic drug-resistance testing involves *in-vitro* cultivation of a reference HIV-1 strain and of a patient-derived HIV-1 strain with different concentrations of a tested drug compound. HIV-1 resistance (or hypersus-

ceptibility) to the drug compound is quantified as the minimum fold difference in drug concentrations that is required to elicit a 50% inhibition of the viral replication of patient-derived HIV-1 when compared to the reference strain. For example, if the drug concentration required to elicit 50% inhibition of a reference strain is $0.02\,\mu$M, and the patient-derived strain requires $0.1\,\mu$M in order to reach 50% inhibition, the fold difference is 5. This number is called fold-change in the 50% inhibitory drug concentration (FC) or resistance factor (RF) [284–290].

First-generation phenotypic resistance tests use PBMCs isolated from the patient for assessing drug resistance. These PBMCs are subsequently cultivated with uninfected, HIV-1-permissive cells [284, 285]. However, this approach exhibits a number of drawbacks. First, it is technically difficult to culture virus from PBMCs [290]. Second, the use of PBMCs from different patients introduces variability into the measurements. Third, distinct viruses present differential replication kinetics, which has to be compensated with assay modifications [284]. Fourth, at least four weeks of *in-vitro* virus cultivation are required for obtaining results [284, 285]. Cultivation of the virus involves sequential passaging of the clinical isolate, which in turn, influences the results of the experiment. Most importantly, the *in-vitro* virus-cultivation procedure may select for unrepresentative viral populations [285]. Fifth, HIV-1 integrated into PBMCs may be in itself an unrepresentative viral population, as drug resistant HIV-1 variants may be produced by other cells in the body while not being produced by PBMCs [290–292].

The second generation of phenotypic drug-resistance tests addresses these issues by using replication-deficient viral vectors that lack the genes that are the targets of antiretroviral drugs. After isolation of HIV-1 RNA from the patient's blood, replication-competent viral vectors are created from the replication-deficient viral vectors and from the isolated viral RNA via genetic recombination. Specifically, the viral genes that the replication-deficient viral vectors lack are provided by the patient-derived viral RNA . Subsequently, the newly created, replication-competent viral vectors are inserted into HIV-1 permissive cells, and cultured *in-vitro* with different concentrations of the drug in question [286, 287]. While the second generation of phenotypic drug-resistance tests (e.g. Antivirogram®) solved many issues concerning the first-generation of tests, other issues remained. First, in second generation tests, measurement of viral replication is performed indirectly with cell-viability assays. Therefore, the extent of viral replication is correlated with its cytopathic effects [286, 287]. Furthermore, in second generation tests, cell viability is very frequently determined by using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) [287], which is reduced by viable cells to formazan, resulting in a purple coloration. However, MTT has been shown to have a cytotoxic effect on the cells whose viability is being tested [293]. Second, a turn-around time of several weeks remains necessary for obtaining results. This time is mainly spent in *in-vitro* cultivation of the virus, which can potentially select for certain viral populations, as mentioned above. The non-specific number of viral replication cycles occurring during *in-vitro* virus cultivation introduces additional variability into the measurement.

Third-generation phenotypic drug resistance tests (e.g. PhenoSense®) represent an improvement over second-generation tests in the following respects. Viral vectors used in third-generation phenotypic drug resistance tests do not only lack the genes that are targets of the tested antiretroviral drugs, but also lack an *env* gene (Section 1.1.1), as it is replaced by a luciferase reporter gene. For production of viral particles with the viral vectors, recombination with genes provided by the HIV-1 variant infecting the patient is not sufficient. Since the *env* gene in the viral vectors has been replaced with another gene, these can only produce replication-competent viral particles if the viral env proteins are present in the cell in which they assemble. Third-generation phenotypic drug resistance tests are designed in such a way that the env proteins are only present in transfected cells,

but not in cells that are infected by the viral particles produced in the transfected cells. Therefore, replication will be limited to one cycle. This decreases turn-around time and diminishes inter-assay variability, when compared to multiple-cycle assays. Viral replication is quantified by measuring the amount of luciferase produced in the cell, which is more direct and precise than using cell-viability assays [288]. Furthermore, luciferase-based quantification of viral replication avoids potentially cytotoxic effects associated with cell-variability assays.

When compared to genotypic drug-resistance tests (Section 3.2), phenotypic drug-resistance tests present the advantage that they inherently consider all possible patterns of viral mutations. Depending on their pattern, HIV-1 mutations can elicit drug resistance or drug hypersusceptibility. Furthermore, the FC is easier to interpret (Section 3.5) than a list of HIV-1 mutations. Interpretation of HIV-1 mutations requires prior knowledge of their correlation with drug resistance. However, in contrast to genotypic drug-resistance testing, phenotypic resistance tests may underestimate resistance *in-vivo* since susceptible and resistant viral strains may coexist in the tested sample and appear in the consensus sequence of the gene in question [294]. Furthermore, viral strains with mutations that do not directly cause resistance, but are strongly associated with the emergence of drug resistance, may be deemed susceptible by *in-vitro* phenotypic drug-resistance assays. If the respective drugs are taken by patients harboring these strains, resistant variants may promptly emerge, preventing the drugs from inhibiting viral replication [295]. Last but not least, certain patterns of drug resistance mutations are associated with a loss in viral replication capacity. For this reason, the procedure used for testing drug resistance phenotypically can potentially select viruses that have fewer drug-resistance mutations and have therefore higher replication capacities.

3.2 Genotypic Drug-Resistance Testing: Past Work

Interpretation of the genotype of HIV-1 with respect to drug resistance is called genotypic drug-resistance testing. In order to perform this interpretation, knowledge on the relationship between HIV-1 mutations and drug resistance is required. This knowledge can be attained through the analysis of associations of HIV-1 mutational patterns and drug resistance. While HIV-1 mutational patterns can be measured through sequencing, drug resistance can be observed *in-vivo* during therapy failure and through *in-vitro* testing (Section 3.1). The least complex form of genotypic drug-resistance interpretation is given by drug-resistance mutation tables [296]. These tables list drug-resistance mutations for each drug. For some drugs, mutations are classified into minor and major drug resistance mutations. While major drug resistance mutations cause drug-resistance by themselves, minor drug-resistance mutations may only act when they occur in certain patterns, which may include major drug-resistance mutations. Some minor drug-resistance mutations have the effect of restoring the loss in replicative capacity that is caused by some major drug-resistance mutations.

Genotypic drug-resistance interpretation systems afford more complex and also accurate estimation of drug-resistance. Rules-based genotypic drug-resistance interpretation systems apply a series of expert derived-rules for producing a drug-resistance score or classifying the genotype into one of several categories describing drug resistance. Experts produce rules for genotypic drug-resistance interpretation systems based on genotype-phenotype pairs (GPPs), HIV-1 genotypes obtained during or preceding therapy failure, and other experimental observations that allow for the association of HIV-1 mutations to drug resistance. An example of such a rule follows: If reverse-transcriptase residue K101 is substituted by 101P in the genotype of a patient's viral strain, then none of the currently available NNRTIs will be effective in inihibiting the replication of that viral strain. Popular rules-based genotypic drug-resistance interpretation systems include ANRS

(http://www.hivfrenchresistance.org/), GRADE [297], HIVdb [298], and REGA (http://rega.kuleuven.be/). Genotypic drug-resistance interpretation can also be performed with data-driven genotypic drug-resistance interpretation systems. Data-driven genotypic drug-resistance interpretation systems are trained on pairs consisting of a genotype and a measurement informative of drug resistance, such as the FC value. geno2pheno$_{\text{[resistance]}}$ (http://www.geno2pheno.org) [299] is a popular data-driven genotypic drug-resistance interpretation systems that is trained on GPPs. Phenotypic drug-resistance testing is associated with several advantages and disadvantages (Section 3.1), and therefore, genotypic drug-resistance interpretation systems trained solely on GPPs can potentially inherit these advantages and disadvantages. In the following, I present a method for deriving a genotypic drug-resistance interpretation system that uses GPPs and also HIV-1 genotypes derived after exposure to antiretroviral drugs. Once programmed, the system does not require expert intervention for training.

3.3   Inference of Drug Exposure from Clinical Data

In this section, I present a method with which a data-driven genotypic drug-resistance interpretation system was developed. The system does not require quantitative phenotypic resistance measurements for training and can make use of HIV-1 data generated during routine medical practice. I consider important that such a system allow for automatic retraining without intervention by human experts. I expect that such a tool, in conjunction with a frequently updated, large clinical HIV-1 database, will be easy to keep up to date and will interpret HIV-1 genotypes with a reduced risk of expert bias. This section additionally provides an analysis of the predictive value of polymorphisms located at positions in the amino-acid sequence where established drug-resistance mutations do not occur and an identification of *in-vivo* genetic drug footprints through the novel method. The term *in-vivo genetic drug footprint* refers to a set of amino-acid substitution weights indicating the propensity of a drug to select for the respective substitution in patients treated with the drug. At the same time, these genetic drug footprints are used as a means for interpretation of the predictions of the method.

3.3.1   Materials and Methods

In the following, I describe the datasets and methods that were used for developing the interpretation system described in this section.

Drugs Considered in this Method

The interpretation system described in this section considers the following antiretroviral drugs: 3TC, ABC, AZT, d4T, ddC, ddI, FTC, TDF, delavidrine (DLV), EFV, etravirine (ETR), NVP, RPV, amprenavir (APV), ATV, DRV, FPV, IDV, LPV, NFV, SQV, TPV, RAL, and EVG. Other antiretroviral drugs were not considered due to insufficient data.

Nucleotide Sequences and Drug-Exposure Information in PRRT and IN

The PRRT dataset was constructed by pooling 70,304 HIV-1 PR and RT nucleotide sequences from two sources: 37,799 sequences from the EuResist Integrated Database (EIDB; http://www.euresist.org; downloaded on April 11th, 2014)[300], 9,627 of which were derived from drug-naïve patients (short: drug-naïve sequences), and 32,506 drug-naïve sequences from the Los Alamos National Laboratory Sequence Database

(LANLSD; http://www.hiv.lanl.gov/; downloaded on March 31st, 2015). Among the sequences in the PRRT dataset derived from therapy-experienced patients (short: drug-exposed sequences), 18,328 sequences were derived from patients whose complete drug history was available at the time of sequencing. The IN dataset includes a total of 5,523 IN nucleotide sequences with the following characteristics: 3,382 sequences were extracted from EIDB, 1,240 of which are drug-naïve, 397 have been exposed to an INI and possibly other drugs, and 1,745 have been exposed only to drugs whose target is different from IN. The complete drug history is available for 1,432 of the drug-exposed IN sequences. Additionally, 3,782 drug-naïve IN sequences from LANLSD (downloaded on March 31st, 2015) were added to the IN dataset. In both PRRT and IN datasets, drug exposure lasting for less than 30 days was disregarded. The inclusion criteria applied to the sequences were the following. (1) Alignment with the MutExt software (http://www.schuelter-gm.de/mutext.html) must not have produced an error due to low sequence similarity to the reference sequence, (2) at most 10% of the residues of the considered protein regions could not be determined by the sequencing procedure (considered protein regions are listed in the section entitled *Subtype Determination, Sequence Alignment and Encoding*, which is part of Section 3.3.1), (3) the amino-acid sequence resulting from nucleotide translation was enforced to be unique within the dataset, unless drug exposure differed between duplicates. The order of appearance of the sequences in the dataset determined which duplicate sequence was excluded, with sequences appearing first preempting inclusion of sequences appearing later. Older RT sequences not covering amino-acid positions 221-230 were excluded as well.

Nucleotide Sequences in Naive_PRRT and Naive_IN

Transmitted drug resistance (TDR) in PI- and RTI-naïve sequences was defined as the presence of at least one mutation in the list of drug resistance mutations for surveillance of transmitted HIV-1 drug resistance [279]. Since the list of transmitted drug-resistance mutations only contains PR and RT mutations, TDR in IN sequences was defined in terms of the presence of an INI drug-resistance mutation in the 2013 IAS list [296]. Following the methodology used for establishing the list of transmitted drug-resistance mutations, INI drug-resistance mutations with a prevalence greater than 0.5% among sequences from the LANLSD in IN were not regarded as indicative of TDR [301]. The Naive_PRRT and Naive_IN were created by randomly sampling 2,500 LANLSD sequences without TDR from the PRRT and IN datasets, respectively. These datasets are used by our web service for z-score calculation. I do not refer to them further, in this work.

Definition of Treatment Episode

In the context of the analysis of treatment success, the notion of a treatment-change episode (TCE) has been established [302]. In summary, a TCE documents relevant clinical parameters concerning a change in the drug compounds of cART. Treatment episodes (TEs) differ from TCEs in that no treatment change is required, i.e. TEs encompass first-line therapies as well. In this analysis, a TE consists of a baseline PR and RT genotype, a list of drug compounds used in a therapy, follow-up VLs and, optionally, a baseline VL. The baseline genotype and the baseline VL must have been obtained no earlier than 90 days before treatment initiation, in line with previously developed definitions of the standard datum of the EIDB [300, 302]. However, baseline genotypes for first-line therapies are exempt from this requirement, as the virus has not been subject to selective pressure by drug therapy. In the presence of multiple data points, baseline measurements obtained at the date closest to therapy initiation are preferred. TEs containing drugs not considered in this analysis or with unboosted PIs

(except for nelfinavir) do not satisfy the TE definition. Furthermore, VL measurements were constrained to those not reaching a lower limit of quantification greater or equal than 50 HIV-1 RNA copies per milliliter of blood serum.

Definition of Treatment Success of the EuResist Standard Datum

According to the definition of the EuResist Standard Datum, treatment success is determined with a follow-up VL, and optionally, a baseline VL, as follows. A follow-up VL must have been obtained between four and twelve weeks after therapy initiation, preferring the VL measurement whose measurement date is closest to eight weeks after therapy start. TEs in which this follow-up VL is below 400 HIV-1 RNA copies per milliliter of blood serum or comparison with the baseline VL shows an at least 100-fold reduction in the VL are labeled as successes. TEs for which a baseline VL is available, the follow-up VL is above 400 HIV-1 RNA copies per milliliter of blood serum, and the VL reduction is less than 100-fold are labeled as failures [300, 302].

Definition of Number of Aviremic Semesters

Treatment success, as defined above, is concerned with short-term therapeutic success or failure. In the following, I describe the NAS, a measure that accounts for long-term therapeutic success. By definition, a therapy semester is one of possibly many contiguous, non-overlapping periods of 26 weeks during a therapy. The first therapy semester begins on the first day of the therapy. For computing the NAS, follow-up VL measurements are grouped by the therapy semester during which they were performed, and VL measurements for each semester are averaged. A therapy semester is considered aviremic if its mean VL is less than some threshold in the mean number of HIV-1 RNA copies per milliliter of blood serum. For determining the success of antiretroviral therapy, a threshold of 50 HIV-1 RNA copies per milliliter of blood serum has become standard [228, 280] *. However, one can choose a higher threshold in order to accommodate blips [278] into the aviremic semesters (Section 4.1.2). The NAS can be right-censored. Specifically, if a therapy is still ongoing when it is stored in the EIDB or a therapy semester is devoid of VL measurements, the NAS might be right-censored, as the missing VL measurements could lead to an increase of this quantity. Furthermore, therapy switches while a patient's VL is suppressed can also be regarded as censoring, as therapy continuation could have led to further aviremic semesters. I do not compute the NAS for therapies for which the number of therapy semesters without a VL measurement is greater than 10% of the total number of recorded therapy semesters, or for therapies lasting less than four weeks. Figure 3.1 shows an example for the computation of the NAS following treatment initiation. For the analysis presented in Section 3.3, a threshold of 50 HIV-1 RNA copies per milliliter of blood serum is used for computing the NAS. Furthermore, therapy switches while a patient's average VL is under this VL threshold are regarded as right censoring.

---

*For defining therapeutic success, the EuResist Standard Datum uses a threshold of 400 HIV-1 RNA copies per milliliter of blood serum in conjunction with the extent of VL decay, eight weeks after initiation of treatment. However, a threshold of 50 HIV-1 RNA copies per milliliter of blood serum has become standard. There are two reasons for the use of this higher VL threshold in the definition of the EuResist Standard Datum. (1) For many patients, eight weeks of chemotherapy are not sufficient to reduce the VL below 50 HIV-1 RNA copies per milliliter of blood serum. Therefore, the use of a threshold of 50 HIV-1 RNA copies per milliliter of blood serum (disregarding the extent of VL decay) would be inadequately stringent, as the VL may decrease further during the course of therapy. (2) The EuResist Integrated Database includes older VL measurements with a lower limit of quantification of 400 HIV-1 RNA copies per milliliter of blood serum. The use of a VL threshold below 400 makes exclusion of these older VL measurements necessary.
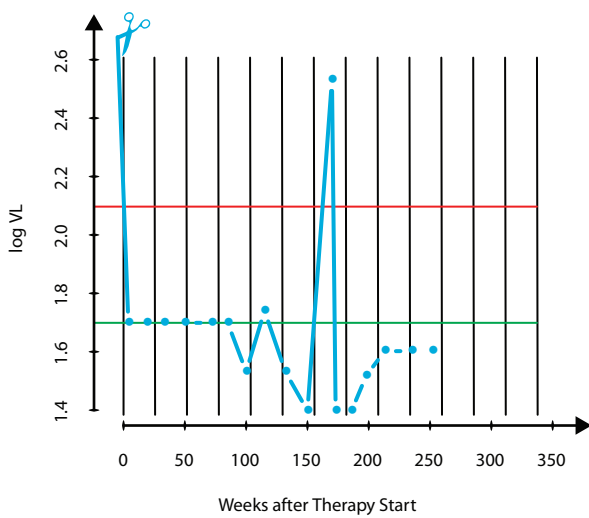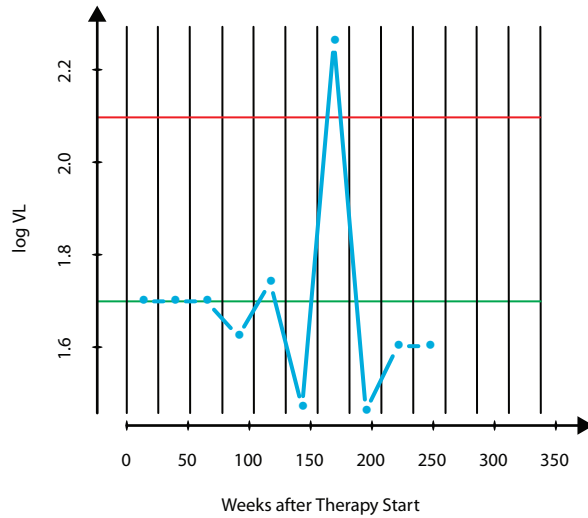
(a) Viral-Load Trajectory   **(b)** Semester-Wise Averaged Viral-Load Trajectory

Figure 3.1: Computation of Aviremic Semesters (Example). For computing the number of aviremic semesters, the viral-load trajectory of a patient is divided into contiguous periods of 26 weeks (therapy semesters). The first period begins on the first day of therapy. The viral-load trajectory of a patient on antiretroviral therapy is depicted in Figure (a). The value of viral load prior to therapy initiation is truncated in Figure (a), which is indicated by the scissors. Vertical black lines delimit therapy semesters. Viral-load measurements for each semester are averaged (Figure (b)), and the number of aviremic semesters is computed as a function of the chosen viral-load threshold. For the therapy in Figure (b), a VL threshold of 50 HIV-1 RNA copies per milliliter of blood serum (green line) would yield 8 aviremic semesters, while a VL threshold of 125 HIV-1 RNA copies per milliliter of blood serum (red line) would yield 9 aviremic semesters. The number of aviremic semesters can be right-censored if any therapy semester is devoid of viral-load measurements or if the therapy is still ongoing when this measure is computed. Furthermore, a number of aviremic semesters can be considered to be right-censored if the therapy was interrupted while the viral load was below the chosen detection threshold.

— 50 HIV-1 RNA copies per milliliter of blood serum   — 125 HIV-1 RNA copies per milliliter of blood serum
— Viral-load trajectory

Definition of Treatment-Failure Episode

TEs with baseline sequences indicating drug resistance to some of the therapy's drug compounds are rare in EIDB. Therefore, I resort to nucleotide sequences measured after the initiation of a therapy, which often implies therapeutic failure, as a certain degree of viral replication is required for sequencing. Towards this end, I define a treatment-failure episode (TFE) as follows. A TFE consists of a list of drugs used during a therapy and an HIV-1 sequence obtained during the therapy. The therapy is required to have lasted at least four weeks, and the genotype is required to have been obtained no earlier than four weeks before therapy stop. In TFEs, I assume that the treating clinician changed the therapy based on the results of the genotypic drug resistance test.

Treatment Episodes in EuResistTE

The EuResistTE dataset was constructed by extracting a total of 9,201 TEs from the EIDB [300]. These TEs were constructed according to the definition of treatment episodes and labeled as successes or failures according to the definition of the EuResist standard datum (see above) [302]. In summary, each TE includes a PR and RT baseline sequence, the compounds that were prescribed to the patient, a baseline and a follow-up VL, and a binary label indicating whether the therapy was successful or not. Additionally, the NAS (see above) was computed for each TE, unless a VL measurement was available for less than 10% of the recorded therapy semesters. To facilitate performance comparison, only therapies including the following antiretroviral drugs were considered: 3TC, ABC, AZT, d4T, ddI, FTC, TDF, EFV, ETR, NVP, APV, ATV, DRV, FPV, IDV, LPV, NFV, SQV, TPV, and RTV as a boosting agent.

Treatment-Failure Episodes in TFE

The treatment-failure episode dataset TFE was constructed by extracting a total of 2,454 EIDB PR and RT genotypes and corresponding therapy compounds obtained during therapy failure, as defined above. Constraints on the investigated drugs were identical to those applied for obtaining EuResistTE (see above). Therapies were required to have lasted at least four weeks, and the genotype was required to have been obtained no earlier than four weeks before therapy stop.

Datasets for the Training and Testing of Models for Predicting Drug Exposure

For the purpose of training and testing models for predicting drug exposure, several datasets were created. These datasets were extracted from PRRT, IN, EuResistTE, and TFE, which are described above. In order to construct training and test sets that are patient-wise disjoint, a dataset of test patients P was iteratively created. With the exception of P, the names given to the datasets described in the following consist of two parts. The first part is an abbreviation related to the purpose for which the dataset was created. The second part, in subscript, indicates the dataset from which it was extracted. The therapy-pause dataset TP_{PRRT} was extracted from PRRT. TP_{PRRT} was created in order to test the capability of drug-exposure models to detect drug exposure in viral sequences that were measured during therapy pauses. Since TP_{PRRT} is a test dataset, patients with sequences in TP_{PRRT} were added to P. For the purpose of testing the capability of drug-exposure models to detect drug exposure in arbitrary HIV-1 sequences, I created the test sets T_{PRRT} and T_{IN}. In order to assign approximately 10% of the number of sequences in PRRT and IN to T_{PRRT} and T_{IN}, respectively, while keeping training and development sets patient-wise disjoint, sequences included in T_{PRRT} and T_{IN} were defined to be those originating from patients P. Starting with the patients with sequences in TP_{PRRT}, P was successively enlarged with

patients from PRRT and IN. In each iteration of the enlargement of P, one patient from PRRT and IN was selected at random and appended to P. The enlargement of P induced the enlargement of T_{PRRT} and T_{IN} in turn. Iterative enlargement of P was repeated until the number of PRRT and IN sequences derived from the patients in P was approximately 10% of the number of sequences in PRRT and IN, respectively. Note that TP_{PRRT} is a subset of T_{PRRT}. In order to train models for predicting drug exposure, the development sets D_{PRRT} and D_{IN} were created. D_{PRRT} and D_{IN} contain the sequences in PRRT and IN, respectively, that are not included in T_{PRRT} or T_{IN}, respectively. In order to estimate cutoffs for drug-exposure models (Section 3.4) and in order to test the capability of drug-exposure models of predicting therapeutic success, I created development and test sets containing TEs, as follows. The baseline sequences of the TEs in EuResistTE partially overlap with the sequences in T_{PRRT}, T_{IN}, D_{PRRT}, and D_{IN}. I created a development dataset $D_{EuResistTE}$ which includes all TEs with baseline sequences in D_{PRRT}. The $T_{EuResistTE}$ dataset was created for testing purposes and it contains all TE baseline sequences not included in D_{PRRT}. Furthermore, I created a dataset $T_{EuResistTETP}$ which includes TEs whose baseline sequence was measured during a therapy pause. $T_{EuResistTETP}$ contains TEs from $T_{EuResistTE}$ whose baseline sequences are also included in TP_{PRRT}. Determination of cutoffs for drug exposure models did not only require TEs, but also TFEs. However, no TFE including a sequence selected for testing purposes should be used for cutoff determination. Thus, I created the D_{TFE} dataset with the sequences in TFE which are included in D_{PRRT}.

Genotype-Phenotype Pairs in Pheno, D_{Pheno}, and T_{Pheno}

For the purpose of determining cutoffs for drug-exposure models (Section 3.4), as well as testing the capability of drug-exposure models to predict drug resistance, a total of 7,597 GPPs were downloaded from the HIV Drug Resistance Database [294] on April 15, 2015 (Pheno dataset). The phenotypic drug-resistance assays used for producing the phenotypes were constrained to Antivirogram® [287] and PhenoSense® [288]. The genotypes are provided in the form of substitutions with respect to the reference sequence *consensus B* [294]. 3,323 GPP quantify PI resistance, 3,477 RTI resistance, and 797 INI resistance. The T_{Pheno} dataset was created from the Pheno dataset by randomly sampling approximately 10% of the GPP. The rest of the GPPs in Pheno were assigned to the D_{Pheno} dataset.

Independent Test Sets HIVdbExposure and HIVdbTCE

For the purpose of generating independent test sets, the TCE repository in the HIV Drug Resistance Database was downloaded in its entirety on November 21, 2013. [294]. A total of 1,441 sequences with drug-exposure information could be extracted from the repository (HIVdbExposure dataset). For creating the HIVdbTCE dataset, the EuResist Standard Datum definition was applied to therapies in HIVdbTCE whose drug compounds are investigated in this study (with the exception of ddC and RAL for the sake of performance comparison). The NAS was computed for the therapies in HIVdbTCE, excluding those with VL measurements reaching a lower limit reaching a lower limit of quantification greater than 50 HIV-1 RNA copies per milliliter of blood serum.

Phenotypic Resistance Cutoffs

In this study, two sets of phenotypic resistance cutoffs were used. The drug-independent cutoff set categorizes RFs into susceptible and resistant with the values one and ten. Specifically, RFs smaller or equal to one are

classified as susceptible, while those greater or equal to ten are classified as resistant. RFs between one and ten are not used for training. The set of clinically relevant cutoffs for PhenoSense GPPs was obtained from the HIVdb website [294] and is composed as follows. 3TC: 3 and 20; ABC: 3 and 6; AZT: 3 and 10; d4T: 1.5 and 2; ddI: 1.5 and 2; TDF 1.5 and 4; all NNRTIs: 3 and 10; and all INIs: 4 and 20. The clinically relevant cutoff set discretizes PhenoSense GPPs into the classes susceptible, intermediate, and resistant.

Subtype Determination, Sequence Alignment, and Encoding

The subtype distribution in the PRRT and IN datasets was determined with the COMET subtyping tool [303, 304]. Nucleotide sequences in PRRT and IN were aligned to the wild-type reference strain HXB2 and translated, using MutExt (http://www.schuelter-gm.de). The resulting amino-acid sequences, along with amino-acid sequences in the Pheno dataset, were represented vectorially with a binary encoding. The vectorial representation considers substitutions, deletions, and the presence of insertions within the following HXB2 amino-acid positions: PR 3-99, RT 40-230, and IN 30-260. The presence of deletions and insertions was encoded for each amino-acid position, while the sequence of a specific insertion was not encoded. For each nucleotide sequence set, three versions of the vectorial representation were produced: V_{full} considers all amino acids, $V_{noIASPos}$ disregards amino-acid positions at which drug-resistance mutations in the IAS 2013 list [296] occur, and $V_{onlyIASPos}$ only considers such amino-acid positions.

Creation of Exposure and ExposurePheno Cross-Validation Sets

For the purpose of cross-validation and training of models for predicting drug exposure, the Exposure and ExposurePheno cross-validation sets were created. Exposure cross-validation sets only consist of nucleotide sequences with drug-exposure information, while ExposurePheno cross-validation sets additionally include nucleotide sequences from GPPs. D_{PRRT} and D_{IN} were used for constructing the cross-validation sets $Exposure_{drug}$ for $drug \in$ {ABC, AZT, d4T, ddC, ddI, 3TC/FTC, TDF, EFV, ETR, NVP, RPV, ATV, DRV, APV/FPV,IDV, LPV, NFV, SQV, TPV, RAL} which contain an equal number of sequences exposed and not exposed to a certain drug. Sequences not exposed to the drug were randomly selected from the development set, as they were in excess; these sequences were required to have been derived from patients whose complete drug exposure history is recorded. Where possible, half of the sequences not exposed to the drug were drug-naïve, and half of them were exposed to some other drug. A cross-validation set $Exposure_{naivePPRT}$ containing an equal number of drug-naïve and drug-experienced PR and RT sequences was constructed as well. An $Exposure_{naiveIN}$ cross-validation set was not created due to the fact that only a sufficient number of RAL-exposed IN sequences was available. The $ExposurePheno_{drug}$ cross-validation sets were created from the $Exposure_{drug}$ cross-validation sets, with additional supplementation of some genotypes from the D_{Pheno} dataset. Specifically, genotypes with corresponding RFs classified as resistant via the drug-independent cutoffs were treated as drug-exposed sequences while those with corresponding RFs classified as susceptible were treated as sequences not exposed to the drug in question (see *Phenotypic Resistance Cutoffs*). Genotypes with corresponding RFs between the two cutoffs were not used for training. This procedure incremented the number of available drug-exposed sequences and allowed for the creation of the cross-validation sets $ExposurePheno_{RPV}$ and $ExposurePheno_{EVG}$, as the number of available drug-exposed sequences for RPV and EVG was very low. Cross-validation sets for DTG could not be created, as neither a sufficient number of resistant phenotypes nor a sufficient number of drug-exposed sequences were available.

Training and Selection of Models for Predicting Drug Exposure

For performing five repetitions of a 10-fold cross validation, each cross-validation set was randomly partitioned five times into ten folds. Each fold contained an equal proportion of sequences with and without exposure to the drug in question. The partitions were used to cross validate linear SVCs [305] (Section 2.2.2), discriminating between sequences with and without exposure to a certain drug. Three versions of each model were trained: the first one is based on V_{full}, the second one is based on $V_{noIASPos}$, and the third one is based on $V_{onlyIASPos}$. The vectorial representation used to train each drug-specific model was constrained to the vector elements describing the drug's target protein (PR, RT or IN). Each cross validation was performed with a certain value for the regularization parameter γ for the SVC, specifically, $\gamma \in \{2^{-8}, 2^{-7}, \ldots, 2^2\}$. Performance was measured in terms of the area under the receiver operating characteristic curve (AUC) [306, 307]. The signed distance to the classification hyperplane (also called decision value) was used as a score for predicting drug exposure. Thus, I call such decision values drug-exposure scores (DES). For each cross-validation set and vectorial representation, the model with the lowest value of γ whose average performance was not significantly lower than the best average performance was selected (Benjamini-Hochberg-corrected Wilcoxon signed-rank test [308] with a significance threshold of 0.05). Finally, each cross-validation set and vectorial representation was used without partitioning to train a final SVC with the selected value of γ. I refer to these SVCs as the final drug-exposure models, and I group models by the cross-validation set they originated from (Exposure or ExposurePheno) and the vectorial representation that was used to train them (V_{full}, $V_{noIASPos}$, or $V_{onlyIASPos}$). I use the name of a cross-validation set in conjunction with the name of an encoding (in subscript) to refer to the models trained with the respective cross-validation set and encoding, e.g. $Exposure_{full}$.

Comparison of Performance to geno2pheno[resistance]

The performance of continuous DES was compared to that of geno2pheno[resistance] version 3.3 (http://www.geno2pheno.org, accessed on March 31st, 2015) [299]. The output of geno2pheno[resistance] includes a prediction of the RF, an estimated probability of resistance, and a susceptible-intermediate-resistant (SIR) classification for each drug and nucleotide sequence. Comparison of the performance of DES with other popular drug-resistance interpretation systems for which only discrete SIR predictions are available, can be found in Section 3.4. In this section, I only compare DES to continuous geno2pheno[resistance] predictions (RFs and probability of resistance). Since each interpretation system uses its own alignment program, performance comparison was constrained to the set of sequences which could be aligned without errors by all interpretation systems (including those mentioned in Section 3.4). Furthermore, the drug ddC was also excluded from performance comparison, as it is not supported by other interpretation systems any more.

Assessment of Performance

Sequences in T_{PRRT}, T_{IN}, TP_{PRRT}, T_{Pheno}, $T_{EuResistTE}$, HIVdbTCE, and HIVdbExposure were interpreted with the final drug-exposure models and geno2pheno[resistance]. Performance was quantified in terms of AUC (unless stated otherwise). The capability of the systems to predict drug exposure was assessed with DES and predicted RFs calculated with T_{PRRT}, T_{IN}, TP_{PRRT}, and HIVdbExposure. For assessing each system's capability to predict therapy success, a score for each therapy in $T_{EuResistTE}$ and HIVdbTCE and for each interpretation system was calculated as follows. DES were normalized by fitting a sigmoidal function to each of the models, as described before [309]. The sigmoidal function aims at estimating the probability that a certain genotype

belongs to the class of genotypes exposed to a certain drug, given the DES for that genotype and drug (short: probability of exposure to a certain drug; POE). The complements of POEs (i.e. one minus the POEs) for the drugs that were used in a therapy were added in order to calculate scores for that therapy. For performance comparison, the complement of probabilities of resistance [299] from geno2pheno[resistance] were used for calculating a score for each therapy, in the same way as the complements of POEs were used. Scores for each therapy were used for predicting therapeutic success, as defined by the EuResist Standard Datum (see above). Furthermore, concordance of therapy scores with the NAS was assessed for therapies in the $T_{EuResistTE}$ and $HIVdbTE$ datasets in terms of Harrell's concordance index (C; Section 2.2.6) [267]. The capability of the DES to predict drug resistance as assessed by the RFs of GPPs in T_{Pheno} was quantified with the Pearson correlation coefficient [266]. Specifically, the correlation between DES and the log RFs in T_{Pheno} was calculated. Comparison to the performance of geno2pheno[resistance] when predicting drug resistance was not possible due to the fact that genotypes in GPPs were only available in the form of amino-acid sequences. geno2pheno[resistance], however, requires nucleotide sequences for prediction. Performance of DES in predicting phenotypic resistance is compared to other drug-resistance interpretation systems in Section 3.4. Significance values in the Section 3.3.2 section were calculated with a two-sided Wilcoxon signed-rank test [308].

Linear Weights for Drug-Exposure-Score Models

Linear SVCs can be represented as linear functions (Sections 2.2.1 and 2.2.2). Summarizing, DES are of the form

$$\text{DES}(x) = \beta x, \tag{3.1}$$

where $\beta \in \mathbb{R}^p$ is a vector of feature weights, and $x \in \mathbb{R}^p$ is a vector encoding a nucleotide sequence, according to one of the used encodings, V_{full}, $V_{onlyIASPos}$, or $V_{noIASPos}$ (*Subtype Determination, Sequence Alignment, and Encoding* above). β is a linear combination of the vectors in the training set, and it results from mathematical optimization with Lagrange multipliers:

$$\beta = \sum_{i=1}^{n} \alpha_i y_i x_i, \tag{3.2}$$

where $y_i \in \{-1, +1\}$ are the class labels in the training set (not exposed or exposed, respectively), $x_i \in \mathbb{R}^p$ are the encoded nucleotide sequences in the training set, and $\alpha_i \in \mathbb{R}$ are Lagrange multipliers resulting from mathematical optimization (Sections 2.2.1 and 2.2.2). For interpretation of DES, the vector β was computed for each of the drug-exposure models. For the sake of simplicity, I call the coefficients contained in β *linear weights*.

3.3.2 Results

In the following, the results of data preprocessing and drug-exposure model training are presented. First, I describe the results of the data preprocessing steps which yield the datasets on which drug-exposure models are trained. Then, I describe the resulting datasets. Finally, I present the performance of the drug-exposure models in predicting drug exposure, drug resistance, and therapeutic success. These performances are compared to those of RFs predicted by geno2pheno[resistance].

Table 3.1: Number of Nucleotide Sequences by Subtype and Dataset. Nucleotide sequences in the PRRT, IN, and HIVdbExposure datasets were subtyped with the COMET subtyping tool. Sequence counts for the ten most frequent subtypes are tabulated below.

Subtype	PRRT	IN	HIVdbExposure
B	42,634	2,721	1,377
C	6,243	1293	1
A1	3,704	166	1
G	3,223	270	0
02_AG	3,010	66	1
01_AE	4,275	596	1
D	1,169	53	1
F1	971	69	0
06_cpx	312	89	0
07_BC	651	4	0
Other	4,112	196	2
Total	70,304	5,523	1,384

PREPROCESSING OF DRUG-EXPOSURE DATASETS AND RESULTING COMPOSITION

Prior to alignment, 48,666 EIDB sequences were extracted from the database. The alignment procedure assigned 38,754 sequences to the PRRT dataset and 6,214 sequences to the IN dataset. PRRT and IN were further complemented with 36,774 and 5,262 sequences, respectively, from LANLSD. The number of sequences in PRRT was reduced to 75,239 sequences after excluding sequences with more than 10% undetermined residues. After removal of duplicate sequences, PRRT included a total of 70,304 sequences (approximately 93% of the initially included sequences). After excluding sequences with more than 10% undetermined residues, the number of sequences in IN was reduced to 7,076. After duplicate removal, 5,523 sequences (approximately 48%) were left. Sequences in PRRT, IN and HIVdbExposure were subtyped. The number of sequences per subtype for these datasets can be seen in Table 3.1. Table 3.2 shows the number of sequences in the datasets D_{PRRT}, D_{IN}, T_{PRRT}, T_{IN}, TP_{PRRT} and, HIVdbExposure by drug exposure. In D_{PRRT}, 37,557 sequences are therapy naïve, of which 3,757 (10.0%) present TDR. A total of 1,917 sequences in D_{IN} are therapy naïve, of which 97 (approximately 5%) present TDR. T_{PRRT} contains 2,056 therapy-naïve sequences, among which 219 (approximately 11%) present TDR, while T_{IN} contains 154 therapy naïve sequences with 5 (3%) presenting TDR.

PREPROCESSING OF THERAPY-CHANGE-EPISODE DATASETS AND RESULTING COMPOSITION

Application of the definition of the EuResist Standard Datum to clinical HIV-1 data yielded the EuResistTE ($n = 9,201$) and the HIVdbTCE ($n = 1,000$) datasets. Figure 3.2(a) depicts the most frequent therapies in the EuResistTE, while Figure 3.2(b) does so for the TEs in HIVdbTCE. The construction of the D_{PRRT} and D_{IN} datasets induced the $D_{EuResistTE}$ ($n = 7,551$) and $T_{EuResistTE}$ ($n = 1,650$) datasets; $T_{EuResistTE}$ includes 619 TEs whose baseline sequences are not included in PRRT. $D_{EuResistTE}$ contains 4,202 (55.6%) first-line therapies among which 415 present TDR in their baseline sequences (9.9%), while $T_{EuResistTE}$ contains 313 first-line therapies (19.0%) among which 44 (14.1%) present TDR in their baseline sequences. Among the NAS computed for the therapies in $T_{EuResistTE}$, 1,017 were censored and 550 uncensored. In HIVdbTCE, 77 NAS were censored, while 346 were uncensored. Figure 3.3(a) displays a histogram of the NAS for the therapies in

Table 3.2: Number of Sequences by Dataset and Drug Exposure. The numbers of sequences by drug exposure for the development and test datasets are tabulated below. Columns including the abbreviation Comp. in their headers indicate the numbers of sequences from a certain dataset and with a certain drug exposure whose complete drug exposure is known. The complete drug exposure for all sequences from the HIVdbExposure dataset is available. 3TC: lamivudine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, FTC: emtricitabine, TDF: tenofovir diisoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, APV: amprenavir, ATV: atazanavir, DRV: darunavir, FPV: fosamprenavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	D_{PRRT}	D_{PRRT} Comp.	D_{IN}	D_{IN} Comp.	T_{PRRT}	T_{PRRT} Comp.	T_{IN}	T_{IN} Comp.	TP_{PRRT}	TP_{PRRT} Comp.	HIVdbExposure
3TC	20,730	13,416	525	394	4,191	2,543	262	151	390	70	0
ABC	7,482	4,560	295	229	1,839	1,028	164	103	163	30	301
AZT	18,542	12,184	441	336	3,895	2,405	222	135	372	68	1,075
d4T	13,335	8,079	259	197	2,956	1,764	141	80	250	31	998
ddC	4,007	2,341	57	45	1,114	750	52	40	71	6	297
ddI	12,113	7,398	227	173	2,725	1,657	123	73	197	23	722
FTC	4,580	3,258	359	266	900	595	162	112	52	8	59
TDF	9,546	6,058	479	356	1,933	1,192	211	130	119	13	219
DLV	118	56	5	2	96	58	26	18	6	1	73
EFV	9,673	6,228	301	238	2,168	1,310	168	110	194	35	400
ETR	255	169	62	51	145	94	70	48	4	3	1
NVP	8,405	5,044	232	178	1,836	1,054	123	74	179	31	508
RPV	5	4	2	1	0	0	0	0	0	0	0
APV	1,240	615	48	37	463	216	41	31	38	3	192
ATV	3,444	2,293	230	166	833	510	131	79	39	6	52
DRV	916	587	152	95	328	200	111	68	16	6	4
FPV	1,028	621	79	58	381	211	52	26	28	5	20
IDV	9,466	5,965	184	150	2,134	1,433	112	84	144	20	737
LPV	8,516	5,293	332	244	2,156	1,315	180	104	142	22	147
NFV	7,540	4,669	137	104	1,698	1,018	101	68	113	17	706
SQV	6,187	3,646	166	125	1,638	951	91	41	136	21	428
TPV	643	345	71	58	246	153	62	39	9	0	5
EVG	10	1	3	2	0	0	0	0	0	0	0
RAL	650	448	223	171	251	156	116	80	7	3	0
Naïve	37,577	37,577	1,917	1,917	2,056	2,056	154	154	3	3	0
Total	61,163	53,098	2,579	2,408	6,641	4,862	444	326	441	84	1,384

(a) EuResistTE
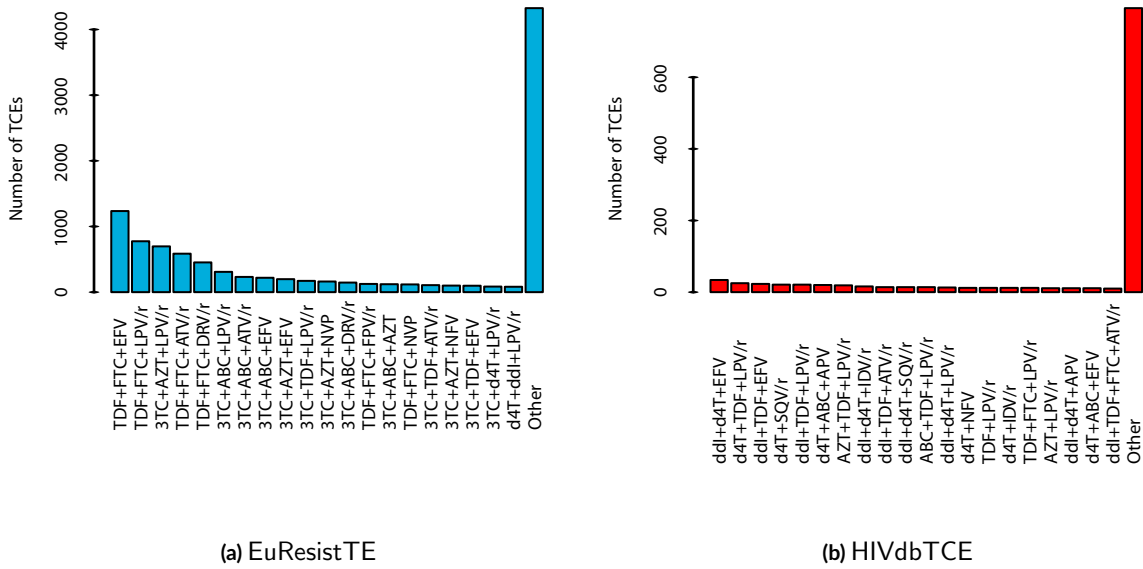
(b) HIVdbTCE

Figure 3.2: Drug-Combination Counts for Therapies in EuResistTE and HIVdbTCE. The counts of the 20 most-frequent drug combinations in EuResistTE (a) and HIVdbTCE (b) datasets are displayed above.

$T_{\text{EuResistTE}}$, while Figure 3.3(b) shows a histogram for the NAS in HIVdbTCE. Drug-wise and overall therapy-success proportions for the therapies in $D_{\text{EuResistTE}}$, $D_{\text{EuResistTE}} \cup D_{\text{TFE}}$, $T_{\text{EuResistTE}}$, and HIVdbTCE can be found in Table 3.3.

Preprocessing of Datasets of Genotype-Phenotype Pairs and Resulting Composition

The Pheno dataset was split into the D_{Pheno} and T_{Pheno} datasets. Subsequently, the RF cutoffs one and ten were applied to the GPPs in the datasets in order to classify them as resistant or susceptible. The resulting compositions of D_{Pheno} and T_{Pheno} are displayed in Tables 3.4 and 3.5, respectively.

Cross-Validation Performances

After creation of the Exposure_{drug}, $\text{Exposure}_{\text{naivePRRT}}$, and $\text{ExposurePheno}_{drug}$, cross-validation sets, ten repetitions of a five-fold cross validation were performed with each cross-validation set and encoding for a series of values for the SVM γ parameter (Sections 2.2.2 and 3.3.1). One value of the γ parameter was chosen for each cross-validation set, encoding, and drug. The mean drug-wise cross-validation performances (AUC) for the chosen values of γ ranged between 0.61 and 0.99. Individual performances are displayed in Tables 3.6 and 3.7. Note that two values for the mean performance across drugs are provided in the tables. The first mean value, to which the term *common drugs* is appended, only considers performances for drugs for which both Exposure and ExposurePheno models were trained. The second mean value, to which the term *all drugs* is appended, considers all drugs for which predictions were available. Thus, means for *common drugs* allow for performance comparison between Exposure and ExposurePheno models. The maximal model-wise mean performance was higher for models trained on ExposurePheno cross-validation sets ($\mu = 0.82$; $\sigma = 0.05$), than for models trained on Exposure cross-validation sets ($\mu = 0.79$; $\sigma = 0.07$).

Table 3.3: Drug-Wise and Overall Therapy-Success Proportions. Therapy-success proportions for therapy episodes including a certain drug are tabulated below for $D_{\mathsf{EuResistTE}}$, $D_{\mathsf{EuResistTE}} \cup D_{\mathsf{TFE}}$, $T_{\mathsf{EuResistTE}}$, and $\mathsf{HIVdbTCE}$. Overall therapy-success rates for each of the datasets can be found below as well. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir.

	$D_{\mathsf{EuResistTE}}$	$D_{\mathsf{EuResistTE}} \cup D_{\mathsf{TFE}}$	$T_{\mathsf{EuResistTE}}$	$\mathsf{HIVdbTCE}$
3FTC	0.84	0.71	0.79	0.65
ABC	0.81	0.65	0.71	0.59
AZT	0.82	0.64	0.77	0.71
d4T	0.65	0.44	0.58	0.56
ddI	0.65	0.48	0.63	0.59
TDF	0.84	0.74	0.78	0.71
EFV	0.89	0.75	0.83	0.66
ETR	0.81	0.75	0.67	1
NVP	0.77	0.50	0.73	0.67
AFPV	0.73	0.62	0.62	0.43
ATV	0.83	0.74	0.78	0.80
DRV	0.88	0.84	0.80	1
IDV	0.68	0.48	0.57	0.55
LPV	0.82	0.74	0.75	0.66
NFV	0.65	0.41	0.73	0.48
SQV	0.65	0.52	0.66	0.58
TPV	0.51	0.38	0.56	1
Overall	0.81	0.67	0.76	0.62

Table 3.4: Number of Phenotypes by Drug in the D_{Pheno} dataset. The numbers of phenotypes by drug in the D_{Pheno} dataset are tabulated below. Phenotypes were measured with the Antivirogram® or PhenoSense® assays. Drug-independent resistance-factor cutoffs one and ten were used for categorizing phenotypes into susceptible and resistant. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	Antivirogram	PhenoSense	Susceptible	Resistant	Total
3TC	905	1,546	346	1,362	2,451
ABC	840	1,473	531	186	2,313
AZT	855	1,567	801	773	2,422
d4T	889	1,573	1,031	60	2,462
ddC	821	451	371	47	1,272
ddI	891	1,575	654	59	2,466
TDF	633	1,234	850	33	1,867
DLV	1,016	1,638	794	1,091	2,654
EFV	1,106	1,652	924	1,127	2,758
ETR	363	476	304	156	839
NVP	1,170	1,653	772	1,447	2,823
RPV	91	176	62	75	267
ATV	774	1,134	401	978	1,908
DRV	282	629	400	178	911
FPV	1,088	1,695	917	859	2,783
IDV	1,151	1,734	782	1,229	2,885
LPV	1,040	1,468	665	1,279	2,508
NFV	1,185	1,780	483	1,584	2,965
SQV	1,181	1,741	985	1,039	2,922
TPV	742	854	584	191	1,596
EVG	97	598	112	137	695
RAL	97	630	336	148	727

Table 3.5: Number of Phenotypes by Drug in the $\mathsf{T}_{\text{Pheno}}$ dataset. The numbers of phenotypes by drug in the $\mathsf{T}_{\text{Pheno}}$ dataset are tabulated below. Phenotypes were measured with the Antivirogram® or PhenoSense® assays. Drug-independent resistance-factor cutoffs one and ten were used for categorizing phenotypes into susceptible and resistant. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	Antivirogram	PhenoSense	Susceptible	Resistant	Total
3TC	115	166	37	158	281
ABC	107	166	60	25	273
AZT	107	165	92	88	272
d4T	110	168	122	6	278
ddC	105	46	38	5	151
ddI	111	168	72	7	279
TDF	87	132	87	4	219
DLV	126	169	81	125	295
EFV	141	171	105	136	312
ETR	43	52	36	15	95
NVP	146	175	82	170	321
RPV	14	21	13	10	35
ATV	85	131	42	115	216
DRV	22	79	50	20	101
FPV	110	193	88	105	303
IDV	125	194	76	142	319
LPV	113	172	76	151	285
NFV	127	199	48	189	326
SQV	129	195	105	119	324
TPV	80	106	56	22	186
EVG	17	61	9	11	78
RAL	17	65	36	21	82

Table 3.6: Mean Cross-Validation Performance for Drug-Exposure Prediction, Exposure Models. For each group of cross-validation sets, encoding, and tested γ parameter, five repetitions of a 10-fold cross validation were performed. Each performance figure shown below corresponds to the selected γ parameter. Performances were calculated in terms of AUC, and performances were averaged for drugs common to all interpretation systems and for all models. The highest performance in each row is underlined. Standard deviations are shown in parentheses. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddl: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir, CD: drugs common to all models, AM: all models, NA: not available, SD: standard deviation.

	$\text{Exposure}_{\text{onlyIASPos}}$	$\text{Exposure}_{\text{noIASPos}}$	$\text{Exposure}_{\text{full}}$
3FTC	0.83 (0.01)	0.79 (0.01)	0.88 (0)
ABC	0.71 (0.01)	0.7 (0.01)	0.74 (0.01)
AZT	0.8 (0.01)	0.78 (0.01)	0.86 (0.01)
d4T	0.79 (0.01)	0.78 (0.01)	0.83 (0.01)
ddI	0.79 (0.01)	0.77 (0.01)	0.83 (0.01)
ddC	0.81 (0.01)	0.81 (0.01)	0.85 (0.01)
TDF	0.69 (0.01)	0.66 (0.01)	0.72 (0.01)
EFV	0.75 (0.01)	0.68 (0.01)	0.78 (0.01)
ETR	0.76 (0.07)	0.77 (0.06)	0.8 (0.05)
DLV	0.89 (0.06)	0.84 (0.08)	0.9 (0.06)
NVP	0.78 (0.01)	0.73 (0.01)	0.81 (0.01)
AFPV	0.82 (0.02)	0.74 (0.02)	0.83 (0.02)
ATV	0.65 (0.02)	0.61 (0.02)	0.67 (0.02)
DRV	0.65 (0.03)	0.65 (0.03)	0.68 (0.03)
IDV	0.76 (0.01)	0.69 (0.01)	0.78 (0.01)
LPV	0.69 (0.01)	0.65 (0.01)	0.71 (0.01)
NFV	0.77 (0.01)	0.68 (0.01)	0.79 (0.01)
SQV	0.81 (0.01)	0.72 (0.01)	0.83 (0.01)
TPV	0.88 (0.03)	0.8 (0.04)	0.89 (0.03)
RAL	0.65 (0.06)	0.64 (0.07)	0.67 (0.07)
Naïve PRRT	0.87 (0)	0.82 (0.01)	0.9 (0.00)
Mean CD (SD)	0.76 (0.06)	0.72 (0.06)	0.79 (0.07)
Mean AM (SD)	0.77 (0.07)	0.73 (0.07)	0.8 (0.07)

Table 3.7: Mean Cross-Validation Performance for Drug-Exposure Prediction, $\mathrm{ExposurePheno}$ **Models.** For each group of cross-validation sets, encoding and tested γ parameter, five repetitions of a 10-fold cross validation were performed. Each performance figure shown below corresponds to the selected γ parameter. Performances were calculated in terms of AUC, and performances were averaged for drugs common to all interpretation systems and for all models. The highest performance in each row is underlined. Standard deviations are shown in parentheses. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir, CD: drugs common to all models, AM: all models, NA: not available, SD: standard deviation.

	$\mathrm{ExposurePheno_{onlyIASPos}}$	$\mathrm{ExposurePheno_{noIASPos}}$	$\mathrm{ExposurePheno_{full}}$
3FTC	0.84 (0.01)	0.79 (0.01)	0.88 (0.01)
ABC	0.73 (0.01)	0.71 (0.01)	0.76 (0.01)
AZT	0.81 (0.01)	0.78 (0.01)	0.85 (0.01)
d4T	0.79 (0.01)	0.78 (0.01)	0.83 (0.01)
ddI	0.79 (0.01)	0.77 (0.01)	0.83 (0.01)
ddC	0.8 (0.02)	0.8 (0.01)	0.84 (0.01)
TDF	0.7 (0.01)	0.69 (0.01)	0.73 (0.01)
EFV	0.77 (0.01)	0.68 (0.01)	0.8 (0.01)
ETR	0.87 (0.04)	0.72 (0.05)	0.88 (0.04)
DLV	0.95 (0.01)	0.7 (0.03)	0.96 (0.01)
NVP	0.79 (0.01)	0.72 (0.01)	0.82 (0.01)
RPV	0.95 (0.06)	0.67 (0.15)	0.94 (0.06)
AFPV	0.86 (0.01)	0.75 (0.02)	0.87 (0.01)
ATV	0.71 (0.02)	0.67 (0.02)	0.75 (0.01)
DRV	0.76 (0.03)	0.71 (0.03)	0.78 (0.03)
IDV	0.77 (0.01)	0.7 (0.01)	0.8 (0.01)
LPV	0.73 (0.01)	0.66 (0.01)	0.75 (0.01)
NFV	0.8 (0.01)	0.7 (0.01)	0.82 (0.01)
SQV	0.82 (0.01)	0.72 (0.02)	0.84 (0.01)
TPV	0.91 (0.02)	0.81 (0.03)	0.92 (0.02)
RAL	0.86 (0.03)	0.78 (0.05)	0.89 (0.03)
EVG	0.98 (0.02)	0.8 (0.08)	0.99 (0.02)
Mean CD (SD)	0.79 (0.06)	0.73 (0.04)	0.82 (0.05)
Mean AM (SD)	0.82 (0.08)	0.74 (0.05)	0.85 (0.07)

(a) EuResistTE
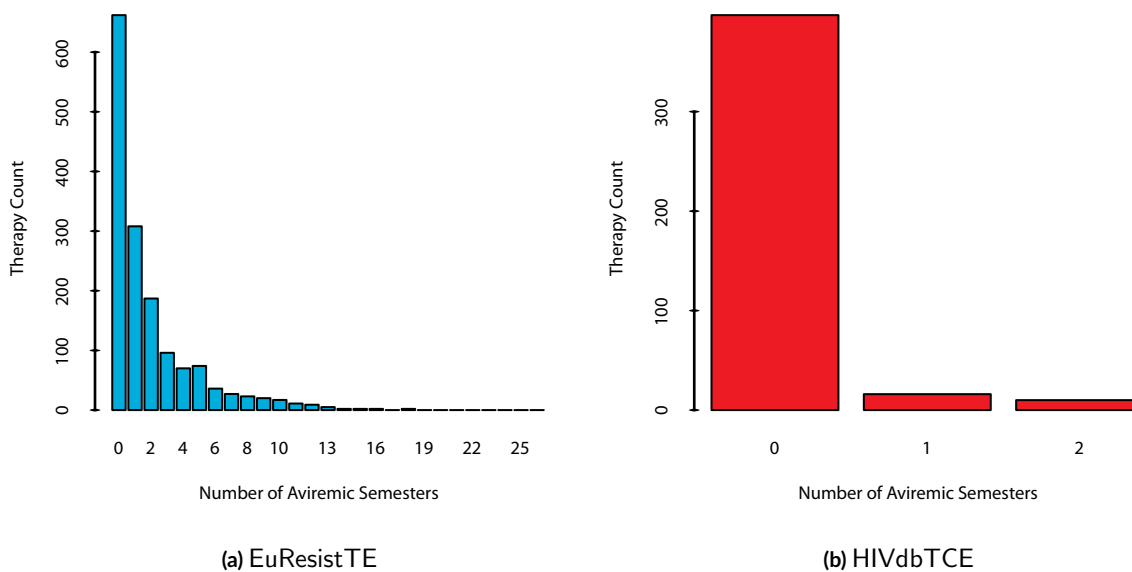
(b) HIVdbTCE

Figure 3.3: Histogram of the Numbers of Aviremic Semesters. For each therapy-change episode in $\mathsf{T_{EuResistTE}}$ and $\mathsf{HIVdbTCE}$, the number of therapy semesters with an average viral load below 50 copies per milliliter of blood serum was quantified (aviremic semesters). Database records on ongoing therapies, therapy semesters without viral load measurements, and therapy changes despite viral suppression result in a right-censored quantification of the number of aviremic semesters. Among the numbers of aviremic semesters computed for the therapies in $\mathsf{T_{EuResistTE}}$ (a), 1,017 were censored and 550 uncensored. A total of 77 $\mathsf{HIVdbTCE}$ (b) numbers of aviremic semesters were censored, while 346 were uncensored.

Assessment and Comparison of Performance Using the Test Sets

The performance of the final drug-exposure models was assessed and compared to that of geno2pheno$_{[resistance]}$. After the application of cutoffs for discretizing DES, performance was also compared to other popular drug-resistance interpretation systems (Section 3.4). Among the 7,275 PR and RT nucleotide sequences that were available for testing, 23 (<0.01%) were not processed by either geno2pheno$_{[resistance]}$ or another drug-resistance interpretation system to which DES are compared (*Interpretation Systems for Performance Comparison* in Section 3.4.1), due to low similarity to the reference sequence. For the sake of performance comparison, these sequences were excluded. In the following, mean model-group-wise performances are stated. In order to be able to compare the different models, these means were calculated only with the performances of the drugs that are common to Exposure and ExposurePheno models, as well as to the drug-resistance interpretation systems that are used for performance comparison in this chapter. Note that in the tables found in the following, two values for the mean performance across drugs are provided. The first mean value, to which the term *common drugs* is appended, only considers performances for drugs common to all drug-exposure model groups and drug-resistance interpretation systems used for performance comparison. The second mean value, to which the term *all drugs* is appended, considers all drugs for which predictions were available. Thus, means for *common drugs* allow for performance comparison between different model groups and drug-resistance interpretation systems. p-values were calculated with a two-sided Wilcoxon signed-rank test.

Table 3.8: Drug-Exposure Prediction Performance (AUC) on T_{PRRT} **and** T_{IN} **for Continuous Drug-Exposure Scores,** Exposure Models. Drug-exposure prediction performance on sequences in the T_{PRRT} and T_{IN} datasets was calculated for each final drug-exposure-score model group. The highest performance in each row is underlined. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, CD: drugs common to all models, AM: all models, NA: not available, SD: standard deviation.

	Exposure$_{full}$	Exposure$_{noIASPos}$	Exposure$_{onlyIASPos}$
3FTC	0.85	0.77	0.82
ABC	0.73	0.69	0.72
AZT	0.84	0.76	0.81
d4T	0.86	0.79	0.82
ddC	0.84	0.79	0.82
ddI	0.85	0.78	0.82
TDF	0.71	0.66	0.69
DLV	0.86	0.62	0.88
EFV	0.76	0.69	0.74
ETR	0.75	0.65	0.75
NVP	0.77	0.7	0.75
AFPV	0.8	0.69	0.8
ATV	0.66	0.6	0.65
DRV	0.69	0.63	0.67
IDV	0.79	0.7	0.78
LPV	0.75	0.66	0.73
NFV	0.76	0.66	0.74
SQV	0.82	0.71	0.81
TPV	0.85	0.81	0.86
RAL	0.62	0.55	0.69
Naïve	0.89	0.82	0.87
Mean CD (SD)	0.78 (0.06)	0.7 (0.06)	0.76 (0.06)
Mean AM (SD)	0.78 (0.07)	0.7 (0.07)	0.77 (0.07)

Performance of Drug-Exposure and Drug-Resistance Prediction Using the Test Sets

DES performances when predicting drug-exposure on the T_{PRRT}, T_{IN}, and HIVdbExposure datasets can be seen in Tables 3.8, 3.9, and 3.10, while that of geno2pheno$_{[resistance]}$ is shown in Table 3.11. In the following, p-values quantify the difference in the AUC distributions between a given drug-exposure model group and geno2pheno$_{[resistance]}$. The best mean performance with lowest standard deviation (SD) on the T_{PRRT} dataset could be attained by the Exposure$_{full}$ models ($\mu = 0.78$; $\sigma = 0.06$), while the performance of geno2pheno$_{[resistance]}$ was lower ($\mu = 0.71$; $\sigma = 0.07$; $p < 10^{-4}$). The model trained with the ExposurePheno$_{RAL}$ cross-validation set and the V_{full} encoding showed the best performance on the T_{IN} dataset (AUC = 0.71). DES obtained from Exposure$_{full}$ models performed best in discriminating therapy-naïve sequences from therapy-experienced sequences in the T_{PRRT} dataset (AUC = 0.89). On the HIVdbExposure dataset, the best mean performance with lowest SD could be attained with the Exposure$_{full}$ models ($\mu = 0.76$; $\sigma = 0.09$), while geno2pheno$_{[resistance]}$ achieved a lower mean performance ($\mu = 0.74$; $\sigma = 0.14$; $p = $

Table 3.9: Drug-Exposure Prediction Performance (AUC) on T_{PRRT} and T_{IN} for Continuous Drug-Exposure Scores, ExposurePheno Models. Drug-exposure prediction performance on sequences in the T_{PRRT} and T_{IN} datasets was calculated for each final drug-exposure-score model group. The highest performance in each row is underlined. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, CD: drugs common to all models, AM: all models, NA: not available, SD: standard deviation.

	$ExposurePheno_{full}$	$ExposurePheno_{noIASPos}$	$ExposurePheno_{onlyIASPos}$
3FTC	0.85	0.77	0.82
ABC	0.73	0.69	0.72
AZT	0.84	0.76	0.81
d4T	0.85	0.79	0.82
ddC	0.84	0.79	0.81
ddI	0.85	0.78	0.82
TDF	0.7	0.65	0.69
DLV	0.83	0.46	0.83
EFV	0.76	0.69	0.73
ETR	0.74	0.62	0.72
NVP	0.76	0.69	0.74
AFPV	0.8	0.72	0.79
ATV	0.63	0.61	0.62
DRV	0.69	0.67	0.66
IDV	0.79	0.71	0.76
LPV	0.74	0.68	0.72
NFV	0.76	0.68	0.74
SQV	0.82	0.72	0.8
TPV	0.85	0.78	0.84
RAL	0.71	0.55	0.69
Mean CD (SD)	0.77 (0.06)	0.71 (0.06)	0.75 (0.06)
Mean AM (SD)	0.78 (0.06)	0.69 (0.09)	0.76 (0.06)

Table 3.10: Drug-Exposure Prediction Performance (AUC) on HIVdbExposure **for Continuous Drug-Exposure Scores.** Drug-exposure prediction performance on sequences in the HIVdbExposure dataset was calculated for each final drug-exposure-score model group. The highest performance (with the lowest standard deviation) in each row is underlined. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, SD: standard deviation.

	Exposure$_{full}$	Exposure$_{noIASPos}$	Exposure$_{onlyIASPos}$
3FTC	0.74	0.56	0.72
ABC	0.71	0.61	0.7
AZT	0.62	0.61	0.62
d4T	0.64	0.62	0.63
ddI	0.72	0.63	0.71
TDF	0.66	0.55	0.64
EFV	0.84	0.56	0.8
ETR	0.97	0.83	1
NVP	0.78	0.59	0.77
AFPV	0.8	0.54	0.81
ATV	0.68	0.5	0.69
DRV	0.83	0.79	0.95
IDV	0.77	0.6	0.77
LPV	0.66	0.54	0.66
NFV	0.76	0.55	0.76
SQV	0.76	0.56	0.77
TPV	0.9	0.79	0.91
Mean (SD)	0.76 (0.09)	0.61 (0.1)	0.76 (0.11)
	ExposurePheno$_{full}$	ExposurePheno$_{noIASPos}$	ExposurePheno$_{onlyIASPos}$
3FTC	0.75	0.56	0.73
ABC	0.71	0.62	0.7
AZT	0.63	0.62	0.62
d4T	0.65	0.62	0.64
ddI	0.72	0.63	0.71
TDF	0.67	0.56	0.64
EFV	0.85	0.56	0.85
ETR	0.96	0.72	0.97
NVP	0.79	0.61	0.76
AFPV	0.81	0.59	0.82
ATV	0.62	0.57	0.59
DRV	0.94	0.73	0.91
IDV	0.78	0.6	0.77
LPV	0.66	0.57	0.66
NFV	0.76	0.55	0.76
SQV	0.77	0.56	0.77
TPV	0.94	0.73	0.94
Mean (SD)	0.76 (0.11)	0.61 (0.06)	0.75 (0.11)

Table 3.11: Drug-Exposure Prediction Performance (AUC) on T_{PRRT} **and** HIVdbExposure **for Predicted Resistance Factors**. Drug-exposure prediction performance on sequences in the T_{PRRT} and HIVdbExposure datasets was calculated with resistance factors predicted by $geno2pheno_{[resistance]}$. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, SD: standard deviation.

	T_{PRRT}	HIVdbExposure
3FTC	0.79	0.76
ABC	0.69	0.65
AZT	0.72	0.69
d4T	0.74	0.64
ddI	0.78	0.55
TDF	0.63	0.52
EFV	0.7	0.83
ETR	0.58	0.98
NVP	0.7	0.74
AFPV	0.76	0.8
ATV	0.6	0.56
DRV	0.68	0.95
IDV	0.73	0.77
LPV	0.72	0.64
NFV	0.7	0.74
SQV	0.75	0.76
TPV	0.83	0.96
Mean (SD)	0.71 (0.07)	0.74 (0.14)

0.46). Table 3.12 shows the performance of the drug-exposure-score models in predicting drug exposure on sequences derived during therapy pauses ($\mathsf{TP_{PRRT}}$ dataset), while the performance of geno2pheno$_{[resistance]}$ on the same dataset is shown in Table 3.13. The best mean performances on $\mathsf{TP_{PRRT}}$ could be attained with models trained with the $\mathsf{V_{full}}$ encoding ($\mu = 0.61$; $\sigma = 0.08$), while geno2pheno$_{[resistance]}$ displayed a lower performance ($\mu = 0.59$; $\sigma = 0.10$; $p = 0.2247$). Tables 3.14 and 3.15 show the correlation of the continuous DES with the logarithmized resistance factors from the Pheno$_\mathsf{T}$ dataset. When testing the models with Antivirogram® resistance factors, ExposurePheno$_{full}$ models show the highest mean correlation ($\mu = 0.46$; $\sigma = 0.20$). The best mean correlation with experimentally measured PhenoSense® resistance factors was attained by the ExposurePheno$_{full}$ and ExposurePheno$_{onlyIASPos}$ models ($\mu = 0.51$; $\sigma = 0.17$).

Performance of Prediction of Therapy Success

Performance in predicting therapy success on the $\mathsf{T_{EuResistTE}}$, $\mathsf{T_{EuResistTETP}}$, and HIVdbTCE datasets are displayed in Table 3.16. Performance in predicting short-term therapeutic success was assessed using dichotomous therapy-success labels and the AUC. Performance in predicting long-term therapeutic success was assessed using the NAS and Harrel's concordance index. On the $\mathsf{T_{EuResistTE}}$ dataset, DES models with the $\mathsf{V_{full}}$ and $\mathsf{V_{onlyIASPos}}$ encodings could attain the highest AUC (AUC $= 0.71$), while DES models with the $\mathsf{V_{full}}$ encoding could attain the best concordance ($C = 0.65$). The highest AUC on the $\mathsf{T_{EuResistTETP}}$ dataset was achieved by the Exposure$_{onlyIASPos}$ models (AUC $= 0.74$), while the highest concordance was achieved by the ExposurePheno$_{full}$ models ($C = 0.63$). geno2pheno$_{[resistance]}$ and ExposurePheno$_{onlyIASPos}$ models performed best in terms of AUC on the HIVdbTCE dataset (AUC $= 0.64$), while the best concordance was achieved by geno2pheno$_{[resistance]}$ ($C = 0.61$).

Linear Weights for Drug-Exposure-Score Models

Linear SVCs used for obtaining DES were represented as linear functions. With these functions, the calculation of the DES amounts to the addition of the weights corresponding to the substitutions, insertions and deletions in a genotype. An excerpt of the ten smallest and ten largest linear DES weights for selected drugs from ExposurePheno$_{full}$ models can be found in Table 3.17.

3.3.3 Discussion

DES models constitute a data-driven interpretation system for HIV-1 PR, RT, and IN sequences. The interpretations provided by this system can be used to address three questions: (1) Was a sequence exposed to a certain drug? (2) Is the HIV-1 variant from which the sequence was derived resistant against a certain drug? and (3) Will a certain drug be useful as a component of a therapy against an HIV-1 variant? In the following, I will refer to these three questions by the number I assigned to them above.

Different versions of DES models were trained and tested. Specifically, I assessed the value of including GPPs into the models' training sets (ExposurePheno models), as opposed to training the models solely on genotypes with drug-exposure information (Exposure models). Furthermore, I used three different encodings for the genotype sequences used to train the models: $\mathsf{V_{full}}$ considers all amino acids, $\mathsf{V_{noIASPos}}$ disregards amino-acid positions at which drug-resistance mutations in the IAS 2013 list [296] occur, and $\mathsf{V_{onlyIASPos}}$ only considers such amino-acid positions. While Exposure DES models show a comparatively high performance in predicting

Table 3.12: Drug-Exposure Prediction Performance (AUC) T_{TPPRRT} **for Continuous Drug-Exposure Scores.** Drug-exposure prediction performance on T_{TPPRRT} (therapy-pause sequences) dataset was calculated for each final drug-exposure-score model group. The highest performance in each row is underlined. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, CD: drugs common to all models, AM: all models, NA: not available, SD: standard deviation.

	$Exposure_{full}$	$Exposure_{noIASPos}$	$Exposure_{onlyIASPos}$
3FTC	0.45	0.47	0.41
ABC	0.61	0.59	0.61
AZT	0.63	0.65	0.63
d4T	0.66	0.64	0.61
ddI	0.68	0.61	0.67
TDF	0.52	0.47	0.52
EFV	0.54	0.54	0.56
ETR	0.79	0.65	0.73
NVP	0.58	0.53	0.59
AFPV	0.6	0.62	0.61
ATV	0.55	0.5	0.51
DRV	0.53	0.53	0.53
IDV	0.64	0.61	0.62
LPV	0.58	0.53	0.62
NFV	0.63	0.66	0.59
SQV	0.63	0.65	0.6
TPV	0.73	0.74	0.81
Mean (SD)	0.61 (0.08)	0.59 (0.08)	0.6 (0.09)
	$ExposurePheno_{full}$	$ExposurePheno_{noIASPos}$	$ExposurePheno_{onlyIASPos}$
3FTC	0.44	0.45	0.4
ABC	0.61	0.6	0.6
AZT	0.61	0.62	0.56
d4T	0.66	0.63	0.61
ddI	0.69	0.62	0.66
TDF	0.5	0.44	0.53
EFV	0.51	0.54	0.53
ETR	0.74	0.85	0.71
NVP	0.57	0.53	0.61
AFPV	0.61	0.61	0.67
ATV	0.54	0.53	0.54
DRV	0.63	0.63	0.42
IDV	0.62	0.62	0.62
LPV	0.62	0.52	0.64
NFV	0.62	0.63	0.61
SQV	0.64	0.64	0.6
TPV	0.74	0.72	0.8
Mean (SD)	0.61 (0.08)	0.6 (0.1)	0.59 (0.1)

Table 3.13: Drug-Exposure Prediction Performance (AUC) $\mathsf{T_{TPPRRT}}$ **for Predicted Resistance Factors.** Drug-exposure prediction performance on $\mathsf{T_{TPPRRT}}$ (therapy-pause sequences) dataset was calculated for resistance factors predicted with geno2pheno[resistance]. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, SD: standard deviation.

3FTC	ABC	AZT	d4T	ddI	TDF	EFV	ETR	NVP
0.46	0.61	0.48	0.54	0.63	0.55	0.6	0.79	0.55
AFPV	ATV	DRV	IDV	LPV	NFV	SQV	TPV	Mean (SD)
0.58	0.49	0.48	0.61	0.58	0.62	0.65	0.82	0.59 (0.1)

drug exposure and therapeutic success (questions (1) and (3)), their correlation with GPP (question (2)) is inferior to that of ExposurePheno models. The training sets of ExposurePheno DES models included genotypes from GPPs with RFs less or equal than one or greater or equal than ten. For *training* these models, clinically relevant categorization of GPPs into susceptible and resistant is not intended, since drug susceptibility may be given even if the virus has mutated due to the drugs. Instead, I aimed at discriminating fully susceptible GPPs from those that have developed at least some extent of resistance (questions (1), (2), and (3)). The inclusion of genotypes from GPP allowed for the training of models for two additional drugs (EVG and RPV). When comparing the performance of the different amino-acid encodings I tested, it can be seen that the $\mathsf{V_{noIASPos}}$ encoding always attains the worst performance. However, the performance of models trained with this encoding is well above that of a random prediction. Thus, substitutions occurring at other protein residues than those at which IAS drug-resistance mutations can be present are informative of drug exposure (question (1)), drug resistance (question (2)), and therapeutic success (question (3)). Models trained with the $\mathsf{V_{onlyIASPos}}$ encoding display a performance similar to that of models trained with the $\mathsf{V_{full}}$ encoding. However, the performance of $\mathsf{V_{full}}$ models was slightly better in many cases [†]. Therefore, I conclude that the additional information contained in $\mathsf{V_{full}}$, as compared to the $\mathsf{V_{onlyIASPos}}$, is highly similar to the information already contained in $\mathsf{V_{onlyIASPos}}$. Still, this additional information can be used to attain a slight increase in performance. The use of expert-derived mutation tables for crafting the $\mathsf{V_{onlyIASPos}}$ encoding represents a potential source of bias and could prevent the exploitation unestablished, yet predictive resistance patterns. However, using $\mathsf{V_{full}}$ instead of $\mathsf{V_{onlyIASPos}}$ makes the models less robust with respect to sequencing errors (Section 5.3). Nonetheless, I think that the benefits of the $\mathsf{V_{full}}$ encoding (higher performance, decreased bias) outweigh the disadvantage of decreased robustness to sequencing errors. For this reason, I prefer ExposurePheno$_{\mathsf{full}}$ models over all other models.

Correlation of ExposurePheno DES with log RFs (question (3)) is weak to strong, depending on the drug in question (Tables 3.14 and 3.15). The correlation is sufficient for predicting the SIR label of GPPs discretized with clinically relevant cutoffs (Section 3.4; question (2)). Furthermore, it should be taken into account that the between-assay correlation of PhenoSense and Antivirogram is weak ($r = 0.36$) [282].

Performance assessment of ExposurePheno$_{\mathsf{full}}$ models shows their validity and utility. In comparison to geno2pheno[resistance], DES performance was superior or comparable, depending on the task and the dataset. DES models exhibit superior performance when predicting cumulative, long-term therapeutic success (C-index; Table 3.16; question (3)). The models present a strong difference in performance when predicting therapeutic

---

[†]Computing a p-value for the higher performance of $\mathsf{V_{full}}$ models when compared to $\mathsf{V_{onlyIASPos}}$ models is challenging due to the fact that the comparison refers the to tasks of predicting drug exposure, drug resistance, and therapeutic success at the same time. Note that different performance measures were used for assessing performance for these different tasks.

Table 3.14: Correlation of Continuous DES with Antivirogram Resistance Factors. With each final drug-exposure-score (DES) model group, DES were calculated for Antivirogram® genotype-phenotype pairs in T_{Pheno}. The correlations between DES and resistance factors are tabulated below. The highest correlation in each row is underlined. AD: All drugs; CD: Drugs Common to all DES model groups; NA: Not Available; SD: Standard Deviation. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, TR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	$Exposure_{full}$	$Exposure_{noIASPos}$	$Exposure_{onlyIASPos}$
3FTC	0.49	0.15	0.43
ABC	0.41	0.11	0.5
AZT	0.1	0.22	0.13
d4T	0.19	0.07	0.16
ddI	0.15	0.03	0.08
TDF	0.15	0.31	0.11
DLV	0.19	0.06	0.19
EFV	0.31	0.05	0.31
ETR	0.09	-0.13	-0.08
NVP	0.37	0.09	0.41
AFPV	0.64	0.27	0.64
ATV	0.4	0.22	0.39
DRV	0.55	0.12	0.4
IDV	0.55	0.41	0.62
LPV	0.6	0.42	0.61
NFV	0.09	0.06	0.08
SQV	0.59	0.32	0.61
TPV	0.16	0.17	0.12
RAL	0.32	0.21	0.75
Mean CD (SD)	0.34 (0.2)	0.17 (0.14)	0.35 (0.24)
Mean AD (SD)	0.33 (0.19)	0.17 (0.14)	0.34 (0.24)
	$ExposurePheno_{full}$	$ExposurePheno_{noIASPos}$	$ExposurePheno_{onlyIASPos}$
3FTC	0.55	0.15	0.52
ABC	0.51	0.15	0.55
AZT	0.13	0.24	0.14
d4T	0.3	0.09	0.19
ddI	0.22	0.06	0.14
TDF	0.23	0.34	0.23
DLV	0.71	0.2	0.68
EFV	0.42	0.05	0.38
ETR	0.47	0.29	0.45
RPV	0.81	-0.31	0.8
NVP	0.53	0.11	0.51
AFPV	0.74	0.29	0.75
ATV	0.64	0.3	0.64
DRV	0.58	0.1	0.45
IDV	0.65	0.41	0.65
LPV	0.62	0.45	0.63
NFV	0.09	0.05	0.08
SQV	0.68	0.35	0.67
TPV	0.25	0.18	0.24
RAL	0.63	-0.07	0.67
EVG	0.52	0.18	0.52
Mean CD (SD)	0.46 (0.2)	0.2 (0.14)	0.44 (0.22)
Mean AD (SD)	0.49 (0.21)	0.17 (0.17)	0.47 (0.22)

Table 3.15: Correlation of Continuous DES with PhenoSense Resistance Factors. With each final drug-exposure-score (DES) model group, DES were calculated for PhenoSense® genotype-phenotype pairs in T_{Pheno}. The correlations between DES and resistance factors are tabulated below. The highest correlation in each row is underlined. AD: All drugs; CD: Drugs Common to all DES model groups; NA: Not Available; SD: Standard Deviation. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, TR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	$Exposure_{full}$	$Exposure_{noIASPos}$	$Exposure_{onlyIASPos}$
3FTC	_0.65_	0.27	0.58
ABC	0.65	0.48	_0.66_
AZT	0.17	_0.38_	0.17
d4T	_0.55_	0.51	_0.55_
ddI	_0.37_	0.24	0.32
TDF	0.1	0.07	_0.13_
DLV	0.05	-0.04	_0.07_
EFV	_0.56_	0.18	0.54
ETR	_0.33_	0.2	0.28
NVP	0.58	0.11	_0.62_
AFPV	_0.5_	0.18	0.48
ATV	_0.39_	_0.39_	0.35
DRV	_0.47_	0.27	0.44
IDV	_0.39_	0.14	0.38
LPV	_0.57_	0.26	0.56
NFV	0.45	0.18	_0.46_
SQV	0.44	0.13	_0.46_
TPV	0.12	0.04	_0.17_
RAL	0.14	-0.09	_0.7_
Mean CD (SD)	0.41 (0.18)	0.22 (0.15)	_0.44 (0.17)_
Mean AD (SD)	0.39 (0.19)	0.21 (0.16)	_0.42 (0.19)_
	$ExposurePheno_{full}$	$ExposurePheno_{noIASPos}$	$ExposurePheno_{onlyIASPos}$
3FTC	_0.69_	0.27	_0.69_
ABC	_0.68_	0.49	0.67
AZT	0.2	_0.42_	0.18
d4T	0.55	0.46	_0.56_
ddI	_0.41_	0.23	0.37
TDF	0.14	0.11	_0.17_
DLV	0.58	0.16	_0.6_
EFV	_0.72_	0.22	0.69
ETR	_0.58_	0.45	0.51
RPV	_0.62_	0.33	0.51
NVP	_0.79_	0.2	0.75
AFPV	0.52	0.17	_0.54_
ATV	0.48	0.43	_0.52_
DRV	0.58	0.28	_0.59_
IDV	_0.41_	0.17	0.39
LPV	0.55	0.3	_0.56_
NFV	_0.47_	0.22	_0.47_
SQV	_0.5_	0.28	_0.5_
TPV	_0.32_	0.06	0.3
RAL	0.58	0.38	_0.71_
EVG	0.45	0.1	_0.48_
Mean CD (SD)	_0.51 (0.17)_	0.29 (0.13)	_0.51 (0.17)_
Mean AD (SD)	_0.52 (0.16)_	0.27 (0.13)	0.51 (0.16)

Table 3.16: **Performance of Therapy-Success Prediction for Drug-Exposure Scores and Predicted Resistance Factors**. Performances for the prediction of therapy success with drug-exposure scores and resistance factors predicted with geno2pheno$_{[resistance]}$ were calculated on the $T_{EuResistTE}$, $T_{EuResistTETP}$, and HIVdbTCE datasets. Below, performances in terms of the area under the receiver-operating-characteristic curve (AUC) and Harrell's concordance index (C) are tabulated separately. The highest performance in each column is underlined.

	$T_{EuResistTE}$ (AUC)	$T_{EuResistTE}$ (C)	$T_{EuResistTETP}$ (AUC)
geno2pheno$_{[resistance]}$	0.68	0.59	0.66
Exposure$_{full}$	0.71	0.65	0.72
Exposure$_{noIASPos}$	0.62	0.58	0.56
Exposure$_{onlyIASPos}$	0.71	0.63	0.74
ExposurePheno$_{full}$	0.71	0.65	0.73
ExposurePheno$_{noIASPos}$	0.63	0.59	0.59
ExposurePheno$_{onlyIASPos}$	0.71	0.63	0.73

	$T_{EuResistTETP}$ (C)	HIVdbTCE (AUC)	HIVdbTCE (C)
geno2pheno$_{[resistance]}$	0.53	0.64	0.61
Exposure$_{full}$	0.62	0.62	0.56
Exposure$_{noIASPos}$	0.55	0.58	0.54
Exposure$_{onlyIASPos}$	0.61	0.63	0.57
ExposurePheno$_{full}$	0.63	0.63	0.56
ExposurePheno$_{noIASPos}$	0.57	0.58	0.55
ExposurePheno$_{onlyIASPos}$	0.62	0.64	0.56

success with baseline sequences obtained during therapy pauses. Thus, DES use can be recommended when prediction of drug exposure (question (1)), drug resistance (question (2)), or therapy success (question (3)) is required. I propose the use of DES in two applications. The first application is the drug-wise interpretation of the genotype with respect to drug exposure (question (1)) and drug resistance (question (2)). This application can be useful when optimizing a therapy for a patient by hand (question (3)), but also as a tool facilitating the evaluation of other studies concerning HIV-1. DES are correlated with drug exposure and drug resistance, with *very high* DES indicating both drug exposure and drug resistance. Depending on the concrete application, the use of DES cutoffs in order to translate them into clinically meaningful categories related to drug exposure and drug resistance may be required. Discrimination between drug exposure and drug resistance can be achieved by determining different sets of DES cutoffs for drug exposure and for drug resistance (Section 3.4). The second application that I propose is the use of DES as input features for training models that predict therapeutic success based on a genotype and a drug combination (Section 4.2; question (3)).

Table 3.17 shows an excerpt of the linear mutation weights from which ExposurePheno$_{full}$ DES are calculated. For a particular drug, the DES for an HIV-1 genotype amounts to the sum of the individual weights for the substitutions in the genotype and the model's intercept. DES for a particular drug exhibit a positive correlation with the probability of exposure to the drug and to the fold-change in the drug's IC50 value. Insights into the inference that DES perform in order to predict drug exposure (question (1)) and drug resistance (question (2)) can be obtained from Table 3.17. For instance, among the weights for 3TC and FTC with the largest values, weights scoring NNRTI drug-resistance mutations can be found. While there is no data supporting the contribution of NNRTI drug-resistance mutations to phenotypic resistance against 3TC or FTC, they present

Table 3.17: **Linear Weights for Selected DES Models**. Support vectors in selected $\mathsf{ExposurePheno_{full}}$ DES models translated into linear functions, yielding linear weights for each considered substitution. Below, the ten smallest and ten largest weights for selected DES models are displayed. 3FTC: lamivudine or emtricitabine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, RPV: rilpivirine, DRV: darunavir, EVG; elvitegravir

3FTC		TDF		EFV		RPV		DRV		EVG	
65K	-0.6	134G	-0.66	190G	-1.17	181Y	-1.15	43K	-0.49	148Q	-0.77
185D	-0.57	113N	-0.61	188Y	-0.89	101K	-0.84	84I	-0.38	155N	-0.7
211D	-0.51	210Y	-0.48	100L	-0.5	138E	-0.83	50L	-0.34	92E	-0.64
210Y	-0.48	77F	-0.43	227F	-0.41	100L	-0.52	11V	-0.34	66T	-0.38
205L	-0.47	103Q	-0.4	101K	-0.25	227F	-0.45	50I	-0.32	140G	-0.34
228N	-0.38	163A	-0.39	101N	-0.22	230M	-0.38	35Q	-0.25	97T	-0.26
181Y	-0.38	126K	-0.38	77L	-0.22	118I	-0.35	87R	-0.25	113I	-0.2
178F	-0.38	110N	-0.38	208F	-0.22	173K	-0.34	42W	-0.25	101I	-0.16
184M	-0.37	145C	-0.38	230M	-0.21	179V	-0.31	33L	-0.24	143Y	-0.15
40I	-0.37	69P	-0.38	117S	-0.2	74L	-0.31	73A	-0.24	212E	-0.13
215F	0.66	69E	0.61	103S	0.42	118V	0.34	84V	0.32	113V	0.17
215I	0.68	122K	0.62	179E	0.42	181V	0.36	73T	0.32	234V	0.21
65R	0.76	106M	0.63	230L	0.48	230L	0.38	89V	0.36	232N	0.25
106M	0.85	142M	0.63	190E	0.6	173I	0.38	54M	0.37	140S	0.26
190S	0.87	70E	0.78	190Q	0.6	138K	0.41	47A	0.38	97A	0.3
103N	0.9	65Ins	0.98	190S	0.81	100I	0.52	74P	0.44	148H	0.32
215Y	0.93	69Ins	1.04	101P	1.05	190E	0.54	69N	0.51	66I	0.33
70R	0.97	214F	1.09	188L	1.07	181I	0.7	50V	0.54	148R	0.4
184I	1.57	184I	1.26	106M	1.17	227C	0.74	33F	0.55	92Q	0.73
184V	1.81	65R	1.4	103N	1.82	101P	1.1	82F	0.64	155H	0.9

evidence for possible 3TC or FTC resistance and exposure, as 3TC and FTC are co-administered with NNR-TIs very frequently. The resistance-driven failure of a therapy including 3TC/FTC and a NNRTI will likely select both 3TC/FTC and NNRTI drug-resistance mutations. Among the negative weights for 3TC and FTC, wild-type mutations can be found. These weights present a mechanism with which DES can be lowered when counter-evidence for drug exposure and drug resistance is found in the genotype. Furthermore, linear mutation weights define *in-vivo drug fingerprints*, i.e. mutation weights that show which mutations are likely to be present in the viral genotypes of patients who have failed a certain drug, and which mutations are not likely to present in these viral genotypes.

DES can be automatically derived from therapy history and GPPs. The method demonstrates that the problem of determining the drug susceptibility of an HIV-1 variant (questions (2) and (3)) can be reduced to the determination of its distance to a classification boundary in a mathematical space encoding the genotype. Furthermore, it could be shown that models trained on nucleotide HIV-1 sequences from patients with known drug history have a comparatively high performance in predicting drug resistance, even if no GPPs are used. When GPPs are used, they are categorized into susceptible and resistant with drug-independent cutoffs. In no case do I resort to RFs in training, which affords, for instance, merging of data from different phenotypic assays. Thus, quantitative measurements of drug resistance are not necessary for producing a quantitative output that is sufficient for predicting clinically-relevant degrees of resistance (Section 3.4; question (2)). The methodology has a comparatively high performance in predicting drug exposure (question (1)) and short- and long-term therapeutic success (question (3)), as well as a sufficiently strong correlation with phenotypic drug resistance (question (2)). In conjunction with a frequently updated database with HIV-1 data from routine clinical practice, such as the EIDB, DES models can be automatically updated on a regular basis. Thus, these models allow for a reduction on the dependency on hard-to-obtain GPPs for offering a publicly available data-driven genotypic drug-resistance interpretation system that is kept up to date. While regularly updatable interpretation systems are evidently the appropriate method for accounting for the growing richness of clinical data, innovative procedures may have to be put in place for adequate certification of such systems. DES models for PIs and RTIs have been made available on the geno2pheno[resistance] server http://www.geno2pheno.org. After a sequence has been submitted for prediction, the tab labeled *Drug Exposure* must be selected in order to view DES predictions. On the website, mutations with the highest influence on the prediction are displayed (see above and *Linear Weights for Drug-Exposure-Score Models*). These are ordered by the magnitude of their influence. Mutations colored in red increase DES, while those colored in green decrease it.

3.4   Determination of Cutoffs for Drug-Exposure Models

In Section 3.3, I present a method for deriving a genotypic drug-resistance interpretation system from clinical data. The models used in the system are trained on clinical HIV-1 genotypes along with the antiretroviral drug-use history of the patient from which the sequenced HIV-1 variant was isolated. The training set of some of these models was supplemented with genotypes from GPPs. Furthermore, three different amino-acid encodings were tested. The best performances could be obtained with models whose training sets include GPPs (ExposurePheno models) and which were trained with the V_{full} encoding (*Subtype Determination, Sequence Alignment, and Encoding* in Section 3.3.1). In this section, I present four methods for deriving DES cutoffs which aim at translating DES into clinically meaningful categories. The following goals are addressed by each method: (1) prediction of drug-exposure, (2) prediction of phenotypic *in-vitro* resistance, and (3) prediction of

therapy-success (two methods are presented for this last goal). Since ExposurePheno$_\mathsf{full}$ models present superior performance (Section 3.3.2), along with other desirable characteristics, I only derive cutoffs for these models, in this chapter.

3.4.1 MATERIALS AND METHODS

In the following, I present four methods for obtaining DES cutoffs. Each method determines a set of cutoffs for each final ExposurePheno$_\mathsf{full}$ drug-exposure model (Section 3.3). The following goals are addressed by each set of cutoffs: (1) prediction of drug-exposure, (2) prediction of phenotypic *in-vitro* resistance, and (3) and (4) prediction of therapy success. Each set of cutoffs includes a lower and an upper cutoff for the corresponding drug-exposure models.

(1) CUTOFFS MAXIMIZING THE PERFORMANCE OF THE PREDICTION OF DRUG EXPOSURE (DE-MAX CUTOFFS)

ExposurePheno$_\mathsf{full}$ cross-validation sets were interpreted with the corresponding final models that were trained on them (which is also called *calculation of reinsertion predictions*). For each cross-validation set, an upper and a lower cutoff were estimated such that the AUC of the drug-exposure prediction is maximized. I call these cutoffs the DEMax cutoffs, and they allow for the discretization of a DES for a drug into the categories *unexposed* (U), *possible exposure* (PE) and *exposed* (E). A detailed description of the procedure with which DEMax cutoffs were determined follows. Function (3.3) was defined for discretization of a value $\delta_s \in \mathbb{R}$ associated to a sequence s by using the lower and upper cutoffs $c_L, c_U \in \mathbb{R}$.

$$\text{discretize}(c_L, c_U, \delta_s) = \begin{cases} 1, \text{if} \delta_s < c_L \\ 2, \text{if} c_L \leq \delta_s \leq c_U \\ 3, \text{if} c_U < \delta_s \end{cases} \tag{3.3}$$

Let $\Delta_{drug} \in \mathbb{R}^n$ be a vector of DES predicting the drug exposure of each of n sequences s to $drug$, and let $E_{drug} \in \{0, 1\}^n$ be the corresponding vector of class labels, indicating whether each sequence s was exposed to the drug or not. Application of cutoffs c_L, c_U and function (3.3) to a vector of DES Δ_{drug} results in the discrete DES vector discretize(cL, cU, Δ_{drug}). For each bootstrap replicate, an upper and a lower cutoff c_L and c_U were determined as

$$\underset{c_L, c_U}{\arg \max} \text{AUC}(\text{discretize}(c_L, c_U, \Delta_{drug}), E_{drug}), \tag{3.4}$$

where AUC(discretize(c_L, c_U, Δ_{drug}), E_{drug}) is the AUC quantifying the performance of discretize(c_L, c_U, Δ_{drug}) in predicting exposure to $drug$ for each sequence with a DES in Δ_{drug}. ExposurePheno$_\mathsf{drug}$ cross-validation sets were interpreted with the corresponding final models that were trained on them. Two thousand bootstrap replicates of the DES of each cross-validation set were created. For each bootstrap replicate and the corresponding class labels, an upper and a lower cutoff were determined by AUC maximization (3.4). The resulting 2,000 upper and 2,000 lower cutoffs for each final drug-exposure model were averaged to yield the final set of cutoffs, i.e. those that are used for discretization of DES. I call these cutoffs the DEMax cutoffs. If a DES for a drug is less than the lower cutoff for that drug, then I discretize that DES as *unexposed* (U). If a DES is greater or equal

than the lower cutoff, but less or equal than the upper cutoff, I discretize that DES as *possible exposure* (PE). Finally, if a DES is greater than both cutoffs, then I discretize that DES as *exposed* (E).

(2) Phenotypically Guided Cutoffs for Prediction of Phenotypic *in-vitro* Drug Resistance (Pheno Cutoffs)

Clinically-relevant cutoffs were used for discretizing PhenoSense GPPs in D_{Pheno} into the categories *susceptible* (S), *intermediate* (I) or *resistant* (R), henceforth called the true labels. The genotypes associated with these GPPs were interpreted with the final ExposurePheno$_{full}$ drug-exposure models. For each drug, an upper and a lower DES cutoff yield predicted GPP labels. These cutoffs, which I call Pheno cutoffs, are determined such that the sum of the penalties quantifying the differences between the true labels and the predicted labels is minimized. An individual penalty equals one, if the true label was R and the predicted label was S. If the true label is I, and the predicted label S, the penalty equals 0.75. All other differences between true and predicted labels were penalized with the value 0.5, while the equality of true and predicted labels was not penalized. Pheno cutoffs allow for discretization of a DES for a drug as *susceptible* (S), *intermediate* (I) or *resistant* (R). Further details on the cutoff-determination procedure, including the rationale for choosing the penalty values follow.

The error matrix $E \in \mathbb{R}^{3 \times 3}$ (3) was defined for penalizing the misclassification of a discretized value δ_s with label $l \in \{1, 2, 3\}$ and predicted label $\hat{l} \in \{1, 2, 3\}$

$$E_{(l, \hat{l})} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 1 & 0.75 & 0 \end{pmatrix} \tag{3.5}$$

The rationale for choosing the values of the error matrix follows. Diagonal entries are zero, as correct classification incurs no penalty. From a clinical perspective, the worst kind of misclassification that can occur is the classification of a resistant viral strain (label 3) as susceptible (label 1), since the prescription of a therapy including a thus misclassified compound could compromise the susceptibility of all compounds in the therapy. Therefore, this kind of misclassification was assigned the maximum penalty, i.e. one. Misclassification of a resistant strain as intermediate (label 2) deserves a smaller penalty, since surpassing the lower cutoff indicates a clinically-relevant decrease in susceptibility, albeit implying that some susceptibility is given. Therefore, this kind of misclassification was assigned the penalty 0.75. All other types of misclassifications are considered equally undesirable, but less severe than the first two, and were assigned the penalty 0.5. Clinically-relevant cutoffs were used to discretize PhenoSense GPPs in D_{Pheno} with function (3.3), yielding their labels. The genotypes s associated with these GPPs were interpreted with the final ExposurePheno$_{full}$ drug-exposure models. For each drug involved in a GPP, 2,000 bootstrap replicates of the PhenoSense GPPs in D_{Pheno} were sampled. In order to assign to each of the three classes the same weight in this procedure, each bootstrap replicate was constructed using an equal number of GPPs with each label. For each drug, this number was equal to the maximum number of GPPs with a certain label. Each bootstrap replicate was used to determine a lower and an upper cutoff $\hat{c}_L, \hat{c}_U$ which minimizes the sum of the penalties $E_{(l,\hat{l})}$ for each label $l = \text{discretize}(c_L, c_U, RF_s)$ with corresponding prediction $\hat{l} = \text{discretize}(\hat{c}_L, \hat{c}_U, DES_s)$ for a resistance factor RF and a drug-exposure score(s) DES associated with genotype s. The resulting 2,000 cutoff pairs for each drug and drug-exposure-model group were averaged, yielding the final phenotypically guided cutoffs. If a DES for a drug is less than both cutoffs for that drug, then I discretize that DES as *susceptible* (S). If a DES is greater or equal than the lower cutoff, but less or equal than

the upper cutoff, I discretize that DES as *intermediate* (I). Finally, if a DES is greater than both cutoffs, then I discretize that DES as *resistant* (R).

(3) Cutoffs Maximizing the Performance of the Prediction of Therapy-Success (ThMax Cutoffs)

Let $D = (d_1, \ldots, d_m) \in \{0, 1\}^{1 \times m}$ be a vector indicating which of the m available drugs were used in a TE. For a given vector of lower (c_{iL}) and upper (c_{iU}) cutoffs $C = (c_{1L}, c_{1U}, \ldots, c_{mL}, c_{mU}) \in \mathbb{R}^{2m}$, I score the susceptibility of an HIV-1 variant s to a certain drug with index i and DES $\delta_{s,i}$ with the function

$$\text{susceptibility}(c_{iL}, c_{iU}, \delta_{s,i}) = \begin{cases} 1, \text{if } \delta_{s,i} < c_{iL} \\ 0.5, \text{if } c_{iL} \leq \delta_{s,i} \leq c_{iU} \\ 0, \text{if } c_{iU} < \delta_{s,i} \end{cases} \tag{3.6}$$

The genetic susceptibility score (GSS) for a TE baseline sequence s with DES $\Delta_s = \{\delta_{s,1}, \ldots, \delta_{s,m}\} \in \mathbb{R}^m$ is computed with the function

$$GSS(C, D, \Delta_s) = \sum_{i \in \{i | d_i = 1\}} \text{susceptibility}(c_{iL}, c_{iU}, \delta_{s,i}) \tag{3.7}$$

Let $D \in \{0, 1\}^{n \times m}$ be the matrix of vectors D indicating the compounds used in a certain set of TEs of size n with corresponding matrices of DES vectors $\Delta \in \mathbb{R}^{n \times m}$. Then, $GSS(C, D, \Delta) \in \mathbb{R}^n$ is the vector containing the GSS for each of the TEs obtained with cutoff vector C. Let $S \in \{0, 1\}^n$ be the vector containing the success labels for the TEs. The function calculating the AUC is $AUC(GSS(C, D, \Delta), S)$. I determine the cutoff set C as

$$\arg\max_C AUC(GSS(C, D, \Delta), S). \tag{3.8}$$

ThMax cutoffs were determined for each final model group by maximization of the AUC for a GSS predicting therapy success, as follows. $D_{\mathsf{EuResistTE}}$ TE baseline sequences and D_{TFE} sequences were evaluated with the final drug-exposure models. Simulated annealing was used for maximization of the AUC (3.8) of 2,000 bootstrap replicates of $D_{\mathsf{EuResistTE}}$. For this purpose, the R [306] package GenSA [310] was used. The maximum number of calls to the objective function was constrained to 1,500,000. This yielded 2,000 cutoff sets for each final model group, which were averaged for obtaining the final sets of cutoffs [311].

(4) Therapy-Success Probability Cutoffs (ThSucc Cutoffs)

Drug-wise therapy-success probabilities conditioned on the DES for the corresponding drug were estimated with the TEs in $D_{\mathsf{EuResistTE}}$, and failing therapies in D_{TFE}. A DES for a drug is discretized as *low success probability* (L) if the corresponding therapy-success probability is less or equal than 0.45. If the estimated therapy-success probability is greater than 0.45 but less than 0.7, the DES is discretized as *intermediate success probability* (IP). Last, if the estimated therapy-success probability is greater or equal than 0.7, the DES is assigned to the category *high success probability* (H). I call the DES corresponding to success probabilities of 0.45 and 0.7 ThSucc cutoffs. The rationale for the choice of these probability thresholds is given in Section 3.4.2. In the following, further details on this cutoff-selection method are given. Let TE_{drug} be a sequence

of TEs of length n in which $drug$ is used, for $drug \in \{$3TC, ABC, AZT, d4T, ddI, FTC, TDF, EFV, ETR, NVP, APV, ATV, DRV, FPV, IDV, LPV, NFV, SQV, TPV$\}$. Each TE in TE_{drug} includes a success label $S \in \{0, 1\}$, indicating whether the therapy was successful or not. Interpretation of the baseline sequence of each TE with the drug-exposure model for drug results in a DES δ. The probability of therapeutic success given DES δ is then estimated as

$$P(S = 1 \mid \delta, drug) = \frac{f_{1,drug}(\delta)\pi_1}{f_{1,drug}(\delta)\pi_1 + f_{0,drug}(\delta)\pi_0} \tag{3.9}$$

where $f_{1,drug}(\delta)$ is the therapy-success density, $f_{0,drug}(\delta)$ the therapy-failure density, and π_1 and π_0 are prior probabilities for therapy success and therapy failure, respectively (Section 2.1.2). Let TFE_{drug} be the set of therapy-failure sequences for which the failing therapy involves $drug$. KDE was used for computing $f_{1,drug}(\delta)$ by using $\mathsf{TE}_{drug} \subset \mathsf{D}_{\mathsf{EuResistTE}}$ with $S = 1$ (Section 2.3.1). For this purpose, drug-exposure scores δ were calculated with the final $\mathsf{ExposurePheno}_{\mathsf{full}}$ drug-exposure model for $drug$. Analogously, a therapy-failure density for drug, $f_{0,drug}(\delta)$, was estimated with the TEs in TE_{drug} for which $S = 0$, along with the failing therapies in $\mathsf{TFE}_{drug} \subset \mathsf{D}_{\mathsf{TFE}}$. KDE was performed with the R [306] package ks [312]. The bandwidth matrix for KDE was selected with the plug-in method, allowing for derivatives up to the second order [312]. Prior probabilities π_1 and π_0 were set to the corresponding success and failure proportions, respectively. For increasing the smoothness of the therapy-success estimates, a logistic curve was fit to the conditional therapy-success probabilities obtained with KDE, such that

$$P(S = 1 \mid \delta, drug) = \frac{a}{1 + \exp(-b(\delta - c))} \tag{3.10}$$

with $a \in [0, 1]$ and $b, c \in \mathbb{R}$. Logistic-curve parameters a, b, and c, were determined with the minipack.lm [313] library of the R programming language. Logistic curves were fit for each possible value of drug on 2,000 bootstrap replicates of $\mathsf{D}_{\mathsf{EuResistTE}}$, yielding 2,000 logistic curves per drug. Each logistic curve was used for cutoff determination. For each curve, the lower cutoff was selected at the DES where $P(S = 1 \mid drug) = 0.7$, and the upper cutoff at the DES where $P(S = 1 \mid drug) = 0.45$. As mentioned above, the rationale for the choice of these probability thresholds is given in Section 3.4.2. Lower or upper cutoffs could not be obtained from logistic curves with values for a lower than 0.7 or lower than 0.45, respectively. In these cases, the lower or upper cutoff equaled the minimum DES for drug in the bootstrap replicate. Lower and upper cutoffs were averaged across bootstrap replicates, yielding the final set of cutoffs.

INTERPRETATION SYSTEMS FOR COMPARISON OF PERFORMANCE

HIValg, a program for rules-based interpretation of HIV-1 sequences, was downloaded from http://hivdb.stanford.edu [294] for performance comparison. XML rule-definition files for the following interpretation systems were obtained from HIV-GRADE (http://www.hiv-grade.de): ANRS 09/2012 (http://www.hivfrenchresistance.org/), GRADE 06/2013 [297], HIVdb 6.0.6 [298], and REGA 8.0.2 (http://rega.kuleuven.be/). HIValg outputs a (discrete) SIR prediction for each drug, rule set and nucleotide sequence. Furthermore, geno2pheno[resistance] version 3.3 (http://www.geno2pheno.org, accessed on March 31st, 2015) [299] was also used for performance comparison. The output of geno2pheno[resistance] includes a RF prediction and a SIR classification for each drug and nucleotide sequence. For performance assessment with the

AUC, SIR classifications were converted to an integer score via $R \rightarrow 2$, $I \rightarrow 1$, and $S \rightarrow 0$. Since each interpretation system uses its own alignment program, performance comparison was constrained to the set of sequences which could be aligned without errors by all programs. Furthermore, the drug ddC was also excluded from performance comparison, as it is not supported by these systems any more. Since the downloadable version of HIValg does not support predictions for drugs whose target is IN, the HIV-GRADE website was used for performing predictions for INIs.

Assessment of Performance Using Test Sets

Sequences in T_{PRRT}, T_{IN}, TP_{PRRT}, T_{Pheno}, $T_{\text{EuResistTE}}$, HIVdbTCE, and HIVdbExposure were interpreted with the final $\text{ExposurePheno}_{\text{full}}$ drug-exposure models, the four rule sets (via HIValg and the HIV-GRADE website), and geno2pheno$_{[\text{resistance}]}$. For discretization of DES, the cutoff sets obtained with the procedures described in above were used. In the following, performance was quantified in terms of AUC [307] unless stated otherwise. Significance values in the Methods section were calculated with a two-sided Wilcoxon signed-rank test [308].

Assessment of Performance in Predicting Therapy Success

Each system's capability to predict therapy success was assessed with $T_{\text{EuResistTE}}$ and HIVdbTCE. For this purpose, a GSS for each therapy was calculated. Specifically, SIR classifications for the drugs in each regimen were converted to integer via $S \rightarrow 1$, $I \rightarrow 0.5$ and $R \rightarrow 0$. Analogously, DES for the drugs in each regimen were discretized based on their respective cutoff sets. The GSS for each therapy consisted of the sum of the individual integer scores for each drug in the regimen. Concordance of each system's GSS with the NAS was assessed for therapies in the $T_{\text{EuResistTE}}$ and HIVdbTCE datasets in terms of Harrell's concordance index (C; Section 2.2.6) [267].

Assessment of Performance in Predicting Resistance in Genotype-Phenotype Pairs

The capability of discrete DES and of the other interpretation systems to predict drug resistance as assessed by the RFs of GPPs in T_{Pheno} was quantified as follows. I computed drug-wise misclassification rates of the tested systems when predicting the SIR label derived from RFs from PhenoSense GPPs and PhenoSense clinically-relevant cutoffs. In addition to computing misclassification rates with all GPPs, susceptible-to-resistant and resistant-to-susceptible misclassification rates were computed for each drug and tested interpretation system.

3.4.2 Results

DES cutoffs were estimated with four different methods for each final $\text{ExposurePheno}_{\text{full}}$ drug-exposure model: DEMax, pheno, ThMax, and ThSucc. DEMax cutoffs can be seen in Table 3.18, pheno cutoffs are shown in Table 3.19, ThMax cutoffs are displayed in Table 3.20, and ThSucc cutoffs in Table 3.21. Drug-wise therapy-success proportions for $D_{\text{EuResistTE}}$, $D_{\text{EuResistTE}} \cup D_{\text{TFE}}$, $T_{\text{EuResistTE}}$, and HIVdbTCE can be seen in Table 3.3. Therapy-success rates in $D_{\text{EuResistTE}} \cup D_{\text{TFE}}$ form two groups (Figure 3.4 and Table 3.3). The first group includes drugs with low success rates, with an average success rate of 0.46. The second group includes drugs with higher success rates, with an average success rate of 0.72. These two average success rates were rounded to the next 0.05 and used as cutoffs for estimated success probabilities, leading to the ThSucc cutoffs (*Therapy-Success Probability Cutoffs* in Section 3.4.1).

Table 3.18: DEMax cutoffs for $\text{ExposurePheno}_{\text{full}}$ models. The final DEMax cutoffs were obtained by averaging the cutoffs obtained with 2,000 bootstrap replicates of the cross-validation sets. SD: Standard Deviation; 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	Lower Cutoff (SD)	Upper Cutoff (SD)
3FTC	-0.9 (0)	0.22 (0.07)
ABC	-0.8 (0.02)	0.68 (0.07)
AZT	-0.9 (0.01)	0.58 (0.06)
d4T	-0.77 (0.07)	0.49 (0.13)
ddI	-0.83 (0.05)	0.61 (0.11)
ddC	-0.69 (0.05)	0.8 (0.05)
TDF	-0.75 (0.05)	0.54 (0.12)
EFV	-1 (0)	-0.5 (0.07)
ETR	-0.81 (0.04)	0.81 (0.15)
DLV	-0.9 (0.02)	0.34 (0.24)
NVP	-0.9 (0)	0.87 (0.05)
RPV	-0.49 (0.19)	0.11 (0.11)
AFPV	-0.9 (0)	0.79 (0.05)
ATV	-0.9 (0.01)	0.46 (0.14)
DRV	-0.91 (0.03)	0.34 (0.3)
IDV	-0.9 (0)	0.18 (0.12)
LPV	-0.91 (0.03)	0.61 (0.33)
NFV	-1 (0)	-0.08 (0.12)
SQV	-0.9 (0.02)	0.55 (0.11)
TPV	-0.9 (0)	0.7 (0.08)
RAL	-0.88 (0.04)	0.11 (0.13)
EVG	-0.53 (0.42)	-0.2 (0.06)
Naïve PRRT	-0.51 (0.12)	0.89 (0.02)

Table 3.19: pheno cutoffs for $\mathrm{ExposurePheno_{full}}$ **models.** The final pheno cutoffs were obtained by averaging the cutoffs obtained with 2,000 bootstrap replicates of the cross-validation sets. SD: Standard Deviation; 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	Lower Cutoff (SD)	Upper Cutoff (SD)
3FTC	0.1 (0.07)	1.95 (0.11)
ABC	-0.05 (0.17)	0.66 (0.27)
AZT	0.53 (0.16)	0.59 (0.16)
d4T	-0.24 (0.07)	0.24 (0.1)
ddI	-0.16 (0.11)	0.19 (0.29)
TDF	-0.51 (0.13)	0.11 (0.46)
EFV	-0.65 (0.02)	0.87 (0.08)
ETR	-0.55 (0.41)	0.9 (0.02)
DLV	-0.74 (0.11)	0.79 (0.14)
NVP	-0.79 (0.04)	0.07 (0.25)
RPV	-0.43 (0.09)	0.69 (0.22)
AFPV	-0.07 (0.07)	1.4 (0.17)
ATV	-0.42 (0.15)	0.99 (0.05)
DRV	0.55 (0.16)	1.9 (0.02)
IDV	-0.36 (0.16)	1.36 (0.08)
LPV	0.32 (0.06)	1.18 (0.08)
NFV	-0.81 (0.03)	0.84 (0.04)
SQV	-0.1 (0.04)	1.44 (0.08)
TPV	-0.67 (0.11)	0.97 (0.02)
RAL	0.04 (0.29)	1.04 (0.06)
EVG	-0.4 (0.02)	0.89 (0.01)

Table 3.20: ThMax cutoffs for $\text{ExposurePheno}_{\text{full}}$ **models.** The final ThMax cutoffs were obtained by averaging the cutoffs obtained with 2,000 bootstrap replicates of the cross-validation sets. SD: Standard Deviation; 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir.

	Lower Cutoff (SD)	Upper Cutoff (SD)
3FTC	-1.01 (0.42)	0.12 (0.44)
ABC	-1.49 (0.2)	-0.62 (0.6)
AZT	-1.33 (0.12)	-0.64 (0.55)
d4T	-2.11 (0.48)	-1.83 (0.45)
ddI	-1.17 (0.13)	-1.03 (0.2)
TDF	-1.36 (0.46)	0.4 (0.9)
EFV	-0.6 (0.15)	-0.09 (0.35)
ETR	-0.81 (0.54)	-0.41 (0.68)
NVP	-0.93 (0.15)	-0.48 (0.2)
AFPV	-1.15 (0.26)	0.51 (0.54)
ATV	-1.01 (0.25)	0.78 (0.68)
DRV	0.14 (0.7)	1.53 (1)
IDV	-1.15 (0.17)	0.1 (0.64)
LPV	-0.77 (0.44)	1.26 (0.14)
NFV	-1.14 (0.04)	-0.58 (0.39)
SQV	-1.06 (0.22)	-0.37 (0.36)
TPV	-2.02 (0.32)	-0.18 (1.31)

Table 3.21: ThSucc cutoffs for $\text{ExposurePheno}_{\text{full}}$ **models.** The final ThSucc cutoffs were obtained by averaging the cutoffs obtained with 2,000 bootstrap replicates of the cross-validation sets. SD: Standard Deviation; 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir.

	Lower Cutoff (SD)	Upper Cutoff (SD)
3FTC	0.43 (0.05)	2.03 (0.08)
ABC	-0.26 (0.06)	0.62 (0.07)
AZT	-0.15 (0.09)	1.3 (0.1)
d4T	-2.5 (0.39)	0.59 (0.21)
ddI	-3.4 (1.19)	1.03 (0.15)
TDF	0.12 (0.04)	1.14 (0.07)
EFV	-0.28 (0.15)	0.05 (0.23)
ETR	-0.04 (0.49)	0.98 (0.42)
NVP	-0.62 (0.06)	-0.26 (0.13)
AFPV	-0.52 (0.31)	0.77 (0.16)
ATV	0.31 (0.27)	1.29 (0.14)
DRV	0.75 (0.38)	1.62 (0.75)
IDV	-1.36 (0.28)	0.31 (0.19)
LPV	0.18 (0.08)	1.19 (0.09)
NFV	-1.19 (0.04)	-0.12 (0.13)
SQV	-1.36 (0.57)	0.35 (0.21)
TPV	-2.5 (0.39)	-0.03 (1.13)
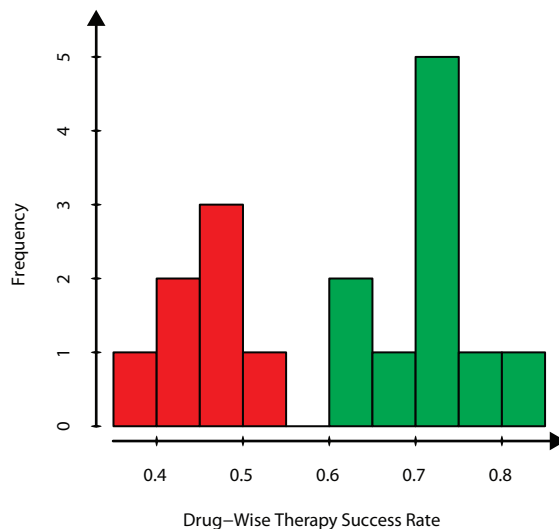
Figure 3.4: Histogram of Drug-Wise Therapy Success Rates in $D_{\text{EuResistTE}} \cup D_{\text{TFE}}$. Drug-wise therapy success rates in $D_{\text{EuResistTE}} \cup D_{\text{TFE}}$ were calculated. The histogram shows the distribution of the success rates of 17 drug compounds. This histogram makes two groups of drugs apparent: drugs with low success rates and drugs with high success rates.
🟥 Low-success-rate therapy group   🟩 High-success-rate therapy group

ASSESSMENT AND COMPARISON OF PERFORMANCE USING DISCRETE PREDICTIONS

The performance of the final $\mathsf{ExposurePheno_{full}}$ drug-exposure models was assessed with the continuous DES (Section 3.3.2) and after discretization with the estimated cutoff sets. Where possible, performance was compared to that of popular drug-resistance interpretation systems. In the following, performances averaged over individual drug performances are stated. In order to be able to compare the different models, these means were calculated only with the drug-wise performances of the drugs that are common to all model groups and drug resistance interpretation systems used for performance comparison in this chapter. Note that in some tables found in the following, two values for the mean performance across drugs are provided. The first mean value, to which the term *common drugs* is appended, only considers performances for drugs common to all drug-exposure model groups and drug-resistance interpretation systems used for performance comparison. The second mean value, to which the term *all drugs* is appended, considers all drugs for which predictions were available. Thus, means for *common drugs* allow for performance comparison between different model groups and drug-resistance interpretation systems. p-values were calculated with a two-sided Wilcoxon signed-rank test.

PERFORMANCE OF DISCRETE DRUG-EXPOSURE AND DRUG-RESISTANCE PREDICTIONS

Performances of prediction of drug exposure calculated via application of DEMax to the DES in $\mathsf{T_{PRRT}}$, $\mathsf{T_{IN}}$ datasets are displayed in Table 3.22, along with the performances of HIVdb, REGA, ANRS, GRADE, and discretized geno2pheno[resistance] predictions. The best average performance for drug-exposure prediction with a discrete score on the $\mathsf{T_{PRRT}}$ dataset was attained by discrete DES ($\mu = 0.75$; $\sigma = 0.06$). Among the drug-resistance interpretation systems I compare to, HIVdb, REGA, and GRADE performed best ($\mu = 0.70$; $\sigma = 0.06$), although GRADE's mean performance is lower than that of the DES models ($p < 10^{-4}$). On the $\mathsf{T_{IN}}$ dataset, discrete DES predictions, HIVdb, REGA, and GRADE attained the best performance (AUC =

Table 3.22: Performance of Prediction of Drug-Exposure (AUC) with Discrete Output, T_{PRRT} and T_{IN} Datasets. For each final ExposurePheno$_{full}$ DES model, drug-exposure prediction performances on sequences in the T_{PRRT} and T_{IN} datasets were calculated after DES discretization with DEMax cutoffs. Discrete HIVdb, REGA, ANRS, GRADE and geno2pheno$_{[resistance]}$ (g2p) performances were calculated as well. The highest performance in each row is underlined. AM: All Models; CD : Drugs Common to all interpretation systems; NA: Not Available; SD: Standard Deviation, 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddC: zalcitabine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir.

	ExposurePheno$_{full}$	HIVdb	REGA	ANRS	GRADE	g2p
3FTC	<u>0.83</u>	0.73	0.74	0.71	0.72	0.72
ABC	<u>0.7</u>	0.68	0.66	0.65	0.66	0.68
AZT	<u>0.81</u>	0.72	0.72	0.71	0.73	0.72
d4T	<u>0.82</u>	0.77	0.76	0.75	0.77	0.71
ddC	<u>0.8</u>	NA	NA	NA	NA	NA
ddI	<u>0.82</u>	0.77	0.76	0.66	0.77	0.74
TDF	<u>0.68</u>	0.62	0.6	0.6	0.64	0.61
DLV	<u>0.79</u>	NA	NA	NA	NA	NA
EFV	<u>0.73</u>	0.69	0.69	0.69	0.7	0.68
ETR	<u>0.73</u>	0.71	0.69	0.62	0.71	0.58
NVP	<u>0.74</u>	0.69	0.68	0.66	0.69	0.67
AFPV	<u>0.77</u>	0.75	0.73	0.72	0.74	0.73
ATV	<u>0.63</u>	0.58	0.58	0.58	0.57	0.58
DRV	<u>0.67</u>	0.64	0.66	0.5	0.66	0.65
IDV	<u>0.76</u>	0.7	0.69	0.7	0.7	0.69
LPV	<u>0.71</u>	0.67	0.67	0.67	0.67	0.67
NFV	<u>0.73</u>	0.7	0.7	0.69	0.69	0.7
SQV	<u>0.79</u>	0.73	0.73	0.69	0.73	0.73
TPV	<u>0.82</u>	<u>0.82</u>	0.77	0.5	0.73	0.77
RAL	<u>0.7</u>	<u>0.7</u>	<u>0.7</u>	<u>0.7</u>	0.69	NA
Mean CD (SD)	<u>0.75 (0.06)</u>	0.7 (0.06)	0.7 (0.05)	0.65 (0.07)	0.7 (0.05)	0.68 (0.05)
Mean AM (SD)	<u>0.75 (0.06)</u>	0.7 (0.06)	0.7 (0.05)	0.66 (0.07)	0.7 (0.05)	NA (NA)

Table 3.23: Performance of Prediction of Drug-Exposure (AUC) with Discrete Output, $\mathsf{HIVdbExposure}$ **dataset.** For each final $\mathsf{ExposurePheno_{full}}$ drug-exposure-score (DES) model, drug-exposure prediction performances on sequences in the HIVdbExposure dataset were calculated after DES discretization with DEMax cutoffs. Discrete HIVdb, REGA, ANRS, GRADE and $\mathrm{geno2pheno_{[resistance]}}$ (g2p) performances were calculated as well. The highest performance in each row is underlined. SD: Standard Deviation, 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir.

	$\mathsf{ExposurePheno_{full}}$	HIVdb	GRADE	REGA	ANRS	g2p
3FTC	0.59	0.75	0.76	0.75	0.75	0.75
ABC	0.66	0.66	0.66	0.67	0.66	0.62
AZT	0.56	0.66	0.67	0.66	0.65	0.67
d4T	0.6	0.65	0.62	0.63	0.61	0.62
ddI	0.68	0.68	0.66	0.66	0.66	0.55
TDF	0.62	0.56	0.56	0.56	0.59	0.51
EFV	0.8	0.81	0.8	0.77	0.77	0.79
ETR	0.94	0.98	0.97	0.89	0.92	0.98
NVP	0.77	0.74	0.74	0.74	0.71	0.72
AFPV	0.78	0.76	0.78	0.76	0.73	0.78
ATV	0.6	0.54	0.57	0.57	0.56	0.54
DRV	0.89	0.89	0.59	0.94	0.5	0.83
IDV	0.74	0.74	0.74	0.75	0.74	0.75
LPV	0.64	0.62	0.62	0.65	0.63	0.62
NFV	0.71	0.71	0.71	0.71	0.74	0.72
SQV	0.74	0.75	0.76	0.73	0.7	0.73
TPV	0.93	0.86	0.77	0.85	0.5	0.94
Mean	0.72 (0.12)	0.73 (0.11)	0.7 (0.1)	0.72 (0.1)	0.67 (0.11)	0.71 (0.13)

0.70). Performances of discretized prediction of drug-exposure on $\mathsf{HIVdbExposure}$ can be seen in Table 3.23. HIVdb displayed the best mean performance on the $\mathsf{HIVdbExposure}$ dataset ($\mu = 0.73$; $\sigma = 0.11$), while the performance of the DES models was lower ($\mu = 0.72$; $\sigma = 0.12$; $p = 0.89$). Table 3.24 shows the performance of discretized DES in predicting drug exposure in $\mathsf{TP_{PRRT}}$. For comparison, the drug-wise performances of other genotype interpretation methods can be found in Table 3.24 as well. The best average drug-wise performance was achieved by the DES models discretized with DEMax cutoffs ($\mu = 0.58$; $\sigma = 0.08$). Among the drug-resistance interpretation systems I compare to, HIVdb displayed the best mean performance ($\mu = 0.56$; $\sigma = 0.10$; $p = 0.0093$).

Performance in predicting the SIR class on PhenoSense GPPs discretized with clinically relevant cutoffs can be seen in Tables 3.25 and 3.26. The lowest mean misclassification rate was attained by REGA ($\mu = 0.23$; $\sigma = 0.10$), while $\mathsf{ExposurePheno_{full}}$ models with pheno cutoffs performed worse ($\mu = 0.28$; $\sigma = 0.1$; $p = 0.02667$). The lowest mean R to S misclassification rates were obtained with REGA ($\mu = 0.01$; $\sigma = 0.03$), while the mean R to S misclassification rate of $\mathsf{ExposurePheno_{full}}$ with pheno cutoffs models was insignificantly higher ($\mu = 0.02$ $\sigma = 0.03$; $p = 0.5992$). The lowest S to R misclassification rate was attained by HIVdb ($\mu = 0.03$; $\sigma = 0.03$), while the $\mathsf{ExposurePheno_{full}}$ models with pheno cutoffs displayed an insignificantly higher S to R misclassification rate ($\mu = 0.08$; $\sigma = 0.11$; $p = 0.1424$).

Table 3.24: Performance of Drug-Exposure Prediction for Discrete Drug-Exposure Scores and Other Interpretation Systems, TP_{PRRT} Dataset. Drug-exposure prediction performances on the TP_{PRRT} dataset were calculated for each final ExposurePheno$_{full}$ drug-exposure-score (DES) model, after discretization with DEMax cutoffs. Discrete HIVdb, REGA, ANRS, GRADE and geno2pheno$_{[resistance]}$ (g2p) performances were calculated as well. The highest performance in each row is underlined. SD: Standard Deviation, 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir.

	ExposurePheno$_{full}$	ANRS	HIVdb	GRADE	REGA	g2p
3FTC	0.41	<u>0.49</u>	0.41	<u>0.49</u>	0.41	0.41
ABC	0.56	0.53	0.56	0.53	0.54	<u>0.58</u>
AZT	<u>0.61</u>	0.5	0.53	0.53	0.51	0.52
d4T	<u>0.6</u>	0.56	0.57	<u>0.6</u>	0.56	0.56
ddI	<u>0.64</u>	0.57	0.61	0.62	0.62	0.61
TDF	0.5	0.47	0.49	0.51	0.47	<u>0.51</u>
EFV	0.53	0.54	0.55	0.55	0.55	<u>0.56</u>
ETR	0.77	0.47	<u>0.89</u>	0.75	0.83	0.71
NVP	<u>0.59</u>	0.51	0.56	0.56	0.56	0.53
AFPV	0.55	0.57	<u>0.59</u>	0.57	0.56	0.56
ATV	<u>0.5</u>	0.49	0.44	0.45	0.44	0.45
DRV	<u>0.58</u>	0.5	0.48	0.52	0.48	0.49
IDV	<u>0.6</u>	0.55	0.54	0.54	0.53	0.54
LPV	<u>0.56</u>	0.51	0.51	0.51	0.52	0.53
NFV	<u>0.6</u>	0.58	0.58	0.56	0.57	0.57
SQV	0.59	<u>0.59</u>	0.6	0.58	<u>0.59</u>	0.58
TPV	<u>0.74</u>	0.5	0.6	0.47	0.45	0.57
Mean (SD)	<u>0.58 (0.08)</u>	0.53 (0.04)	0.56 (0.1)	0.55 (0.07)	0.54 (0.09)	0.55 (0.07)

Table 3.25: Misclassification Rates of SIR-Discretized PhenoSense Genotype-Phenotype Pairs (1). The misclassification rate (M) was calculated as the fraction of discordant susceptible-intermediate-resistant (SIR) label pairs obtained with PhenoSense GPPs and DES discretized with pheno cutoffs. The R to S (R → S) misclassification rate states the fraction of resistant-labeled GPPs which were predicted susceptible. Conversely, the S to R (S → R) misclassification rate states the fraction of susceptible-labeled GPPs which were predicted resistant. Misclassification rates for HIVdb and ANRS were computed in the same way. In Table 3.26, misclassification rates for REGA and GRADE are shown. AM: All Models; CD: Common Drugs; NA: Not Available; SD: Standard Deviation. 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	ExposurePheno$_{full}$			HIVdb			ANRS		
	M	R → S	S → R	M	R → S	S → R	M	R → S	S → R
3FTC	0.31	0	0.04	0.13	0	0.02	0.15	0.01	0.02
ABC	0.35	0.02	0.13	0.36	0.05	0.04	0.33	0.05	0.04
AZT	0.35	0.04	0.33	0.28	0.02	0.06	0.27	0	0.15
d4T	0.39	0.09	0.18	0.26	0	0.07	0.35	0	0.24
ddI	0.41	0.1	0.25	0.4	0.1	0.03	0.42	0.29	0.09
TDF	0.52	0	0.26	0.36	0	0.01	0.38	0	0.08
DLV	0.2	0.04	0.01	0.22	0	0.07	NA	NA	NA
EFV	0.14	0.02	0.03	0.18	0.02	0.05	0.2	0.02	0.21
ETR	0.28	0	0.03	0.24	0	0.03	0.22	0.17	0.03
RPV	0.32	0.2	0	NA	NA	NA	0.21	0	0.17
NVP	0.19	0.04	0.06	0.1	0.04	0.06	0.13	0.01	0.15
AFPV	0.25	0	0.03	0.3	0	0.03	0.29	0.18	0.16
ATV	0.21	0	0.02	0.22	0	0.02	0.26	0.14	0.03
DRV	0.14	0	0	0.19	0	0	0.18	1	0
IDV	0.27	0	0.01	0.27	0.02	0.02	0.34	0	0.18
LPV	0.22	0	0.02	0.31	0.05	0	0.33	0	0.01
NFV	0.16	0	0	0.11	0	0.1	0.2	0.03	0.05
SQV	0.28	0	0	0.26	0	0.04	0.28	0.07	0.11
TPV	0.23	0	0	0.3	0	0	0.34	1	0
RAL	0.15	0	0.02	NA	NA	NA	NA	NA	NA
EVG	0.2	0	0	NA	NA	NA	NA	NA	NA
Mean CD (SD)	0.28 (0.1)	0.02 (0.03)	0.08 (0.11)	0.25 (0.09)	0.02 (0.03)	0.03 (0.03)	0.27 (0.08)	0.17 (0.32)	0.09 (0.08)
Mean AM (SD)	0.27 (0.1)	0.03 (0.05)	0.07 (0.1)	0.25 (0.09)	0.02 (0.03)	0.04 (0.03)	0.27 (0.08)	0.16 (0.31)	0.1 (0.08)

Table 3.26: Misclassification Rates of SIR-Discretized PhenoSense Genotype-Phenotype Pairs (2). The misclassification rate (M) was calculated as the fraction of discordant susceptible-intermediate-resistant (SIR) label pairs obtained with PhenoSense GPPs and the genotypic drug-resistance interpretation systems REGA and GRADE. The R to S (R → S) misclassification rate states the fraction of resistant-labeled GPPs which were predicted susceptible. Conversely, the S to R (S → R) misclassification rate states the fraction of susceptible-labeled GPPs which were predicted resistant. Misclassification rates for DES, HIVdb, and ANRS are shown in Table 3.26. AM: All Models; CD: Common Drugs; NA: Not Available; SD: Standard Deviation; 3FTC: lamivudine or emtricitabine, ABC: abacavir, AZT: zidovudine, d4T: stavudine, ddI: didanosine, TDF: tenofovir disoproxil fumarate, DLV: delavirdine, EFV: efavirenz, ETR: etravirine, NVP: nevirapine, RPV: rilpivirine, AFPV: amprenavir or fosamprenavir, ATV: atazanavir, DRV: darunavir, IDV: indinavir, LPV: lopinavir, NFV: nelfinavir, SQV: saquinavir, TPV: tipranavir, RAL: raltegravir, EVG: elvitegravir.

	REGA			GRADE		
	M	R → S	S → R	M	R → S	S → R
3FTC	0.14	0	0.02	0.15	0.01	0.02
ABC	0.36	0	0.04	0.36	0	0.08
AZT	0.26	0	0.08	0.3	0	0.09
d4T	0.27	0	0.11	0.33	0	0.19
ddI	0.38	0.07	0.07	0.41	0.05	0.15
TDF	0.44	0.1	0.08	0.53	0	0.29
DLV	NA	NA	NA	NA	NA	NA
EFV	0.2	0.02	0.19	0.19	0.02	0.06
ETR	0.16	0	0	0.2	0	0.03
RPV	NA	NA	NA	0.32	0	0
NVP	0.1	0.04	0.06	0.09	0.03	0.06
AFPV	0.18	0	0	0.27	0	0.04
ATV	0.16	0.02	0.03	0.24	0.02	0.05
DRV	0.19	0	0.02	0.13	0	0
IDV	0.3	0	0.1	0.29	0	0.03
LPV	0.23	0	0	0.29	0	0.01
NFV	0.11	0	0.08	0.11	0.01	0.08
SQV	0.26	0	0.07	0.26	0	0.06
TPV	0.25	0	0.04	0.27	0.27	0
RAL	NA	NA	NA	NA	NA	NA
EVG	NA	NA	NA	NA	NA	NA
Mean CD (SD)	0.23 (0.1)	0.01 (0.03)	0.06 (0.05)	0.26 (0.11)	0.02 (0.06)	0.07 (0.08)
Mean AM (SD)	0.23 (0.1)	0.01 (0.03)	0.06 (0.05)	0.26 (0.11)	0.02 (0.06)	0.07 (0.07)

Table 3.27: Performance of Prediction of Therapy-Success for Therapies in $\mathsf{T}_{\mathsf{EuResistTE}}$, $\mathsf{TP}_{\mathsf{PRRT}}$, and HIVdbTCE. Therapy-Success (EuResist Standard Datum) was predicted for therapies in the $\mathsf{T}_{\mathsf{EuResistTE}}$, $\mathsf{TP}_{\mathsf{PRRT}}$, and HIVdbTCE datasets with a genetic susceptibility score based on discretized drug-exposure score predictions and that of the interpretation methods HIVdb, GRADE, REGA, ANRS, and $\mathrm{geno2pheno}_{[\mathrm{resistance}]}$ (g2p). Drug-exposure score predictions were discretized with ThMax and ThSucc cutoffs. The highest performance in each row is underlined.

| | ExposurePheno$_{\mathsf{full}}$ | | HIVdb | GRADE | REGA | ANRS | g2p |
	ThMax	ThSucc					
$\mathsf{T}_{\mathsf{EuResistTE}}$ AUC	<u>0.72</u>	0.71	0.7	0.69	0.69	0.69	0.71
$\mathsf{T}_{\mathsf{EuResistTE}}$ C-Index	<u>0.68</u>	0.66	0.64	0.63	0.63	0.62	0.64
$\mathsf{TP}_{\mathsf{PRRT}}$ AUC	<u>0.69</u>	0.68	0.64	0.64	0.63	0.66	0.67
$\mathsf{TP}_{\mathsf{PRRT}}$ C-Index	0.65	<u>0.66</u>	0.6	0.61	0.61	0.59	0.63
HIVdbTCE AUC	0.63	0.64	<u>0.66</u>	0.62	0.64	0.63	0.64
HIVdbTCE C-Index	<u>0.63</u>	0.59	0.57	0.53	0.57	0.55	0.59

PERFORMANCE OF DISCRETE PREDICTIONS FOR THERAPY SUCCESS

DES discretized with ThMax and ThSucc cutoffs were used for calculating a GSS for the compounds used in the therapies recorded in $\mathsf{T}_{\mathsf{EuResistTE}}$, $\mathsf{TP}_{\mathsf{PRRT}}$, and HIVdbTCE. Performance in predicting short-term therapeutic success was assessed using dichotomous therapy-success labels and the AUC. Performance in predicting long-term therapeutic success was assessed using the NAS and Harrel's concordance index. The performances and concordances for predicting therapeutic success for the TEs are displayed in Table 3.27. For comparison, GSS were calculated with other drug-resistance interpretation methods. The resulting performances and concordances can be seen in Table 3.27 as well. On the $\mathsf{T}_{\mathsf{EuResistTE}}$ dataset, the best performance and concordance could be attained with ThMax-discretized ExposurePheno$_{\mathsf{full}}$ DES (AUC = 0.72; $C = 0.68$). Among the drug-resistance interpretation systems I compare to, $\mathrm{geno2pheno}_{[\mathrm{resistance}]}$ displayed the best performance and concordance (AUC = 0.71; $C = 0.67$). The distribution of the GSS calculated with $\mathrm{geno2pheno}_{[\mathrm{resistance}]}$'s predictions is significantly different from that of the GSS calculated with the predictions of the ExposurePheno$_{\mathsf{full}}$ models with ThMax discretization ($p < 10^{-15}$). On the $\mathsf{TP}_{\mathsf{PRRT}}$ dataset, ThMax-discretized ExposurePheno$_{\mathsf{full}}$ DES could attain the best performance (AUC = 0.69), while ThSucc-discretized ExposurePheno$_{\mathsf{full}}$ DES could attain the best concordance ($C = 0.66$). Among the drug-resistance interpretation systems I compare to, $\mathrm{geno2pheno}_{[\mathrm{resistance}]}$ could attain the best performance and concordance (AUC = 0.67; $C = 0.63$). The difference of the GSS distributions estimated with $\mathrm{geno2pheno}_{[\mathrm{resistance}]}$ and ThMax-discretized ExposurePheno$_{\mathsf{full}}$ DES is statistically significant ($p < 10^{-16}$). The best performance on HIVdbTCE is displayed by HIVdb (AUC = 0.66), while the best concordance was displayed by the ExposurePheno$_{\mathsf{full}}$ models with ThMax discretization ($C = 0.63$). The difference of the GSS distributions estimated with HIVdb and ThMax-discretized ExposurePheno$_{\mathsf{full}}$ DES is statistically significant ($p < 10^{-16}$).

3.4.3 DISCUSSION

As mentioned in Section 3.3.3, DES models form a data-driven interpretation system for HIV-1 PR, RT, and IN sequences. The interpretations provided by this system can be used to address three questions: (1) Was a sequence exposed to a certain drug? (2) Is the HIV-1 variant from which the sequence was derived resistant against a certain drug? and (3) Will a certain drug be useful as a component of a therapy against an HIV-1 variant? While continuous DES can be used as an input for higher-level models (Section 4.2), discretization of DES into clinically meaningful categories facilitates their use by experts.

Four different cutoff sets were estimated for solving three different tasks: (1) prediction of drug exposure, (2) prediction of drug resistance, and (3) prediction of therapy success. This maximized predictive performance. On average, ThMax cutoffs are smaller than DEMax cutoffs which, in turn, are smaller than ThSucc cutoffs. Pheno cutoffs are largest, on average. Thus, evidence for drug exposure does not necessarily entail phenotypic drug resistance. At the same time, evidence indicating a certain degree of phenotypic drug resistance (or lack thereof) is not sufficient for predicting the usefulness of a certain drug compound in an antiretroviral regimen. In addition to HIV-1 resistance to the compounds used in a therapy, the determinants of therapeutic success include compound tolerability and potency. This is reflected in distinct drug-wise therapy-success proportions (Table 3.3).

Both ThMax and ThSucc cutoffs were estimated for predicting therapeutic success. However, the methods used for estimating each of these two sets of cutoffs differ in important aspects. First, ThMax cutoffs are optimized to maximize predictive performance on the training set, while ThSucc cutoffs are selected at two values of an estimated DES-dependent success probability. These two values of the DES-dependent success probability were chosen on the basis of empirical drug-wise therapy-success rates (Figure 3.4 and Table 3.3). ThMax cutoffs are slightly better, on average, at estimating therapeutic success. However, the categories into which DES are discretized by ThMax cutoffs could be deemed devoid of clinical meaning by clinical experts, for the following reasons. In the context of expert-guided antiretroviral therapy optimization, a DES above the clinically relevant upper cutoff for a drug means that no clinical benefit is to be expected from including that drug in a therapy (see introduction of Section 3.5). Many upper ThMax Cutoffs (Table 3.20) are negative, while DES are positively correlated with the probability of exposure and with phenotypic drug resistance (Section 3.3.2). Since DES for each drug are identical to the decision values of the SVCs of the individual models, negative DES suggest no drug exposure, positive DES indicate drug exposure, and DES equal to zero lie on the classification boundary (Section 2.2.1). Thus, negative upper cutoffs might be regarded as too low. A further hint at the discordance of ThMax cutoffs with the categories experts are used to can be found in the fact that ThMax cutoffs are, on average, smaller than all other cutoffs, including pheno cutoffs. Pheno cutoffs are optimized to emulate expert-defined SIR categories, and ThSucc cutoffs are closest to pheno cutoffs, on average. Selection of cutoffs at specific values of estimated probabilities of therapeutic success, as done with ThSucc cutoffs, results in categories that are more close to established SIR categories. Second, ThMax cutoffs are optimized for all drugs *jointly*, while ThSucc cutoffs are *individually* estimated for each drug. This characteristic of ThMax cutoffs can foster further deviance from established resistance categories, since thresholds of frequently co-administered drugs can interact with each other. For these reasons, and in light of the small differences in performance, I prefer ThSucc cutoffs over ThMax cutoffs.

In my opinion, the use of three different cutoff sets for solving these three different tasks is justified. ThSucc cutoffs are trained to produce a GSS that reflects the probability of short-term therapeutic success. Thus, information on drug resistance as well as on general drug efficacy is encoded in these cutoffs. Values for discretizing therapy-success probabilities were determined by analyzing therapy-success proportions (Table 3.3 and Figure 3.4). For this reason, these result from the therapy-success rates that are currently feasible, and therefore, these values need to be periodically revised as new antiretroviral drugs are introduced and further evidence for drug efficaciousness is obtained. ThSucc cutoffs were estimated by using a dataset composed of TEs and therapy-failure sequences. Therapy-failure sequences were included in order to compensate for the lack of TEs in which a drug is prescribed although resistance against this drug is present. Correlation of DES with log RFs is weak to strong, depending on the drug in question (Tables 3.14 and 3.15). However, the correlation is sufficient

for predicting the SIR label of GPPs discretized with clinically relevant cutoffs (Table 3.25). While a mean mis-classification rate of 27% seems high, most of errors arise from misclassification of intermediate-labeled GPPs, for which the clinical relevance is uncertain [283]. On average, only 3% of resistant-labeled GPPs are predicted to be susceptible, which is on par with rules-based drug-resistance interpretation systems (Tables 3.25 and 3.26). Misclassification of susceptible-labeled GPPs as resistant is higher for both drug-exposure models and rules-based interpretation systems. However, it is known that certain mutations do not result in *in-vitro* phenotypic resistance, although their presence precludes response to antiretroviral chemotherapy [314].

Performance assessment of the models shows their validity and utility. In comparison to other drug-resistance interpretation systems, discretized DES performance was comparable or significantly superior, depending on the test set. Discretized DES exhibit superior performance when predicting cumulative, long-term therapeutic success (C-index; Table 3.27). The models present a strong difference in performance when predicting therapeutic success with baseline sequences obtained during therapy pauses. Thus, discretized DES use can be recommended when prediction of drug exposure, drug resistance, or therapy success is required. In Section 3.3, I present a novel approach with which a data-driven genotypic drug-resistance interpretation system can be automatically derived from therapy history and GPPs. This is complemented by the methods presented in Section 3.4, with which clinically-meaningful categories from DES can be obtained. Drug-exposure models, along with DEMax and pheno cutoffs, have been integrated into the geno2pheno[resistance] prediction system (http://www.geno2pheno.org).

3.5 Determination of Cutoffs for Predictions of Phenotypic Drug Resistance

In Section 3.4, methods were presented for estimating cutoffs for models predicting drug exposure. In this section, I present a method for estimating clinically relevant cutoffs for geno2pheno[resistance]. This method presents some similarity to ThSucc cutoffs (Section 3.4). In the following, I mention two differences between the method presented in this section and the method for determination of ThSucc cutoffs. First, the method presented in this section uses a weighting procedure in order to correct for the effect of backbone compounds in cART, as well as for the fact that some therapies fail to succeed even if full susceptibility of HIV-1 to the drug compounds is given. Second, the method presented in this section makes use of the form of the sigmoid function for deciding where cutoffs should be selected. In this way, the need for manual selection of cutoff values for estimated probabilities of success is circumvented.

Drug resistance of HIV-1 can be measured *in-vitro* with phenotypic resistance tests (Section 3.1). The output of these tests is the FC between a certain HIV-1 variant and a reference strain. This quantity is called the FC or RF and is also referred to as phenotype, in this context. Statistical models trained on GPPs afford accurate prediction of the RF, given a genotype [299, 315]. The utility of phenotypes in optimizing cART has been established [316–319]. However, phenotypes require interpretation with respect to the *in-vivo* activity of the tested variant, i.e. the range of possible RFs for a drug (called dynamic range) has to be divided into a suitable number of clinically meaningful intervals [318–323]. These intervals are drug-dependent and defined in terms of cutoffs of the RF. Historically, refinement of the interpretation of the RF has undergone several iterations. Initially, a cutoff per drug was defined in terms of the reproducibility of the phenotypic tests (technical cutoffs) as, for instance, mentioned in [324]. Later, the distribution of RFs of (samples from) therapy-naïve patients was used for defining a susceptible-to-resistant RF cutoff for each drug (biological cutoffs) [325, 326]. Finally, it was recognized that the RF, an *in-vitro* measurement, requires explicit translation for its intended *in-vivo* application,

namely the prediction of the suppression of the VL. Efforts in defining clinically relevant cutoffs (short: clinical cutoffs) gave rise to the notion of the division of the dynamic range into the idealized categories susceptible, intermediate, and resistant, using two cutoffs per drug. Susceptible indicates full drug activity, intermediate decreased drug activity, and resistant no drug activity [318, 321, 324]. While drug compounds in cART act in concert, the RF quantifies the activity of a drug *in-vitro* and in isolation. Thus, RF cutoff determination with clinical data requires correcting for the activity of the backbone compounds [318, 321]. This correction can be achieved by using hard-to-obtain (pseudo-) monotherapy data for cutoff determination. Specifically, the addition of an examined drug compound to a failing regimen allows for observing the activity of the drug with reduced influence of the backbone [318, 319, 327]. Correction for the backbone activity can also be achieved mathematically [311, 322, 328]. However, no methodology exists for calculating cutoffs that produce clinically meaningful SIR categories without requiring the expert selection of thresholds for drug activity.

In this section, I establish a novel methodology for calculating clinically relevant phenotypic resistance cutoffs from routine clinical data. Although drug resistance is a continuum, I aim at calculating cutoffs with which the dynamic range can be divided into intervals that best approximate the SIR categories, as defined above. I correct for the activity of the backbone compounds mathematically, as well as for lack of therapeutic success in spite of full susceptibility against the drug compounds in cART.

3.5.1 Materials and Methods

In the following, I present the methods used for estimating cutoffs for the RF predictions of geno2pheno$_{[resistance]}$.

Drugs Considered in this Analysis

The following antiretroviral drugs are considered in this analysis: 3TC, ABC, AZT, d4T, ddI, FTC, TDF, EFV, NVP, APV, ATV, DRV, FPV, IDV, LPV, NFV, SQV, TPV. Other antiretroviral drugs were excluded due to insufficient representation in the EIDB or lack of a model for genotypic drug-resistance interpretation in geno2pheno$_{[resistance]}$ [299] version 3.4.

Data Sources

A total of 36,744 PR and RT sequences from treatment-naïve patients were obtained from LANLSD [329] (http://www.hiv.lanl.gov/; downloaded on March 31st, 2015). HIV-1 data from routine clinical practice was obtained from two sources: the EIDB (http://www.euresist.org; downloaded April 11th, 2014) and the HIVdb TCE repository [330] (http://hivdb.stanford.edu; downloaded November 21st 2013). The EIDB contains data from 66,254 patients, including HIV-1 genotypes, VL measurements, CD4 counts, and compounds used in antiretroviral therapies. The HIVdb TCE repository stores 1,527 TCEs from four data sources, including 58 TCEs from the EIDB. In the context of the HIVdb TCE repository, a TCE documents relevant clinical parameters concerning a change in the drug compounds of cART.

Distribution of Fold-Change for Therapy-Naïve Patients

Nucleotide sequences from LANLSD were aligned, translated, and interpreted with geno2pheno$_{[resistance]}$ 3.4, a data-driven genotypic drug-resistance interpretation system that predicts the RF from the genotype for various PIs and RTIs. Sequences resulting in an alignment error or producing a warning due to missing important sequence regions were discarded. Since I require an RF distribution for susceptible virus variants, sequences

containing at least one major drug resistance mutation [296] were excluded as well, since therapy-naïve patients may carry drug resistant variants [279, 295]. The remaining sequences were used for calculating a therapy-naïve RF distribution for the drugs considered in this study.

Calculation of Probabilities of Susceptibility

Using geno2pheno[resistance] 3.4, all nucleotide sequences in EIDB were aligned. Sequences resulting in an alignment error or producing a warning due to missing important sequence regions were discarded. The remaining PR and RT sequences were translated and interpreted. The resulting RF predictions were used for fitting a sigmoid function for approximating the probability of susceptibility given a predicted RF, as described before (*probability of resistance*; [299]). Briefly, a two-component Gaussian-mixture model is fitted to the RF predictions for each drug. These two Gaussians represent the susceptible and resistant viral populations, respectively. A sigmoid function is then fitted to approximate the probability of resistance, given an RF, that can be calculated with these two Gaussians. I define the probability of susceptibility as one minus the probability of resistance.

Calculation of Probabilities of Success

After application of the TE and therapy success definitions (Section 3.3.1) to EIDB, approximately 10% of the resulting TE were selected at random and set aside for testing purposes. The remaining TEs were merged with TFEs (*Definition of Treatment Failure Episode* in Section 3.3.1) obtained from EIDB, labeled as failures. The resulting dataset was used for estimating RF-conditional probabilities of success for each drug, as follows. TE baseline genotypes were interpreted with geno2pheno[resistance], resulting in predicted RFs for the drugs in each TE. For each drug considered in this analysis, weighted KDE (Section 2.3.1) was applied to the predicted RFs of success-labeled TEs containing the drug in question. Specifically, let $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ be a vector of RFs for a drug and $W = (w_1, \ldots, w_n) \in R_0^+$ a vector of weights for these RFs. Weighted KDE estimates the probability density of an RF distribution at x as

$$\hat{f}(x) = \frac{1}{\sum_{k=1}^n w_k} \sum_{k=1}^n w_k \phi(x - x_k), \tag{3.11}$$

where ϕ is the probability density function of the normal distribution. When estimating the probability of success, I correct for the activity of the backbone compounds of a therapy by down-weighting RFs from success-labeled TEs with the probability of susceptibility of the backbone compounds. The probability of susceptibility of the involved drugs enter (3.11) through an appropriate definition of the weights w_i. For some success-labeled TE containing c drugs, let $d \in \mathbb{R}$ be the predicted RF for the drug in question, let $b = (b_1, \ldots, b_{c-1}) \in \mathbb{R}^{c-1}$ be the predicted RFs for the backbone compounds, and let $\mathrm{p}_{s,i}(x)$ be the probability of susceptibility for a certain drug indexed by i, given its RF. The weights for success-labeled RFs are given by

$$\mathrm{w}_s(d, b) = \frac{1}{2 - \mathrm{p}_{s,c}(d) + \sum_{i=1}^{c-1} \mathrm{p}_{s,i}(b_i)} \tag{3.12}$$

The constant, 2, is used in order to avoid negative weights and division by zero. In the same way, for each drug considered in this analysis, weighted KDE was applied to failure-labeled TEs containing the drug in question. A minority of failure-labeled TEs present small RFs for their drug compounds. Therefore, I correct for lack of success in the absence of resistance by down-weighting failures with the probability of susceptibility of their

drug compounds. Specifically, let $t = (t_1, \ldots, t_c) \in \mathbb{R}^c$ be the predicted RFs for the drug compounds, and let $p_{s,i}(x)$ be the probability of susceptibility for a certain drug given its RF. The weights for success-labeled RFs are given by

$$w_f(t) = \frac{1}{1 + \sum_{i=1}^{c} p_{s,i}(t_i)} \tag{3.13}$$

As in (3.12), 1 is a constant used for avoiding division by zero. For each considered drug, probability densities for RF distributions of success and failure-labeled TEs, $\hat{f}_{s,i}(x)$ and $\hat{f}_{f,i}(x)$, respectively, were obtained with weighted KDE (3.11). For this purpose, weights were calculated with the weight functions described above, $w_s(d, b)$ and $w_f(t)$, for success and failure densities, respectively. $\hat{f}_{s,i}(x)$ and $\hat{f}_{f,i}(x)$ were used for calculating probabilities of success, as follows. The probability of success for a therapy including some drug with RF x is given by

$$P_i(x) = \frac{0.5 \hat{f}_{s,i}(x)}{0.5 \hat{f}_{s,i}(x) + 0.5 \hat{f}_{f,i}(x)}. \tag{3.14}$$

In (3.14), 0.5 is the prior probability (Section 2.1.2). For obtaining a non-informative prior, I chose to give both classes equal weight and thus this term cancels. For a certain drug, I denote probabilities of success for all TE containing the drug with $P_i(x)$. Aiming at noise reduction and analytical determination of cutoffs, I fit the following sigmoid function to the probabilities of success. Let $a, b, c, d \in \mathbb{R}_0^+$ be the parameters of the sigmoid function

$$\hat{P}_i(x) = \frac{a - d}{1 + \exp(-b(x - c))} + d. \tag{3.15}$$

I use non-linear least squares in conjunction with the Levenberg-Marquardt algorithm [331] for fitting the parameters a, b, c, and d to the probabilities of success. (In my experience, the Levenberg-Marquardt algorithm is more robust with respect to different starting solutions than the Gauss-Newton algorithm.) For determining the susceptible-to-intermediate (lower) and the intermediate-to-resistant (upper) cutoffs, I use the roots of the third derivative of (3.15), as these are located at the extrema of the curvature of $\hat{P}_i(x)$. Thus, they are located at RF values at which the probability of success is significantly reduced or marginal, respectively. Cutoff determination for each drug was performed with 1,000 bootstrap replicates [256] of each subset of TEs including that drug. However, for some drugs, this procedure selected the lower cutoff at an RF below the 95th percentile of the RF distribution of therapy-naïve patients. In order to avoid what could be interpreted as overcalling of intermediate resistance, the lower cutoff for a drug was selected either at the smaller root of the third derivative of (3.15) or at the 95th percentile of the RF distribution of therapy-naïve patients for that drug, whichever is larger [327].

Assessment and Comparison of Performance

I assess the performance of the cutoffs calculated with the procedures I used on two test datasets: the first test dataset includes TEs from EIDB which were not used for cutoff determination (approximately 10% of the total). The second test dataset contains TEs extracted from the HIVdb TCE Repository by applying the TE definition to it. For each TE, a GSS was calculated. Specifically, RFs predicted for the drug compounds in the TEs were obtained with geno2pheno[resistance], subsequently producing SIR labels via the cutoffs. Each drug was assigned an integer score, depending on its SIR label: S $\rightarrow$ 1, I $\rightarrow$ 0.5, R $\rightarrow$ 0. The GSS for a TE amounted to the sum of its integer scores. Performance of the GSS in predicting therapeutic success was calculated in terms of AUC.

For performance comparison, the baseline genotypes of the TEs in the test datasets were interpreted with

HIVdb v.6.0.6 [298], resulting in SIR labels for the drug compounds in the TEs. GSS was calculated as described above, and the performance of the GSS in predicting therapeutic success was quantified in terms of AUC as well. p-values were calculated with a two-sided Wilcoxon signed-rank test [308]. Whenever multiple testing was performed, p-values were corrected using the Benjamini-Hochberg method [332].

3.5.2 RESULTS

Among 36,744 nucleotide sequences of the PR and RT genes downloaded from LANLSD, 43 ($< 1\%$) PR and 75 ($< 1\%$) RT sequences were discarded due to alignment problems or sequence-quality warnings. Further 860 (2.3%) PR and 680 (1.8%) RT sequences were discarded because they contained some major drug-resistance mutation. In addition, a total of 74,764 nucleotide sequences from EIDB were submitted to geno2pheno$_{[resistance]}$ for interpretation, some of which did not correspond to either the PR or RT genes. Of these, 21,199 (28%) were discarded due to alignment problems or since they triggered a sequence-quality warning. RF percentiles for the resulting RF distributions for therapy-naïve patients, along with the corresponding probabilities of susceptibility, can be found in Table 3.28. On average, RFs at the 95th percentile of the RF distribution of therapy-naïve patients have a probability of susceptibility of 0.87 with a standard deviation of 0.13. Baseline characteristics for the TEs and TFEs extracted from EIDB and the HIVdb TCE repository can be found in Table 3.29, while the drug-compound distribution for these datasets can be found in Table 3.30. Upper and lower RF cutoffs obtained with the procedure described in Section 3.5.1 are displayed in Table 3.31. Lower cutoffs for the drugs ddI, TDF, NVP, and NFV were replaced by the 95th percentile of the RF distribution of therapy-naïve patients, as they were smaller than this percentile. For performance assessment and comparison, GSS for the TEs in two test sets were computed with predicted RFs, discretized with the obtained cutoffs, as well as with HIVdb discrete predictions. On the EIDB test set, this method and HIVdb performed equally well (AUC = 0.68). However, this method ($\mu = 2.55$; $\sigma = 0.88$) produced lower GSS than HIVdb ($\mu = 2.66$; $\sigma = 0.75$; $p < 10^{-16}$). This method's integer scores for NRTIs ($\mu = 0.83$; $\sigma = 0.35$; vs. $\mu = 0.86$; $\sigma = 0.30$), NNRTIs ($\mu = 0.92$; $\sigma = 0.24$; vs. $\mu = 0.95$; $\sigma = 0.20$), and PIs ($\mu = 0.88$; $\sigma = 0.30$; vs. $\mu = 0.92$; $\sigma = 0.23$) were lower than those produced with HIVdb (corrected $p < 0.0043$). On the HIVdbTCE test set, this method (AUC = 0.63) and HIVdb (AUC = 0.65) performed comparably well. On average, this method ($\mu = 1.82$; $\sigma = 0.97$) produced lower GSS than HIVdb ($\mu = 2.18$; $\sigma = 0.88$; $p < 10^{-16}$) on the HIVdbTCE test set. Integer scores of this method for NRTIs ($\mu = 0.16$; $\sigma = 0.33$; vs. $\mu = 0.21$; $\sigma = 0.37$), NNRTIs ($\mu = 0.19$; $\sigma = 0.39$; vs. $\mu = 0.19$; $\sigma = 0.39$), and PIs ($\mu = 0.06$; $\sigma = 0.23$; vs. $\mu = 0.07$; $\sigma = 0.24$) were also lower than those produced with HIVdb (corrected $p < 0.043$).

3.5.3 DISCUSSION

In Section 3.5, I present a method for calculating clinically-relevant cutoffs for the RF without the need for expert intervention. The method offers the following advantages. (1) The cutoffs produced with this method are tightly coupled with the clinically meaningful definition of lower and upper RF cutoffs for producing SIR categories (see the introduction of Section 3.5). I achieve this by providing a mathematical procedure for deriving the relevant cutoffs. Specifically, the method chooses cutoffs at the curvature extrema of a sigmoid function approximating the probability of success (Figure 3.5). By doing so, I circumvent the need for manual selection of (arbitrary) drug activity thresholds [322]. (2) The method corrects for the activity of backbone drug compounds, but does not achieve this by using coarse, discrete weights, as with GSS [322]. Instead, I correct

Table 3.28: Fold Changes and Probabilities of Susceptibility at Percentiles of the Distribution of Therapy-Naïve Patients. RF: resistance factor; POS: probability of susceptibility; 3FTC: lamivudine or emtricitabine; ABC: abacavir; AZT: zidovudine; d4T: stavudine; ddI: didanosine; TDF: tenofovir disoproxil fumarate; EFV: efavirenz; NVP: nevirapine; AFPV: amprenavir or fosamprenavir; ATV: atazanavir; DRV: darunavir; IDV: indinavir; LPV: lopinavir; NFV: nelfinavir; SQV: saquinavir; TPV: tipranavir.

		5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
3FTC	RF	1.2	1.4	1.7	2	2.5
	POS	1	1	1	1	1
ABC	RF	1	1.1	1.3	1.4	1.7
	POS	1	1	1	0.98	0.75
AZT	RF	0.5	0.9	1.3	1.8	2.9
	POS	1	1	1	1	0.94
d4T	RF	0.9	1.1	1.1	1.2	1.4
	POS	1	1	1	0.99	0.91
ddI	RF	1.1	1.2	1.4	1.5	1.8
	POS	1	0.99	0.97	0.92	0.62
TDF	RF	0.8	1	1.1	1.3	1.7
	POS	1	0.99	0.97	0.9	0.56
EFV	RF	0.3	0.8	1.1	1.6	2.7
	POS	1	1	1	1	0.95
NVP	RF	0.3	0.7	1.4	2.3	5.3
	POS	1	1	1	0.98	0.82
AFPV	RF	0.7	0.9	1.1	1.4	1.9
	POS	1	1	1	1	0.96
ATV	RF	0.8	1	1.2	1.5	2.2
	POS	1	1	1	1	0.93
DRV	RF	0.6	0.8	1.1	1.4	2
	POS	1	1	1	1	0.92
IDV	RF	0.7	0.9	1.2	1.5	2.3
	POS	1	1	1	1	0.95
LPV	RF	0.6	0.7	0.9	1.2	1.7
	POS	1	1	1	1	0.99
NFV	RF	0.7	1	1.2	1.6	2.6
	POS	1	1	1	1	0.94
SQV	RF	0.6	0.8	1	1.3	1.9
	POS	1	1	1	0.99	0.92
TPV	RF	0.5	0.7	1	1.4	2.2
	POS	1	1	1	0.98	0.82

Table 3.29: Therapy Characteristics. For therapies in the development and test sets, the following characteristics are summarized below: baseline viral load, baseline CD4 count, number of recorded past treatment lines, number of first-line therapies, drug-class composition, and genotypic resistance by drug class. IQR: interquartile range; NA: not available; NNRTI: non-nucleoside reverse-transcriptase inhibitor; NRTI: nucleotide/nucleoside reverse-transcriptase inhibitor; PI: protease inhibitor; TE: therapy episode; TFE: therapy-failure episode; VL: viral load.

		EuResist TE Development Set	EuResist TFE Development Set	EuResist TE Test Set	HIVdbTCE TE Test Set
n		8,855	1,589	807	917
Log Baseline VL	Available	7,016 (82%)	NA	765 (95%)	917 (100%)
	Mean	4.4	NA	4.5	4.3
	Median	4.6	NA	4.6	4.2
	IQR	5.2 - 3.8	NA	5.2 - 3.8	4.7 - 3.7
Baseline CD4 Count	Available	6,849 (77%)	NA	670 (83%)	917 (100%)
	Mean	283.4	NA	298.1	290.9
	Median	257	NA	281	258
	IQR	410 - 83	NA	418 - 120	398 - 141
Number of Recorded Past Treatment Lines	Mean	2.8	4.4	2.5	3.5
	Median	1	3	0	3
	IQR	4 - 0	6 - 1	4 - 0	5 - 2
Number of First-Line Therapies		3,315 (37%)	0 (0%)	379 (47%)	0 (0%)
Compound Frequency	NRTI	8,703 (98%)	1,539 (97%)	792 (98%)	905 (99%)
	NNRTI	2,933 (34%)	602 (38%)	264 (33%)	398 (43%)
	PI	5,531 (62%)	754 (47%)	532 (66%)	714 (78%)
Genotypic Resistance	NRTI	3,420 (39%)	1,169 (74%)	243 (30%)	843 (92%)
	NNRTI	2,691 (30%)	809 (51%)	195 (24%)	461 (50%)
	PI	1,558 (18%)	529 (33%)	111 (14%)	553 (60%)
	No Resistance	4,499 (51%)	252 (16%)	485 (60%)	26 (3%)
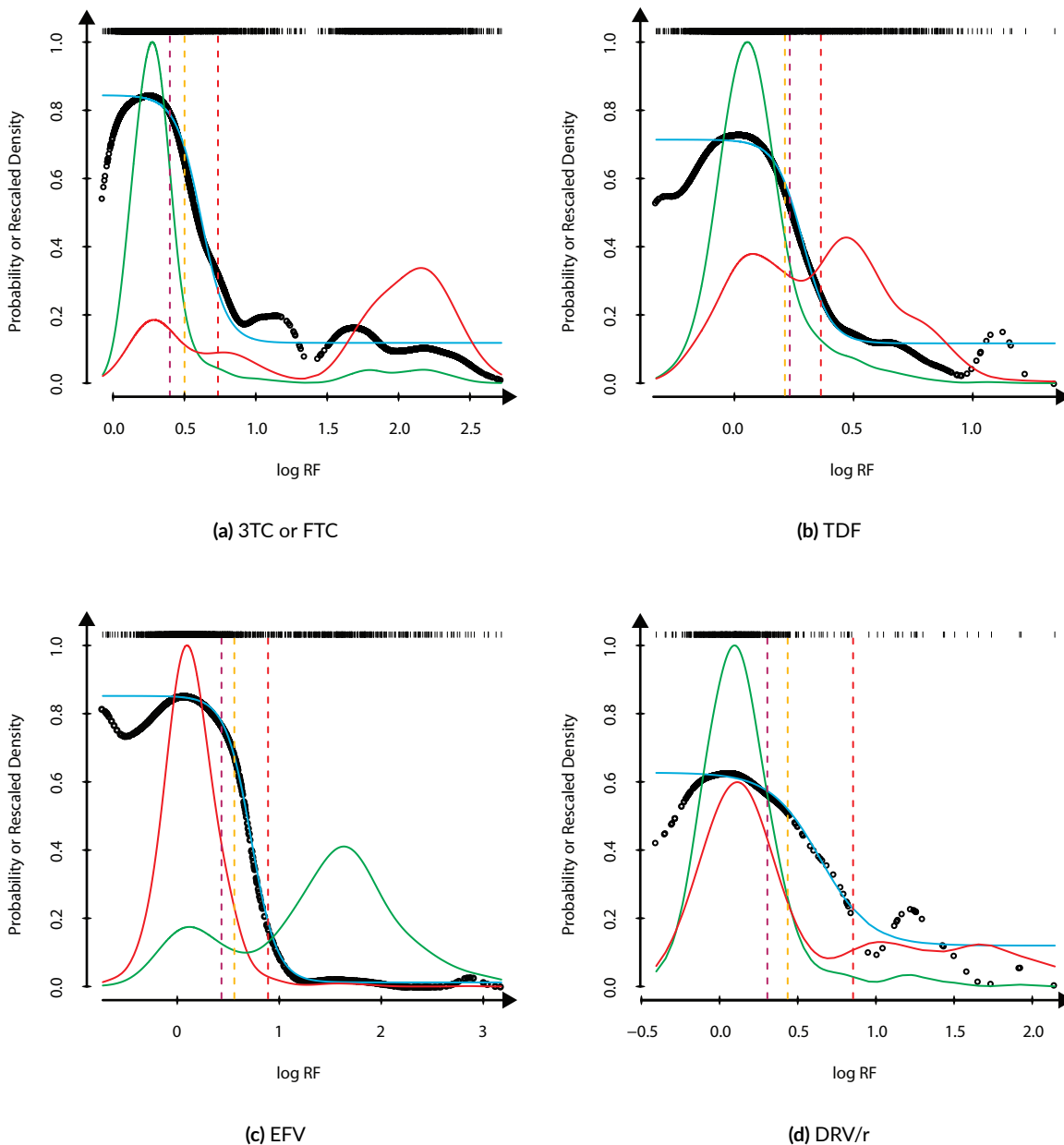
Figure 3.5: Density and Probability Plots (Selected Compounds). Weighted kernel density estimation was used for estimating a success and a failure density for each drug compound. These densities were used for calculating RF-dependent empirical success probabilities, to which a sigmoid function was fitted. A lower and an upper RF cutoff were chosen at the roots of the third derivative of the sigmoid function. However, the lower cutoff was replaced by the 95th percentile of the RF distribution of therapy-naïve patients, if it was lower than this percentile. Success and failure densities, as well as empirical and sigmoid success probabilities are plotted above for the drug compounds lamivudine or emtricitabine (a), tenofovir disoproxil fumarate (b), efavirenz (c), and ritonavir-boosted darunavir (d). 95th percentiles of the RF factor distribution, as well as lower and upper cutoffs determined with the roots of the third derivative are indicated by vertical dashed lines. Resistance-factor density is also depicted as a rug plot at the top of each individual plot. Note that success and failure densities are rescaled to the interval between zero and one in these plots.

Table 3.30: Drug Compound Distributions. The numbers of drug compounds in the development and test sets are tabulated below. TE: therapy episode; TFE: therapy-failure episode; 3FTC: lamivudine or emtricitabine; ABC: abacavir; AZT: zidovudine; d4T: stavudine; ddI: didanosine; TDF: tenofovir disoproxil fumarate; EFV: efavirenz; NVP: nevirapine; AFPV: amprenavir or fosamprenavir; ATV: atazanavir; DRV: darunavir; IDV: indinavir; LPV: lopinavir; NFV: nelfinavir; SQV: saquinavir; TPV: tipranavir

	EuResist TE Development Set	EuResist TFE Development Set	EuResist TE Test Set	HIVdbTCE TE Test Set
3FTC	6,069	1,093	674	484
ABC	1,445	313	157	265
AZT	1,756	525	181	169
d4T	833	463	90	407
ddI	928	396	99	338
TDF	3,669	442	411	307
EFV	1,831	357	204	287
NVP	502	246	60	111
AFPV	336	64	31	117
ATV	926	98	117	76
IDV	136	65	14	103
LPV	2,306	258	244	223
NFV	344	188	39	72
SQV	205	70	27	115
TPV	55	13	7	3
DRV	569	27	63	6

for backbone activity in success-labeled TEs with fine-grained RF weights that depend on the probabilities of susceptibility of all therapy compounds. (3) In failure-labeled TEs, I correct for failures not associated with resistance, which is especially important since I use data from routine clinical practice for calculating the cutoffs.

The computation of RF weights for an antiretroviral therapy with three drug compounds is exemplified in Table 3.32. For failure-labeled TEs, weights decrease with the sum of the probabilities of susceptibility for all compounds in the therapy. The maximum weight, 1, is awarded to RFs from TEs for which the sum of probabilities of susceptibility is equal to zero. Thus, the influence of RFs from failure-labeled TEs is the greater the less susceptible the virus is to the administered drug compounds, which corrects for non-resistance associated failure. In contrast, for calculating RF weights for success-labeled TEs, the probability of susceptibility of the drug in question is subtracted from the sum of probabilities of susceptibility for the backbone drug compounds. The maximum weight, 1, is awarded to RFs of the drug in question indicating full drug susceptibility, while the virus is not susceptible to the backbone drug compounds. Hence, RFs from success-labeled TEs have a greater influence the higher the susceptibility of the drug in question and the lower the susceptibility of the backbone drug compounds. This corrects for the activity of the backbone compounds. For selected drug compounds, a comparison of the densities obtained with and without RF weighting is shown in Figure 3.6. As can be seen, the weighting procedure had the largest influence on the failure densities. On both success and failure densities, the influence of the weighting procedure is largest for older, less effective drug compounds (ddI and IDV/r in Figure 3.6). The definition of SIR categories takes the activity of a drug into account, and (intentionally) does not account for the differential capacities of antiretroviral drugs to elicit virological success (for instance, Table 3.3). For instance, at the same probability of success, the virus may be classified as susceptible to one drug whereas at the same probability level it is classified as intermediate for another drug. Taking this into consideration, cut-

Table 3.31: Clinically Relevant Phenotypic Resistance Cutoffs. For each considered drug, sigmoid functions were fitted to empirical probabilities of success. Lower and upper clinically-relevant fold-change cutoffs for each drug were selected at the roots of the third derivative of the corresponding sigmoid function (sigmoid cutoffs). Below, these cutoffs are tabulated along with the percentile of the RF distribution of therapy-naïve patients to which they correspond (*Sigmoid Cutoff Percentile*). Lower sigmoid cutoffs corresponding to a percentile of the RF distribution of therapy-naïve patients below the 95th percentile were replaced by the 95th percentile of this distribution (underlined). SD: standard deviation; 3FTC: lamivudine or emtricitabine; ABC: abacavir; AZT: zidovudine; d4T: stavudine; ddI: didanosine; TDF: tenofovir disoproxil fumarate; EFV: efavirenz; NVP: nevirapine; AFPV: amprenavir or fosamprenavir; ATV: atazanavir; DRV: darunavir; IDV: indinavir; LPV: lopinavir; NFV: nelfinavir; SQV: saquinavir; TPV: tipranavir.

	Lower Sigmoid Cutoff (SD)	Lower Sigmoid Cutoff Percentile	Selected Lower Cutoff	Upper Sigmoid Cutoff (SD)	Upper Sigmoid Cutoff Percentile
3FTC	3.18 (1.02)	0.99	3.18	5.44 (1.05)	1
ABC	1.78 (1.03)	0.98	1.78	2.97 (1.1)	1
AZT	3.31 (1.06)	0.97	3.31	7.15 (1.1)	1
d4T	1.53 (1.02)	0.99	1.53	2.12 (1.08)	1
ddI	<u>1.46 (1.09)</u>	<u>0.69</u>	<u>1.82</u>	3.32 (1.2)	1
TDF	<u>1.6 (1.03)</u>	<u>0.92</u>	<u>1.71</u>	2.33 (1.03)	1
EFV	3.66 (1.08)	0.98	3.66	7.81 (1.09)	1
NVP	<u>5 (1.08)</u>	<u>0.94</u>	<u>5.34</u>	13.7 (1.08)	1
AFPV	2.69 (1.09)	0.99	2.69	5.93 (1.19)	1
ATV	2.65 (1.1)	0.98	2.65	4.97 (1.21)	1
IDV	3.23 (1.29)	0.99	3.23	9.48 (2.11)	1
LPV	2.44 (1.11)	0.99	2.44	8.52 (1.3)	1
NFV	<u>2.44 (1.08)</u>	<u>0.93</u>	<u>2.6</u>	5.73 (1.17)	1
SQV	2.64 (1.13)	0.99	2.64	6.06 (1.24)	1
TPV	3.1 (2.82)	0.99	3.1	13.86 (4.69)	1
DRV	3.97 (9.36)	1	3.97	12.87 (27.62)	1

Table 3.32: Example Values for RF Weights. Resistance factor weights were calculated as described in Section 3.5.1. For a combination antiretroviral therapy with three antiretroviral drugs, values for RF weighting are tabulated below. Depending on the probabilities of susceptibility for the compounds of a therapy, resistance-factor weights will assume values between those tabulated below. POS: probability of susceptibility; RF: resistance factor; TE: treatment episode.

Sum of Backbone POS	POS for Drug in Question	RF Weight for Success-Labeled TEs	RF Weight for Failure-Labeled TEs
2	1	$1/3$	$1/4$
2	0	$1/4$	$1/3$
1	1	$1/2$	$1/3$
1	0	$1/3$	$1/2$
0	1	1	$1/2$
0	0	$1/2$	1
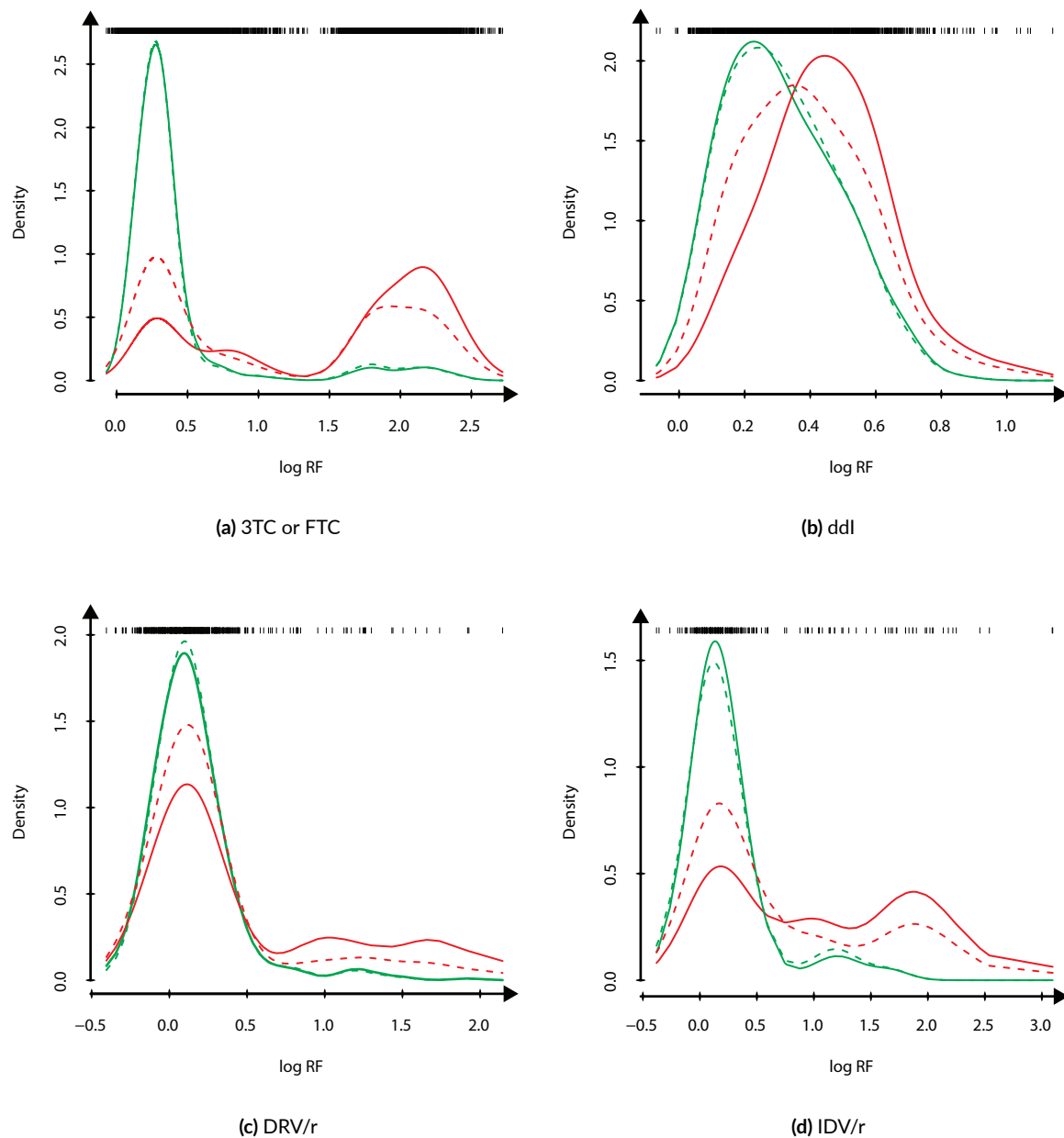
Figure 3.6: Comparison of Weighted and Unweighted Densities (Selected Compounds). Empirical success and failure densities were estimated with weighted kernel density estimation, as described in Section 3.5.1. Above, the weighted densities for the drug compounds lamivudine or emtricitabine (a), didanosine (b), ritonavir-boosted darunavir (c), and ritonavir-boosted indinavir (d) are plotted along with their unweighted versions. Resistance-factor density is depicted as well, as a rug plot at the top of each individual plot.

— Weighted Success Density　— Weighted Failure Density　– – Unweighted Success Density　– – Unweighted Failure Density

off determination via mathematical optimization is hampered by the difficulty in finding an adequate objective function for optimization that is universal for all drugs. In the past, cutoff calculation has been attempted by maximization of a measure quantifying the performance of prediction of therapeutic-success (*Cutoffs Maximizing the Performance of the Prediction of Therapy-Success* in Section 3.4 and [311]). Briefly, cutoffs are iteratively optimized by producing a GSS with a given set of cutoffs. Subsequently these cutoffs are modified such that there is an increase in the AUC. This procedure involves the calculation of cutoffs for all drugs in the dataset simultaneously, and may be less effective in separating the effects of individual drugs. Furthermore, the use of such an objective function is bound to maximize predictive performance in terms of the AUC while producing inconsistencies of the applied SIR categories. I avoid this undesired effect by estimating cutoffs for each drug separately. Nonetheless, the separation of the contributions of individual drugs to therapeutic success remains challenging, as antiretroviral drugs may interact with each other [333, 334]. Lower cutoffs for the drugs ddI, TDF, NVP, and NFV were replaced by the 95th percentile of the RF distribution of therapy-naïve patients, as they were smaller than this percentile (Table 3.31). Before this was done, the lower cutoffs for TDF, NVP, and NFV were above the 91st percentile the RF distribution of therapy-naïve patients, which can be said to be very close to the 95th percentile. The lower cutoff for ddI, however, was originally selected at the 69th percentile of the RF distribution of therapy-naïve patients. Replacement of lower cutoffs below the 95th percentile of the RF distribution of therapy-naïve patients was performed following advice of clinical cooperation partners. They felt that the great majority of therapy-naïve patients without transmitted drug resistance should be awarded *susceptible* predictions for all drugs. I chose to follow this advice for the following reasons. First, the lower cutoffs for the great majority of drugs are above or very close to the 95th percentile of the RF distribution of therapy-naïve patients, such that I felt that following their advice would foster their confidence in geno2pheno$_{[resistance]}$ without significantly altering my findings for these drugs. Second, the data suggest that there are negative differences in the therapy-success rates patients to whom ddI was prescribed and for whom the RF of ddI at baseline is below the 69th percentile of the RF distribution of therapy-naïve patients. ddI is a drug that has fallen into disuse due to its side effects (Section 1.5.5), such that it is nearly impossible to prospectively validate this cutoff in order to find out whether it should be selected at the 69th percentile or at the 95th percentile of the RF distribution of therapy-naïve patients. Thus, I thought that it would be futile to start an undecidable scientific dispute over the correct selection of the lower cutoff of a drug that nobody uses. Except for DRV/r and TPV/r, the standard deviation of the calculated cutoffs is at most 1.5 (Table 3.31). The higher variability of the cutoffs for DRV/r can be explained by the drug's comparatively high barrier to resistance [335]. The dataset that I use contains only few DRV/r-containing TEs with an RF above the upper cutoff (Figure 3.5). Therefore, significantly different numbers of TEs with a high resistance to DRV/r at baseline will be selected across bootstrap replicates, which results in increased variability. The increased variability for the cutoffs for TPV/r is due to the low numbers of TEs containing this drug (Table 3.30). I tested the calculated cutoffs on two test sets. The performance of cutoffs was equal or comparable to that of HIVdb. However, the GSS calculated with the cutoffs were significantly smaller than those calculated with HIVdb. Therefore, the cutoffs have a higher tendency to label an RF with non-susceptible SIR categories than HIVdb. I interpret this to be a result of the strict adherence to SIR categories in the cutoff estimation procedure. Specifically, genotypic drug-resistance interpretation systems produced by experts may take the propensities of individual drugs to elicit therapeutic success more into account. This results in fewer resistant predictions for drugs with a high potency and a high barrier to resistance.

This method for cutoff determination did not lead to a greater predictive performance when compared to

HIVdb; the predictive performance of geno2pheno$_{[resistance]}$ and the cutoffs was *comparable* to that of HIVdb. When interpreting this fact, one should bear in mind that the experts who created and update the rules-based interpretation system *HIVdb* use more sources of information for crafting the rules than are available for training of geno2pheno$_{[resistance]}$. Specifically, the rules of HIVdb are based on HIV-1 nucleotide sequences with drug exposure information, GPPs, and clinical data on the baseline characteristics and outcome of cART. In contrast, geno2pheno$_{[resistance]}$ only uses GPPs. The method presents the following advantages. (1) It provides a valid method for cutoff selection that does not require expert intervention. In the past, retraining of geno2pheno$_{[resistance]}$ almost inevitably entailed calling for a meeting of clinical experts for the purpose of cutoff determination. Thanks to this method for cutoff determination, this laborious and time-consuming step can be now avoided. (2) Intuitive plots of the procedure leading to the selection of cutoffs can be generated. This can foster the confidence of clinical experts in the cutoffs, since the plots substantially facilitate the understanding of the method. (3) In the future, the method could be applied for understanding the role of different drugs in cART.

4

Improving the Prediction of Therapy Success

Over the last 15 years, cART for treating HIV-1 infection has improved remarkably [336–339]. This is due to the continuous acquisition of knowledge on how to treat HIV-1-infected patients [228, 280, 340], the systematic collection of data on the interplay between applied drug therapy and evolutionary response of the virus [294, 329, 341], the introduction of novel drug compounds with improved potency [335], growing control over side-effect profiles [342], increased understanding of the genetic barrier to drug resistance [343, 344], and increased intake convenience [345]. In resource-rich settings, the risk of therapy failure has therefore decreased steadily over the last 15 years [339, 346]. Nowadays, selecting a cART that will achieve initial virologic response has become significantly more effective [207, 347–350]. However, the selection of durable antiretroviral regimens still remains challenging. For this reason, this chapter is devoted to statistical models for predicting the success (or failure) of antiretroviral chemotherapy.

I begin this chapter by reviewing the goals of antiretroviral therapy, from which several therapy-success definitions have been derived. Then, I review past work on computational systems that can support the treating clinician in the process of antiretroviral therapy selection by predicting the success of individual therapy combinations. Lastly, I present a model that represents a further development of these systems.

4.1 Definitions of Therapy Success

A summary on how therapy success is defined from the perspective of a treating physician can be found in the following. This summary is based to a large extent on [3, 228, 229, 280]. In Section 4.1, citations only refer to sources other than these ones.

4.1.1 The Goals of Antiretroviral Therapy

At current, eradication of HIV-1 is not possible for the large majority of infected patients. However, HIV-1 infection can be kept under control with cART, under the premise that treatment is effective and that antiretroviral drugs are continuously taken by the patients for the rest of their lives. Section 1.5 reviews the history of cART, gives an overview on antiretroviral drug compounds, and summarizes current recommendations for

cART. In the following, I will summarize the goals of cART that relate to treatment of chronic HIV-1 infection.

For determining the success of cART, virological, immunological, and clinical criteria are used. The primary marker for determining the virological success of cART is the VL. When treatment is successful, the VL declines in two phases. During the first phase, a rapid decline of the VL can be observed. In the second phase, this decline slows down. The greater the initial decline in VL, the more durable the response to cART. Although patients can profit from any decline in the VL, virological success of a cART is generally defined as a reduction of the VL below 50 HIV-1 RNA copies per milliliter of blood serum. This VL threshold is somewhat arbitrary, as it is based on the detection threshold of wide-spread VL assays. Historically, the VL threshold for determining treatment success has been decreased several times as a result of the availability of more sensitive VL assays and of more potent drug combinations. Currently, reduction of the VL below 50 HIV-1 RNA copies per milliliter of blood serum is often simply referred to as *suppression of the VL*. A patient with HIV-1 infection who has reached this state is often referred to as being *(virologically) suppressed*.

Immunological treatment success is generally defined as an increase in the CD4+ cell count. For determining immunological treatment success, no definitions on the extent of CD4+-cell count increase have been established. Thus, immunological success definitions vary from study to study. These range from an increase in the CD4+-cell count of 50 to 200 cells per microliter of blood (when compared to baseline) or to levels above 200 or 500 cells per microliter of blood (presupposing that they CD4+-cell counts were lower at baseline). Immunological failure is considered to be the lack of increase or the decrease of the CD4+ cell count. Immunological treatment success usually follows virological treatment success. However, the incidence of immunological treatment success is subject to variability among individuals. On the one hand, early treatment initiation is associated with larger increases in CD4+-cell counts. On the other hand, some baseline factors are negatively correlated with a future increase in the number of CD4+ cells:

- Low CD4+-cell count
- High VL
- Old age
- Bad condition of the thymus
- Myelo- and immunosuppressive concomitant therapies
- Autoimmune diseases
- Liver cirrhosis.

When compared to virological success, immunological success is more difficult to influence. The patient profits from any reduction of the VL, such that cART may elicit immunological success even if the VL stays above 50 HIV-1 RNA copies per milliliter of blood serum. Nonetheless, this precludes virological success, per definition.

Clinical treatment success usually follows immunological success. In studies of antiretroviral treatment, clinical treatment success is usually evaluated via clinical end points, i.e. the lack of occurrence of AIDS-defining illnesses (see [1] for a list of these illnesses) or of death. However, other patient-specific characteristics of the outcome of the treatment are important as well, such as the improvement of constitutional symptoms. The lack of disease progression should also be regarded as clinical success. On the contrary, the onset of serious side effects due to medication should be clearly regarded as treatment failure. Less severe side effects are important as well, but are not sufficient to determine clinical failure.

Summarizing, treatment success is determined with virological, immunological, and clinical criteria. Immunological success usually follows virological success. Nevertheless, immunological success is more difficult

to influence than virological success. Clinical success usually follows immunological success. Since a suppressed VL stands at the beginning of a potential virological, immunological and clinical success cascade, the VL is the most important parameter for the determination of treatment success. Thus, the primary goal of antiretroviral therapy is the sustained reduction of the VL below 50 HIV-1 RNA copies per milliliter of blood serum.

4.1.2 Blips and Low-Level Viremia

While using cART, suppression of the VL below 50 HIV-1 RNA copies per milliliter of blood serum may be interrupted by transient, relatively small increases of the VL. For these transient VL increases to be considered blips, they are constrained, by definition, to reach less than 200 HIV-1 RNA copies per milliliter of blood serum. Little is known about the causes of blips. They have been not conclusively associated with compliance, a certain drug combination, resistance, or other clinical data. Their frequency seems to be negatively correlated with early treatment initiation, which is why it has been speculated that they could result from the interaction between immunological mechanisms and latent viral reservoirs (Section 1.1.2). Concomitant infections are known to cause blips, which probably results from the immune activation they cause. While there does not seem to be an association between blips and treatment failure, transient viremia above 200 HIV-1 RNA copies per milliliter of blood serum should be a matter of concern for the treating clinician.

Blips should be distinguished from low-level viremia. Low level viremia is defined as persistent HIV-1 replication leading to a VL below 1,000 (or sometimes constrained to be below 500) HIV-1 RNA copies per milliliter of blood serum. The prognostic value of low-level viremia is not clear, although some publications shed some light on the topic [351–353]. Apparently, low-level viremia is associated with treatment failure (see below), but not with disease progression or mortality. The detection of drug-resistance mutations during low-level viremia seems to be an indicator for imminent treatment failure. Nonetheless, HIV-1 treatment guidelines do not clearly state which procedure a clinician should undertake when low-level viremia is detected in a patient [228, 229].

4.1.3 Specific Features of First-Line Therapies

The first antiretroviral therapy that an HIV-infected person receives presents the best chance of prolonged success. This goal can be hampered by TDR, and factors negatively influencing compliance, such as side effects and the complexity of the drug regimen [354–357]. Therefore, a first-line antiretroviral therapy must be selected under consideration of these determinants. Antiretroviral drug combinations require approval by regulatory institutions in order to be used as first-line therapies. This results in a small number of compound combinations accounting for the majority of first-line antiretroviral therapies (Table 1.11). When TDR is detected, a PI- or DTG-containing therapy is recommended, due to their high genetic barrier [204–206, 358–360]. Otherwise, the low pill count of regimens containing an NNRTI is often preferred (even one pill a day is possible [355]), leaving the PI and INI options open.

The transmission of a resistant viral strain can lead to an earlier therapy failure in first-line antiretroviral therapies. Available aids for selection of first-line antiretroviral therapy are limited. Suitable first-line antiretroviral therapies are recommended by various guidelines, summarized in Table 1.11. Therapies in this table have to be narrowed down by using mutation tables [279, 296] or interpretation systems (Section 3.2) in order to exclude potentially ineffective drug combinations. For the selection of antiretroviral therapy combinations, bioinformatics systems are available (Section 4.2). However, these systems have been designed for therapy-experienced patients and may over- and also underestimate the short- and long-term efficacy of drug combi-

nations in therapy-naïve patients. In the case of TDR, single drug-resistance mutations have been anecdotally reported to be indicative of other drug-resistance mutations that are present in viral minority populations that do not show up in the consensus sequence of the patient sample (Section 5.1) [314, 358, 360–362]. However, sequencing of minority viral populations in patients with single TDR mutations has rarely revealed additional drug-resistance mutations, suggesting that the aforementioned scenario may also be rare [352, 363–366]. In addition to the viral activity of a regimen, its tolerability is of particular importance for the success of a first-line antiretroviral therapy and has to be incorporated into a prediction system.

4.1.4 Modification of Antiretroviral Therapy Due to Failure

The most frequent situations that justify the modification of antiretroviral therapy are threefold:

1. Acute side effects
2. (Potential) long-term toxicity
3. Virological treatment failure.

Even modern cART is associated with a number of toxicities (Section 1.5.1). Common side effects following treatment initiation include mild forms of nausea, diarrhea, allergic reactions, and CNS disorders. Usually, the treating physician can intervene in order to mitigate or even eliminate these initial side effects. Modification of antiretroviral therapy is only justified by severe side effects. These include:

- Severe untreatable diarrhea
- Severe untreatable nausea
- Persistent sleeping disorder
- Polyneuropathy
- Severe anemia
- Severe, progressive muscular weakness
- Pancreatitis
- Lactic acidosis
- Severe allergies
- Renal failure
- Hepatotoxicity
- Jaundice
- Rhabdomyolysis (rapid breakdown of skeletal striated muscle, potentially resulting in kidney damage due to the released breakdown products)
- Depression, psychosis.

Long-term toxicities are a special case of side effects. They are cumulative in nature and appear after long-term use of antiretroviral compounds. The most frequent long-term toxicities are lipodystrophy (Figure 1.14) and dyslipidemia (abnormal quantities of lipids in blood). Since some compounds have a higher propensity to cause long-term toxicities than others (*Improving Combination Antiretroviral Therapy* in Section 1.5.1), treating clinicians may be motivated to modify virologically successful therapies hoping that the patient will be spared from long-term toxicities.

Virological treatment failure justifies modification of antiretroviral therapy. Virological treatment failure is defined as *repeated* VL measurements above 50 HIV-1 RNA copies per milliliter of blood serum. However,

transient viremia (below 1,000 HIV-1 RNA copies per milliliter of blood serum) is not considered virological failure (Section 4.1.2). Virological failure can result from lack of treatment adherence by the patient, malabsorption of the drug compounds, insufficient dosing, drug-drug interactions, or by the development of drug resistance by the virus (Section 1.5.3). In any case, therapy failure is a matter of concern, as replication in the presence of subinhibitory concentrations of antiretroviral compounds quickly lead to the development of (a higher degree of) drug resistance, potentially compromising future therapy options. The amount of time a physician may wait before reacting to virological treatment failure depends upon the drug combination used in the failing therapy. While NNRTIs and first-generation INIs are more prone to resistance development, PIs and DTG have a higher genetic barrier to resistance (Section 1.5.3), giving the physician more time to perform an intervention.

4.2 Predicting Therapy Success with Statistical Models

The presence of drug-resistant HIV-1 variants in a patient impedes virologic response to cART (Section 1.5.3) [199, 228, 276, 280]. For this reason, the use of genotypic drug-resistance interpretation systems has become the standard of care when selecting drug compounds for cART prescription (Section 3.2) [367]. These systems interpret the genotypes of the viral genes whose protein products are targets of antiretroviral drugs. Their output consists of one out of maximally six categories indicating various degrees of resistance [297, 298], or a continuous score for drug resistance, such predicted RFs (Chapter 3). Viral variants presenting some degree of resistance against at least one drug compound in cART are likely to obstruct (long-lasting) virologic response. However, the absence of genotypic drug resistance at baseline does not guarantee therapeutic success. Even in the presence of full drug susceptibility, individual drugs and drug combinations differ in their propensities to achieve (long-lasting) therapeutic success (Table 3.3) [354–357]. Drug-resistance interpretation systems do not consider these propensities and rate drugs solely with respect to clinically relevant drug resistance.

Therapy-success prediction systems aim at predicting the suppression of the VL below some threshold by a certain cART *in vivo* [300, 368–375]. Therapy-success prediction systems are trained on clinical HIV-1 data, mainly arising from routine clinical practice. Available therapy-success prediction systems differ in several respects. In the following, I comment on (1) the patient group they target, (2) the input that they require for making a prediction, (3) the type of prediction they deliver, and (4) the predictive performance that has been reported. Some therapy-success prediction systems are constrained by design to predict therapeutic success on certain groups of patients, e.g. pretreated patients with drug-resistance mutations at baseline [370]. The minimal input that is required for obtaining a prediction varies considerably from system to system. Most therapy-success prediction systems require a genotype as an input [300, 368, 371, 374], while some do not require it [376–378]. Further necessary variables may include date of birth and sex of the patient, a recent VL measurement and CD4+-cell count, the complete or partial treatment history of the patient, and a list of drug combinations that are admissible or considered for therapy [368, 370, 374]. The target of prediction of all therapy-success prediction systems is the VL (either as a continuous value or discretized in some manner). However, the point in time, relative to treatment initiation, at which this prediction is targeted, varies from system to system. This point in time might be fixed [300], selected by the user [370], or undefined [373, 374]. One published therapy-success prediction system does not aim at predicting whether the VL will be suppressed at some point in time, but rather, how long will it take for virologic failure to occur [368]. The predictive performance reported by the authors of available therapy-success prediction systems is frequently expressed in terms of the AUC. The

reported performances of different therapy-success prediction systems have been assessed with several different treatment-success definitions and datasets, which makes direct comparison of the performance of the different systems difficult.

4.2.1 Materials and Methods

In the following, I describe five related, yet different models for therapy-success prediction. The similarity of these models consists in the fact that all of them are trained on subsets of the same dataset. The models differ in the definition they use for determining therapeutic success. Specifically, three of these models use only one VL measurement for determining therapeutic success, each one at a different point in time. One of these models uses at most two VL measurements (baseline and follow-up). The remaining model uses all VLs measured during a therapy for quantitatively defining therapeutic success. I validated these models on two independent datasets. While doing so, I tested the capability of the models to predict short-, mid-, and long-term therapeutic success. Furthermore, I discriminate between their performance with therapy-naïve and therapy-experienced patients. When designing these models, I chose to require solely a PR and RT genotype as the minimal input, for the following reason. While further input variables may improve predictive performance of therapy-success prediction systems, they can also represent a limitation for their use, if these variables are unknown. Last but not least, the methods used for training the models are linear, which allows them to deliver interpretable weights.

Drugs Considered in this Study

The following antiretroviral drugs are considered by the models described in this section: 3TC, ABC, AZT, d4T, ddC, ddI, FTC, TDF, EFV, ETR, NVP, APV, ATV, DRV, FPV, IDV, LPV, NFV, SQV, TPV, EVG, DTG, and RAL. Other antiretroviral drugs were excluded due to insufficient representation in the therapies extracted from EIDB [300].

EuResist Integrated Database and HIVdb TCE Repository

EIDB [300] (http://www.euresist.org, downloaded April 11th, 2014) contains HIV-1 data from routine clinical practice. The downloaded version of EIDB contains data from 66,254 patients, predominantly from Europe, but also including 865 patients from Rwanda. Available clinical data types include: HIV-1 genotypes, VL measurements, CD4+-cell counts, compounds used in antiretroviral therapies, among others. The HIVdb TCE Repository [330] (downloaded November 21st, 2013) archives 1,527 TCEs from four data sources, including 58 TCEs from the EIDB. In the context of the HIVdb TCE Repository, a TCE contains relevant clinical parameters documenting the outcome of a change in the drug compounds of cART.

Definition of Treatment Episode

The following definition of treatment episode is essentially the same as the one found in Section 3.3.1. However, the inclusion of INIs made slight modifications of the definition necessary. For the sake of clarity and of the comfort of the reader, I repeat the definition in the following. TEs differ from TCEs [302] in that no treatment change is required, encompassing first-line therapies as well. A TE consists of a set of drug compounds used in a therapy, a baseline RT and PR genotype for that therapy, and a sequence of dated VL measurements obtained during the therapy. TEs in which INIs are used were required to include either an IN genotype or to indicate whether the corresponding patient had used INIs in therapies predating the TE. Inclusion criteria for TEs are

as follows. (1) The therapy must have lasted at least four weeks, because shorter therapies might not allow for sufficient manifestation of the effects of the drugs. (2) The drug compounds in the TEs must not include drugs that are not considered in this study. TEs with unboosted PIs (except for NFV) were excluded as well [334]. (3) Baseline genotypes must have been measured before treatment initiation, choosing the genotype closest to the treatment-initiation date. For treatment-experienced patients, baseline genotypes were required to have been measured no earlier than 90 days before treatment start. In accordance with the EuResist Standard Datum (`http://www.euresist.org`), no treatment lasting for longer than 14 days may have taken place during this period. (4) Baseline genotypes must comply with previously described quality criteria (Section 3.3.1) (5) TE VL measurements must have been performed with an assay whose lower limit of quantification is 50 HIV-1 RNA copies per milliliter of blood serum or lower, as this has become standard [228, 280]. I require that TE VL measurements must have taken place regularly after treatment initiation. A minimum VL-monitoring frequency of once per semester (26 weeks) is desired [228, 280], with a tolerance of at most 10% of the total number of recorded therapy semesters not having a VL measurement.

Dichotomization of Therapies into Successes and Failures

I dichotomize therapies into successes and failures with two different definitions. The first definition stems from the Standard Datum of the EIDB (Section 3.3.1). This definition requires one follow-up VL, and optionally, one baseline VL. Follow-up VLs must have been obtained between four and twelve weeks after therapy initiation, preferring VL measurements whose measurement date is closest to eight weeks after therapy start. With regard to the time at which they were measured, baseline VLs must fulfill the same requirements as baseline genotypes (see (3) above). TEs with a follow-up VL below 400 HIV-1 RNA copies per milliliter of blood serum or presenting at least 100-fold reduction in the VL are labeled as successes. TEs for which a baseline VL is available, the follow-up VL is above 400 HIV-1 RNA copies per milliliter of blood serum, and the VL reduction is less than 100 fold are labeled as failures [300, 302]. I call these success and failure labels the EuResist labels.

The second definition I use is solely based on VLs obtained at a specific number of weeks following treatment initiation. A week-w VL, $w \in \mathbb{N}$, is defined as a VL obtained during the course of a therapy, no earlier than $w - 4$ weeks, and no later than $w + 4$ weeks after its initiation. If several VL measurements are performed during this period of time, the measurement closest to w is chosen. I label a TE as a week w success if a week-w VL measurement is available, and the measured VL is below some threshold t in the number of HIV-1 RNA copies per milliliter of blood serum. Conversely, if the week-w VL is above t, then I label the TE as a week-w failure. I call such labels week-w labels with threshold t.

Computation of the Numbers of Aviremic Therapy Semesters

The NAS is a measure that accounts for long-term therapeutic success. This measure is described in detail in Section 3.3.1. In summary, VL measurements performed during a therapy are averaged for each therapy semester (comprising 26 weeks). A threshold for the average VL is then used to dichotomize therapy semesters into viremic or aviremic. Counting of aviremic semesters results in the NAS. Two criteria are used for determining right-censoring of therapies: the intent-to-treat (ITT) and the on-treatment (OT) censoring criteria. An NAS is marked as right-censored by the ITT criterion if the therapy has not been terminated at the time when it was stored in the EIDB or no VL measurements are available for some therapy semester. The OT censoring criterion additionally marks an NAS as right-censored if it was terminated while the VL was below the selected threshold.

Definition of Therapy Failure Episode

Therapy-failure episodes have been defined in Section 3.3.1. For the sake of the comfort of the reader, I repeat the definition in the following. A TFE consists of a set of drug compounds used during a therapy, and a PR and an RT (and optionally also an IN) genotype obtained during the therapy. The therapy is required to have lasted at least four weeks, and the genotype is required to have been obtained no earlier than four weeks before therapy stop.

Training of Drug-Exposure Models

Statistical models for interpreting genotypes with respect to drug exposure and drug resistance were computed as described in Section 3.3. In summary, these models are trained to discriminate between HIV-1 genotypes derived from patients who have taken a certain drug and those who have not taken that drug. Before the models are trained, HIV-1 genotypes from routine clinical practice are complemented with genotypes from GPPs, treating genotypes with RFs below a certain threshold as not drug exposed, and those with RFs above another threshold as drug exposed. Interpretation of a genotype with one of these models results in a number called DES which relates to a specific drug. In order to normalize DES, a sigmoidal function was fitted to each one of the models, as described before [309]. The sigmoidal function aims at estimating the probability that a certain genotype belongs to the class of genotypes exposed to a certain drug, given the DES for that genotype and drug (short: probability of exposure to a drug; POE). One drug-exposure model delivers the probability that a certain genotype was obtained from a treatment naïve patient (short: probability of naïvety; PON). As described in the following, POEs and the PON were used as input features in models for the prediction of therapeutic success.

Preparation of Development and Test Set

The TE definition was applied to EIDB and HIVdbTCE. A random selection of approximately 10% of the TEs from EIDB was set aside as a test set, which I call EuResist$_T$. The random selection was performed such, that the inclusion of a TE from a patient in EuResist$_T$ triggered the inclusion of all available TEs from that patient. TEs containing the drugs ddC, RPV, APV, EVG, and DTG were not eligible for inclusion in EuResist$_T$, as ddC and APV are not produced anymore, and as RPV, EVG, and DTG are insufficiently represented in the EIDB. Thus, I train on TEs including these drugs in order to exploit the information on co-administered drugs that the TEs contain, but I do not test on them or provide predictions for them (except for an example for discussion). A further, independent test set comprised TEs in the HIVdb TCE repository not arising from EIDB. I call this set of TEs HIVdbTCE. The development set EuResist$_D$ was created from all TEs from EIDB which are not included in EuResist$_T$. A dataset containing therapy failures, TF, was created by applying the TFE definition to EIDB while excluding patients with TE in EuResist$_T$.

Drug-exposure models were trained as described in Section 3.3. However, in order to avoid overtraining [256], cross-validation partitions were constructed such, that the HIV-1 genotypes contained in the different folds are patient-wise disjoint, i.e. that all genotypes derived from a given patient are contained in the same fold. Furthermore, genotypes from patients with TEs in EuResist$_T$ were excluded from the development sets of the drug-exposure models. POEs and the PON for each RTI and PI were calculated for the genotypes in EuResist$_D$ and TF. POEs for INIs were calculated for TEs including an IN genotype. Since EuResist$_D$, TF, and the datasets on which drug-exposure models were trained are not patient-wise disjoint, POE calculation

was performed with cross-validation models, such that the prediction model was not trained on any sequence derived from the patient associated to a given TE or TFE. Since some TEs included INIs without including an IN baseline genotype, INI-use history was used as a surrogate for POEs to INIs. However, INI-use history was predicted for TEs including an IN baseline genotype, assuming drug exposure if the POE for the drug in question was greater than 50%. For training, therapy drugs and INI history were encoded in binary.

TRAINING OF MODELS FOR PREDICTION OF DICHOTOMOUS LABELS

TEs were dichotomized into successes and failures, using EuResist labels, and week-w labels with threshold t, for $w \in \{8, 24, 48\}$ and for $t \in \{50, 100, 200, 400, 800, 1600\}$. TFEs in TE were labeled as failures and merged with EuResist$_D$. Subsequently, SVCs (Section 2.2.2) were trained to predict success or failure labels with probabilistic outputs [309]. The input features of these models consist of the binary encoding for the therapy drugs and for INI use history, as well as the POEs and the PON. For each of the used success-failure definitions, EuResist$_D$ $\cup$ TF was randomly partitioned into ten folds, such that each fold was patient-wise disjoint. The partition was used to calculate the cross-validation predictions for therapy-success labels. In order to incorporate interactions of the input features, SVMs with a polynomial kernel (Section 2.2.3) were used. All combinations of the following two SVM parameters were tested: the degree of the polynomial kernel $p \in \{1, 2, 3\}$ and the complexity parameter $\gamma \in \{2^{-14}, 2^{-13}, \ldots, 2^{-2}\}$ (Sections 2.2.3 and 2.2.2).

TRAINING OF MODELS FOR PREDICTION OF THE NAS

For quantifying cumulative therapeutic success, the NAS for each TE in EuResist$_D$ was calculated, using the ITT censoring criterion, and the following average VL thresholds: 50, 100, 200, 400, 800, and 1,600 HIV-1 RNA copies per milliliter of blood serum. TFEs in TE were assigned zero uncensored aviremic semesters, and merged with EuResist$_D$. SVMs for the regression of right-censored data (Section 2.2.5) were trained to predict the NAS. The same input features encoding the therapy drugs, the INI use history, the POEs and the PONs were used as in the models for prediction of dichotomous labels (above). EuResist$_D$ $\cup$ TF was randomly partitioned into ten folds, such that each fold was patient-wise disjoint. The partition was used to calculate the cross-validation predictions for the NAS. As in the models for predicting dichotomous labels, SVMs with a polynomial kernel (Section 2.2.3) were used. All combinations of the following three SVM parameters were tested: the degree of the polynomial kernel $p \in \{1, 2, 3\}$, the complexity parameter $\gamma \in \{2^{-14}, 2^{-13}, \ldots, 2^{-2}\}$, and the insensitivity parameter $\epsilon \in \{2^{-10}, 2^{-9}, \ldots, 2^{-2}\}$ (Sections 2.2.3 and 2.2.5). Furthermore, the regression of logarithmized NAS was tested as well.

MODEL SELECTION

For model selection, Harrell's concordance index (C; Section 2.2.6) [267, 268, 368] was applied for assessing cross-validated predictions, i.e. even models trained on dichotomous success-failure labels were assessed using the NAS. This yielded cross-validated performances for each parameter combination and therapy success definition used for training. Among all trained models, one was selected for each therapy success definition. In order to avoid overtraining (Section 2.1.1), models were selected as follows. For each therapy-success definition, p-values comparing the performance of the best performing parameter combination with that of all other parameter combinations were computed with a one-sided Wilcoxon's signed-rank test [308]. After p-value correction with the Benjamini-Hochberg method [332], models performing significantly worse than the best

performing model were discarded. The significance threshold used was 0.05. Among the remaining models, the model with the minimum p and minimum γ (and maximum ϵ, if applicable) parameters was selected.

Training of Final Models

With the procedure described above, parameters for training each final model were selected. Final models were trained with the entire development set. In light of the obtained cross-validation performances, and since a VL threshold of of 50 copies per milliliter of blood serum has become standard (Section 4.1) [228, 280], I only trained final models for this VL threshold, in addition to the models trained with the EuResist labels, which are created using a fixed VL threshold (Section 3.3.1). For final model training, POEs and PON were not calculated on cross-validation models, but on drug-exposure models trained on sequences from patients not in EuResist$_T$. Furthermore, the SVM was not trained with a polynomial but with a linear kernel in order to afford interpretability. The polynomial kernel was replaced by explicitly calculating all interactions of the input features, such that the inner products in the enlarged feature space are equivalent to the output values of the polynomial kernel in the original feature space. In the following, I include an abstract example of how the feature interactions implicitly considered by a polynomial kernel can be explicitly calculated [256]. Let our example feature be two-dimensional with inputs x_1 and x_2. Consider the polynomial kernel of degree two, $k(x, x') = (\langle x, x', \rangle + 1)^2$. Then,

$$
\begin{aligned}
k(x, x') &= (\langle x, x' \rangle + 1)^2 \\
&= (1 + x_1 x_1' + x_2 x_2')^2 \\
&= 1 + 2x_1 x_1' + 2x_2 x_2' + (x_1 x_1')^2 + (x_2 x_2')^2 + 2x_1 x_1' x_2 x_2'.
\end{aligned}
\tag{4.1}
$$

In order to be able to compute the inner product (4.1) with the standard scalar product $\langle \cdot, \cdot \rangle$ as the outermost function, a transformation of the input feature space is necessary. This transformation is given by

$$
h(x_1, x_2) = (1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1 x_2)^{\mathrm{T}},
\tag{4.2}
$$

since $\langle h(x_1, x_2), h(x_1', x_2') \rangle = (\langle x, x' \rangle + 1)^2$. By proceeding in the fashion exemplified above, I used a suitable transformation of the input features in order to be able to use linear SVMs for the final models, while maintaining the higher performance of models with a polynomial kernel. In this way, I could extract interpretable, linear weights from the models.

Assessment and Comparison of Performance

Performance of the final models was assessed on EuResist$_T$ and HIVdbTCE. For performance comparison, the baseline sequences of the TEs in these datasets were interpreted with the HIVdb drug-resistance interpretation system version 6.0.6. [298]. The interpretations were used for calculating a GSS for each TE. Specifically, SIR predictions for each drug compound in the TEs were converted to a number via $S \rightarrow 1$, $I \rightarrow 0.5$, and $R \rightarrow 0$. The drug RAL was always assigned the score one, as no RAL-including TEs had an IN baseline genotype available. The addition of the numeric scores for all compounds in a TE yielded the GSS for that TE.

For calculating performance figures for the final models and GSS, the NAS, EuResist labels, and week-w labels with threshold 50 for $w \in \{8, 24, 48\}$ were calculated for the TEs in EuResist$_T$ and HIVdbTCE. Censoring labels for NAS were calculated with both the ITT and the OT criteria. Performance in predicting the

NAS was assessed with Harrell's concordance index (C; Section 2.2.6) [267, 268, 368]. Furthermore, the AUC [307] was used for calculating the performance of predicting EuResist labels and week-w labels. These figures were calculated on all TEs, as well as on the subsets containing either TEs with at least one major drug-resistance mutation [296] in their baseline genotypes or containing none, respectively. Furthermore, performance figures were also computed on TEs from treatment-naïve and treatment-experienced patients (including those for whom the number of previous therapy lines is uncertain), separately.

Ranking of Drug Combinations

The behavior of the final model for predicting the NAS when ranking drug combinations was explored, as follows. A list of 12 drug combinations was constructed with three *NRTI backbones*, FTC+TDF, 3TC+ABC, and 3TC+AZT, which were complemented with each of the drug compounds EFV, DRV/r, IDV/r, and RAL [*]. Twelve NASs were predicted for each TE in EuResist$_T$ by replacing the drug combination used in the TE with each one of the 12 drug combinations in the list. Ranks for drug combinations are calculated by sorting the resulting predicted numbers of aviremic semesters (pNAS). The drug combination with the highest NAS has the highest rank, by this definition. For analyzing the results of drug-combination ranking for first-line therapies, TDR mutations [279] in the baseline genotypes were quantified by drug class.

Assessment of Statistical Significance

Differences in the distributions of continuous variables of development and test sets were assessed statistically using a two-sided Wilcoxon rank-sum test [308]. For quantifying statistical association between two discrete variables, a two-sided Fisher's exact test [379] was used. Where necessary, p-values were corrected for multiple testing with the Benjamini-Hochberg method [332].

4.2.2 Results

In the following, I present the results of training and testing different models for predicting therapeutic success. The models differ in the definitions that they use for defining therapeutic success. Since some therapy-success definitions require that VL measurements be obtained a certain number of weeks after treatment initiation, some therapy-sucess definitions may constrain the number of samples that can be used to train them. For this reason, models trained with EuResist and week-w labels only use a subset of the TEs in EuResist$_D$.

Summary of the Composition of the Datasets

In the following, I report on analyses describing the compositions of EuResist$_D$, TF, EuResist$_T$, and HIVdbTCE. The distribution of the most frequent subtypes can be seen in Table 4.1 [304]. The fraction of non-B subtypes in the datasets is 29% (EuResist$_D$), 20% (TF), 31% (EuResist$_T$), and 1% (HIVdbTCE). Histograms for the most frequent therapies are plotted in Figure 4.1, while counts for the most frequent compounds in the therapies can be found in Table 4.2. The most frequent drug combinations in the datasets are the following: TDF+FTC+EFV (EuResist$_D$), 3TC+AZT+NVP (TF), TDF+FTC+EFV (EuResist$_T$), and d4T+ddI+EFV (HIVdbTCE). As can be seen, therapies including fewer than three drug compounds were also included in the datasets. A summary of other characteristics of the datasets, including therapy baseline parameters, as well as

---

[*]Therapy combinations in ranking were selected for purposes of the exposition. Lists of hunderts of therapies are possible.

Table 4.1: Subtype Distribution. Subtypes in EuResist$_D$, TF, EuResist$_T$ and HIVdbTCE were determined with the COMET subtyping tool. The numbers of sequences with the ten most frequent subtypes are tabulated below.

Subtype	EuResist$_D$	TF	EuResist$_T$	HIVdbTCE
B	8,043	1,579	834	577
C	712	76	80	0
02_AG	573	66	58	0
01_AE	512	34	52	0
A1	470	49	44	1
G	174	37	16	0
F1	180	28	32	0
D	121	23	23	1
06_cpx	61	8	2	0
42_BF	32	3	5	0
Other	474	76	66	1
Total	11,352	1,979	1,212	580

Table 4.2: Most Frequent Drug Compounds. The most frequent drug compounds in EuResist$_D$, TF, EuResist$_T$ and HIVdbTCE are tabulated below.

	EuResist$_D$	TF	EuResist$_T$	HIVdbTCE
3TC	5,005	1,137	539	255
ABC	2,110	410	210	192
AZT	2,457	605	297	94
d4T	1,005	477	118	278
ddC	18	18	0	0
ddI	1,279	458	173	213
FTC	4,534	290	447	33
TDF	6,235	690	623	172
EFV	2,718	424	279	141
ETR	157	14	25	2
NVP	693	327	79	77
RPV	31	1	0	0
APV	123	34	0	90
ATV	1,468	135	151	37
DRV	1,083	47	117	5
FPV	417	54	43	12
IDV	157	65	19	45
LPV	3,628	354	402	124
NFV	401	203	55	51
SQV	261	79	26	69
TPV	104	26	11	2
DTG	16	0	0	0
EVG	27	0	0	0
RAL	510	36	62	7
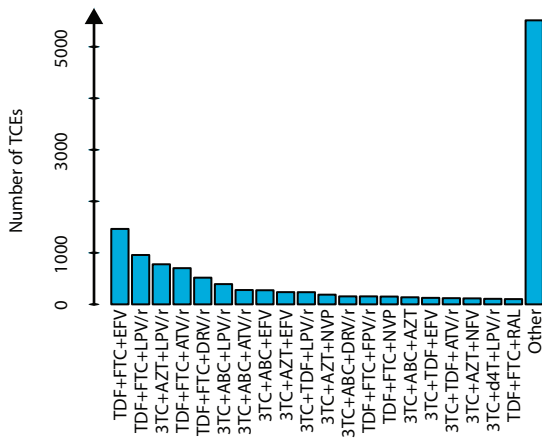
(a) EuResist$_D$
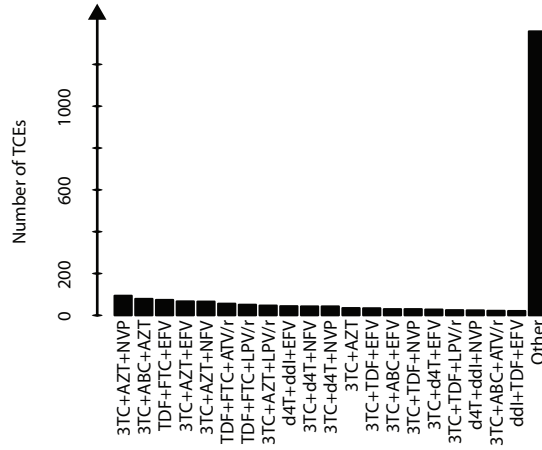
(b) TF
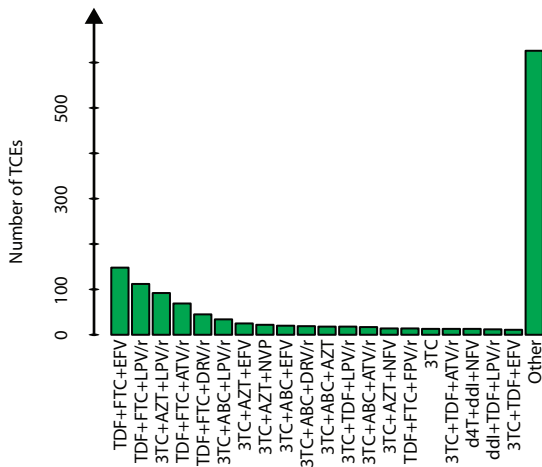
(c) EuResist$_T$
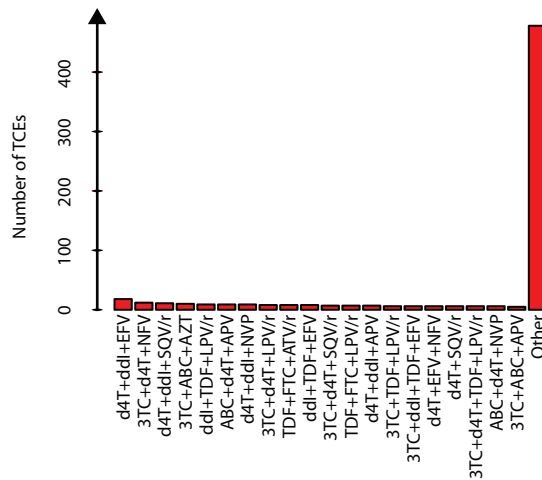
(d) HIVdbTCE

Figure 4.1: Therapy Distribution. Histograms for the 20 most frequent therapies in EuResist$_D$ (a), TF (b), EuResist$_T$ (c), and HIVdbTCE (d) are shown above.

outcomes, can be found in Table 4.3. In the following, I state the most striking differences as quantified in Table 4.3. The distributions of the numbers of recorded past treatment lines differ significantly among all four datasets (corrected $p < 0.02$), as do the proportions of genotypes presenting major drug resistance mutations (corrected $p < 0.02$). Among EuResist$_D$, EuResist$_T$, and HIVdbTCE, the distributions of therapy durations differ significantly (corrected $p < 0.01$). There is a large difference in therapy success rates between therapies extracted from EIDB and those in HIVdbTCE. For week 48, the difference between therapy-success-rate distributions in EuResist$_D$ and EuResist$_T$ is statistically significant (corrected $p < 10^{-3}$), however, this difference is much stronger when these datasets are compared to HIVdbTCE (corrected $p < 10^{-32}$). As expected, the number of available therapy VLs for weeks 8, 24, and 48 decrease with increasing week numbers. Furthermore, therapy-success rates for the aforementioned measurement weeks rise with increasing week numbers. Therapies in HIVdbTCE are of a much shorter duration than those in EuResist$_D$ and EuResist$_T$ (Table 4.3), and present significantly lower NAS (Figure 4.2; $\mu_{\text{EuResist}_D} = 2.52$; $\sigma_{\text{EuResist}_D} = 3.09$; $\mu_{\text{EuResist}_T} = 2.15$; $\sigma_{\text{EuResist}_T} = 2.89$; $\mu_{\text{HIVdbTCE}} = 0.12$; $\sigma_{\text{HIVdbTCE}} = 0.42$; corrected $p < 10^{-15}$).

RESULTS OF THE MODEL-SELECTION PROCEDURE

Cross-validation performances of the selected models for different VL thresholds can be seen in Table 4.4. The best cross-validation performance was attained by the model trained with week 48 labels and a VL threshold of 800 HIV-1 RNA copies per milliliter of blood serum ($\bar{C} = 0.8930$; $\sigma = 0.0200$). For VL thresholds between 50 and 400 HIV-1 RNA copies per milliliters of blood serum, performance increases as these thresholds are increased. Past 400 HIV-1 RNA copies per milliliters of blood serum, performance continues to increase for week 8 and week 24 labels. For week-48 labels and NAS, performance decreases as VL thresholds exceed 800 and 400 HIV-1 RNA copies per milliliters of blood serum, respectively. For week-w labels, performance increases as w increases. However, performances cannot be directly compared between different therapy-success definitions, as these definitions induce cross-validation sets of different compositions. The reason for differential cross-validation set composition is the differential availability of week-w VLs in the TEs, for different values of w (Table 4.3). As w increases, the number of training and test instances decreases.

RESULTS OF THE ASSESSMENT OF PERFORMANCE

For reasons explained in Section 4.2.3, I decided to only assess the performance of final models that use a VL threshold of 50 HIV-1 RNA copies per milliliter of blood serum, in addition to assessing the performance of the final model trained with EuResist labels. p-values for differences in performance are not given because of the lack of an appropriate statistical test. Performance of therapy-success prediction on the EuResist$_T$ and HIVdbTCE test sets is shown in Figure 4.3. Since both AUC [380] and Harrell's concordance index [267] express predictive performance in terms of a probability, in the following, I average performances across both datasets and all therapy-success definitions. pNAS could attain the best average performance ($\mu = 0.71$; $\sigma = 0.07$), followed by GSS ($\mu = 0.62$; $\sigma = 0.05$). The average performance of predicted EuResist ($\mu = 0.61$; $\sigma = 0.09$), week 48 ($\mu = 0.61$; $\sigma = 0.0.08$), week 24 ($\mu = 0.60$; $\sigma = 0.09$), and week 8 ($\mu = 0.58$; $\sigma = 0.07$) labels was lower. When therapies with drug-resistance mutations at baseline are evaluated separately (Figure 4.4(a)), pNAS attains the best average performance ($\mu = 0.72$; $\sigma = 0.07$), followed by GSS ($\mu = 0.63$; $\sigma = 0.05$) and predicted EuResist labels ($\mu = 0.63$; $\sigma = 0.10$). The average performance of predicted week 48 ($\mu = 0.60$; $\sigma = 0.07$), week 24 ($\mu = 0.60$; $\sigma = 0.08$), and week 8 ($\mu = 0.58$; $\sigma = 0.08$) labels is lower. On

**Table 4.3: Therapy Characteristics**. For therapies in EuResist$_D$, TF, EuResist$_T$, and HIVdbTCE, the following characteristics are summarized below: baseline viral load, baseline CD4+-cell count, number of recorded past treatment lines, number of first-line therapies, therapy drug-class composition, genotypic resistance by drug class, therapy duration, and therapy success figures at weeks 8, 24, and 48. Virologic suppression was defined as a viral load below 50 HIV-1 RNA copies per milliliter of blood serum. IN: integrase; INI: integrase inhibitor; IQR: interquartile range; NA: Not Available; NNRTI: non-nucleoside reverse-transcriptase inhibitor; NRTI: nucleotide/nucleoside reverse-transcriptase inhibitor; PI: protease inhibitor; VL: viral load.

		EuResist$_D$	TF	EuResist$_T$	HIVdbTCE
n		11,352	1,979	1,212	580
Log Baseline VL	Available	10,413 (92%)	NA	1,105 (91%)	580 (100%)
	Mean	4.4	NA	4.4	4.4
	Median	4.6	NA	4.6	4.4
	IQR	5.2 - 3.7	NA	5.2 - 3.8	4.9 - 3.9
Baseline CD4 Count	Available	9,123 (80%)	NA	989 (82%)	580 (100%)
	Mean	287.6	NA	277.8	268.9
	Median	263	NA	257	229.5
	IQR	410 - 100	NA	390 - 105	370.5 - 111.8
Number of Recorded Past Treatment Lines	Mean	2.6	4.9	3.4	4.7
	Median	0	4	1	4
	IQR	4 - 0	7 - 1	5 - 0	6 - 2
Compound Frequency	NRTI	11,070 (96%)	1,905 (96%)	1,170 (97%)	571 (98%)
	NNRTI	3,590 (31%)	761 (38%)	383 (32%)	220 (38%)
	PI	7,471 (66%)	941 (48%)	805 (66%)	435 (75%)
	INI	551 (5%)	34 (2%)	62 (5%)	7 (1%)
INI-including Therapies with Baseline IN Genotype		111 (20%)	18 (53%)	12 (15%)	0 (0%)
First-Line Therapies		4,828 (43%)	NA	426 (35%)	0 (0%)
Genotypic Resistance	NRTI	3,767 (33%)	1,482 (75%)	456 (38%)	507 (87%)
	NNRTI	3,483 (30%)	1,114 (56%)	399 (33%)	323 (56%)
	PI	1,804 (16%)	697 (35%)	232 (19%)	378 (65%)
	INI	4 (<1%)	10 (1%)	1 (< 1%)	NA
	No Resistance	6,203 (55%)	310 (16%)	594 (49%)	18 (3%)
Therapy Duration (Weeks)	Mean	98.1	NA	91.1	40.2
	Median	67	NA	60	52
	IQR	138 - 31	NA	129 - 26	52 - 28
Week 8 VLs	Available	9,142 (81%)	NA	964 (80%)	425 (73%)
	Suppressed	3,024 (33%)	NA	303 (31%)	21 (5%)
Week 24 VLs	Available	5,862 (52%)	NA	586 (48%)	300 (52%)
	Suppressed	4,014 (68%)	NA	362 (62%)	25 (8%)
Week 48 VLs	Available	2,392 (21%)	NA	245 (20%)	252 (43%)
	Suppressed	1,770 (74%)	NA	156 (64%)	32 (13%)
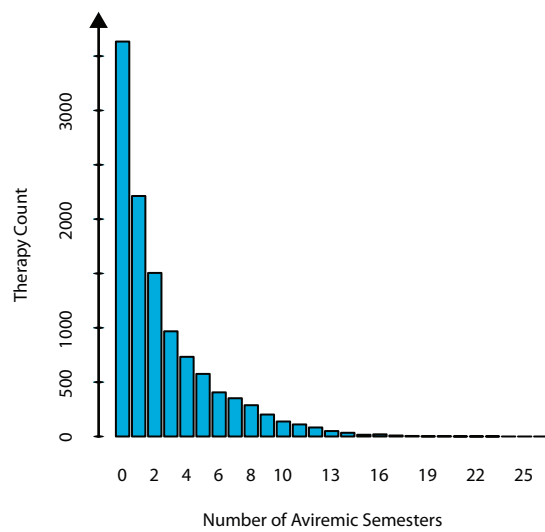
(a) EuResist$_D$
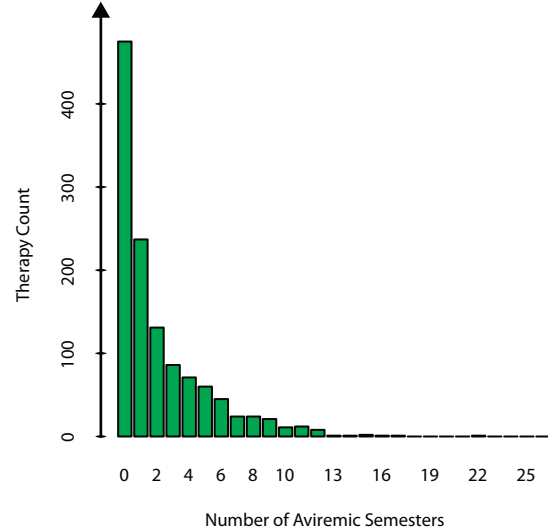
(b) EuResist$_T$
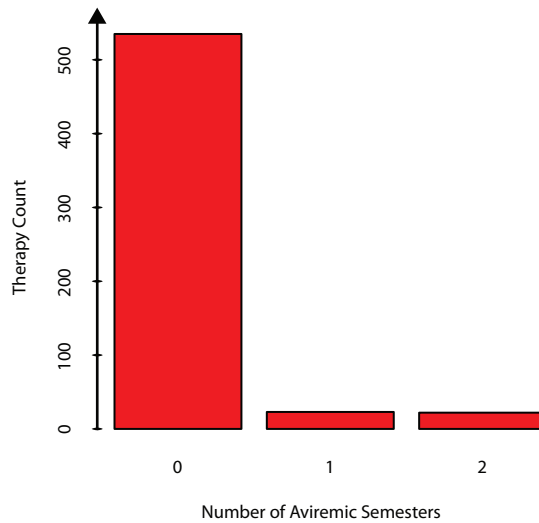
(c) HIVdbTCE

Figure 4.2: Numbers of Aviremic Semesters. Histograms of the numbers of aviremic semesters with a viral load threshold of 50 HIV-1 RNA copies per milliliter of blood serum are displayed above. For EuResist$_D$ (a), EuResist$_T$ (b), and HIVdbTCE (c), the intention-to-treat censoring criterion marked 4,516 / 11,352, 422 / 1,235, and 85 / 581 numbers of aviremic semesters as censored, respectively. In contrast, the on-treatment censoring criterion marked 7,951 / 11,352, 767 / 1,235, and 121 / 581 numbers of aviremic semesters as censored, respectively. All therapy-failure episodes in TF were assigned zero uncensored aviremic semesters.

Table 4.4: Cross-Validation Performances for Chosen Models (C). Four different treatment-success definitions were used to produce four groups of models, each group of models trained on one treatment success definition. The treatment success definitions used by each group of models are the following: EuResist labels, week 8 labels, week 24 labels, week 48 labels, and the number of aviremic semesters. Each group contains models trained with different parameters. Parameters were selected for models trained on the following VL thresholds: 50, 100, 200, 400, 800, and 1,600 HIV-1 RNA copies per milliliter of blood serum. The cross-validation performances of the selected models are tabulated below. All models were evaluated by their capacity to predict the number of aviremic semesters. Note that performances are not directly comparable between different therapy-success definitions, Performances are stated in terms of mean Harrell's concordance index (C). NAS: Number of Aviremic Semesters; NA: Not Available; SD: Standard Deviation.

Viral Load Threshold (cp / ml)	EuResist Labels Performance (SD)	Week 8 Labels Performance (SD)	Week 24 Labels Performance (SD)	Week 48 Labels Performance (SD)	NAS Performance (SD)
50	NA	0.6691 (0.0379)	0.8126 (0.0162)	0.8547 (0.0109)	0.7544 (0.0152)
100	NA	0.6695 (0.0231)	0.8375 (0.0084)	0.8791 (0.0249)	0.7565 (0.0166)
200	NA	0.6927 (0.0268)	0.8606 (0.0115)	0.8851 (0.0235)	0.7591 (0.0160)
400	0.8101 (0.0231)	0.7218 (0.0335)	0.8689 (0.0148)	0.8896 (0.0228)	0.7605 (0.0156)
800	NA	0.7592 (0.0309)	0.8749 (0.0129)	0.8930 (0.0200)	0.7588 (0.0157)
1600	NA	0.7982 (0.0225)	0.8766 (0.0154)	0.8918 (0.0211)	0.7568 (0.0155)
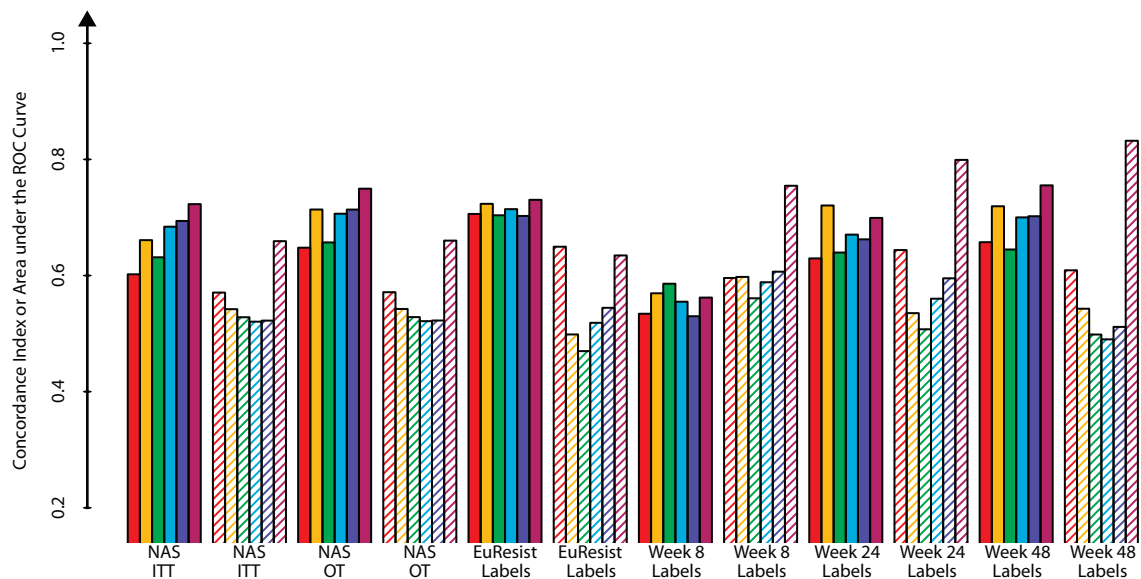
Figure 4.3: Performance of Therapy-Success Prediction (1). Six different therapy-success definitions were applied to $\mathsf{EuResist_T}$ and to $\mathsf{HIVdbTCE}$: the number of aviremic semesters (NAS) with intent-to-treat (ITT) or on-treatment (OT) censoring criterion, EuResist labels, week 8 labels, week 24 labels, and week 48 labels. Six different models were used for predicting therapeutic success according to each of these definitions: Support Vector Machines (SVMs) trained for predicting the NAS, EuResist labels, week 8 labels, week 24 labels, or week 48 labels, as well as a genetic susceptibility score (GSS) based on the HIVdb rule set. Above, performance on all therapies in the datasets is shown for each prediction model and therapy-success definition. Performance in predicting dichotomous therapy-success labels are stated in terms of the area under the receiver-operating characteristic (ROC) curve, while performance in predicting the NAS is stated in terms of Harrell's concordance index.

Solid bars: $\mathsf{EuResist_T}$　　　　　　　　　　　　　　　　　　　　　　Shaded bars: $\mathsf{HIVdbTCE}$

■ GSS　　　　　　　■ SVM Predicting EuResist Labels　　　　　　　■ SVM Predicting Week 8 Labels
■ SVM Predicting Week 24 Labels　　　■ SVM Predicting Week 48 Labels　　　■ SVM Predicting the NAS
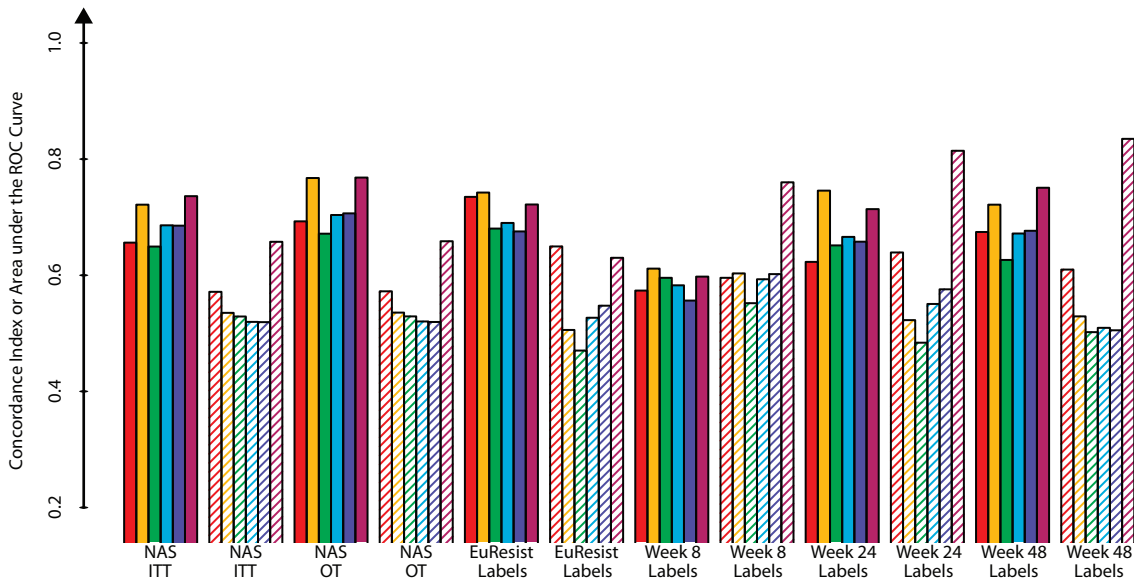
(a) Therapies with Drug Resistance at Baseline
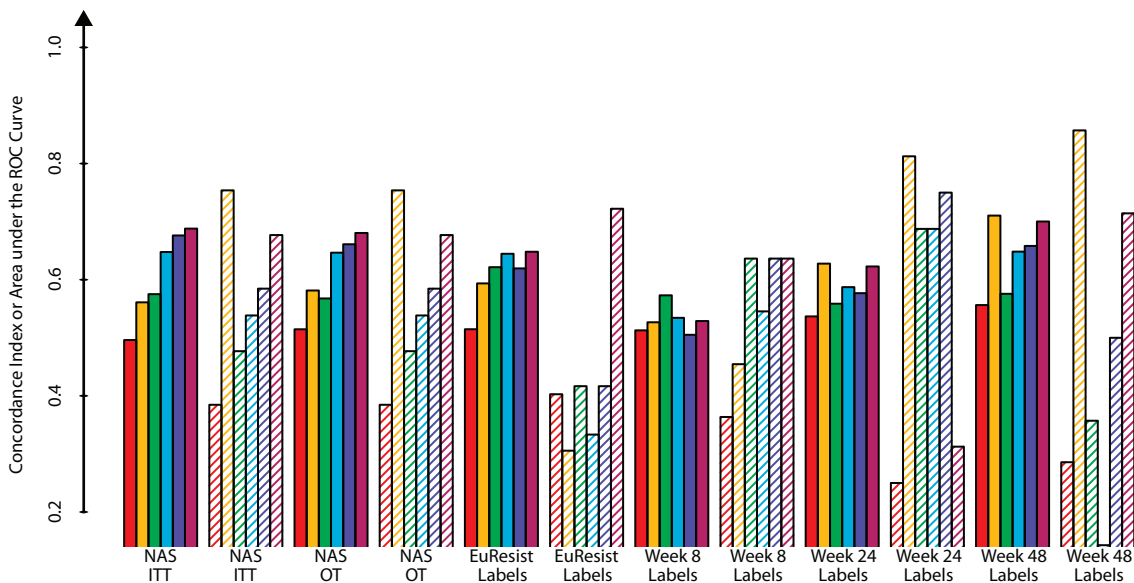
(b) Therapies with No Drug Resistance at Baseline

Figure 4.4: Performance of Therapy-Success Prediction (2). Six different therapy-success definitions were applied to $EuResist_T$ and to $HIVdbTCE$: the number of aviremic semesters (NAS) with intent-to-treat (ITT) or on-treatment (OT) censoring criterion, EuResist labels, week 8 labels, week 24 labels, and week 48 labels. Six different models were used for predicting therapeutic success according to each of these definitions: Support Vector Machines (SVMs) trained for predicting the NAS, EuResist labels, week 8 labels, week 24 labels, or week 48 labels, as well as a genetic susceptibility score (GSS) based on the HIVdb rule set. Above, performance on therapies with (a) or without (b) resistance at baseline is shown for each prediction model and therapy-success definition. Performance in predicting dichotomous therapy-success labels are stated in terms of the area under the receiver-operating characteristic curve (ROC), while performance in predicting the NAS is stated in terms of Harrell's concordance index. Of note: $HIVdbTCE$ only contains 18 TEs without drug-resistance mutations at baseline.

Solid bars: $EuResist_T$ Shaded bars: $HIVdbTCE$

■ GSS   ■ SVM Predicting EuResist Labels   ■ SVM Predicting Week 8 Labels

■ SVM Predicting Week 24 Labels   ■ SVM Predicting Week 48 Labels   ■ SVM Predicting the NAS

therapies with no drug-resistance mutations at baseline (Figure 4.4(a)), pNAS ($\mu = 0.63$; $\sigma = 0.11$) and EuResist labels ($\mu = 0.63$; $\sigma = 0.16$) could attain the best average performance. The average performance predicted week 48 ($\mu = 0.60$; $\sigma = 0.09$), week 8 ($\mu = 0.54$; $\sigma = 0.09$), and week 24 ($\mu = 0.54$; $\sigma = 0.16$) labels, as well as GSS ($\mu = 0.43$; $\sigma = 0.10$) could attain was lower. Average performance on first-line therapies (Figure 4.5(a)) was highest with pNAS ($\mu = 0.61$; $\sigma = 0.06$), followed by week 48 labels ($\mu = 0.60$; $\sigma = 0.07$). The average performance of predicted week 24 ($\mu = 0.57$; $\sigma = 0.05$), EuResist ($\mu = 0.56$; $\sigma = 0.03$), and week 8 ($\mu = 0.56$; $\sigma = 0.01$) labels, as well as that of GSS ($\mu = 0.51$; $\sigma = 0.02$) is lower. On therapies of therapy-experienced patients (Figure 4.5(b)), pNAS could attain the best average performance ($\mu = 0.71$; $\sigma = 0.06$), followed by GSS ($\mu = 0.63$; $\sigma = 0.04$). The average performance of predicted EuResist ($\mu = 0.62$; $\sigma = 0.09$), week 24 ($\mu = 0.60$; $\sigma = 0.07$), week 48 ($\mu = 0.60$; $\sigma = 0.06$), and week 8 ($\mu = 0.58$; $\sigma = 0.07$) labels could attain was lower. Of note: all patients in HIVdbTCE are treatment-experienced. Note that these results are discussed in Section 4.2.3.

## Results of the Ranking of Drug Combinations

A list of 12 drug combinations was constructed with the drug compounds EFV, DRV/r, IDV/r, and RAL. Each of these compounds was appended to each of the NRTI backbones FTC+TDF, 3TC+ABC, and 3TC+AZT. Twelve NAS predictions were made for each TE in EuResist$_T$, by replacing the drug combinations used in the TEs with each of the combinations in the list. The pNAS for all drug combinations ranged from -0.3 to 8.7 (Figure 4.6(a)). Among the top-ranked drug combinations, pNAS ranged from 0.5 to 8.8 (Figure 4.6(b)). The drug combination FTC+TDF+EFV was most frequently in the top rank, while drug combinations with 3TC+AZT as a backbone and those including IDV/r were never in the top rank (Figure 4.6(c)). The drug combination FTC+TDF+DRV/r was most frequently in the top rank of TEs with drug-resistance mutations in their baseline genotypes (Figure 4.7(a)), while FTC+TDF+EFV was the top-ranking drug combination for TEs without drug-resistance mutations in their baseline genotypes (Figure 4.7(a)). The association between DRV/r-containing drug-combinations being ranked at the top of the list and the presence of drug-resistance mutations at baseline is significant ($p < 10^{-15}$). In contrast, the frequency with which other drug combinations were ranked at the top of the list for TEs without drug-resistance mutations in their baseline genotypes was very low. FTC+TDF+EFV was most frequently in the top rank of both first-line therapies (Figure 4.7(c)) and therapies on treatment-experienced patients (Figure 4.7(d)). However, while drug combinations other than FTC+TDF+EFV were ranked at the top of the list for 450 (57%) TEs of treatment-experienced patients, FTC+TDF+EFV was ranked at the top for 395 (93%) of first-line therapies. Among the baseline genotypes of first-line therapies in which FTC+TDF+EFV was ranked at the top of the list, 22 (6%) contained NRTI, 1 (<1%) contained NNRTI, and 7 (2%) contained PI TDR mutations. DRV/r-containing drug combinations were ranked at the top of the list in 24 (6%) TEs. Among the baseline genotypes of these TEs, 11 (46%) contained NRTI, 11 (46%) contained NNRTI, and 2 (8%) contained PI TDR mutations. For first line therapies, the association between a DRV/r-containing drug combination being ranked at the top of the list and TDR mutations being present in the baseline genotypes is significant ($p < 10^{-13}$).

Results of the Analysis of Misclassifications

An excerpt of TEs from EuResist$_T$ with the largest or smallest pNAS contradicting their week 8, week 24, and week 48 labels is shown in Tables 4.5, 4.6, and 4.7, respectively. Contradiction occurs when a large NAS is
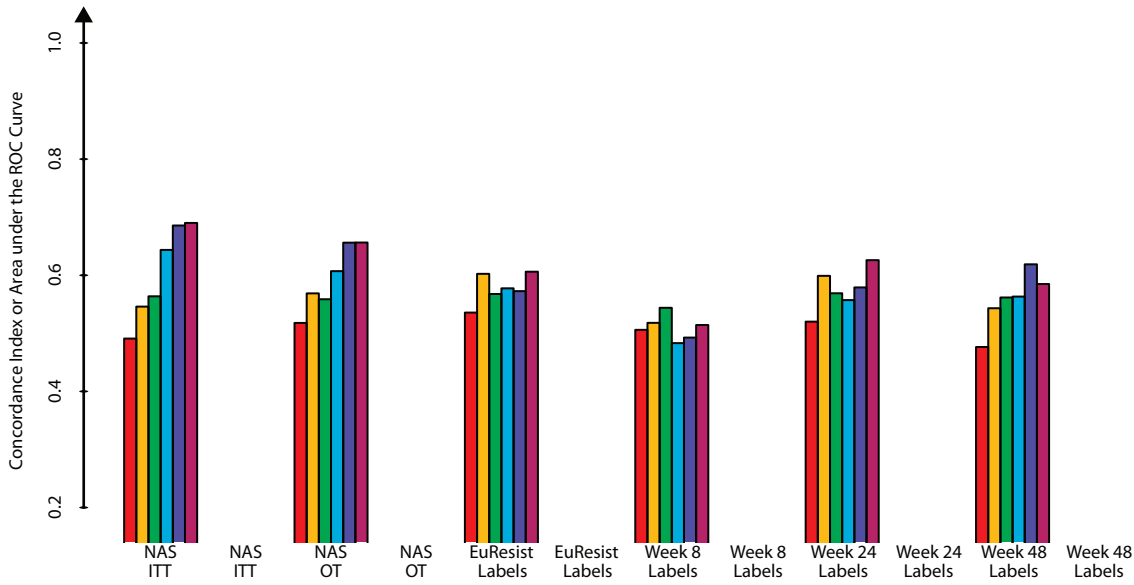
(a) First-Line Therapies
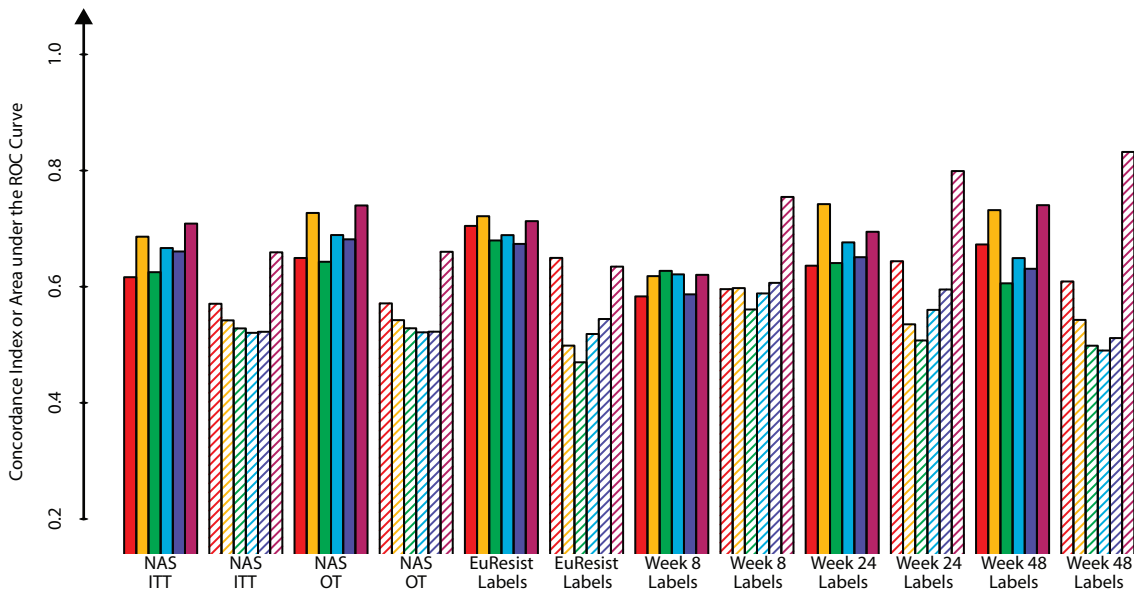
(b) Therapies on Treatment-Experienced Patients

Figure 4.5: Performance of Therapy-Success Prediction (3). Six different therapy-success definitions were applied to $EuResist_T$ and to $HIVdbTCE$: the number of aviremic semesters (NAS) with intent-to-treat (ITT) or on-treatment (OT) censoring criterion, EuResist labels, week 8 labels, week 24 labels, and week 48 labels. Six different models were used for predicting therapeutic success according to each of these definitions: Support Vector Machines (SVMs) trained for predicting the NAS, EuResist labels, week 8 labels, week 24 labels, or week 48 labels, as well as a genetic susceptibility score (GSS) based on the HIVdb rule set. Above, performance on first-line therapies with (a) or therapies on treatment-experience patients (b) resistance at baseline is shown for each prediction model and therapy-success definition. Performance in predicting dichotomous therapy-success labels are stated in terms of the area under the receiver-operating characteristic (ROC) curve, while performance in predicting the NAS is stated in terms of Harrell's concordance index. Note that $HIVdbTCE$ does not contain first-line therapies.

Solid bars: $EuResist_T$ Shaded bars: $HIVdbTCE$

- ■ GSS
- ■ SVM Predicting EuResist Labels
- ■ SVM Predicting Week 8 Labels
- ■ SVM Predicting Week 24 Labels
- ■ SVM Predicting Week 48 Labels
- ■ SVM Predicting the NAS
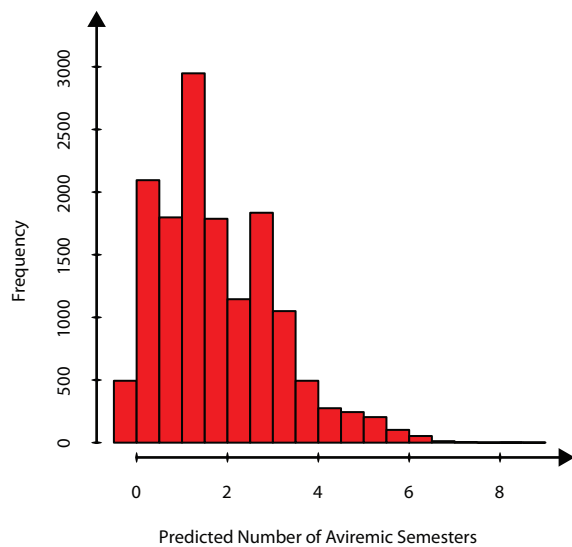
(a) pNAS of All Ranked Drug Combinations
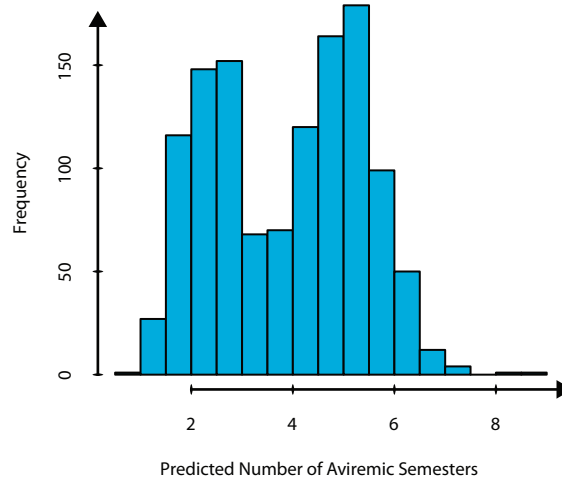
(b) pNAS of Drug Combinations in Top Ranks
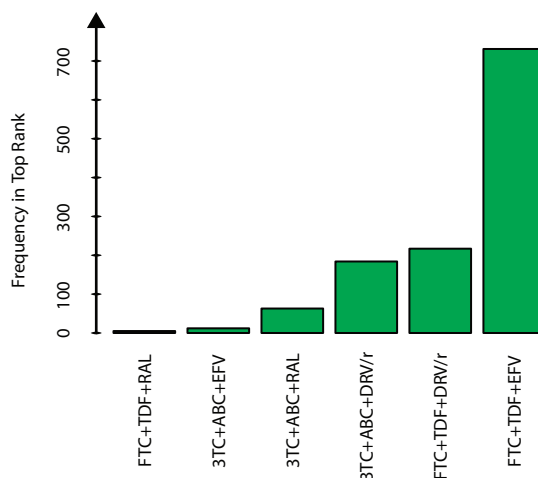
(c) Frequency of Drug Combinations in Top Ranks

Figure 4.6: Ranking of Drug Combinations (1). A list of 12 drug combinations was constructed using three *NRTI backbones*, FTC+TDF, 3TC+ABC, and 3TC+AZT in addition to the drug compounds EFV, DRV/r, IDV/r, and RAL. Each of the drug combinations in the list was used for predicting the NAS for each TE in EuResist$_T$. Histograms of the resulting pNAS for all drug combinations (a) and drug combinations in the top rank (b) are shown above. The distribution of drug combinations in the top ranks is shown as well (c). Drug combinations with AZT+3TC as the NRTI backbone and drug combinations including IDV/r were never at the top rank and are therefore not plotted in (c).
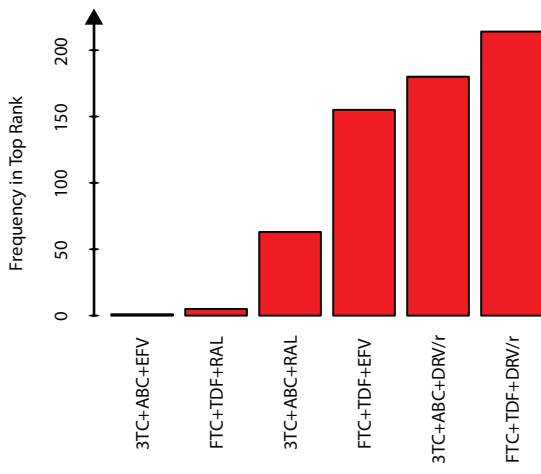
(a) Therapies with Drug Resistance at Baseline
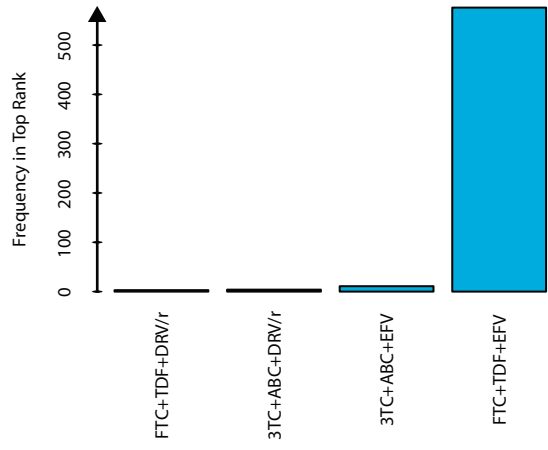
(b) Therapies with No Drug Resistance at Baseline
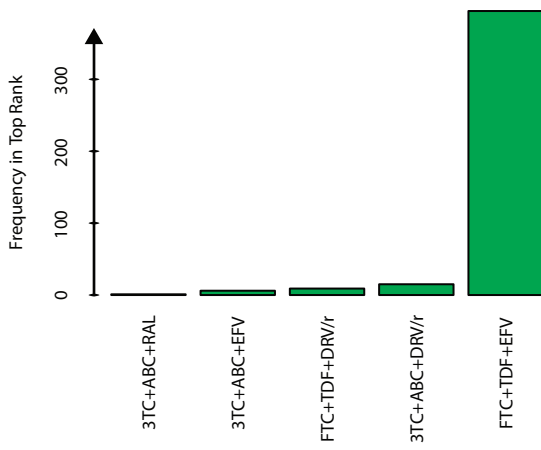
(c) First-Line Therapies
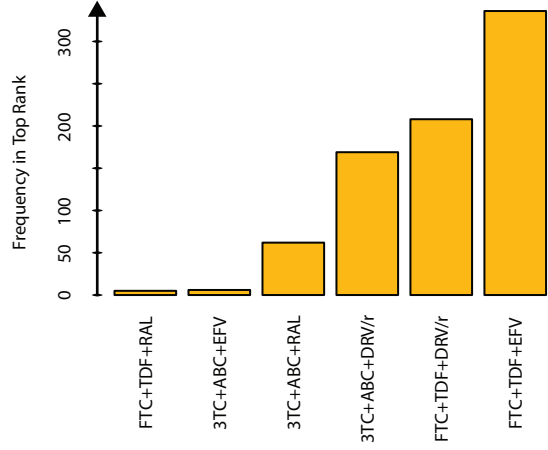
(d) Therapies on Treatment-Experienced Patients

Figure 4.7: Ranking of Drug Combinations (2). A list of 12 drug combinations was constructed using three *NRTI backbones*, FTC+TDF, 3TC+ABC, and 3TC+AZT, in addition to the drug compounds EFV, DRV/r, IDV/r, and RAL. Each of the drug combinations in the list was used for predicting the NAS for each TE in EuResist$_T$. The distributions of drug combinations in the top ranks for treatments with (a) and without (b) major drug-resistance mutations at baseline, first-line therapies (c), and therapies on treatment-experienced patients (d) are shown above. Drug combinations that were never at the top ranks of a specific subset of therapies are not shown in the corresponding plots. Drug combinations with AZT+3TC as the NRTI backbone and drug combinations including IDV/r were never at a top rank and are therefore not plotted above.

Table 4.5: Selected Misclassifications (1). Treatment episodes with week-8 success and failure labels were ordered by their predicted number of aviremic semesters. The three successes with the lowest predicted number of aviremic semesters were selected. In the same way, the three failures with the highest predicted number of aviremic semesters were selected. The corresponding treatment episodes are tabulated below. Displayed therapy viral loads are limited to those measured during the first year of therapy. ITT: intention to treat; NAS: number of aviremic semesters; pNAS: predicted number of aviremic semesters; VL: viral load; /r: ritonavir boosting dose; 3TC: lamivudine; ABC: abacavir; AZT: zidovudine; ddI: didanosine; EFV: efavirenz; FTC: emtricitabine; NFV: nelfinavir; LPV: lopinavir; TDF: tenofovir disoproxil fumarate.

| TE Number	1	2	3	4	5	6
Week 8 Label	Failure	Failure	Failure	Success	Success	Success
NAS	1	11	3	0	0	1
pNAS	6.8	6.5	6.4	-0.1	0	0
Censored (ITT)	Yes	Yes	Yes	No	No	No
Drug Compounds	TDF, FTC, EFV	TDF, FTC, EFV	TDF, FTC, EFV	3TC, AZT, EFV	ABC, LPV/r, SQV	TDF, 3TC, ddI, TDF, NFV
Protease Substitutions	10I, 19I, 35D, 37N, 63P, 67Y, 71I, 93L	31, 57K, 62V, 63P, 93L	31, 10I, 13V, 35D, 36I, 37N, 63P, 74PT, 93L	31, 13V, 36I, 37N, 41K, 69K, 89M	31, 10V, 33F, 37N, 46IL, 54V, 61N, 62V, 63P, 64M, 71V, 74P, 82A, 90M, 93L	31, 13V, 20KR, 35D, 36I, 37DN, 41K, 57K, 69K, 89IM
Reverse-Transcriptase Substitutions	122K, 202V, 214F	200V, 202V, 211A, 20R, 207A, 214F	162C, 33AP, 121Y, 17DE, 210M, 214F, 11R, 20R, 28A, 60I, 83K, 135T, 207HQ, 211KM, 184V, 207A, 211KR, 214F	6D, 11T, 35T, 39K, 103N, 108I, 121H, 158S, 122P, 173A, 177E, 166R, 177E, 184V, 190AG, 203A, 207E, 211KR, 70R, 98S, 108I, 123E, 135T, 184V, 214F, 219Q	35T, 67G, 69D, 70R, 98S, 108I, 122P, 123E, 135T, 166R, 177E, 184V, 203A, 207E, 214F, 219Q	35T, 41LM, 60I, 67DN, 69NT, 121H, 135T, 138A, 70KR, 177E, 173S, 174K, 179I, 184V, 200A, 207S, 211NS, 219KQ
Therapy V Ls (Week:VL)	6:331, 15:50, 25:50, 43:50	4:575, 21:110, 31:40, 44:40	12:201, 8:157, 26:40, 44:40	0:247000, 2:447, 6:50, 15:134, 23:50, 45:50	15:40, 6:50, 10:50, 19:50, 28:50	6:50, 21:1040, 0:100, 11:49

Table 4.6: Selected Misclassifications(2). Treatment episodes with week-24 success and failure labels were selected. The three successes with the lowest predicted number of aviremic semesters were selected. In the same way, the three failures with the highest predicted number of aviremic semesters were selected. The corresponding treatment episodes are tabulated below. Displayed therapy viral loads are limited to those measured during the first year of therapy. ITT: intention to treat; NAS: number of aviremic semesters; pNAS: predicted number of aviremic semesters; VL: viral load; /r: ritonavir boosting dose; 3TC: lamivudine; ABC: abacavir; AZT: zidovudine; d4T: stavudine; DRV: darunavir; EFV: efavirenz; ETR: etravirine; ddI: didanosine; FTC: emtricitabine; LPV: lopinavir; TDF: tenofovir disoproxil fumarate; TPV: tipranavir.

TE Number	2	7	8	4	9	10
Week 24 Label	Failure	Failure	Failure	Success	Success	Success
NAS	11	2	1	5	1	2
pNAS	6.5	6.2	5.5	-0.1	0	0
Censored (ITT)	Yes	No	No	No	No	No
Drug Compounds	TDF, FTC, EFV	TDF, FTC, EFV	ETR, DRV/r, RAL	3TC, AZT, EFV	3TC, ABC, AZT, ddI, TPV	3TC, d4T, TDF, LPV/r
Protease Substitutions	3I, 57K, 62V, 63P, 93L	3I, 12P, 13V, 19I, 20R, 35D, 36I, 37E, 41K, 45R, 57K, 69K,70KR, 74K, 89M	3I, 10F, 37N, 63P, 71V, 82T, 84V	3I,13V,36I,37N, 41K, 69K, 89M	10I, 13V, 20I, 32L, 36I, 37N, 54V, 58E, 60E, 63P, 71I, 72V, 82A, 90M, 93L	3I, 35D, 37DE, 63A, 93L
Reverse-Transcriptase Substitutions	20R, 162C, 207A, 214F	35T, 60I, 122K, 135M, 162A, 171Y, 173S, 174K,177E,179I, 207A, 211S, 214F	41L, 67N, 69D, 102KQ, 165I, 184MV, 200A, 207E, 214F,215Y	6D, 11T, 35T, 39K, 103N, 108I, 121H, 158S, 173A, 177E, 184V, 190AG, 207A, 211KR, 214F	67G, 69N, 70R, 98S,103N,122K, 123E, 135M, 142V, 184V, 188L, 196E, 200I, 219Q, 228H	122K, 178M, 207KQR, 211KR, 214F
Therapy V Ls (Week:VL)	4:575, 12:201, 21:110, 31:40, 44:40, 56:40	4:1000, 15:31, 25:160, 31:50, 35:41, 44:33	5:6400, 7:461, 12:193, 19:188, 27:54, 37:50	0:247000, 2:447, 6:50, 10:50, 15:134, 19:50, 23:50, 28:50, 45:50	1:890, 13:50, 20:50	0:205827, 7:1164, 13:193, 22:50, 26:50, 36:50, 40:50, 47:50

Table 4.7: Selected Misclassifications (3). Treatment episodes with week-48 success and failure labels were ordered by their predicted number of aviremic semesters. The three successes with the lowest predicted number of aviremic semesters were selected. In the same way, the three failures with the highest predicted number of aviremic semesters were selected. The corresponding treatment episodes are tabulated below. Displayed therapy viral loads are limited to those measured during the first year of therapy. ITT: intention to treat; NAS: number of aviremic semesters; pNAS: predicted number of aviremic semesters; VL: viral load; /r: ritonavir boosting dose; 3TC: lamivudine; ABC: abacavir; AZT: zidovudine; d4T: stavudine; DRV: darunavir; EFV: efavirenz; ETR: etravirine; ddI: didanosine; FTC: emtricitabine; LPV: lopinavir; RAL: raltegravir; TDF: tenofovir disoproxil fumarate; TPV: tipranavir.

TE Number	11	12	13	4	10	14
Week 48 Label	Failure	Failure	Failure	Success	Success	Success
NAS	0	2	0	5	2	2
pNAS	6.1	5.5	5	-0.1	0	0.1
Censored (ITT)	No	Yes	Yes	No	No	No
Drug Compounds	TDF, FTC, EFV	TDF, FTC, EFV	ETR, DRV/r, RAL	3TC, AZT, EFV, LPV/r	3TC, d4T, TDF, LPV/r	ABC, ddI, LPV/r
Protease Substitutions	3I, 13V, 35D, 36I, 37N, 41K, 61E, 69K, 89M	3I, 10I, 13V, 20R, 36I, 37N, 41K, 43R, 69Q, 89M	3I, 4P, 10I, 13V, 16A, 19I, 20R, 32I, 33F, 34EK, 35D, 36I, 37N, 43T, 46I, 47V, 50V, 53FL, 63C, 71V, 73S, 82A, 93L	3I, 13V, 36I, 37N, 41K, 69K, 89M	3I, 35D, 37DE, 63A, 93L	3I, 10F, 13V, 20T, 35D, 36I, 37D, 41K, 46I, 63P, 69K, 72V, 84V, 89T
Reverse-Transcriptase Substitutions	6D, 35T, 39K, 43E, 165I, 171Y, 173I, 174K, 177E, 178M, 179I, 200E, 207A, 211S, 214F	21IV, 35T, 39A, 40D, 122K, 135V, 162A, 173T, 174K, 177E, 200A, 207E, 211K, 214F	20R, 30R, 67N, 70R, 135L, 162A, 179I, 184V, 196K, 200A, 203K, 207E, 211K, 214F, 214F, 223Q	6D, 11T, 35T, 39K, 103N, 108I, 121H, 158S, 173A, 177E, 184V, 190AG, 207A, 211KR	122K, 178M, 207KQR, 211KR, 214F	4H, 8LV, 11T, 28K, 32E, 35T, 39K, 41L, 43E, 49R, 67N, 101E, 121Y, 135T, 162C, 169D, 172K, 173A, 174K, 177E, 190S, 196E, 207A, 210W, 21IN, 214F, 215Y
Therapy V Ls (Week:VL)	8:228, 18:47, 31:46, 50:660	13:40, 22:40, 48:1317	3:2960, 32:120, 44:56	4:118, 16:1286, 24:79, 33:41, 36:32, 45:80	0:247000, 6:50, 15:134, 23:50, 45:50	2:447, 10:50, 19:50, 28:50, 36:50, 47:50

predicted for a failure-labeled TE or a small NAS is predicted for a success-labeled TE. Thus, the three largest and three smallest pNAS were selected for each week and contradicting label instance. In the following, TE numbers are stated as displayed in the aforementioned tables. TEs with numbers 1, 2, 3, 7, 8, 11, 12, and 13 are all labeled as failures and do not present resistance mutations relevant to their drug compounds. In all of them, suppression of the VL is attained at some other point in time. TEs with numbers 4 and 9 are labeled as successes. Viral suppression is attained in these TEs in spite of the presence of several drug-resistance mutations that are relevant to their drug compounds. TEs with numbers 5, 6, 10 and 14 are labeled as successes as well, but do not present resistance mutations that are relevant to their drug compounds. Nonetheless, they attain at most two ITT uncensored aviremic semesters. Note that these results are discussed in Section 4.2.3.

Description of Linear Model Weights

Among the 1,128 features (including feature interactions) on which the model for predicting the NAS was trained, 1,014 (90%) had non-zero weights (Figure 4.8(a)). Predictions are calculated by multiplying the value of the feature (binary drug indicator, binary past INI-use indicator, or POE) by the value of the weight, and subsequently adding the resulting values. Features can be classified as either drug-dependent or solely sequence-dependent. Solely sequence-dependent features are those which do not involve drugs and whose value does not change given a baseline genotype. In contrast, drug-dependent features change their value according to the therapy for which the prediction is required. Figure 4.8(b) displays a graphical representation of a subset of drug-dependent weights.

4.2.3 Discussion

In this chapter, I present a method for deriving a genotypic interpretation system that can predict which cART can be expected to be more durable for suppressing the VL before it needs to be replaced by another cART. For this purpose, a novel measure for retrospectively assessing a cART was employed, the NAS.

Quantifying Virologic Success with NAS

In clinical settings, virologic success of cART is determined with VL measurements (Section 4.1). Among the parameters used for making clinical decisions, the VL is one of the most important ones. After treatment initiation, each VL measurement evaluates the performance of cART. The measurement of a VL above 50 HIV-1 RNA copies per milliliter of blood serum may trigger the decision of changing the treatment, depending on the time since treatment initiation and other patient-related factors. Conversely, VLs below 50 HIV-1 RNA copies per milliliter of blood serum speak for the decision of continuation of treatment. Indeed, the gold standard for determining therapeutic success of cART is the dichotomization of VLs measured at a certain week after initiation of treatment [228, 280]. Probably, this has motivated researchers designing therapy-success prediction systems to use dichotomous, cross-sectional VL measurements in definitions of treatment success for the *training* of models. While I believe that the *testing* of models for predicting therapeutic success must be aligned with established definitions of therapeutic success, other definitions of treatment success may be more appropriate for training of the models. Specifically, the testing of models must allow for comparability with previously performed analyses from the literature. The training of models, however, can be carried out in such a way that their performance is increased, including the use of more appropriate therapy-sucess definitions. In the following, I elaborate on this.
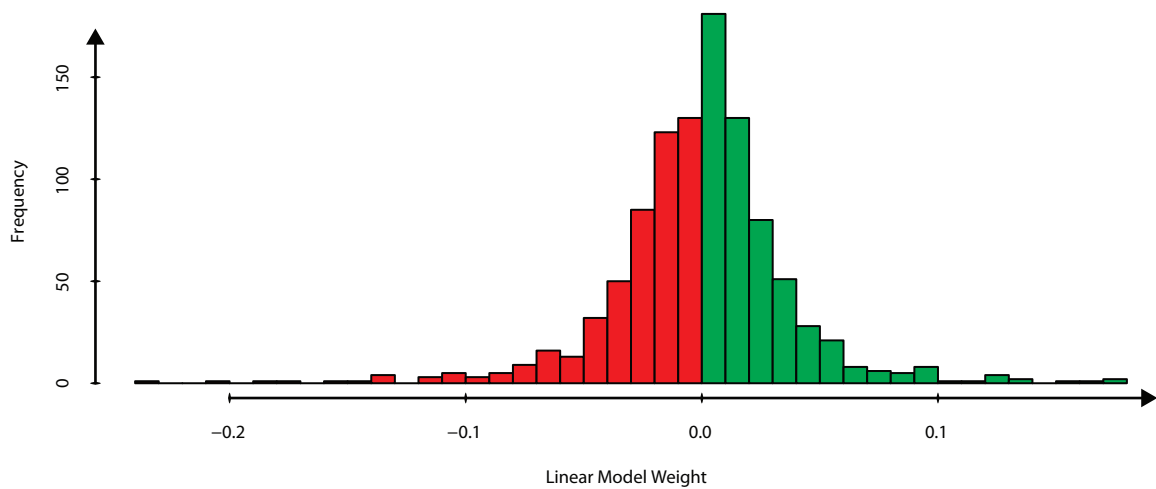
(a) Histogram of Non-Zero Model Weights

drug.d4T	
drug.EFV x POE.3FTC	
drug.NVP x POE.3FTC	
drug.FTC x drug.RAL	
drug.FPV x POE.AFPV	
drug.d4T x PON	
drug.APV x drug.LPV	
drug.AZT x PON	
drug.DRV x POE.TPV	
drug.ddl x drug.DRV	
drug.LPV x PON	
drug.EFV x POE.NVP	
drug.TDF x PON	
drug.3TC x POE.3FTC	
drug.FTC x POE.AZT	
drug.LPV x POE.AFPV	
drug.RAL x past.RAL	
drug.FTC x POE.3FTC	
drug.LPV x POE.LPV	
drug.EFV x POE.EFV	

Green boxes (positive):
drug.d4T x POE.3FTC
drug.ATV x POE.NVP
drug.FPV
drug.3TC x drug.FPV
drug.LPV x POE.ddC
drug.LPV x POE.3FTC
drug.FPV x POE.EFV
drug.TDF x drug.FTC
drug.ATV x POE.3FTC
drug.ATV x POE.EFV
drug.DRV x POE.AZT
drug.DRV
drug.3TC x PON
drug.DRV x POE.3FTC
drug.NVP x PON
drug.ATV
drug.FTC x PON
drug.LPV
drug.EFV
drug.FTC

(b) Largest Model Weights

Figure 4.8: Linear Model Weights. Linear weights were extracted from the final linear Support Vector Machine trained on the number of aviremic semesters. Among a total of 1,128 feature weights, 1,1014 (90%) were non-zero. In Figure (a), a histogram of non-zero feature weights is depicted. In Figure (b), an excerpt of the model weights that depend on the therapy drugs are graphically represented. The size of the boxes in Figure (b) is proportional to the absolute value of the weights. The 20 negative weights with the largest absolute value are represented by red boxes, while the 20 largest positive weights are represented by green boxes. Features can be either drugs, e.g. drug.EFV; probabilities of exposure, e.g. POE.EFV; the probability that the sequence is naïve, PON; or drug-use history, e.g. past.RAL. Interactions between features are denoted by a cross, e.g. drug.EFV x POE.EFV. 3TC: lamivudine; 3FTC: lamivudine or emtricitabine; AFPV: amprenavir or fosamprenavir; APV: amprenavir; ATV: Atazanavir; AZT: Zidovudine; d4T: stavudine; ddC: zalcitabine; DRV: darunavir; EFV: efavirenz; FPV: fosamprenavir; FTC: emtricitabine; LPV: lopinavir; NVP: nevirapine; RAL: raltegravir; TDF: tenofovir; TPV: tipranavir.

■ Negative Weights   ■ Positive Weights

In this chapter, I use the NAS as a success definition for training a model for predicting the success of cART. The arguments in favor of using the NAS in the context of data from routine clinical practice are the following. (1) Since the NAS averages VLs by treatment semester, it corrects for differential intervals of VL monitoring. Depending on the country of treatment, treating physician, patient, and treatment stage, VLs may be monitored in intervals from one to six months [228, 280]. (2) The NAS is robust to transient increases of the VL that are not due to virologic failure, as well as to slow decrease of the VL after treatment initiation (Section 4.1.2). Treatment guidelines recommend a change of treatment if the VL rebounds to or remains above 50 to 500 HIV-1 RNA copies per milliliter of blood serum 24 weeks after treatment initiation [228, 280]. Among 4,000 TEs in EuResist$_D$ with a NAS greater than two, 990 (25%) TEs present two consecutive VL measurements below 50 HIV-1 RNA copies per milliliter of blood serum after which the VL rebounds above 50 HIV-1 RNA copies per milliliter of blood serum with subsequent re-suppression of the VL below 50 HIV-1 RNA copies per milliliter of blood serum. 483 (49%) of these transient rebounds consist of VLs greater than 200 HIV-1 RNA copies per milliliter of blood serum, and 171 (17%) consist of VLs greater than 1,000 HIV-1 RNA copies per milliliter of blood serum. Furthermore, in 1,064 (27%) TEs with a NAS greater than two, the first of two consecutive VL measurements below 50 HIV-1 RNA copies per milliliter of blood serum occurred later than 24 weeks after treatment initiation. (3) The NAS indirectly considers the potency, side-effect profile, and the adherence-fostering characteristics of a cART. Specifically, less potent therapies take longer to suppress the VL, resulting in a smaller NAS. Under the assumption that physicians undertake everything within their possibilities in order to offer their patients the best possible therapy, a cART with an unfavorable side-effect profile will tend to be replaced by a more promising alternative, resulting in a smaller NAS. Furthermore, if patients tend to be more adherent to some therapies than to others, adherence-fostering therapies should incur in less virologic failure due to resistance development and result in less transient VL rebounds, which results in a greater NAS.

Comparison of NAS to Other Therapy-Success Definitions

As mentioned above, a significant proportion of the TEs in EuResist$_D$, which were generated in clinical routine, document therapies that have elicited a considerable degree of viral suppression (at least two aviremic semesters) although the treating physician did not follow treatment guidelines. In the context of training models for predicting the success of antiretroviral therapies, the use of few VL measurements for determining treatment success is not robust to transient viremia and cannot simultaneously consider treatment potency, side-effect profile, and propensity to success. Furthermore, treatment-success definitions based on VLs measured at a certain number of weeks after treatment initiation constrain the number of available training samples. These issues are solved by using the NAS. Improvements in model training have been achieved with a model that considers all VLs measured during the first year of therapy [370]. The parameters of this model include the number of the week after treatment initiation at which the VL was measured. This approach presents the following disadvantages. First, only the initial year of therapy is considered, which makes the model blind for therapies that work (or do not work) for a longer time. Second, prediction of therapeutic success with this model requires the selection of a target-week number. Both of these disadvantages can be avoided by using the NAS. A further advancement in model training was attained by regressing the area under the VL curve [381]. While cumulative VL seems to have a prognostic value [382, 383], the area under the VL curve fails to correct for differential intervals of VL monitoring and can potentially penalize long-lasting therapies. Specifically, the more frequently a therapy is monitored, the higher the chance that a transient increase in the VL is observed. Furthermore, the longer a

therapy lasts, the longer it has the chance to blip. The NAS considers cumulative VL while correcting for differential VL monitoring intervals. Additionally, penalization of long-lasting therapies due to transient viremia can be avoided by selecting an appropriate VL threshold. Another approach used by a further therapy-success prediction system quantifies therapeutic success with time it takes for the VL to rebound [368]. However, again, this measure in not robust to transient increases of the viral load.

Discussion of Model Selection

Model parameters were selected using cross-validation, yielding parameters for each tested treatment-success definition and VL threshold. The cross-validation performance of all tested treatment-success definitions was higher with VL thresholds above 50 HIV-1 RNA copies per milliliter of blood serum. Nonetheless, I decided to only train final models with a VL threshold of 50 HIV-1 RNA copies per milliliter of blood serum, in addition to a final model trained with EuResist labels, which use a fixed VL threshold, for the following reasons. Both a VL threshold of 50 HIV-1 RNA copies per milliliter of blood serum [228, 280] and the the EuResist labels [300, 302] have been established as treatment-success definitions. Regarding week-8 labels, a threshold of 50 HIV-1 RNA copies per milliliter of blood serum is probably too low, as patients with high VLs at baseline may require longer for attaining virologic suppression. Instead, a higher VL threshold and the consideration of the extent of VL reduction is more appropriate. EuResist labels use both a higher VL threshold for week-8 viral loads and consider the extent of VL reduction. Regarding week-24 and week-48 labels, VL thresholds above 50 HIV-1 RNA copies per milliliter of blood serum are regarded as unsuitable, since treatment guidelines prescribe VL suppression 24 weeks after treatment initiation, and since these definitions determine therapeutic success with only one VL measurement. Thus, although week-w labels with higher VL thresholds are easier to predict correctly, I think that they deviate from clinical definitions of therapeutic success in such a way that would be deemed unacceptable by clinical experts. With regard to NAS as a treatment-success definition, higher VL thresholds lead to only slightly higher performances, such that a deviation from the standard VL threshold cannot be justified. When evaluating the cross-validation performances of the selected models, one should bear two things in mind. First, the number of training, and therefore also test samples, decreases as the target VL week increases (Table 4.3). Second, virological suppression increases with the target VL week (Table 4.3), as failing therapies are interrupted at earlier time points. Thus, cross-validation performances for week-w labels with higher values of w neither account nor correct for failing therapies that are interrupted prior to week w. For these reasons, I prefer models trained with NAS and a VL threshold of 50 HIV-1 RNA copies per milliliter of blood serum threshold over models trained with other treatment-success definitions, in spite of their higher cross-validation performances.

Discussion of the Prediction of Therapy Success

As mentioned in Section 4.2.2, the test sets used for assessing the performance of pNAS in predicting therapeutic success, EuResist$_T$ and HIVdbTCE, differ significantly from the development sets EuResist$_D$ and TF in several ways. The most striking differences between these datasets can be found in the numbers of recorded past treatment lines, the duration of the therapies, the frequency of drug-resistance mutations in the baseline genotypes, and the therapy success rates (Table 4.3). In spite of all these differences, pNAS generalize well. Overall performance in predicting ITT NAS on EuResist$_T$ is very close to cross-validation performance, with performances on the subsets of EuResist$_T$ not decreasing below 0.69 (Figures 4.3, 4.4, and 4.5). Overall per-

formance in predicting ITT NAS on HIVdbTCE is comparatively lower. However, treatment length in this dataset is very often 52 weeks and never greater than 52 weeks, which is suggestive of undeclared censoring of the therapy VLs after 52 weeks. Inspection of performances in predicting therapeutic success for a certain target week uncovers high performances comparable or higher than those attained on EuResist$_T$ (Figures 4.3, 4.4, and 4.5). When performance figures are averaged across all therapy-success definitions, pNAS average performance is higher than that of GSS or of almost all models trained with other therapy-success definitions, regardless of the test data set and subset. The only exception to this was the equal average performance of the model trained with EuResist labels on therapies without drug-resistance mutations at baseline. Models trained with EuResist labels showed a high performance on 18 HIVdbTCE TEs, but not on 594 EuResist$_T$ TEs without drug-resistance mutations at baseline. pNAS performance remains comparatively high even when tested on treatment-naïve patients or on patients with no resistance mutations on the baseline genotype, which is a factor common to the majority of treatment-naïve patients. One therapy-success prediction system reports a performance of 0.87 (AUC) on 375 TEs from HIVdbTCE [371]. This performance figure was computed on a TE set which smaller than the one I use for performance assessment ($n = 580$). Furthermore, performance was computed on all available VL measurements at once, as the model allows for selection of the desired target week. Therefore, direct comparison to pNAS is not possible.

Discussion of the Ranking of Drug Combinations

Ranking of a list of 12 drug combinations showed strong preference of FTC+TDF+EFV over other drug combinations when no drug-resistance mutations are present in the baseline genotypes. FTC+TDF+EFV is most frequently prescribed as a single-tablet regimen and the regimen has shown superior short- and long-term efficacy in patients without drug-resistance mutations at baseline, including therapy-naïve patients [345, 384]. However, EFV has a low genetic barrier to resistance, such that single mutations in HIV-1 can render the drug ineffective [385]. In the presence of drug-resistance mutations, FTC+TDF+DRV/r and 3TC+ABC+DRV/r were significantly preferred by the model over other drug combinations. In light of the high potency and genetic barrier to drug resistance of DRV/r [335], its use in patients with drug-resistant HIV-1 can be generally said to be more appropriate than the use of EFV. Furthermore, IDV/r-containing drug combinations were never ranked at the top of the list. Being a boosted PI, IDV/r has a higher barrier to drug resistance than NNRTIs, such as EFV [185]. However, the unfavorable side-effect profile of IDV/r [162] has discouraged its use. This has been captured by the models trained with NAS. With respect to the backbones, 3TC+AZT-containing drug combinations were never ranked at the top of the list. This is in line with clinical studies suggesting the inferiority of 3TC+AZT as an NRTI backbone when compared to FTC+TDF and 3TC+ABC [386, 387]. With regard to first-line therapies, the correctness of ranking of FTC+TDF+EFV at the top of the list in TEs with TDR mutations at baseline is disputable, due to EFV's low genetic barrier to drug resistance. Specifically, with Sanger bulk sequencing, single drug-resistance mutations have been anectodally reported to be indicative of further drug resistance mutations in minority viral populations. In contrast, studies in which minority viral populations have been sequenced do not report additional drug-resistance mutations than those revealed by Sanger bulk sequencing (Section 4.1.3). However, ranking of FTC+TDF+EFV at the top of the list only occurred in a minority of first-line therapies with TDR mutations at baseline. DRV/r-containing drug combinations were significantly more often ranked at the top of the list when TDR mutations were present at baseline.

Discussion of the Analysis of Misclassifications

Analysis of (apparent) misclassifications with the most extreme pNAS (Tables 4.5, 4.6, and 4.7) revealed the following facts. All apparently misclassified failure-labeled TEs do not present relevant drug-resistance mutations in their baseline genotypes. Relevant drug-resistance mutations are those which are likely to preclude therapeutic success. Furthermore, all TEs presented viral suppression at some point in time other than the one used for label determination. Along with the rates of transient VL rebound mentioned above, this is a reason against the use of a fixed week for dichotomization of TEs into success and failure, for training the therapy-success prediction system. Success-labeled misclassified TEs are of two kinds: (1) TEs with numbers 4 and 9 present drug-resistance mutations which are relevant for their drug compounds. Success of TE number 4 could be attributed to the resensitizing effect of RT mutation 184V on AZT, along with the fixating effects 3TC has on this mutation [333]. Among the TEs in EuResist$_D$, 333 (3%) present RT mutation 184V while including AZT and 3TC in their drug compounds. The mean NAS for this TE subgroup is 1.65 with a median of 0 and an interquartile range of 2-0, confirming that drug regimens including AZT and 3TC do not perform well, in general, when the baseline genotype presents 184V [388]. Five drug compounds are used in TE 9, including ddI [389] and TPV/r [390], which have several disadvantages (Section 1.5). Although the regimen suppresses the VL , it has a brief duration, and attains one aviremic semester. (2) TEs with numbers 5, 6, 10 and 14 do not present relevant drug resistance mutations at baseline. However, none of them could attain more than 2 ITT uncensored aviremic semesters.

Interpreting Therapy-Success Predictions

The model for predicting the NAS features interpretability through the use of a linear SVM. The resulting linear model weights can be interpreted to be a statistical summary of the relationship between the input features in EuResist$_D$ ∪ EuResist$_{TF}$ and the NAS. In my view, many of the model weights depicted in Figure 4.8(b) are directly interpretable since they are in line with basic knowledge on antiretroviral therapy. For instance, I deem intuitive that the inclusion of a drug in a cART (e.g. drug.FTC) or the interaction of the inclusion of a drug with the PON (e.g. drug.FTC x PON) obtain positive weights, since the drugs have antiretroviral activity that can be reduced in the case of drug resistance. In the same manner, it is intuitive that interactions of a drug with the POE for the baseline genotype and that drug (e.g. drug.EFV x POE.EFV) or a drug exhibiting cross-resistance (e.g. drug.EFV x POE.NVP) obtain a negative weight. The POE is correlated with drug resistance (Section 3.3), and drug resistance can preclude antiviral activity of the drugs. Other model weights are less intuitive and require examination of the development set for satisfying interpretation. For example, the interaction of DRV, a potent PI [335], with the PON resulted in a negative weight (not shown in Figure 4.8(b)). In the development set, the baseline genotypes of 270 DRV-containing TEs have a PON above 80%, with a mean NAS of 1.81, a median of 1, and an IQR of 3-0. In contrast, 377 DRV-including TEs have a PON below 20%, with a mean NAS of 2.48, a median of 2, and an IQR of $4 - 0$ ($p = 0.028$ for the difference in the NAS). In contrast, 840 EFV-including TEs have a PON above 80%, with a mean NAS of 3.7, a median of 3, and an IQR of 6-1. EFV-including TEs with a PON below 20% ($n = 838$) present a mean NAS of 1.2, with a median of 0, and an IQR of $1 - 0$ ($p < 10^{-15}$ for the difference in the NAS). Thus, analysis of the development set could deliver a justification for the negative weight of the interaction between DRV and the PON.

Individual therapy-success predictions can be interpreted by using the linear weights of the model as well. As mentioned in Section 4.2.2, input features can be classified as either drug-dependent or solely sequence-

(a) 4.5 Predicted Aviremic Semesters
(b) -0.1 Predicted Aviremic Semesters
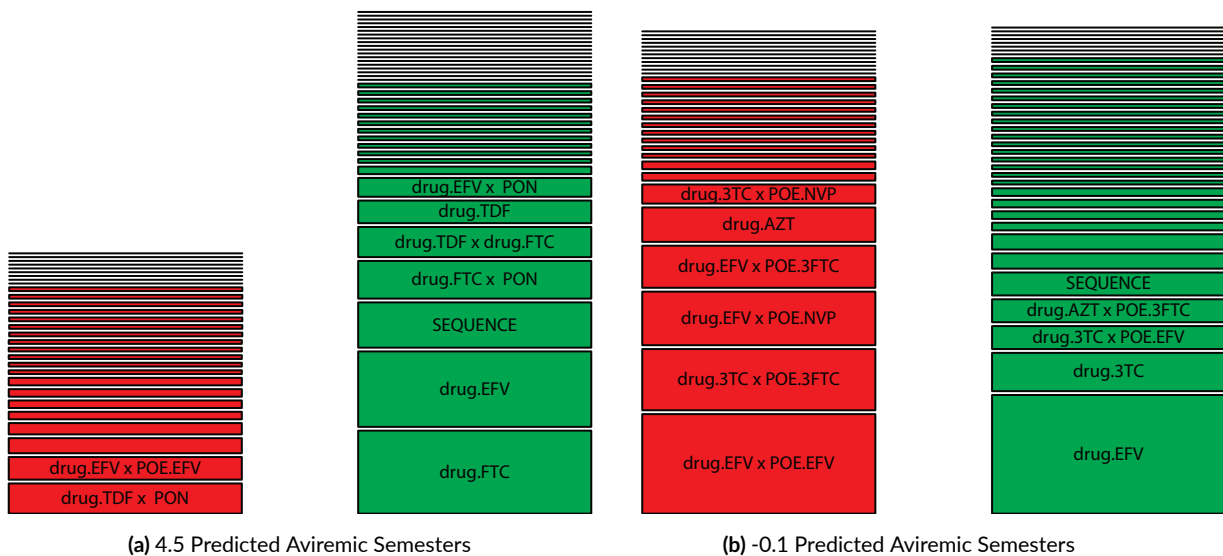
Figure 4.9: Interpretation of Therapy-Success Predictions. Above, a graphical depiction of the component values for prediction of the NAS is shown. Red rectangles represent negative component values, while green rectangles represent positive component values. The areas of the rectangles are proportional to the absolute values of the component values. Only the largest rectangles are labeled in order to avoid incurring in excessive detail for the interpretation of the prediction. Component values solely dependent on features related to the genotype are summarized in the form of a *sequence* component value. Figure (a) depicts the interpretation of a prediction for a treatment episode with a baseline genotype with no major drug-resistance mutations and the drug compounds FTC, TDF, and EFV. Figure (b) depicts the interpretation of a prediction for a treatment episode in which the drug-resistance mutations 103N, 184V, and 190A are present in the baseline genotype and the drug combination 3TC+AZT+EFV is used. More details on the interpretation of the prediction can be found in the main text. Feature interactions are denoted by a cross. 3TC: lamivudine; 3FTC lamivudine or emtricitabine; AZT: zidovudine; EFV: efavirenz; FTC: emtricitabine; POE: probability of exposure; PON: probability of naïvety; TDF: tenofovir disoproxil fumarate.

🟥 Negative Component Values   🟩 Positive Component Values

dependent. Solely sequence-dependent input features are those which do not involve drugs and whose value does not change given a baseline genotype and information on past INI-use. In contrast, drug-dependent features change their value according to the therapy for which the prediction is required. pNAS is calculated by multiplying the value of each input feature by the value of the corresponding model weight, and subsequently adding the resulting *component* values. The component values resulting from multiplication of input features with model weights can be used for interpretation of the prediction. I propose the following interpretation. Component values resulting from multiplication of solely sequence-dependent features with their corresponding model weights should be added and thus summarized in a solely sequence-dependent component value. This *sequence* value can be interpreted to be an offset representing the easiness or difficulty of eliciting therapeutic success with a given baseline sequence. Drug-dependent component values can be displayed individually, thus exhibiting individual components of the prediction. Positive and negative component values can be represented by green and red rectangles, respectively, whose area is proportional to the absolute values. Two examples for the interpretation of predictions of the NAS for TEs in EuResist$_T$ can be found in Figure 4.9. Figure 4.9(a) depicts the interpretation for a prediction for the drug combination FTC+TDF+EFV and a baseline genotype with no major drug-resistance mutations. For this TE, 4.5 aviremic semesters were predicted. The sequence-

specific component value for this prediction is positive, and is represented by a green rectangle. The drugs FTC, TDF, and EFV, as well as the interaction between FTC and TDF, are always assigned positive component values, and are represented by green rectangles. Additionally, positive component values for FTC and EFV result from multiplication of the PON with the corresponding model weights. A negative component value results from the multiplication of the PON with a model weight relating to TDF. Finally, the multiplication of the POE for EFV with the corresponding model weight results in a negative weight. Component values with small absolute values are represented with unlabeled rectangles in Figure 4.9. In this way, their influence is graphically represented without incurring in excessive detail for interpretation of the prediction. In Figure 4.9(a), one can see that the positive component values outweigh the negative component values, resulting in a prediction of several aviremic semesters. The interpretation for a prediction for the drug combination 3TC+AZT+EFV and a baseline genotype with major drug-resistance mutations is shown in Figure 4.9(b). The drug-resistance mutations 103N, 184V, and 190G are indicative of drug resistance against 3TC and EFV. For this TE, -0.1 aviremic semesters where predicted. As in Figure 4.9, a positive sequence component value is shown, albeit smaller. Furthermore, positive weights result from the inclusion of the drug 3TC, and from the multiplication of POEs with some of the included drugs. The positive weights, however, are outweighed by the negative weights, resulting in a slightly negative prediction for the NAS. Note that a negative prediction of the NAS is tantamount to a prediction of the NAS equal to zero, i.e. the therapy is predicted to fail. Negative component values decrease the pNAS due to resistance to 3TC and EFV, which is expressed in the model in terms of the POEs. Due to cross-resistance (Section 1.5.3) between EFV and NVP, the POE for NVP also decreases the pNAS. The negative component value for AZT can be attributed to a generally comparatively bad efficacy of the drug, which can be compensated in certain situations. Specifically, it is known that resistance to FTC and 3TC can increase susceptibility to AZT [333], which is accounted for in the positive component value considering the use of AZT and resistance to FTC or 3TC. Predictions for POEs and the POE are interpretable as well, due to the fact that these are calculated with linear SVMs as well (Section 3.3).

Interaction of the User with Therapy-Success Predictions

Therapy-success prediction systems are often made available through a web service. The typical *modus operandi* of these systems is the following [302, 370]. The user inputs patient baseline information into system and selects a list of drug combinations. Using the patient baseline information, the system produces a prognostic score for each drug combination in the list. The output of the system is a table containing the drug combinations sorted by the prognostic score, which is included in the table as well. In addition to the ranking of lists of drug combinations, I propose the following mode of operation for therapy-success prediction systems. Users should be allowed to interactively obtain therapy-success predictions, such that they can incrementally compose different drug combinations. This can be especially useful when composing a therapy for a patient whose drug options are limited, e.g. due to a combination of viral drug resistance and restrictions with respect to possible drug combinations. Specifically, the treating clinician can input into the system the drug combinations the patient can take and view a prediction for the therapeutic prospects of these drug combinations. In order to achieve a clear arrangement of therapy alternatives, trees should be used for representing drug combinations with drug compounds in common (Figure 4.10). In these trees, parent nodes represent the drug compounds that the combinations have in common. Furthermore, predictions performed with adequate drug-combination lists can be visualized with either a *list view* or a *tree view*. In the context of patients with many therapy options,
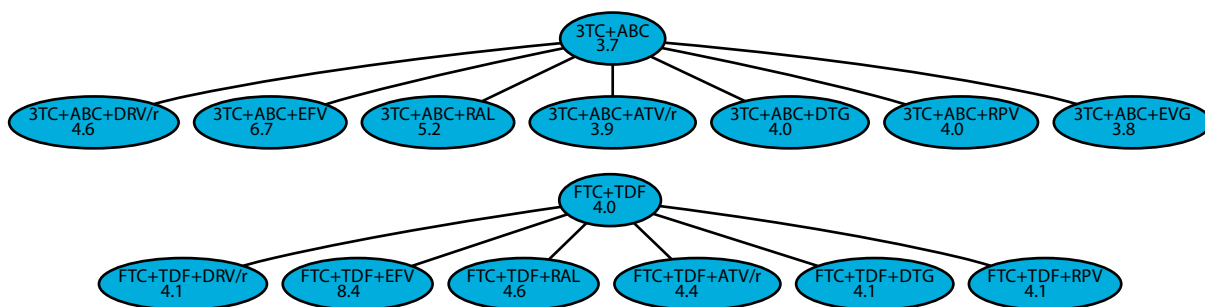
Figure 4.10: Tree Visualization for Therapy Success Predictions. Above, therapy-success predictions for a baseline genotype without major drug resistance mutations and several different drug combinations are depicted in the form of trees. Drug combinations containing drug compounds in common are grouped in a tree, with parent nodes representing the compounds that the combinations have in common. pNAS are displayed in each node of the tree. When parent nodes contain *NRTI backbones*, a direct comparison of the predicted efficacy of the backbones is possible. Furthermore, interactions between the backbones and further compounds are depicted as well. Tree visualization for therapy-success prediction can be constructed from a list of drug combinations or by the user, manually. For discussion in the main text, non-validated predictions for the drug compounds dolutegravir (DTG) and rilpivirine (RPV) are shown as well. /r: ritonavir as a boosting agent; 3TC: lamivudine, ABC: abacavir, ATV: atazanavir; DRV: darunavir, DTG: dolutegravir; EFV: efavirenz; RAL: raltegravir; RPV: rilpivirine; TDF: tenofovir disoproxil fumarate

the tree view is a form of visualization of the prediction that can be especially useful when the user desires to compare different NRTI backbones (Figure 4.10).

Concluding Remarks

In the following, I state some important remarks regarding the use of models for predicting the NAS. (1) The NAS is a quantity that assumes non-negative integer values. However, I do not enforce this property when computing the pNAS, such that these can also assume negative and non-integer values. Note that a negative prediction of the NAS is tantamount to a prediction of the NAS equal to zero, i.e. the therapy is predicted not to be successful at all. (2) Harrell's concordance index was used for selecting model parameters and assessing the performance of the model. While the concordance index quantifies concordance between the NAS and the pNAS, it does not consider their possible numerical deviation (Section 2.2.6)). Therefore, the pNAS is a prognostic score correlated to the NAS without necessarily predicting its exact value. For this reason, neither integrality nor positivity are enforced when computing the pNAS. (3) The pNAS is estimated with an SVM for the regression of right censored data. SVM ϵ-regression can be phrased as a regularized regression with an error-insensitivity parameter (Section 2.2.5). Due to regularization, function coefficients (model weights) are estimated with the smallest possible value. This characteristic of the regression method that I employ has consequences for novel drug compounds. Specifically, the model can be expected to have a bias towards drug combinations that have worked well in the past, since in the development set, many TEs including these drug combinations have high NAS. Due to their performance in clinical studies, some novel drug compounds may be expected to be more efficacious than older compounds. From the perspective of the development set, however, drug combinations including novel drug compounds have not had the time to show whether they can be associated to large NAS or not. Even if most of the NAS for novel drug combinations are marked as censored, regularization will not allow for large pNAS. An example for this can be found in Figure 4.10. In the figure, non-validated pNAS for the drug compounds DTG and RPV are included. These novel drug compounds are said to be superior than EFV [202, 229, 356, 391]. However, pNAS for drug combinations including these drugs are lower than similar drug combinations including EFV. Thus, the availability of predictions for drug

combinations including novel drug compounds might need to be deferred until sufficient evidence for the efficacy or these novel drug compounds can be found in the development set (or not). There is no proxy for this evidence since the long-term performance of novel compounds is unknown. (4) The model for predicting the NAS does not include a mechanism for lowering the value of the prediction if too many drug compounds are included in a drug combination. Standard cART includes three different drug compounds. Even though some patients might require more than three drug compounds in order to attain viral suppression, therapy with more than three drug compounds is associated with therapeutic failure, mainly due to the toxicities of the drug compounds (Section 1.5). However, this shortcoming of the models can be easily counteracted by normalizing the pNAS with respect to therapies consisting of three drug compounds in which full drug susceptibility of the virus against all drug compounds is given.

In this work, a novel, quantitative measure for therapeutic success is presented, the NAS. Retrospective evaluation of cART with the NAS presents a number of advantages. (1) It provides an evaluation of the performance of a drug combination that considers the entire duration of the therapy. (2) The measure is robust to transient viremia. (3) The measure implicitly considers side effects and adherence-fostering characteristics of the different drug combinations. (4) NAS is anticorrelated to cumulative viral load, which has a negative prognostic value [382, 383]. In contrast, the established method for determining virologic success dichotomizes a therapy into success and failure with a VL obtained at a certain point in time. When several VLs are available for a therapy, dichotomization can result in discordant labels for one therapy. In this chapter, I present a method for predicting the NAS. With this method, pNAS, a prognostic score for therapy success can be calculated. pNAS not only considers determinants of long-term therapeutic success that are related to drug-resistance development, but also to other characteristics of the drug combinations. This fact manifests itself in the sustained performance on both TEs with and without drug-resistance mutations at baseline. Furthermore, pNAS incorporates POEs and the PON, which are derived from a methodology for predicting therapeutic history from the viral genotype. It has been shown that therapeutic history is a strong predictor of therapeutic success that can significantly boost the performance of a therapy-success prediction system [302, 370], and that therapeutic history alone is sufficient for accurately predicting therapeutic success [376, 378]. However, complete therapeutic history may not be available for every patient, and if it is available, it might not always be efficiently communicated to diagnostic laboratories, where patient samples are analyzed and diagnostic reports written. For this reason, I believe that improvement of therapy-success prediction systems will not be achieved by disregarding the genotype, but by making better use of the genotype. pNAS affords interpretability which is important in the context of health care, as it can both foster confidence in the interpretation system and explain (apparently) implausible predictions.

5

Assessing the Robustness of Genotype Interpretation Systems

Nobody wants to be responsible for an erroneous diagnosis or for a failed therapy, as this could result in detrimental consequences for the patient. The use of statistical models in a biomedical context dictates increased requirements for the robustness of these models. The robustness of a model is the extent by which its output changes when its input presents a certain degree of noise or variability. Since the genotype is the input of all models presented in this work, in this chapter, I present a method for assessing the robustness of genotype interpretation models with respect to sequencing variability. First, I present a summary on the limits of the Sanger sequencing technology. After this, I introduce geno2pheno$_{[coreceptor]}$, a genotype interpretation method whose robustness is assessed in this chapter. Lastly, I present an analysis on the robustness of geno2pheno$_{[coreceptor]}$ and of models for predicting drug exposure from the genotype. This chapter is based to a large extent on [392].

5.1 Variability and Limits of Sanger Nucleotide Sequencing

The term *Sanger sequencing* refers to a family of methods for determining the sequence of nucleotides of a sample of DNA molecules. In Sanger sequencing, DNA is amplified using a mixture of deoxynucleotides and of labeled dideoxynucleotide chain terminators, resulting in amplicons of different lengths. Subsequently, a sequence of nucleotides is resolved by detecting the labeled dideoxynucleotides of the electrophoretically separated amplicons. In the following, I state the typical steps that are carried out when HIV-1 from a clinical isolate is sequenced with the Sanger method. These steps aim at obtaining the nucleotide sequence of a specific region of the genome of HIV-1. They are a summary of the relevant standard operating procedures of the Institute of Virology of the University Clinic of Cologne ([361, 393, 394] and personal communication).

1. Collection of blood samples
2. Isolation of HIV-1 RNA and/or proviral DNA
3. Amplification of the genomic region of interest. Reverse-transcription polymerase chain reaction (RT-PCR) is used in samples containing viral RNA and polymerase chain reaction (PCR) is used in samples

containing proviral DNA. Initial amplification is usually followed by a further amplification with nested PCR

4. Amplification in sequencing PCR with fluorescent-dye-marked dideoxynucleotides
5. Purification of sequencing-PCR products
6. Separation of purified sequencing-PCR products with capillary electrophoresis and subsequent detection of fluorescent-dye-marked dideoxynucleotides
7. Automated base calling with subsequent manual editing of the automatically generated sequence.

The steps listed above describe what is known as a *direct*, *bulk*, or *population* Sanger-sequencing approach. This approach results in a consensus sequence of the most frequent variants in the viral population. Due to the variability of HIV-1, viral RNA and DNA extracted from clinical isolates exists as a mixture of closely related variants (Section 1.4). For this reason, multiple, non-identical DNA templates participate in PCRs performed for sequencing (a region of) the viral genome. Population sequencing with the Sanger method has been validated for the simultaneous detection of mutations in mixtures of HIV-1 variants [395, 396]. Minority populations in the sequenced sample can be reliably detected when they account for more than 20% of the sample. However, minority populations representing more than approximately 10% of the sample may be detected with Sanger sequencing as well [395, 396]. In addition to the inconsistent detection of minority populations accounting for less than 20% of the total population, further variability in replicate sequencing experiments can be observed. In the following, I list possible sources for this variability. (1) Multitemplate PCR is subject to amplification bias. When single mismatches between primer and template are present, amplification efficiency for such primer-template duplexes is substantially decreased [397]. (2) Multitemplate PCR may produce chimeric amplicons. Chimeras may result from template switching during DNA polymerization or incomplete extension of primers. DNA molecules arising from incomplete extension of primers may act, in turn, as primers in PCR cycles [397] following their synthesis. (3) Random events such as misannealing of primers and misincorporation of bases occur during PCR. These random events may produce artifacts. If artifacts are produced at an early cycle in PCR, these may be amplified to large quantities [397]. (4) Deterioration of the capillary used for electrophoresis decreases the resolution of the sequencing instrument [398]. The quality of capillaries for electrophoresis is guaranteed to remain high for a certain number of runs. However, diagnostic laboratories may use the capillaries for a number of runs greater than the number of runs covered by the warranty (personal communication). (5) Determination of the nucleotide consensus sequence depends on the base calling procedure. While automatic base calling generally produces accurate results, manual post-editing of the automatically generated sequence can lead to improved sequencing accuracy. Manual post-editing is reported to be especially accurate in assigning nucleotide mixtures or pure base calls [399].

While Sanger sequencing is still considered the gold standard in terms of combined sequencing accuracy and read length, other sequencing methods exist. Within the last decade, novel high-throughput sequencing methods have emerged, often referred to as *next-generation* or *massively parallel* sequencing methods. These methods can produce up to 2 giga base pairs of data in a single run (Ion Torrent ®; [400]). When compared to Sanger sequencing, massively parallel sequencing methods require significantly less time and monetary resources. In the context of sequencing of heterogeneous viral populations, such as those found in patients with HIV-1 infection, massively parallel sequencing methods present the advantage that they can easily produce a large number of sequence fragments representing the same genomic region. This characteristic allows for the resolution of minority variants accounting for as little as 0.1% of the total viral population [400]. Massively parallel sequencing

methods are currently mainly used in research settings, and not in clinical diagnostics settings, due to a number of hurdles for their wide-scale implementation. First, the technology has high start-up costs. Second, the error rates of massively parallel sequencing are reported to range from approximately 0.1% to 13%, depending on the sequencing plattform [400]. In contrast, error rate of Sanger sequencing ranges from 0.1% to 1% [401]. Due to the increased error rates of some of these sequencing methods, regulatory authorities often require certification of the methods with respect to their ability to differentiate low-level mutant populations from sequencing and amplification errors. Third, the handling of the large amounts of data that these methods produce requires a high degree of bioinformatics expertise. Fourth, the diagnostic value of the additional information provided by massively parallel sequencing methods is still uncertain. Nonetheless, a growing number of laboratories have validated massively parallel sequencing methods for routine diagnostics [400, 402–404].

5.2   Genotypic Determiniation of Tropism

HIV-1 employs two host molecules in order to enter the host cell: the CD4 receptor and a coreceptor (Section 1.1.2). In vivo, either of two coreceptors can be: CCR5 and CXCR4. The capability to use a certain coreceptor is called viral tropism. HIV-1 so-called R5 strains can only use CCR5. CXCR4-capable strains can use either CXCR4, exclusively (X4 strains), or both coreceptors (dual/mixed tropic viruses) [1]. MVC is an antiretroviral drug that inhibits HIV-1 entry into the cell by binding to CCR5, and is thus ineffective against X4-capable strains. Therefore, viral tropism determination must precede MVC prescription [191].

Tropism can be determined phenotypically or genotypically [277, 405–407]. Phenotypic determination in cell cultures is expensive, time-consuming, and requires specialized labs. Furthermore, samples with VLs up to 1000 HIV-1 RNA copies per milliliter of blood serum often yield indefinite results, although proviral DNA testing is performed as well [408]. Genotypic determination of tropism requires sequencing V3 with subsequent computer-based interpretation. Several methods for interpreting sequences in order to determine HIV-1 tropism have been developed [407]. geno2pheno$_{[coreceptor]}$ [409] is an extensively validated bioinformatic method for genotypic determination of tropism [277, 405–407, 410–412]; its use as an alternative to phenotypic determination is recommended by the European and the Austrian-German HIV-treatment guidelines [413–415]. geno2pheno$_{[coreceptor]}$ interprets V3 with an SVM (Section 2.2) trained on GPPs. The output of geno2pheno$_{[coreceptor]}$ is the false-positive rate (FPR) with X4-capable being defined as positive [409]. FPR dichotomization yields a (predicted) viral classification into X4-capable or R5. When the FPR is in a range where MVC antiviral action is considered possible, yet uncertain, the virus is classified as X4-capable. Alternatively, this intermediate FPR range can be translated into an additional *intermediate* category, as is customary for interpretation of drug resistance to other antiretroviral drugs (Section 3.2). Thus, MVC administration with an FPR in the intermediate range could be made dependent on whether other therapy options co-exist, rather than excluding it altogether. stablishment of the most suitable cutoff for FPR dichotomization has been a matter of substantial debate. Currently, there is no universally accepted cutoff.

5.3   Simulation of the Variability Resulting from Sanger Sequencing

The input to geno2pheno$_{[coreceptor]}$ is a V3 sequence (Section 5.2). Therefore, the quality of its predictions depends on the quality of these sequences. With Sanger bulk sequencing, the measured sequence is a consensus of the dominating strains in the viral population. Here, minorities comprising less than 10%-20% of the viral population are unreliably represented, due to the limits of the experimental technology (Section 5.1). X4-capable

minorities may render MVC ineffective. Therefore, some labs have suggested that performing the amplification step of the sequencing procedure in triplicate increases the chances of detecting minorities. Indeed, sequencing errors can be reduced by performing amplification in replicate and mixing of the replicate amplification products prior to sequencing [397]. The use of duplicate amplification with subsequent amplicon mixing and sequencing was reported to increase the specificity and sensitivity of genotypic determination of tropism when compared to phenotypic tropism determination as a gold standard [416]. Replicate amplification is not routinely performed when sequencing the *pol* gene, but suggested when sequencing V3, since *env* is subject to higher selective pressure than *pol*, resulting in higher variability of the gene [410]. Contrary to the established approach of mixing amplification products, treatment guidelines considered performing both amplification and sequencing in triplicate, leading to the production of three nucleotide sequences [410, 414, 417]. Interpretation of these three nucleotide sequences with geno2pheno$_{[coreceptor]}$ was recomended, with subsequent use of the minimum FPR for genotypic determination of tropism, ignoring the potential multiple-testing problem associated with this approach. For this reason, I performed an analysis that addresses two related, previously unresolved questions: (1) How robust is geno2pheno$_{[coreceptor]}$ with respect to sequencing / base-calling variability in terms of change of predicted tropism? (2) What is the influence of undetected minority populations on the predictions of geno2pheno$_{[coreceptor]}$? Both issues are of critical importance for assessing the reliability of geno2pheno$_{[coreceptor]}$ for clinical purposes. In a further analysis included in this chapter, I test the robustness of models for predicting drug exposure (Section 3.3). Since variability in the sequencing procedure is not a matter of concern when interpreting the sequence of the *pol* gene, the results concerning the drug-exposure models are much less detailed than the results concerning geno2pheno$_{[coreceptor]}$. Note that this section is largely based on [392]. Futhermore, I amply quote from this publication.

5.3.1 Materials and Methods

LA Dataset

A dataset of 163,958 HIV-1 V3 nucleotide sequences was downloaded from the LANLSD (http://www.hiv.lanl.gov/) on September, 19th 2013. Nucleotide sequences with duplicate V3 regions were discarded, resulting in the Los-Alamos dataset (LA) which comprises 67,997 nucleotide sequences. Subtypes in LA were determined with COMET [303, 304]. LA was used to create further datasets by altering its sequences *in silico*.

T$_{PRRT}$, T$_{IN}$

The T$_{PRRT}$ and T$_{IN}$ datasets were used for testing several drug-resistance interpretation systems, as described in Section 3.3. T$_{PRRT}$ contains 6,641 PR and RT nucleotide sequences from the EuResist Integrated Database and the Los Alamos National Laboratory database, while T$_{IN}$ contains 444 IN sequences from the aforementioned databases.

Electropherogram$_{V3}$, Electropherogram$_{PRRT}$, and Electropherogram$_{IN}$ Datasets

The Electropherogram$_{V3}$, Electropherogram$_{PRRT}$, and Electropherogram$_{IN}$ datasets were obtained from the Institute of Virology of the University of Cologne. Electropherogram$_{V3}$ arises from genotypic determination of tropism [393] in 164 clinical blood samples. This dataset comprises four electropherograms per blood sample, one for each of the forward sequencing primers Env2 and Env6 and one for each of the reverse sequencing primers Env7 and Env11 (Table 5.1). Furthermore, for each blood sample, an automatically generated

Table 5.1: **Sequencing Primers**. Sequencing primers used for generating the Electropherogram$_{V3}$, Electropherogram$_{PRRT}$, and Electropherogram$_{IN}$ datasets are tabulated below. Genomic coordinates (short: coordinates) are given with respect to the HIV-1 variant HXB2. Arrows between coordinates indicate whether the primers are forward or reverse. Arrows pointing right indicate forward primers. ENV: envelope; IN: integrase; PRRT: protease and reverse transcriptase.

Primer Name	Gene(s)	Nucleotide Sequence	Coordinates
Env2	ENV	GTACAATGYACACATGGAATTAGGC	$6957 \rightarrow 6981$
Env6	ENV	GGCCAGTAGTATCAACTCAAC	$6979 \rightarrow 6999$
Env7	ENV	TGTCCACTGATGGGAGGGGC	$7530 \leftarrow 7549$
Env11	ENV	TACATTGCTTTTCCTACTTTCTGCCAC	$7502 \leftarrow 7528$
A	PRRT	GAGCCAACAGCCCCACC	$2149 \rightarrow 2165$
B	PRRT	CAATGGCCATTGACAGAAG	$2616 \rightarrow 2634$
C	PRRT	GGATCACCAGCAATATTCCA	$3012 \rightarrow 3031$
D	PRRT	GGAACTGTATCCTTTARCTTCCC	$2232 \rightarrow 2254$
F	PRRT	TGGGCCATCCATTCCTGGCTT	$2586 \leftarrow 2606$
G	PRRT	CATCCCTGTGGAAGCACATT	$2988 \leftarrow 3007$
H	PRRT	TCTGCTATTAAGTCTTTTGAT	$3512 \leftarrow 3532$
PRRT-2F	PRRT	GGCTGTTGGAAATGTGGAAAGGA	$2023 \rightarrow 2045$
3p31	IN	ATCCTGTCTACYTGCCACACAA	$5066 \leftarrow 5087$
5'-INT	IN	ATTGGAGGAAATGAACAAGT	$4173 \rightarrow 4192$
IN-F	IN	GGAATTGGAGGAAATGAACAAGTAGATAAA	$4170 \rightarrow 4199$
SEQ1	IN	GAATTTGGSATTCCCTACAATCC	$4641 \rightarrow 4663$
SEQ2	IN	GGATGAATACTGCCATTTGTACTGC	$4752 \leftarrow 4776$

nucleotide sequence for V3 is included. This sequence arises from automatic base calling based on the four electropherograms. Manual post-editing of the sequence based on inspection of the corresponding electropherograms is performed for quality control, and results in a final sequence called manually edited sequence. The Electropherogram$_{PRRT}$ and Electropherogram$_{IN}$ datasets arise from genotypic determination of resistance [361, 394], and present a similar structure to Electropherogram$_{V3}$. The Electropherogram$_{PRRT}$ dataset originates from 3,104 blood samples, while 1,288 blood samples were used for creating the Electropherogram$_{IN}$ dataset, and they comprise eight or five electropherograms per blood sample, respectively. For generating the Electropherogram$_{PRRT}$ dataset, the following sequencing primers were used: A, B, C, D, F, G, H, and PRRT-2F (Table 5.1). The following primers were used in order to generate the Electropherogram$_{IN}$ dataset: 3p31, 5'-INT, IN-F, SEQ1, and SEQ2 (Table 5.1). As in Electropherogram$_{V3}$, an automatically generated and a manually edited nucleotide sequence is included for each blood sample. Electropherogram$_{PRRT}$ contains PR and RT sequences, while Electropherogram$_{IN}$ contains IN sequences.

ESTIMATION OF POSITION-WISE ALTERATION RATES FOR SANGER SEQUENCING

I used the Electropherogram$_{V3}$, Electropherogram$_{PRRT}$, and Electropherogram$_{IN}$ datasets to estimate the variability resulting from Sanger sequencing. For this purpose, I performed automated base calling on each of the 164, 3,104, and 1,288 groups, respectively, of four, eight, or five electropherograms, respectively, that were used to create the automatic sequences. Position-wise alteration rates resulted from comparison of nucleotide sequences obtained from a single electropherogram with the corresponding manually edited sequence. This procedure quantifies what would happen if sequencing was performed with only one electropherogram and without manual post-editing. In the following, I describe this procedure in detail, which was partly inspired by Recall, an automated base-caller [418]. Specifically, I used Phred [401] to preprocess electropherogram files,

yielding a quality file and a polymorphic-base file for each one of them. The quality file contains Phred quality scores for each called base. Among other information, the polymorphic-base file contains, for each nucleotide sequenced, the area of the called base peak in the electropherogram and the area of an "uncalled" base peak (i.e. the second-best candidate peak). Polymorphic-base files were used for performing base calling in each of the electropherograms. In a first step, nucleotide sequences were produced by examining each base call and calling a mixture if the ratio of the areas of the uncalled and the called bases was larger than 0.2 [418]. Otherwise, only the called base was considered. Each of the resulting sequences was aligned to its corresponding manually edited sequence. After alignment, Phred quality files were used to mark regions in the aligned sequences with Phred quality scores below 20 as unsuitable [418]. The resulting annotated sequences are called single-electropherogram sequences. Disagreement between the single-electropherogram sequence and the manually edited sequence is considered an alteration, as the manually edited sequence is a more accurate genotype measurement [399]. Five alteration types were quantified by comparing the single-electropherogram sequences with the manually edited sequences. I use IUPAC ambiguity codes (M, R, W, S, Y, K, V, H, D, B, and N) to refer to ambiguous bases [419]. The following alteration types were quantified.

1. Including ambiguous base: these are alterations that result in an ambiguous base in the single-electropherogram sequence which at least partially includes the definite or ambiguous base contained in the manually edited sequence

2. Excluding ambiguous base: these alterations result in an ambiguous base in the single-electropherogram sequence that does not even partially include the definite or ambiguous base contained in the manually edited sequence

3. Including definite base: alterations resulting in definite bases in the single-electropherogram sequence that are contained in the ambiguous base in the manually edited sequence

4. Excluding definite base: those alterations resulting in definite bases in the single-electropherogram sequence that are not contained in the ambiguous base or are different from the definite base in the manually edited sequence

5. Insertion or deletion (indel): alterations involving an insertion or a deletion in the single-electropherogram sequence with respect to the manually edited sequence.

In the following, I detail the procedure with which alterations were quantified. Manually edited sequences were aligned to the corresponding gene reference sequence. For V3 sequences, *consensus B* (105 nucleotides) was used as a reference sequence, since this reference is used by geno2pheno[coreceptor]. In contrast, HXB2 was used as a reference sequence for PR (297 nucleotides), RT (first 720 nucleotides), and IN (864 nucleotides). I quantified the number of alterations present in single-electropherogram sequences by using (1) the alignment of the single-electropherogram sequence to the manually edited sequence and (2) the alignment of the manually edited sequence to the reference sequence. The first alignment was used to determime the alterations, and the second alignment was used to determine the alignment position number, corresponding to the reference sequence, at which the alteration occurs. I tabulated the numbers of alterations separately for each primer, reference-sequence alignment position, and alteration type. Only segments of the sequences that were not marked as unsuitable were considered in alteration quantification. Sequences with more than 10% of their bases marked as unsuitable due to low Phred scores were excluded as a whole. V3 in HXB2 has the genomic coordinates 7110 - 7217 (108 nucleotides), and typically V3 presents the same or a similar length. The typical read length for Sanger sequencing is 400 to 900 base pairs [420]. Since the sequencing primers used for generating the

Electropherogram$_{V3}$ dataset (Table 5.1) anneal to a region of the HIV-1 genome close to, but outside V3, single-electropherogram sequences typically cover V3 entirely. However, single-electropherogram sequences for the PR and RT genes, as well as for the IN gene, typically only partially cover these genes. For this reason, only the gene regions covered by at least 50% of all single-electropherogram sequences produced with a given primer were used for quantification of alterations.

Let $\epsilon \in$ {Including Ambiguous, Excluding Ambiguous, Including Definite, Excluding Definite, Indel, None} denote an alteration type, let $p \in$ {Env2, Env6, Env7, Env11, A, B, C, D, F, G, H, PRRT-2F, 3p31, 5'-INT, IN-F, SEQ1, SEQ2} denote a primer, let $\beta \in$ {Ambiguous, Definite} denote whether the ith-base in the manually edited sequence is ambiguous or definite, with $i \in \mathbb{N}$ denoting the position number of an alignment. Using the alterations quantified for each primer, alteration type, and reference-sequence alignment position, as described above, alteration rates were calculated in terms of the probabilities of ϵ, conditioned on p, β, and i:

$$P(\epsilon \mid p, \beta, i). \tag{5.1}$$

The procedure described above yielded position-wise alteration rates for each primer and alteration type. Since error rates for Sanger sequencing increase linearly in 3' direction [401, 421], alteration rates were linearly regressed on the nucleotide position number. In order to determine whether a linear function describes the change in the alteration rates well, a t-test was used to assess whether the regression coefficients for the nucleotide position numbers were significantly different from zero. The resulting p-values were corrected using the method by Benjamini and Hochberg [332]. If the difference was significant at the 0.05 level, the linear regression was used for alteration estimation. Otherwise, the mean alteration rate was used. Alteration rates calculated with the Electropherogram$_{V3}$, Electropherogram$_{PRRT}$, and Electropherogram$_{IN}$ datasets are called the EP$_{V3}$, EP$_{PRRT}$, and EP$_{IN}$ alteration rates, respectively.

SangerAlteration$_{V3}$, SangerAlteration$_{PRRT}$, AND SangerAlteration$_{IN}$ DATASETS

In order to simulate the variability that arises from Sanger sequencing, primer-specific sequences were created from each sequence in the LA, T$_{PRRT}$, T$_{IN}$ datasets. Specifically, for sequences in LA, primer-specific sequences were created for each of the primers Env2, Env6, Env7, and Env11. For sequences in T$_{PRRT}$, primer-specific sequences were created for each of the primers A, B, C, D, F, G, H, and PRRT-2F. Primer-specific sequences for the primers 3p31, 5'-INT, IN-F, SEQ1, and SEQ2 were created for sequences in T$_{IN}$. Primer-specific sequence generation was accomplished by introducing alterations of the types described above, according to the EP$_{V3}$, EP$_{PRRT}$, and EP$_{IN}$ alteration rates, respectively, at the nucleotide position in question. All primer-specific sequences derived from a given *original* sequence were combined to one *final* sequence that possibly contains alterations. The details of this procedure are as follows. The defined alteration types allow many possible changes (ambiguous or definite bases) at a certain position. For creating a primer-specific sequence, one change was sampled from the uniform distribution. When the alteration was of the type indel, a deletion or an insertion was chosen with equal probability. Insertions consisted of a definite nucleotide chosen according to the uniform distribution. From each group of primer-specific sequences derived from an original sequence, a single sequence was created in the following manner: bases or indels present at each sequence position in the primer-specific sequences were tabulated. Bases or indels with a relative frequency less than 50% were discarded as putative errors [418]. The remaining bases were used for determining a definite or ambiguous base for the considered position,

or an indel. In case both indels and bases remained, preference was given to the non-indel alteration. Whenever the original sequence was longer than the corresponding reference sequence, the alteration rate for the highest alignment position number was used. The procedure for generating sequences with simulated variability was repeated ten times for each sequence in the LA, T_{PRRT}, T_{IN} datasets. The datasets of sequences with simulated variability are called SangerAlteration$_{V3}$, SangerAlteration$_{PPRT}$, and SangerAlteration$_{IN}$ datasets, respectively.

SE$_{V3}$ Dataset

The single-error dataset (SE$_{V3}$) was created by generating sequences from each LA sequence by systematically exchanging every nucleotide in V3 by each of the 15 possible definite and ambiguous bases, independently of their probability of occurrence. Thus, from each sequence in LA, all possible sequences diverging by one definite or ambiguous base were generated.

M$_{V3}$ and MS$_{V3}$ Datasets

The mixture dataset (M$_{V3}$) was created from sequences in LA containing ambiguities (excluding N). Ambiguities in each of these sequences were combinatorially resolved into all possible sequence alternatives without ambiguities. To avoid combinatorial explosion, sequences that would result in more than 20,000 derived sequences were excluded from this procedure. The mixture-sampling dataset (MS$_{V3}$) was created to simulate a scenario in which sequencing depth is insufficient to resolve all sequence variants in the sample. From each sequence group in M$_{V3}$ derived from the same LA sequence, a certain proportion of sequences was extracted at random by uniform sampling without replacement, and a new sequence was created by retaining positions that are identical among the sequences in the subset and representing differential positions with the corresponding ambiguities. Proportions represent sequencing depth and ranged from 1% to 100% in steps of 2%; each sequence group was sampled 3 x 100 times (100 repetitions that allow for triplicate FPRs).

Quantification of Changes in Predictions

Sequences in LA, SangerAlteration$_{V3}$, SE$_{V3}$, M$_{V3}$, and MS$_{V3}$ were interpreted with geno2pheno$_{[coreceptor]}$. For each sequence interpreted with geno2pheno$_{[coreceptor]}$, the FPR shift was calculated as the difference between the FPR of the altered sequence and that of its unaltered counterpart in LA. When I consider n FPRs from variability-simulation replicates on the same sequence, I call the FPR obtained with the first sequence singleton FPR. For all further sequences, I take the minimum FPR among the first n FPRs and call it nth replicate FPR. The 3rd replicate FPR is also called triplicate FPR. geno2pheno$_{[coreceptor]}$'s FPR was used to determine coreceptor tropism as X4-capable or R5. Four different FPR cutoff sets were used for tropism determination:

- {5, 10}: FPR < 5 $\Rightarrow$ X4-capable, 5 $\leq$ FPR < 10 $\Rightarrow$ Intermediate, FPR $\geq$ 10 $\Rightarrow$ R5
- {5, 15}: FPR < 5 $\Rightarrow$ X4-capable, 5 $\leq$ FPR < 15 $\Rightarrow$ Intermediate, FPR $\geq$ 15 $\Rightarrow$ R5 [415]
- {10} : FPR < 10 $\Rightarrow$ X4-capable, FPR $\geq$ 10 $\Rightarrow$ R5 [414]
- {20}: FPR < 20 $\Rightarrow$ X4-capable, FPR $\leq$ 20 $\Rightarrow$ R5 [414]

According to Austrian-German treatment guidelines, MVC can be effective when a tropism prediction is labeled intermediate albeit with much less certainty than for R5 variants [415]. The probability the tropism predicted by geno2pheno$_{[coreceptor]}$ changed due to the introduced sequence alterations was calculated by sample counting

as $P(T_{A,C} \mid T_{U,C})$, with $T_{A,C}$ denoting the tropism of the altered sequences as determined with cutoff set C, and $T_{U,C}$ denoting the tropism of the unaltered sequences as determined with cutoff set C. The reference sequence used to number V3 nucleotide positions is consensus B (105 nucleotides), which is the reference used by geno2pheno[coreceptor].

All available ExposurePheno$_{\text{onlyIASPos}}$ and ExposurePheno$_{\text{full}}$ drug-exposure models (Section 3.3) were used for interpretation of sequences in SangerAlteration$_{\text{PPRT}}$, T$_{\text{PRRT}}$, SangerAlteration$_{\text{IN}}$, and T$_{\text{IN}}$. Subsequently, the obtained DES were discretized with DEMax cutoffs (Section 3.4). Note that DEMax cutoffs for ExposurePheno$_{\text{onlyIASPos}}$ were different than cutoffs for ExposurePheno$_{\text{full}}$ drug-exposure models. Lastly, changes in the discretized predictions of ExposurePheno$_{\text{onlyIASPos}}$ and ExposurePheno$_{\text{full}}$ models resulting from the introduction of sequence alterations were quantified.

5.3.2 Results

In the following, the results comparative interpretation of unaltered nucleotide sequences and nucleotide sequences with simulated Sanger variability are presented. First, the results for geno2pheno[coreceptor] are presented. Subsequently, I present the results for drug-exposure models. Since errors resulting from sequencing variability are not a matter of concern when interpreting *pol* nucleotide sequences, the results for drug-exposure models are much less detailed than the results for geno2pheno[coreceptor].

Robustness of geno2pheno[coreceptor] in the Presence of Sequencing Alterations

The electropherograms and nucleotide sequences in Electropherogram$_{\text{V3}}$ were used for estimating the EP$_{\text{V3}}$ alteration rates. These rates are shown in Table 5.2. As can be seen, linear regression was only invoked for the estimation of four types of alteration rates: the indel rate for the Env2 primer, and the definite-to-including-ambiguous rates for the primers Env2, Env7, and Env11. Nucleotide sequences in LA were interpreted with geno2pheno[coreceptor], resulting in an FPR per sequence. The resulting FPR distribution in LA is illustrated in Figure 5.1. A peak in the distribution can be seen for FPRs between zero and ten, while the rest of the FPR brackets present a more even distribution. For subtype determination, LA was interpreted with the COMET subtyping tool. The resulting numbers of strains by subtype are tabulated in Table 5.3. Subtype B accounted for most subtypes in LA (52%), followed by subtype C (16%). In LA, 0.24% of the bases are ambiguous, while 99.76% of the bases are definite.

Among the FPR shifts between the altered SangerAlteration$_{\text{V3}}$ sequences and their unaltered counterparts in LA, 79% are equal to zero, 11% are below zero and 10% are above zero (Figure 5.2). Figure 5.3 shows how the probability of obtaining FPRs lower than the singleton FPR increases with the number of amplifications. Finally, Figure 5.4 depicts the probabilities of change of predicted tropism for different cutoff sets (probabilities labeled S for singleton FPRs and T for triplicate FPRs). Figure 5.4 shows that geno2pheno[coreceptor] is more likely to reduce than to raise FPR when sequence alterations are present, slightly favoring a false prediction of R5 viruses as X4-capable over the reverse misprediction (see also Figure 5.2). Since the FPR shifts are both negative and positive, FPR determination in triplicate will always reduce FPRs, as the minimum FPR is selected. When determining FPR in triplicate, there is a 27% chance that the triplicate FPR will be lower than the singleton FPR (Figure 5.3).

The overall average FPR shift between the altered sequences in SE$_{\text{V3}}$ and their unaltered counterparts in LA is -2.22 ($\sigma = 13.87$); 23% of these shifts are zero, 36% are above zero and 40% are below zero (Figure 5.5).

Table 5.2: **Estimated** EP_{V3} **Sequencing Variability Rates.** Position-wise alterations were estimated with the Electropherogram$_{V3}$ dataset. Alteration rates were regressed linearly on the nucleotide position and a t-test was performed to assess whether the regression coefficient was significantly different from zero. p-values were corrected using the Benjamini-Hochberg method. If the coefficient had a significant corrected p-value at the 0.05 level, the linear function was used. Otherwise, the average alteration rate was used. Significant p-values are underlined. NA: not available.

Primer	Original Base	Error Type	Error Rate Position 1	Error Rate Position 105	Corrected p-value
		Excluding Ambiguous	0.0002	0.0002	0.9121
		Excluding Definite	0.0061	0.0061	0.4967
	Ambiguous	Including Ambiguous	0.0096	0.0096	0.9376
		Including Definite	0.3016	0.3016	0.9121
Env2		Indel	0.1835	0.1835	0.9951
		Excluding Ambiguous	0.0002	0.0002	0.2561
	Definite	Excluding Definite	0.0043	0.0043	0.0912
		Including Ambiguous	0.0056	0.0174	0.0005
		Indel	0.0145	0.0077	2×10^{-08}
		Excluding Ambiguous	0	0	NA
		Excluding Definite	0.0037	0.0037	0.8966
	Ambiguous	Including Ambiguous	0.0108	0.0108	0.3463
		Including Definite	0.3361	0.3361	0.9951
Env6		Indel	0.0492	0.0492	0.8326
		Excluding Ambiguous	0	0	NA
	Definite	Excluding Definite	0.0033	0.0033	0.2561
		Including Ambiguous	0.0143	0.0143	0.0778
		Indel	0.0066	0.0066	0.2561
		Excluding Ambiguous	0	0	NA
		Excluding Definite	0.0119	0.0119	0.9121
	Ambiguous	Including Ambiguous	0.039	0.039	0.9376
		Including Definite	0.2328	0.2328	0.9121
Env7		Indel	0.0413	0.0413	0.2561
		Excluding Ambiguous	0	0	NA
	Definite	Excluding Definite	0.0038	0.0038	0.4605
		Including Ambiguous	0.0472	0.0273	0.0313
		Indel	0.0018	0.0018	0.9121
		Excluding Ambiguous	0	0	NA
		Excluding Definite	0.0032	0.0032	0.9121
	Ambiguous	Including Ambiguous	0.0136	0.0136	0.8969
		Including Definite	0.2208	0.2208	0.8578
Env11		Indel	0.1442	0.1442	0.9038
		Excluding Ambiguous	0.0008	0.0008	0.8326
	Definite	Excluding Definite	0.0037	0.0037	0.9121
		Including Ambiguous	0.0509	0.0274	0.0171
		Indel	0.0064	0.0064	0.8326
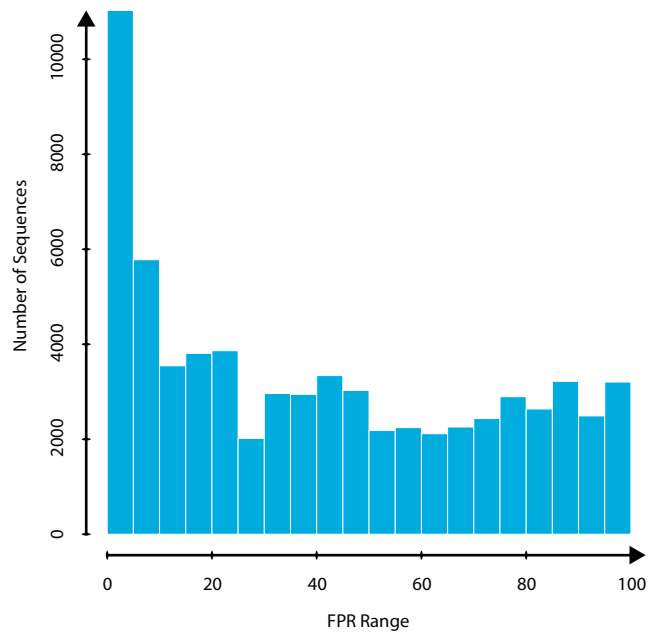
Figure 5.1: Histogram of FPRs in the LA Dataset. The histogram depicts the distribution of FPRs in the LA dataset. Each bar represents a range of FPRs, its height indicates the number of sequences falling within that FPR range.

Table 5.3: Number of Strains by Subtype, Circulating Recombinant Form or Group, LA Dataset. The numbers of strains of each subtype, circulating recombinant form (CRF) or group in the LA dataset are tabulated above.

Subtype, CRF or Group	Number of Strains
B	35,312
C	11,056
A1	10,187
01_AE	4,986
D	2,691
G	1,422
F1	1,308
F2	324
O	226
A2	207
H	121
CPZ	68
J	46
K	27
N	11
P	5
Total	67,997
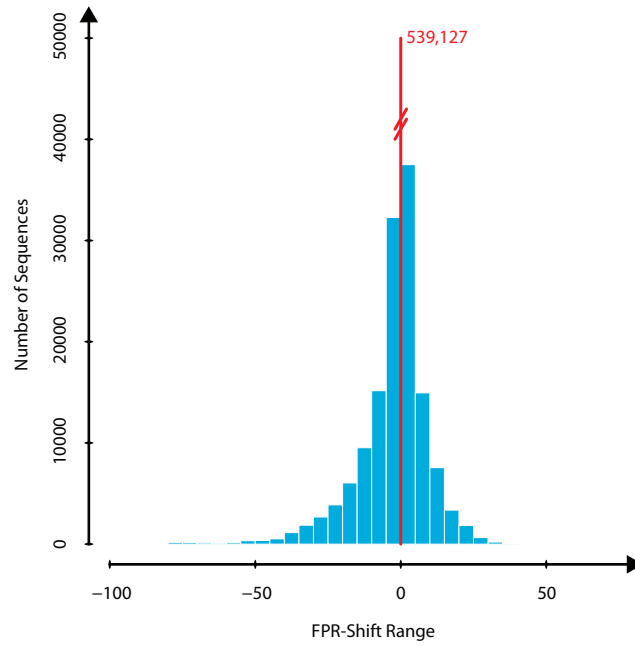
Figure 5.2: Histogram of the FPR Shifts between the LA and $\mathrm{SangerAlteration}_{V3}$ Datasets. The histogram displays the distribution of the shifts in FPR between sequences in the LA dataset and their altered counterparts in the $\mathrm{SangerAlteration}_{V3}$ dataset. Each bar represents a FPR-shift range, its height indicates the number of sequences falling within that FPR-shift range. Zero-valued shifts were excluded from the histogram and are represented by the red peak instead. 79% of the shifts are equal to zero (red peak), 11% are below zero (bars left of red peak) and 10% are above zero (bars right of red peak).
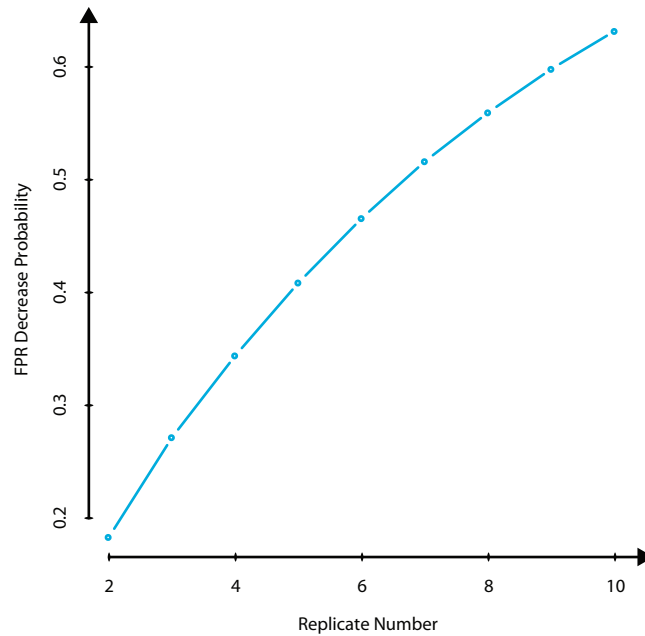
Figure 5.3: Probabilities of Decreasing FPR as More Replicates are Performed. FPRs in SA dataset were compared with their unaltered counterparts in the LA dataset. The probabilities of obtaining nth replicate FPRs lower than the singleton FPRs are plotted with increasing values for n (replicate number) in the chart above.
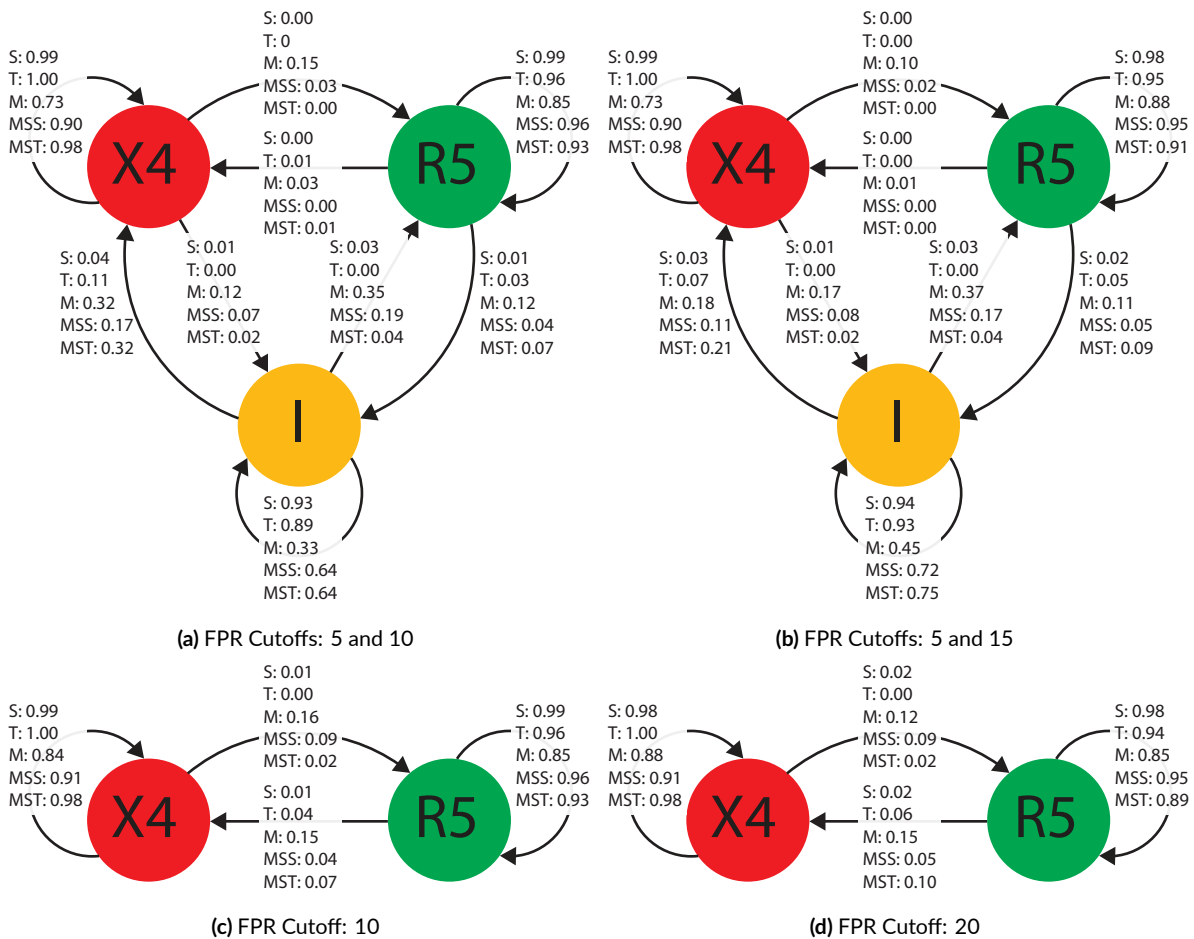
Figure 5.4: Conditional Probabilities of Change in Predicted Tropism. Cutoff sets {5,10} (a), {5,15} (b), {10} (c), and {20} (d) were applied to FPRs in the LA, SangerAlteration$_{V3}$, M$_{V3}$, and MS$_{V3}$ datasets to calculate the conditional probabilities of change in predicted tropism. Results are shown for SangerAlteration$_{V3}$ dataset singleton (S), SangerAlteration$_{V3}$ dataset triplicate (T), M$_{V3}$ dataset (M), MS$_{V3}$ dataset singleton (MSS), and MS$_{V3}$ dataset triplicate (MST) FPRs. There are no replicates on the M$_{V3}$ dataset since its sequences contain no ambiguities. The circles in the figure represent the predicted tropism. The arrows indicate a change in predicted tropism with the head of the arrow pointing towards the change. The arrow labels contain the probabilities for the respective changes in predicted tropism as calculated with the dataset mentioned above.
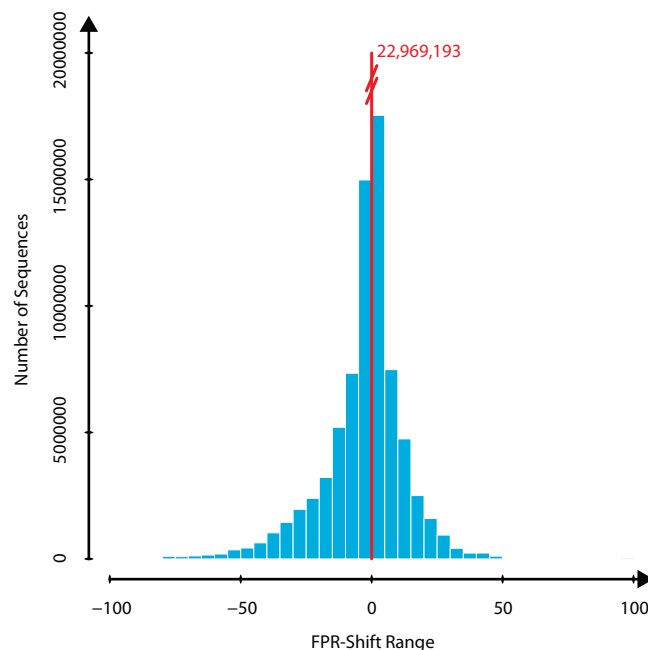
Figure 5.5: Histogram of the FPR Shifts between the LA and SE$_{V3}$ Datasets. The histogram displays the distribution of the shifts in FPR between sequences in the LA dataset and their altered counterparts in the SE$_{V3}$ dataset. Each bar represents a FPR-shift range, its height indicates the number of sequences falling within that FPR-shift range. Zero-valued shifts were excluded from the histogram and are represented by the red peak instead. 23of these shifts are zero (red peak), 40% are below zero (bars left of red peak) and 36% are above zero (bars right of red peak).

Here again we see the tendency of geno2pheno$_{[coreceptor]}$ to reduce FPR when sequence alterations are present. Figure 5.6 shows a plot of FPR shifts averaged by nucleotide position. Alterations in some parts of V3 have a higher propensity for changing coreceptor tropism than others, as expected. The magnitude and sign of the average shifts vary greatly with the nucleotide position. Shifts averaged by nucleotide or ambiguity code can be seen in Table 5.4. Among these average FPR shifts, the smallest equals -4.85, while the largest equals -1.20. Gaps induced an average FPR shift of 0.26.

In LA, 6,133 sequences contained ambiguous bases. The sequences in the M$_{V3}$ dataset were derived from 6,118 of these sequences by resolving their ambiguities. Fifteen sequences were excluded to avoid combinatorial explosion. Among the 6,133 original sequences, 41% resulted in two derived sequences while 59% resulted in more than two sequences (Figure 5.7). Figure 5.8 shows a plot of the average shift in MS$_{V3}$ against the sampling proportion, for triplicate and for singleton FPRs. The magnitude of the shift decreases as the proportion increases, and was zero for proportions of 85% of the sequences or more. The lowest proportion tested, 1%, yielded a mean singleton FPR shift of 0.61 ($\sigma = 10.91$), and a mean triplicate FPR shift of -4.5 ($\sigma = 9.35$). Since conditional change-of-predicted-tropism probabilities decrease as the sampling proportion increases, Figure 5.4 only displays those for sampling 1% of the variants in M$_{V3}$, labeled MSS for singleton FPRs and MST for triplicate FPRs. Change-of-predicted-tropism probabilities calculated with all the sequences in M$_{V3}$ are also displayed in Figure 5.4 (probabilities labeled M).

ROBUSTNESS OF DRUG-EXPOSURE MODELS IN THE PRESENCE OF SEQUENCING ALTERATIONS

The EP$_{PRRT}$ (Tables 5.5 and 5.6) and EP$_{IN}$ (Table 5.7) alteration rates were estimated using the nucleotide sequences and electropherograms in the Electropherogram$_{PRRT}$ and Electropherogram$_{IN}$ datasets, respec-
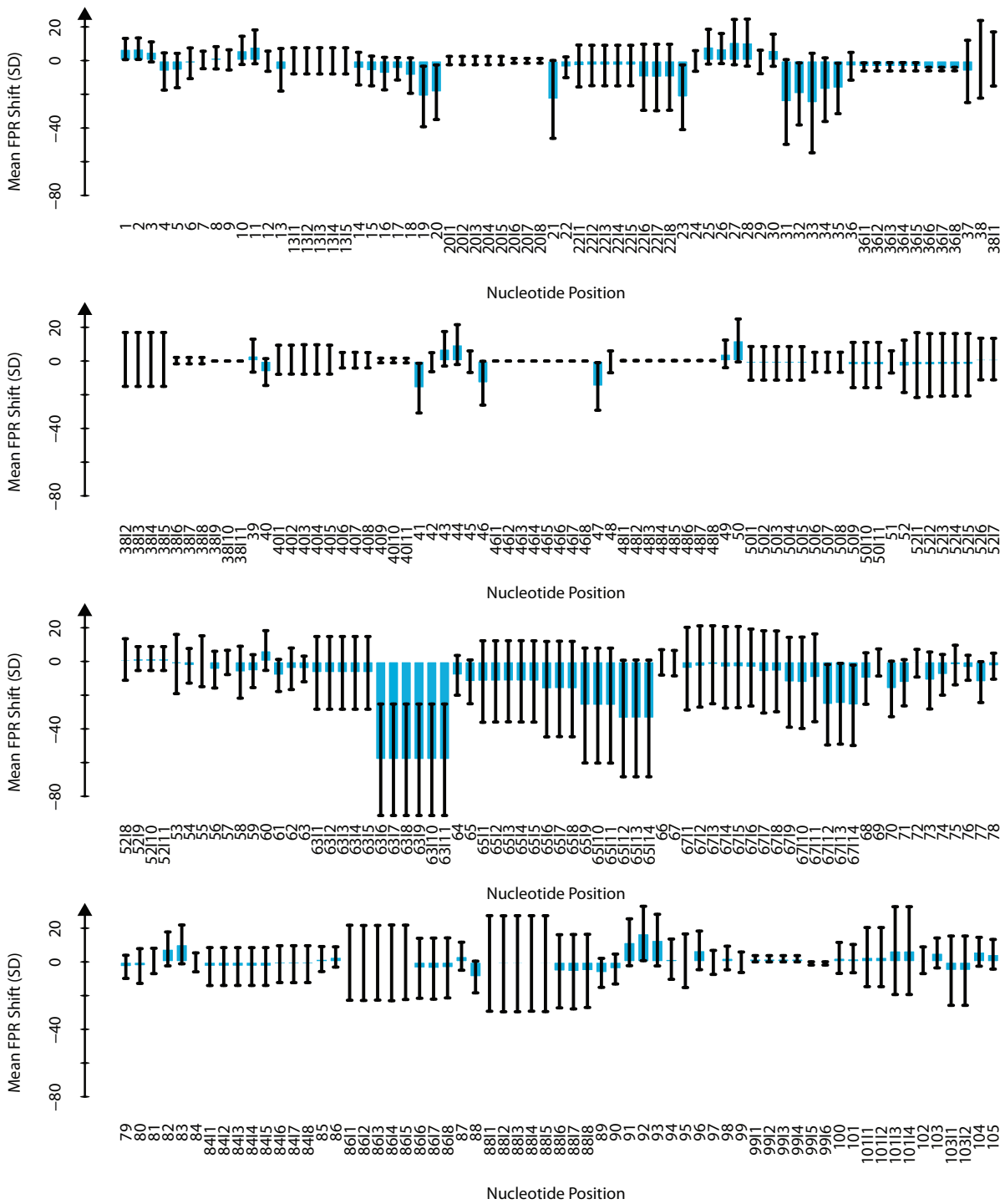
Figure 5.6: Mean FPR Shifts Averaged by Nucleotide Position, SE_{V3} Dataset. The graphic above shows the mean shift in the FPR averaged by nucleotide position. Shifts were calculated by comparing FPRs in the SE_{V3} dataset with their unaltered counterparts in the LA dataset. Each bar represents a nucleotide position, its height indicates the mean FPR shift resulting from an alteration at that position. Error bars show standard deviation. Insertions are labeled with the preceding nucleotide position, the character *I* and an index.

Table 5.4: Mean FPR Shift Averaged by Nucleotide Base, SE_{V3} Dataset. The table shows the mean FPR shift and standard deviation (SD), averaged by substituting nucleotide base after alignment. Shifts were calculated by comparing FPRs in the SE_{V3} dataset with their unaltered counterparts in the LA dataset. Although gaps were not considered when generating the SE_{V3} dataset, some nucleotide alterations were transformed to gaps by the alignment program of $geno2pheno_{[coreceptor]}$.

IUPAC Base	Mean	Standard Deviation
A	-3.33	20.01
C	-1.77	17.98
G	-4.85	17.49
T	-2.40	18.33
B	-2.57	12.98
D	-2.18	12.12
H	-1.77	13.26
K	-2.65	12.94
M	-1.88	13.35
N	-2.19	12.50
R	-2.86	12.72
S	-2.89	13.06
V	-2.42	12.26
W	-1.63	13.12
Y	-1.20	14.25
-	0.26	18.07
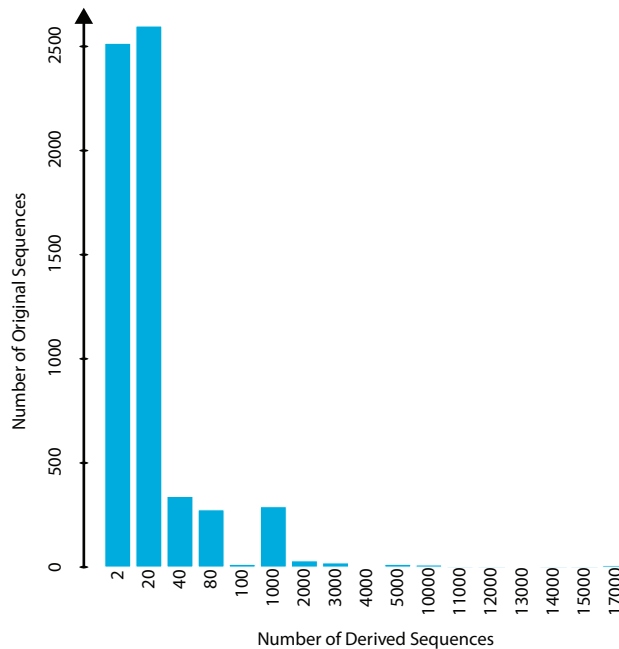
Figure 5.7: Histogram of the Number of Derived Sequences in the M_{V3} Dataset. The M_{V3} dataset was created from sequences in the LA dataset containing ambiguous bases and having an FPR below 20. The ambiguities in each of these sequences were resolved by deriving further sequences with all nucleotide combinations represented by the ambiguities. Each bar represents a range in the number of derived sequences, its height indicates the number of LA sequences that resulted in an amount of derived sequences falling within that range.
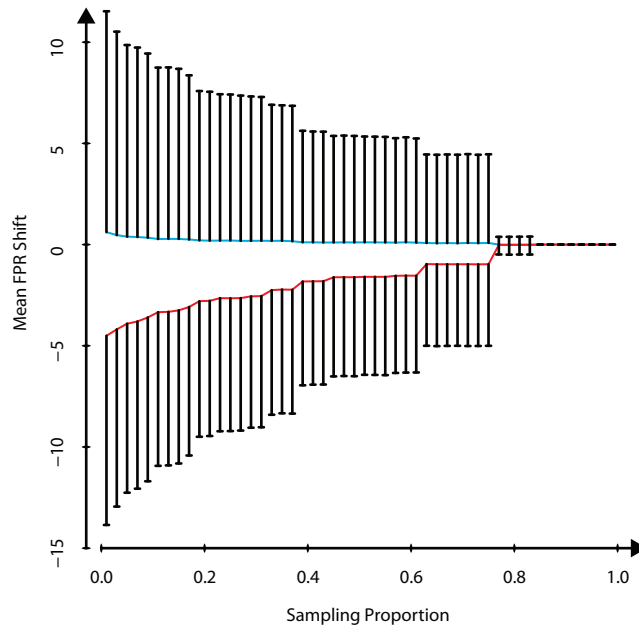
Figure 5.8: Mean Shift in FPR as the Sampling Proportion Increases, MS_{V3} Dataset. For each group of variants in the M_{V3} dataset, a certain proportion of the variants were sampled at random, and a sequence (with ambiguous bases) was constructed from these variants. The sampling was performed 100 x 3 times, i.e. 100 repetitions in triplicate. The figure depicts the mean shift in FPR between the sampled and the unaltered variants. Both singleton (blue line) and triplicate (red line) FPRs were considered. Error bars show standard deviation; they were only plotted in one direction in order to improve the readability of the plot.

— Singleton FPRs — Triplicate FPRs

tively. Using the EP_{PRRT} and EP_{IN} rates, ten nucleotide sequences with *in-silico* simulated Sanger-sequencing variability were generated from each sequence in the T_{PRRT} and T_{IN} datasets (these datasets are described in Section 3.3), respectively. The resulting sequences, called SangerAlteration$_{PRRT}$ and SangerAlteration$_{IN}$, respectively, were interpreted with ExposurePheno$_{onlyIASPos}$ and ExposurePheno$_{full}$ drug-exposure models (Section 3.3). Subsequently, continuous drug-exposure predictions were discretized with DEMax cutoffs. The fraction of predictions that changed as a result of simulated sequencing variability is shown in Table 5.8. ExposurePheno$_{onlyIASPos}$ models presented a higher fraction of unchanged predictions ($\mu = 0.96$; $\sigma = 0.01$) than ExposurePheno$_{full}$ models ($\mu = 0.93$; $\sigma = 0.02$)

5.3.3 Discussion

The analysis presented in Section 5.3 addresses two related questions: the robustness of geno2pheno$_{[coreceptor]}$ with respect to sequence alterations in terms of change of predicted tropism, and the influence of minority populations on the predictions of geno2pheno$_{[coreceptor]}$. I have subjected geno2pheno$_{[coreceptor]}$ to a challenge involving the systematic introduction of simulated sequencing variability into a large set of V3 nucleotide sequences in order to generate four datasets: SangerAlteration$_{V3}$ simulates alterations due to the Sanger sequencing technique. SE_{V3} explores the effects of systematically introduced single nucleotide exchanges. With M_{V3} and MS_{V3} I investigated the influence of simulated viral minorities undetected by bulk sequencing. Additionally, the SangerAlteration$_{PRRT}$ and SangerAlteration$_{IN}$ datasets were created for assessing the robustness of drug-exposure models.

The effect of sequencing X4-associated codons at amino-acid positions 11, 13, 24, 25, and 32 (nucleotide

205

Table 5.5: Estimated EP_{PRRT} **Sequencing Variability Rates (1).** Position-wise alterations were estimated with the Electropherogram$_{\text{PRRT}}$ dataset. Alteration rates were regressed linearly on the nucleotide position and a t-test was performed to assess whether the regression coefficient was significantly different from zero. p-values were corrected using the Benjamini-Hochberg method. If the coefficient had a significant corrected p-value at the 0.05 level, the linear function was used. Otherwise, the average alteration rate was used. Significant p-values are underlined. NA: not available.

Primer	Original Base	Error Type	First Pos.	Last Pos.	Error Rate First Pos.	Error Rate Last Pos.	Corrected p-value
		Excluding Ambiguous	1	436	0.0002	0.0002	1.0
		Excluding Definite	1	436	0.0007	0.0007	0.76
	Ambiguous	Including Ambiguous	1	436	0.0244	0.008	<u>0.03</u>
		Including Definite	1	436	0.144	0.144	0.72
PRRT-2F		Indel	1	436	0.005	0.005	0.07
		Excluding Ambiguous	1	436	0.0001	0.0001	0.69
		Excluding Definite	1	436	0.0012	0.0012	0.84
	Definite	Including Ambiguous	1	436	0.0671	0.0401	<u>$< 10^{-16}$</u>
		Indel	1	436	0	0.0006	<u>$< 10^{-10}$</u>
		Excluding Ambiguous	1	564	0.0001	0.0001	0.29
		Excluding Definite	1	564	0.0008	0.0008	0.69
	Ambiguous	Including Ambiguous	1	564	0.023	0.023	0.09
		Including Definite	1	564	0.2789	0.2789	0.10
A		Indel	1	564	0.0193	0.0193	0.43
		Excluding Ambiguous	1	564	0.0001	0.0001	0.33
		Excluding Definite	1	564	0.0012	0.0008	<u>0.02</u>
	Definite	Including Ambiguous	1	564	0.0599	0.0599	0.12
		Indel	1	564	0.0001	0.0005	<u>$< 10^{-4}$</u>
		Excluding Ambiguous	409	885	0.0001	0.0001	0.26
		Excluding Definite	409	885	0.0002	0.0002	0.84
	Ambiguous	Including Ambiguous	409	885	0.0029	0.0099	<u>0.04</u>
		Including Definite	409	885	0.2511	0.1476	<u>0.02</u>
B		Indel	409	885	0.0166	0.0166	0.17
		Excluding Ambiguous	409	885	0	0.0001	<u>$< 10^{-3}$</u>
		Excluding Definite	409	885	0.0008	0.0014	<u>0.02</u>
	Definite	Including Ambiguous	409	885	0.0196	0.0453	<u>$< 10^{-14}$</u>
		Indel	409	885	0.0008	0.0008	0.39
		Excluding Ambiguous	799	994	0.0002	0.0002	0.81
		Excluding Definite	799	994	0.0005	0.0005	0.29
	Ambiguous	Including Ambiguous	799	994	0.0043	0.0043	0.25
		Including Definite	799	994	0.162	0.162	0.37
C		Indel	799	994	0.0268	0.0268	0.38
		Excluding Ambiguous	799	994	0.0001	0.0001	0.19
		Excluding Definite	799	994	0.0013	0.0013	0.10
	Definite	Including Ambiguous	799	994	0.0195	0.0195	0.78
		Indel	799	947	0.0054	0	<u>0.03</u>

Table 5.6: Estimated EP_{PRRT} Sequencing Variability Rates (2). Position-wise alterations were estimated with the Electropherogram$_{PRRT}$ dataset. Alteration rates were regressed linearly on the nucleotide position and a t-test was performed to assess whether the regression coefficient was significantly different from zero. p-values were corrected using the Benjamini-Hochberg method. If the coefficient had a significant corrected p-value at the 0.05 level, the linear function was used. Otherwise, the average alteration rate was used. Significant p-values are underlined. NA: not available.

Primer	Original Base	Error Type	First Pos.	Last Pos.	Error Rate First Pos.	Error Rate Last Pos.	Corrected p-value
		Excluding Ambiguous	28	643	0	0	NA
		Excluding Definite	28	504	0.0014	0	0.04
	Ambiguous	Including Ambiguous	28	643	0.0086	0.0086	0.37
		Including Definite	28	643	0.1495	0.1495	0.84
D		Indel	28	643	0.0167	0.0167	0.21
		Excluding Ambiguous	28	643	0.0001	0.0001	0.43
		Excluding Definite	28	643	0.002	0.0004	$< 10^{-12}$
	Definite	Including Ambiguous	28	643	0.0294	0.061	$< 10^{-11}$
		Indel	28	643	0.0005	0.001	0.01
		Excluding Ambiguous	1	301	0.0001	0.0001	0.29
		Excluding Definite	82	301	0	0.0026	0.02
	Ambiguous	Including Ambiguous	1	301	0.0041	0.0041	0.42
		Including Definite	1	301	0.1963	0.316	0.03
F		Indel	1	301	0.0472	0.0472	0.36
		Excluding Ambiguous	1	301	0.0002	0	0.01
		Excluding Definite	1	301	0.0014	0.0014	0.21
	Definite	Including Ambiguous	1	301	0.0176	0.0176	0.07
		Indel	83	301	0	0.0175	$< 10^{-2}$
		Excluding Ambiguous	69	715	0.0001	0.0001	0.18
		Excluding Definite	69	715	0.0008	0.0008	0.84
	Ambiguous	Including Ambiguous	69	715	0.0159	0.0062	0.02
		Including Definite	69	715	0.3674	0.3674	0.07
G		Indel	69	682	0.0765	0	$< 10^{-12}$
		Excluding Ambiguous	69	715	0.0001	0	$< 10^{-3}$
		Excluding Definite	69	715	0.0013	0.0005	$< 10^{-12}$
	Definite	Including Ambiguous	69	715	0.0376	0.0224	$< 10^{-24}$
		Indel	69	715	0.0018	0.0018	0.81
	Ambiguous	Excluding Ambiguous	583	993	0	0	NA
		Excluding Definite	583	993	0.0002	0.0002	0.19
		Including Ambiguous	583	993	0.002	0.002	0.09
		Including Definite	583	993	0.0531	0.0531	0.21
H		Indel	583	993	0.0111	0.0111	0.31
		Excluding Ambiguous	583	993	0.0001	0.0001	0.29
		Excluding Definite	583	993	0.0035	0.0008	$< 10^{-10}$
	Definite	Including Ambiguous	583	993	0.0335	0.0225	$< 10^{-2}$
		Indel	583	965	0.0027	0	$< 10^{-21}$

Table 5.7: **Estimated** EP_{IN} **Sequencing Variability Rates**. Position-wise alterations were estimated with the $Electropherogram_{IN}$ dataset. Alteration rates were regressed linearly on the nucleotide position and a t-test was performed to assess whether the regression coefficient was significantly different from zero. p-values were corrected using the Benjamini-Hochberg method. If the coefficient had a significant corrected p-value at the 0.05 level, the linear function was used. Otherwise, the average alteration rate was used. Significant p-values are underlined. NA: not available.

Primer	Original Base	Error Type	First Pos.	Last Pos.	Error Rate First Pos.	Error Rate Last Pos.	Corrected p-value
5'-INT	Ambiguous	Excluding Ambiguous	1	649	0	0	NA
		Excluding Definite	1	649	0.0004	0.0004	0.95
		Including Ambiguous	1	649	0.0044	0.0044	0.33
		Including Definite	1	649	0.1162	0.1162	0.63
		Indel	1	649	0.0162	0.0162	0.71
	Definite	Excluding Ambiguous	1	507	0.0002	0	$\underline{< 10^{-6}}$
		Excluding Definite	1	649	0.0019	0.0008	$\underline{< 10^{-6}}$
		Including Ambiguous	1	649	0.0486	0.0164	$\underline{< 10^{-27}}$
		Indel	1	649	0.0012	0.0033	$\underline{< 10^{-2}}$
IN-F	Ambiguous	Excluding Ambiguous	1	643	0.0001	0.0001	0.46
		Excluding Definite	1	643	0.0001	0.0001	0.86
		Including Ambiguous	1	643	0.0038	0.0038	0.76
		Including Definite	1	643	0.0908	0.0908	0.29
		Indel	1	643	0.003	0.003	0.86
	Definite	Excluding Ambiguous	1	643	0.0002	0	$\underline{< 10^{-2}}$
		Excluding Definite	1	643	0.0014	0.0003	$\underline{< 10^{-5}}$
		Including Ambiguous	1	643	0.0285	0.0285	0.80
		Indel	1	643	0.001	0.001	0.36
3p31	Ambiguous	Excluding Ambiguous	151	806	0	0	NA
		Excluding Definite	151	806	0.0006	0.0006	0.65
		Including Ambiguous	151	806	0.0022	0.0022	0.83
		Including Definite	151	806	0.1844	0.0562	$\underline{< 10^{-4}}$
		Indel	151	721	0.0419	0	$\underline{< 10^{-4}}$
	Definite	Excluding Ambiguous	151	806	0.0001	0.0001	0.65
		Excluding Definite	151	806	0.0024	0.0004	$\underline{< 10^{-18}}$
		Including Ambiguous	151	806	0.0254	0.0392	$\underline{< 10^{-6}}$
		Indel	151	806	0.0031	0.0002	$\underline{< 10^{-16}}$
SEQ1	Ambiguous	Excluding Ambiguous	459	856	0.0002	0.0002	0.31
		Excluding Definite	459	856	0	0	NA
		Including Ambiguous	459	856	0.0005	0.0005	0.48
		Including Definite	459	856	0.0267	0.0267	0.29
		Indel	459	856	0.0006	0.0006	0.27
	Definite	Excluding Ambiguous	459	780	0.0001	0	$\underline{0.03}$
		Excluding Definite	459	823	0.0009	0	$\underline{< 10^{-4}}$
		Including Ambiguous	459	856	0.0069	0.0324	$\underline{< 10^{-4}}$
		Indel	459	763	0.0061	0	$\underline{< 10^{-5}}$
SEQ2	Ambiguous	Excluding Ambiguous	1	496	0	0	NA
		Excluding Definite	1	496	0	0	0.80
		Including Ambiguous	1	496	0.0043	0.0043	0.83
		Including Definite	1	496	0.0492	0.0492	0.48
		Indel	1	496	0.0008	0.0008	0.66
	Definite	Excluding Ambiguous	1	496	0.0001	0.0001	0.29
		Excluding Definite	1	496	0.0008	0.0008	0.29
		Including Ambiguous	1	496	0.0584	0.0584	0.76
		Indel	93	496	0	0.0027	$\underline{< 10^{-11}}$

Table 5.8: Fraction of Unchanged Drug-Exposure Predictions under Simulated Sequencing Variability. From each sequence in T_{PRRT} and T_{IN}, ten sequences with simulated sequencing variability were constructed. After evaluation with each ExposurePheno$_{onlyIASPos}$ and ExposurePheno$_{full}$ drug-exposure model, and discretization with DEMax cutoffs, the fraction of predictions that did not change as a result of the simulated sequence alterations was quantified. SD: Standard Deviation.

	ExposurePheno$_{onlyIASPos}$	ExposurePheno$_{full}$
3FTC	0.97	0.92
ABC	0.97	0.91
AZT	0.98	0.93
d4T	0.97	0.93
ddC	0.98	0.94
ddI	0.97	0.93
TDF	0.96	0.89
DLV	0.97	0.92
EFV	0.95	0.87
ETR	0.95	0.92
NVP	0.97	0.92
RPV	0.98	0.96
AFPV	0.94	0.94
ATV	0.95	0.92
DRV	0.93	0.94
IDV	0.96	0.94
LPV	0.94	0.93
NFV	0.95	0.93
SQV	0.96	0.94
TPV	0.97	0.94
EVG	0.99	0.97
RAL	0.98	0.93
Naïve PRRT	0.95	0.92
Mean AM (SD)	0.96 (0.01)	0.93 (0.02)

position numbers 30-33, 37-39, 69-72, 73-75, and 94-96) more accurately by using specific primers has been studied [422]. The primer sets were specifically designed for subtype B. The study I present extends to other subtypes, as about half of the 67,997 sequences were non-Bs (Table 5.3).

The FPR shifts between the original and the *in-silico* mutated sequences in SE_{V3} have been analyzed (Figure 5.5 and Table 5.4). 9.1% are less than -20 and 11% are greater than 10. These two shift values were chosen to be multiples of 10, close to the 10th and 90th percentile of the shift distribution. The shift cutoff -20 identifies amino-acid positions 7, 8, 11 (nucleotide position numbers 19-21, 22-24, and 31-33), and insertions after amino-acid positions 21, 22 and 23 (nucleotide position numbers 61-63, 64-66, and 67-69) as highly relevant for detection of CXCR4-capable viruses with geno2pheno$_{[coreceptor]}$. Using a shift cutoff of 10, positions 9, 10, 17, 28 and 31 (nucleotide positions 25-27, 28-30, 48-51, 82-84, 91-93) are highly relevant for the detection of R5 viruses with geno2pheno$_{[coreceptor]}$. Mean shifts averaged by nucleotide base are negative, unless the sequence alteration is replaced with a gap by the alignment program. The smallest mean shift is -4.85. These results indicate that the position at which the alteration occurs is far more important than the nucleotide base the alteration consists of, as expected.

Triplicate sequencing is performed with the intention of improving the detection of X4-capable minority variants. In triplicate testing, the lowest FPR is considered the correct one, without comparing the obtained nucleotide sequences. Thus, if the lowest FPR is caused by a nucleotide sequencing error, triplicate testing will bias the prediction towards favoring X4-capable. If the obtained sequences are not manually inspected, errors in the individual sequences remain undetected. Thus, the more sequencing replicates are performed, the higher the chance you get an FPR lower than the first one, even if X4-capable minority variants are absent. In this case, triplicate testing can result in an exclusion of MVC-eligible patients, without increasing the safety of the prediction.

Figure 5.4 shows the influence of introduced variability for change of predicted tropism depending on the FPR cutoffs, for both single and triplicate testing. geno2pheno$_{[coreceptor]}$ is robust when introduced variability is present: for all cutoff sets, the probability of no change of predicted tropism is above 98% (or above 93% if the intermediate tropism prediction is considered) for singleton FPRs. In this scenario, triplicate FPRs raise the probability that predicted X4 capability will not change by up to 2%, but reduce the probability that predicted R5 tropism will not change by up to 4%.

In order to address the study of minorities, I created the M_{V3} and MS_{V3} datasets. M_{V3} contains the variants from which the sequence with ambiguities could have arisen, but might also contain variants absent in the sample. In contrast, MS_{V3} presents a more realistic picture, since a sequence that may contain ambiguities is constructed from a sample of the sequences in M_{V3} representing the limited sequencing depth of an experiment. For the smallest fraction tested, 1%, the probability predicted X4 tropism does not change is 90% with singleton FPRs, and 98% with triplicate FPRs. Thus, triplicate FPRs raise the probability predicted X4 tropism does not change by 9%, at most, for the lowest tested proportion (Figure 5.4).

The lowest value obtained for the probability that there is no change of predicted tropism is 90% with geno2pheno$_{[coreceptor]}$. In order to put this number into context, we mention Trofile's reported sensitivity to detect MVC responders, 92% [412]. geno2pheno$_{[coreceptor]}$ proved to be robust in the presence of sequence alterations and when detectable minorities are missed by bulk sequencing. False R5 predictions were either rare or absent in our analysis, depending on the selected cutoffs. If a change of predicted tropism occurs, it is much more likely that sequence alterations result in false X4-capable predictions than false R5 predictions. This speaks for the safety of MVC prescription to patients with predicted R5 viruses.

Simulated Sanger-sequencing variability was introduced into the test sets of models for predicting drug exposure (Section 3.3). The consequences of this variability were assessed by quantifying the number of discrete drug-exposure predictions that changed as a result of the introduced variability. Both $ExposurePheno_{onlyIASPos}$ and $ExposurePheno_{full}$ drug-exposure models proved to be robust when simulated sequencing variability is present in the nucleotide sequences they interpret. Among all drug-exposure models, the $ExposurePheno_{full}$ for TDF showed the highest fraction of predictions that changed as a result of the simulated sequencing variability (11%; Table 5.8). For predicting drug exposure, $ExposurePheno_{full}$ models consider more residues than $ExposurePheno_{onlyIASPos}$ models. This results in a higher performance of $ExposurePheno_{full}$ models in predicting drug exposure when compared to $ExposurePheno_{onlyIASPos}$ models (Tables 3.9 and 3.10). However, the fact that $ExposurePheno_{full}$ models consider more alignment positions than $ExposurePheno_{onlyIASPos}$ models makes them more susceptible to variability in the input sequences (Table 5.8).

6

Conclusion and Outlook

THIS FINAL, CONCLUDING CHAPTER OF this work is structured as follows. First, I provide a summary of the problem in question and of the state of the art. Next, I argue why the methods described in this work represent an advancement of the state of the art (Section 6.1). Last, I propose research topics related to this work that could further advance the state of the art (Section 6.2).

HIV-1 is a pathogen that causes a deadly infection. Left untreated, HIV-1 infection almost inevitably leads to the death of the human host, who will develop AIDS in the advanced stages of the disease and who will probably succumb to an opportunistic infection (Section 1.2.2). HIV-1 emerged through zoonosis in West-Central Africa between 1853 and the early 1900s. After its emergence, the pathogen, carried by its human hosts, traveled along the Sangha and Congo rivers until it reached the city of Kinshasa. The conditions in this city allowed for the spread of the disease to a large number of persons. Probably, the establishment of the HIV-1 epidemic in Kinshasa was only possible due to iantrogenesis. From Kinshasa, HIV-1 infection spread to the rest of the world, evolving into a pandemic (Section 1.3).

In 1983, HIV-1 was identified as the etiologic agent of AIDS. Over the course of the decades following 1983, outstanding progress was made towards the treatment and the prevention of the disease. Regarding the treatment of the disease, the most important scientific accomplishments are related to the development of antiretroviral drugs and the acquisition of knowledge on how to effectively use antiretroviral drugs for treating HIV-1 infection (Section 1.5.1). cART, the current standard of care in HIV-1 infection, involves the simultaneous intake of several antiretroviral drugs. cART presents the best chances of therapeutic success if the drug combinations used for treatment are tailored to each individual patient. Effective choice of the drug compounds of cART must consider the potential of HIV-1 to develop drug resistance, as well as several patient-specific factors. Molecular diagnostics are essential for the accurate selection of the drug compounds in cART, as well as for monitoring the effectiveness of therapy (Section 1.5.4).

Drug resistance of HIV-1 will preclude therapeutic success if drug compounds are used against which HIV-1 is resistant (Section 1.5.3). For this reason, determination of HIV-1 drug resistance prior to treatment initiation is decisive for therapeutic success. Determination of HIV-1 drug resistance is performed with aid of molecular diagnostics. There are two approaches with which HIV-1 drug resistance can be determined. One approach

involves the *in-vitro* measurement of the replicative capacity of HIV-1 in the presence of different concentrations of an antiretroviral compound. Comparison of the replicative capacity of an HIV-1 strain isolated from a clinical blood sample to that of a reference HIV-1 strain allows for the determination of the FC (Section 3.1). Interpretation of the FC values for the drugs considered for therapy is necessary. Specifically, a decision regarding the prospects that the drug combination has for success must be made on the basis of individual, *in-vitro* resistance measurements for each drug. This decision must take into account the following issues. First, the extent of drug resistance indicated by a certain FC value is drug-specific, i.e. the FC value for a drug may indicate susceptibility of HIV-1 to the drug, while the same value for the FC of another drug may indicate HIV-1 to this other drug. Second, there are differences in the mechanics of the replication of HIV-1 *in-vivo*, as compared to *in-vitro*. Third, although FC values are determined for each drug individually, compounds in cART act in concert (Section 3.5). The second approach for determination of drug resistance involves sequencing the viral genes that are target of antiretroviral drugs. The extent of viral resistance to available antiretroviral drugs can be determined through interpretation of the resulting amino-acid sequences of the viral target proteins. This requires prior knowledge of the association of the mutations of HIV-1 to drug resistance (Section 3.2).

Genotypic drug-resistance interpretation systems facilitate the interpretation of an HIV-1 genotype with respect to drug resistance and with respect to the prospects that a certain drug combination has for therapeutic success. Two generations of these systems exist. The key feature of the first generation of genotypic drug-resistance interpretation systems is that they rate each individual antiretroviral drug with respect to the resistance that an HIV-1 variant may present and do not do so for combinations of drugs. The treating clinician must therefore interpret the ratings for each individual drug in order to decide whether a certain drug combination has good prospects for success. First-generation genotypic drug-resistance interpretation systems can be classified into two types, according to the basis on which they rate drug resistance. The first type, *rules-based*, comprises interpretation systems that operate on set of rules crafted by human experts in order to estimate drug resistance. The second type, *data-driven*, comprises interpretation systems that use a statistical model that was trained on a dataset in order predict drug resistance (Section 3.2). Second-generation genotypic drug-resistance interpretation systems offer a more advanced interpretation in that they rate drug *combinations* with respect to their prospects for therapeutic success. For this reason, their use in the selection of drug combinations for cART is less dependent on interpretation by human experts. I think that *genotypic therapy-success interpretation system* is a more accurate name for this second generation of systems. While several data-driven therapy-success interpretation systems exist, rules-based interpretation systems are not known to me (Section 4.2).

First-generation genotypic drug-resistance interpretation systems are an integral part of the standard of care for treating HIV-1 infection. However, genotypic therapy-success interpretation systems have not yet reached broad acceptance. Genotypic therapy-success interpretation systems described prior to the publication of this work are confronted with a number of obstacles that have prevented them from reaching the bedside. First, for apparently no good reason, their predictions contradict the results of clinical studies on the effectiveness of antiretroviral therapy, as well as the intuition of treating clinicians. Second, the systems do not provide the users with an insight on the basis with on drug combinations are rated. Third, treating clinicians may doubt whether the predictions of a system are robust to the variability that is inherent to nucleotide sequencing. Fourth, certification of the systems by an authority is challenging due to their data-driven nature. Specifically, procedures for certifying the systems are not yet in place. In the same way, it is not clear how retraining of a system, possibly involving the expansion of the capabilities of the system (e.g. providing predictions for previously disregarded drug compounds), should affect a previously granted certification for the system (Section 4.2).

6.1 Conclusion

In the following, I recapitulate the main findings of this work and highlight, how these findings have advanced the state of the art.

In Section 3.3, I present a statistical method that is trained on both GPPs and genotype-therapy-history pairs (GTHPs). The method produces scores that I call DES. DES are correlated with the drug exposure and drug resistance of HIV-1, while simultaneously being predictive for the therapeutic success of cART. In the following, I refer to figures describing the best-performing DES-model variants. The performance of DES in predicting drug exposure was tested on the T_{PRRT} and on the HIVdbExposure datasets. Compared to geno2pheno[resistance] ($\mu = 0.71$ and $\mu = 0.74$, respectively), the average performance of DES on these two datasets is superior ($\mu = 0.78$; $p < 10^{-4}$) or comparable ($\mu = 0.76$; $p = 0.46$), respectively. The average correlation of the method with FC values from GPPs measured with the PhenoSense® assay is 0.51, which is sufficient for accurate classification of HIV-1 genotypes as susceptible or resistant (Section 3.4.2). The performance of DES in predicting therapeutic success was tested on the $T_{EuResistTE}$ and on the HIVdbTCE datasets. Compared to geno2pheno[resistance] (AUC = 0.68 and AUC = 0.64, respectively), the performance of DES was higher (AUC = 0.71) or comparable (AUC = 0.64), respectively. These higher performances represent an improvement of the state of the art.

An important advantage of rules-based interpretation systems is that the rules on which they base their interpretations consider different types of information. Specifically, these rules are crafted under consideration of three sources of information that are associated to HIV-1 genotypes: *in-vitro* measurements of drug resistance, drug use history, and the therapeutic success of drug combinations (e.g. http://hivdb.stanford.edu). The information contained in each of these three sources presents a high pairwise similarity. Nonetheless, there are specific reasons why the exploitation of all three sources of information leads to a better performance in predicting therapeutic success (Sections 3.5, 3.2, 4.1, and 4.2.3). geno2pheno[resistance] is trained on GPPs and therefore only exploits one source of information. DES models are trained on both GPPs and GTHPs. DES feature two advances of the state of the art, which originate from their incorporation of both GPPs and GTHPs. First, the additional information provided to the training sets by GTHPs results in a higher performance. The higher performance of DES when compared to geno2pheno[resistance] (Section 3.3.2) can be attributed to the additional source of information that they exploit. Additionally, the models allow for the exploitation of the information contained in the EuResist database to an extent that had not been accomplished before. This manifests itself in the predictive performance of DES as well as in the *in-vivo* drug fingerprints that were extracted from the models. Furthermore, the *in-vivo* drug fingerprints allow for interpretation of the predictions of the method. Second, DES models are easier to keep up to date than geno2pheno[resistance], since GTHPs are easier to obtain than GPPs and since DES models allow for automatic retraining without the need for expert intervention. While GPPs are currently not widely used in routine clinical practice, molecular diagnostics involving the sequencing of HIV-1 are part of the standard of care. Therefore, GTHPs are constantly being generated.

In Section 4.2, I present a system for predicting the success of cART. This system is trained on TEs, which describe instances of the course of cART when specific factors are given at baseline (i.e. a specific viral genotype and a specific drug combination). For training the system, I used DES as input features and the NAS, a measure for cumulative, long-term therapeutic success, as the prediction target. Furthermore, the system was trained with an SVM for the regression of right-censored data. These characteristics allow the therapy-success prediction system to advance the state of the art in several ways, which I detail in the following. The use of DES

as input features is advantageous. Specifically, the system makes better use of the genotype by using DES, since DES are correlated with both therapeutic history and drug resistance. It is known that information on therapeutic history and drug resistance at baseline is very useful in predicting the outcome of cART (Sections 4.2 and 1.5.4). Thus, the use of DES as input features allows for the incorporation of more information into the predictions of therapeutic success, than that which is contained in the training set of the system. Furthermore, the use of DES in conjunction with TEs allows for the simultaneous use of the three sources of information mentioned above, which eliminates the main competitive advantage of rules-based interpretation systems. The therapy-success prediction system uses the NAS as a prediction target, which has a number of advantages. First, the NAS allows for evaluation of the success of cART based on the course of the entire therapy, and not only based on a snapshot or a number of snapshots. Second, the measure implicitly considers the side-effect profile, the potency, and adherence-fostering characteristics of drug combinations. Third, the measure accommodates transient increases of the VL, rating therapies with these transient increases lower, but without labeling them as failures altogether. Last, the NAS can accommodate differential VL monitoring intervals, as long as the VL is measured at least once every six months, allowing for a maximum of ten percent of the total number of therapy semesters without VL measurements. The models output a predicted version of the NAS called the pNAS. The pNAS has a comparatively high performance in predicting therapeutic success. Specifically, averages of the performances across multiple datasets and multiple test prediction targets indicate that pNAS could attain a higher average performance (0.71) than a GSS obtained with the HIVdb rule set (0.61; Section 4.2.2). pNAS performs comparatively better when used for predicting the NAS and also when used for predicting therapy success at a certain number of weeks after treatment initiation. The analysis of the ranking of drug combinations shows that pNAS are, at least *grosso modo*, in line with the results of clinical studies. More importantly, pNAS are in line with the intuition of treating clinicians. Specifically, the list of drug combinations that was ranked includes currently used drug compounds and drug compounds that are not used any more due to their comparatively low efficacy and unfavorable side-effect profile. pNAS never placed disused drug combinations into the top rank. At the same time pNAS correctly ranked NNRTI-containing drug combinations at the top for patients without drug resistance mutations at baseline. In contrast, when drug-resistance mutations were present at baseline, pNAS correctly ranked a PI-containing drug combination at the top. pNAS can be used for composing a therapy incrementally in an interactive fashion (Section 4.2.3). This can be especially useful in patients with limited drug options. Both DES, which are used for calculation of pNAS, and pNAS themselves are interpretable (Section 4.2.3). This is important, since it can both foster confidence in the interpretation systems and provide an explanation for (apparently) implausible predictions. Last but not least, both DES and pNAS are generated by completely data-driven models that are trained on data from routine clinical practice. The fact that the models are data-driven reduces or even eliminates the risk of expert bias, as compared to rules-based interpretation systems. The training set of the model, which mainly consists of data from routine clinical practice, gives them a comparatively strong anchor to reality, when compared to models that are solely mainly trained on data from *in-vitro* experiments or data from clinical studies.

In Sections 3.4 and 3.5, I present several methods for the translation of therapeutically relevant quantities into clinically meaningful categories (*cutoff determination*). These methods are important since they allow for the interpretation and use of molecular diagnostics by treating clinicians in order to improve antiretroviral treatment. The cutoffs generated with the methods presented in Section 3.4 allow for interpretation of DES with respect to drug exposure, drug resistance, and the prospects a therapy has for success. Furthermore, with aid of the cutoffs, discrimination between drug exposure, drug resistance, and the prospects for therapy success is

made possible. The method presented in Section 3.5 tackles the long-standing problem of cutoff determination for (predicted) phenotypic resistance measurements. This method allows for FC-cutoff determination using data from routine clinical practice without the need for expert intervention. The main advances that are provided by the method, when compared to the state of the art, are two fold. First, the method does not require the *explicit selection of cutoffs for the determination of the cutoffs*. Specifically, previously described methods [322] translate the FC values for each individual drug (which are not comparable between drugs) into a quantity that is comparable between drugs. Cutoff determination for FC values is then achieved by applying (expert-selected) cutoffs to this quantity. The method presented in Section 3.5 selects cutoffs for FC values based on the form of a sigmoid function, such that it does not need cutoffs for selecting the cutoffs. Furthermore, the method produces cutoffs that translate FC values into categories that are in line with the clinical definition of the categories *susceptible*, *intermediate*, and *resistant*. Second, the method can produce intuitive plots that greatly enhance the method's interpretability.

In Section 5.3, I subject two data-driven genotypic interpretation systems, geno2pheno$_{[coreceptor]}$ and DES, to an *in-silico* challenge in order to test their robustness with respect to variability of the input sequence. The most important finding for geno2pheno$_{[coreceptor]}$ concerns the suggested practice of performing genotypic determination of tropism in triplicate and subsequently only considering the prediction with the lowest FPR while discarding the other two predictions. I show that this practice does little for avoiding false R5 predictions, while substantially increasing false X4 predictions. Furthermore, my analysis showed that both geno2pheno$_{[coreceptor]}$ and DES are robust to the variability that is inherent to Sanger sequencing. This addresses the valid concern for obtaining proof of the robustness of systems that are used in a healthcare context.

6.2 Outlook

In the following, I delineate some ideas that could be used in order to improve the accuracy and utility of methods for the prediction of the success of cART.

An important, pending improvement of genotypic drug-resistance interpretation systems concerns their adaptation such that they can interpret nucleotide sequences generated with massively parallel sequencing methods (Section 5.1). When compared to Sanger sequencing, massively parallel sequencing methods present two main advantages. First, they require significantly less time and monetary resources per sequenced sample. (Depending on the sequencing platform, this might only be true if the number of samples available for simultaneous analysis is sufficient for operating the sequencing device at full capacity.) Second, they can resolve minority HIV-1 variants accounting for as little as 0.1% of the population in the patient sample. The are two main challenges that need to be solved in order to adapt genotypic drug-resistance interpretation systems such that they can analyze data produced with massively parallel sequencing methods. First and foremost, it is unclear to which extent minority HIV-1 variants that cannot be detected with Sanger sequencing are clinically relevant [423, 424]. Second, since massively parallel sequencing methods can produce up to 2 giga base pairs of data in a single run (Section 5.1), effective strategies and methods must be implemented in order to compress the data prior to transmission via the internet.

Software used for interpretation of measurements performed in the context of molecular diagnostics can benefit from third-party certification in many ways. Certification of the software results in substantially increased user confidence. This, in turn, has the potential to largely increase the popularity and impact of the software. Furthermore, certification can reduce the risk of releasing versions of the software that contain programming

errors. In the worst of cases, these errors could result in false predictions. However, procedures for certifying statistical models that are trained on datasets are yet to be established. A great advantage of the novel interpretation models described in this work lies in the fact that they can be easily updated, as new data become available. However, this advantage could stand in the way of certification. The following questions are still open. Does retraining of a model invalidate its certification, without regard to the extent with which the training set was modified? If a model needs to be re-certified due to re-training, does the re-certification procedure need to be as extensive as the initial certification procedure? (One has to bear in mind that certification procedures can be very expensive.)

There are a number of settings in antiviral therapy for which currently available genotypic interpretation systems, including the ones presented in this work, were not designed. In the following, I mention some of them. (1) During the course of cART, older drug compounds are often replaced by newer drug compounds if the treating clinician fears that prolonged therapy with the older drug compounds will cause long-term toxicities (Section 1.5.1). The switch from older drug compounds to newer drug compounds is often performed while the VL is suppressed below the level of detection. Undetectability of the VL entails that it will not be possible to obtain *conventional* HIV-1 genotypes from blood, i.e. nucleotide sequences of circulating HIV-1 particles. Thus, in this setting, prediction of therapeutic success must be based on historical genotypes and therapy history. Additionally, the proviral DNA (which is integrated in the host's genome) can be isolated from PBMCs and sequenced. However, one must consider that the predictive value of nucleotide sequences of proviral DNA from PBMCs might be limited, since it may not be representative of resistant viral variants archived in other compartments of the body. (2) With the intent of reducing the risk of long-term toxicity, NRTI-sparing therapy strategies are currently being tested (Section 1.5.5). One therapeutic strategy comprises the initiation of cART with a popular drug combination that includes NRTIs (induction phase), with subsequent interruption of NRTI intake after suppression of the VL below the limit of detection (maintenance phase). In this therapeutic setting, prediction of therapeutic success at baseline must consider both the induction phase and the maintenance phase. The following questions could probably be answered with a therapy-success prediction system that has been tailored for this therapeutic setting. Which are the most-promising drug compounds for the induction phase? Which drug compounds should be used during the maintenance phase? Should drug-resistance criteria be more stringent with respect to the drug compounds used during the maintenance phase? (3) *Hepatitis B virus* and *Hepatitis C virus* are transmitted via the same routes as HIV-1. For this reason, there are a number of patients with HIV-1-hepatitis co-infection. Treatment of HIV-1-hepatitis co-infected patients is especially challenging due to the immunosuppressive effects of HIV-1 infection and due to possible drug-drug interactions resulting simultaneous treatment of HIV-1 infection and hepatitis. A particularity regarding the treatment of HIV-1-hepatitis-B co-infection lies in the fact that some NRTIs used for treatment of HIV-1 infection are also active against *Hepatitis B virus* [425, 426]. Thus, treatment can be optimized to target both viruses. Alike HIV-1 infection, Hepatitis C is treated with a combination of drugs [427]. Thus, successful treatment of Hepatitis C is also dependent upon the optimal selection of a drug combination. (4) No therapy-success interpretation system for therapy against *human immunodeficiency virus type 2* is known to me.

Currently available therapy-success prediction systems mainly consider factors relating to HIV-1 and the drug compounds used in therapy. However, these systems could benefit from consideration of patient-specific factors other than the viral variant harbored by the patient. Patient-specific factors that alter the effect of drug compounds can be grouped into two categories: pharmacogenetic and pharmacoecologic factors (reviewed in [428]). Pharmacogenetic factors comprise variants in the genes of the patients which result in a differential

effect of the drug compounds. These include:

- The human-leukocyte-antigen (HLA) type
- Gene variants that influence the absorption, distribution, metabolization, and excretion (ADME) of drugs
- Gene variants that influence the behavior molecules in the body which transport drugs.

Pharmacoecologic factors influence the effect of drug compounds and are related to the health and behavior of the patient. Examples for pharmacoecologic factors include:

- The lifestyle of the patient
- The adherence of the patient to the medication
- Interactions between coadministered drugs
- Comorbidities that are not related to HIV-1
- Reduced organ function
- Pregnancy

At current, antiretroviral therapy is allegedly the best example for personalized medicine. Interpretation models used in cART have contributed immensely to the personalization of antiretroviral therapy by providing accurate interpretation of the effect of viral genes in cART. I expect that the consideration of further, patient-specific factors will greatly boost the success of antiretroviral therapy. Beyond antiretroviral therapy, the investigation of non-viral patient-specific factors for the personalization of treatment harbor a great potential for the improvement of human health.

A
List of Publications

FIRST-AUTHOR PUBLICATIONS

Pironti, A., Pfeifer, N., Jensen, B., Knops, E., Oette, M., Walter, H., Kaiser, R., Lengauer, T. Prediction of the time an antiretroviral therapy will remain effective — A tool for treatment-naïve and treatment-experienced patients. (In preparation).

Pironti, A., Pfeifer, N., Jensen, B., Kaiser, R., Lengauer, T. Using drug exposure for predicting drug resistance — A data-driven genotypic interpretation tool. (In preparation.)

Pironti, A., Walter, H., Pfeifer, N., Knops, E., Lübke, N., Büch, J., Kaiser, R, Lengauer, T. Determination of phenotypic resistance cutoffs from routine clinical data. *Journal of Acquired Immunodeficiency Syndromes*. (Submitted.)

Pironti, A., Sierra, S., Kaiser, R., Lengauer, T., Pfeifer, N. Effects of sequence alterations on results from genotypic tropism testing. *Journal of Clinical Virology* 2015, 65, pp. 68-73.

CONTRIBUTING-AUTHOR PUBLICATIONS

Heger, E., Theis, A., Remmel, K., Walter, H., Pironti, A., Di Cristanziano, V., Jensen, B., Esser, S., Kaiser, R., Lübke, N. Generation of training data for geno2pheno[integrase] by a modified phenotypic HIV susceptibility assay. *Journal of Virological Methods*. (Submitted.)

Sierra, S., Dybowski, J.N., Pironti, A., Heider, D., Güney, L., Thielen, A., Reuter, S., Esser, S., Fätkenheuer, G., Lengauer, T., Hoffmann, D., Pfister, H., Jensen, B., Kaiser, R. Parameters influencing baseline HIV-1 genotypic tropism testing related to clinical outcome in patients on maraviroc. *PLoS ONE* 2015, 10 (5), art. no. e0125502.

Charpentier, C., Camacho, R., Ruelle, J., Eberle, J., Gürtler, L., Pironti, A., Stürmer, M., Brun-Vézinet, F., Kaiser, R., Descamps, D., Obermeier, M. HIV-2EU - supporting standardized HIV-2 drug-resistance interpretation in Europe: an update. *Clinical Infectious Diseases* 2015, 61 (8), pp. 1346-1347.

Sangeda, R.Z., Theys, K., Beheydt, G., Rhee, S.-Y., Deforche, K., Vercauteren, J., Libin, P., Imbrechts, S., Grossman, Z., Camacho, R.J., Van Laethem, K., Pironti, A., Zazzi, M., Sönnerborg, A., Incardona, F., De Luca, A., Torti, C., Ruiz, L., Van de Vijver, D.A.M.C., Shafer, R.W., Bruzzone, B., Van Wijngaerden, E., Vandamme, A.-M. HIV-1 fitness landscape models for indinavir treatment pressure using observed evolution in longitudinal sequence data are predictive for treatment failure. *Infection, Genetics and Evolution* 2013, 19, pp. 349-360.

Charpentier, C., Camacho, R., Ruelle, J., Kaiser, R., Eberle, J., Gürtler, L., Pironti, A., Stürmer, M., Brun-Vézinet, F., Descamps, D., Obermeier, M. HIV-2EU: Supporting standardized HIV-2 drug resistance interpretation in Europe. *Clinical Infectious Diseases* 2013, 56 (11), pp. 1654-1658.

Obermeier, M., Pironti, A., Berg, T., Braun, P., Däumer, M., Eberle, J., Ehret, R., Kaiser, R., Kleinkauf, N., Korn, K., Kücherer, C., Müller, H., Noah, C., Stürmer, M., Thielen, A., Wolf, E., Walter, H. HIV-GRADE: A publicly available, rules-based drug resistance interpretation algorithm integrating bioinformatics knowledge. *Intervirology* 2012, 55 (2), pp. 102-107.

References

[1] Jay A. Levy. *HIV and the pathogenesis of AIDS*. ASM Press, Washington, D.C, 3rd ed edition, 2007. ISBN 978-1-55581-393-2.

[2] Andrea Rubbert, Georg Behrens, and Mario Ostrowski. Pathogenesis of HIV-1 Infection. In *HIV 2012/2013*. Medizin Fokus Verlag, Hamburg, Germany, 2012. ISBN 978-3-941727-11-3. URL https://hivbook.files.wordpress.com/2011/10/hivbook-2012.pdf.

[3] Christian Hoffmann. ART 2015/2016. In *HIV 2015/16*, pages 64–280. Mdizin Fokus Verlag, Hamburg, Germany, 2015. ISBN 978-3-941727-17-5. URL https://hivbook.files.wordpress.com/2011/10/hiv-2015-16-complete.pdf.

[4] ICTV Taxonomy History, November 2015. URL http://www.ictvonline.org/taxonomyHistory.asp?taxnode_id=20143643&taxa_name=Human%20immunodeficiency%20virus%201.

[5] Abdul A. Waheed and Eric O. Freed. HIV Type 1 Gag as a Target for Antiviral Therapy. *AIDS Research and Human Retroviruses*, 28(1):54–75, January 2012. ISSN 0889-2229, 1931-8405. doi: 10.1089/aid.2011.0230. URL http://www.liebertonline.com/doi/abs/10.1089/aid.2011.0230.

[6] Joëlle V Fritz, Laurence Briant, Yves Mély, Serge Bouaziz, and Hugues de Rocquigny. HIV-1 viral protein r: from structure to function. *Future Virology*, 5(5):607–625, September 2010. ISSN 1746-0794, 1746-0808. doi: 10.2217/fvl.10.47. URL http://www.futuremedicine.com/doi/abs/10.2217/fvl.10.47.

[7] Jane S. Flint, Vincent R. Racaniello, and Robert Krug. *Principles of Virology: Molecular Biology, Pathogenesis, and Control*. American Society for Microbiology, Washington, D.C, September 1999. ISBN 978-1-55581-127-3.

[8] Chukwuka A. Didigu and Robert W. Doms. Novel Approaches to Inhibit HIV Entry. *Viruses*, 4(12):309–324, February 2012. ISSN 1999-4915. doi: 10.3390/v4020309. URL http://www.mdpi.com/1999-4915/4/2/309/.

[9] Robert W. Buckheit Iii and Joel N. Blankson. Immunology of Latent HIV Infection. In Thomas J. Hope, Mario Stevenson, and Douglas Richman, editors, *Encyclopedia of AIDS*, pages 1–7. Springer New York, 2014. ISBN 978-1-4614-9610-6. URL http://link.springer.com/referenceworkentry/10.1007/978-1-4614-9610-6_190-1. DOI: 10.1007/978-1-4614-9610-6_190-1.

[10] M. Sagar. Origin of the Transmitted Virus in HIV Infection: Infected Cells Versus Cell-Free Virus. *Journal of Infectious Diseases*, 210(suppl 3):S667–S673, December 2014. ISSN 0022-1899, 1537-6613.

doi: 10.1093/infdis/jiu369. URL http://jid.oxfordjournals.org/lookup/doi/10.1093/infdis/jiu369.

[11] Thomas C. Quinn, Maria J. Wawer, Nelson Sewankambo, David Serwadda, Chuanjun Li, Fred Wabwire-Mangen, Mary O. Meehan, Thomas Lutalo, and Ronald H. Gray. Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1. *New England Journal of Medicine*, 342(13):921–929, March 2000. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM200003303421303. URL http://www.nejm.org/doi/abs/10.1056/NEJM200003303421303.

[12] Ülgen Semaye Fideli, Susan A. Allen, Rosemary Musonda, Stan Trask, Beatrice H. Hahn, Heidi Weiss, Joseph Mulenga, Francis Kasolo, Sten H. Vermund, and Grace M. Aldrovandi. Virologic and Immunologic Determinants of Heterosexual Transmission of Human Immunodeficiency Virus Type 1 in Africa. *AIDS Research and Human Retroviruses*, 17(10):901–910, July 2001. ISSN 0889-2229, 1931-8405. doi: 10.1089/088922201750290023. URL http://www.liebertonline.com/doi/abs/10.1089/088922201750290023.

[13] M. Laga, A. Manoka, M. Kivuvu, B. Malele, M. Tuliza, N. Nzila, J. Goeman, F. Behets, V. Batter, and M. Alary. Non-ulcerative sexually transmitted diseases as risk factors for HIV-1 transmission in women: results from a cohort study. *AIDS (London, England)*, 7(1):95–102, January 1993. ISSN 0269-9370.

[14] Richard E. Haaland, Paulina A. Hawkins, Jesus Salazar-Gonzalez, Amber Johnson, Amanda Tichacek, Etienne Karita, Olivier Manigart, Joseph Mulenga, Brandon F. Keele, George M. Shaw, Beatrice H. Hahn, Susan A. Allen, Cynthia A. Derdeyn, and Eric Hunter. Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtype A and C HIV-1. *PLoS Pathogens*, 5(1):e1000274, January 2009. ISSN 1553-7374. doi: 10.1371/journal.ppat.1000274. URL http://dx.plos.org/10.1371/journal.ppat.1000274.

[15] D. I. Boeras, P. T. Hraber, M. Hurlston, T. Evans-Strickfaden, T. Bhattacharya, E. E. Giorgi, J. Mulenga, E. Karita, B. T. Korber, S. Allen, C. E. Hart, C. A. Derdeyn, and E. Hunter. Role of donor genital tract HIV-1 diversity in the transmission bottleneck. *Proceedings of the National Academy of Sciences*, 108 (46):E1156–E1163, November 2011. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1103764108. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1103764108.

[16] Yi Liu, Marcel E. Curlin, Kurt Diem, Hong Zhao, Ananta K. Ghosh, Haiying Zhu, Amanda S. Woodward, Janine Maenza, Claire E. Stevens, Joanne Stekler, Ann C. Collier, Indira Genowati, Wenjie Deng, Rafael Zioni, Lawrence Corey, Tuofu Zhu, and James I. Mullins. Env length and N-linked glycosylation following transmission of human immunodeficiency virus Type 1 subtype B viruses. *Virology*, 374(2):229–233, May 2008. ISSN 00426822. doi: 10.1016/j.virol.2008.01.029. URL http://linkinghub.elsevier.com/retrieve/pii/S0042682208000676.

[17] N. F. Parrish, F. Gao, H. Li, E. E. Giorgi, H. J. Barbian, E. H. Parrish, L. Zajic, S. S. Iyer, J. M. Decker, A. Kumar, B. Hora, A. Berg, F. Cai, J. Hopper, T. N. Denny, H. Ding, C. Ochsenbauer, J. C. Kappes, R. P. Galimidi, A. P. West, P. J. Bjorkman, C. B. Wilen, R. W. Doms, M. O'Brien, N. Bhardwaj, P. Borrow, B. F. Haynes, M. Muldoon, J. P. Theiler, B. Korber, G. M. Shaw, and B. H. Hahn. Phenotypic properties of transmitted founder HIV-1. *Proceedings of the National Academy of Sciences*,

110(17):6626–6633, April 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1304288110. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1304288110.

[18] K. Ritola, C. D. Pilcher, S. A. Fiscus, N. G. Hoffman, J. A. E. Nelson, K. M. Kitrinos, C. B. Hicks, J. J. Eron, and R. Swanstrom. Multiple V1/V2 env Variants Are Frequently Present during Primary Infection with Human Immunodeficiency Virus Type 1. *Journal of Virology*, 78(20):11208–11218, October 2004. ISSN 0022-538X. doi: 10.1128/JVI.78.20.11208-11218.2004. URL http://jvi.asm.org/cgi/doi/10.1128/JVI.78.20.11208-11218.2004.

[19] S. Gnanakaran, Tanmoy Bhattacharya, Marcus Daniels, Brandon F. Keele, Peter T. Hraber, Alan S. Lapedes, Tongye Shen, Brian Gaschen, Mohan Krishnamoorthy, Hui Li, Julie M. Decker, Jesus F. Salazar-Gonzalez, Shuyi Wang, Chunlai Jiang, Feng Gao, Ronald Swanstrom, Jeffrey A. Anderson, Li-Hua Ping, Myron S. Cohen, Martin Markowitz, Paul A. Goepfert, Michael S. Saag, Joseph J. Eron, Charles B. Hicks, William A. Blattner, Georgia D. Tomaras, Mohammed Asmal, Norman L. Letvin, Peter B. Gilbert, Allan C. Decamp, Craig A. Magaret, William R. Schief, Yih-En Andrew Ban, Ming Zhang, Kelly A. Soderberg, Joseph G. Sodroski, Barton F. Haynes, George M. Shaw, Beatrice H. Hahn, and Bette Korber. Recurrent signature patterns in HIV-1 B clade envelope glycoproteins associated with either early or chronic infections. *PLoS pathogens*, 7(9):e1002209, September 2011. ISSN 1553-7374. doi: 10.1371/journal.ppat.1002209.

[20] Manish Sagar, Ludo Lavreys, Jared M. Baeten, Barbra A. Richardson, Kishorchandra Mandaliya, Jeckoniah O. Ndinya-Achola, Joan K. Kreiss, and Julie Overbaugh. Identification of modifiable factors that affect the genetic diversity of the transmitted HIV-1 population. *AIDS (London, England)*, 18(4):615–619, March 2004. ISSN 0269-9370.

[21] E. M. Long, H. L. Martin, J. K. Kreiss, S. M. Rainwater, L. Lavreys, D. J. Jackson, J. Rakwar, K. Mandaliya, and J. Overbaugh. Gender differences in HIV-1 diversity at time of infection. *Nature Medicine*, 6(1):71–75, January 2000. ISSN 1078-8956. doi: 10.1038/71563.

[22] T. B. H. Geijtenbeek and Y. van Kooyk. DC-SIGN: A Novel HIV Receptor on DCs That Mediates HIV-1 Transmission. In R. W. Compans, M. D. Cooper, Y. Ito, H. Koprowski, F. Melchers, M. B. A. Oldstone, S. Olsnes, M. Potter, P. K. Vogt, H. Wagner, and Alexander Steinkasserer, editors, *Dendritic Cells and Virus Infection*, volume 276, pages 31–54. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. ISBN 978-3-642-07926-9 978-3-662-06508-2. URL http://link.springer.com/10.1007/978-3-662-06508-2_2.

[23] Elizabeth Miller, Nina Bhardwaj, and Meagan O'Brien. Dendritic cell function in HIV infection. *HIV Therapy*, 3(5):527–537, September 2009. ISSN 1758-4310. doi: 10.2217/hiv.09.34. URL http://www.futuremedicine.com/doi/abs/10.2217/hiv.09.34.

[24] C. Bradley Hare and James O. Kahn. Primary HIV infection. *Current Infectious Disease Reports*, 6(1):65–71, January 2004. ISSN 1523-3847, 1534-3146. doi: 10.1007/s11908-004-0026-1. URL http://link.springer.com/10.1007/s11908-004-0026-1.

[25] R. M. Ribeiro, L. Qin, L. L. Chavez, D. Li, S. G. Self, and A. S. Perelson. Estimation of the Initial Viral Growth Rate and Basic Reproductive Number during Acute HIV-1 Infection. *Journal of Virology*, 84 (12):6096–6102, June 2010. ISSN 0022-538X. doi: 10.1128/JVI.00127-10. URL http://jvi.asm.org/cgi/doi/10.1128/JVI.00127-10.

[26] S. Kassutto and E. S. Rosenberg. Primary HIV Type 1 Infection. *Clinical Infectious Diseases*, 38 (10):1447–1453, May 2004. ISSN 1058-4838, 1537-6591. doi: 10.1086/420745. URL http://cid.oxfordjournals.org/lookup/doi/10.1086/420745.

[27] Luwy Musey, James Hughes, Timothy Schacker, Theresa Shea, Lawrence Corey, and M. Juliana McElrath. Cytotoxic-T-Cell Responses, Viral Load, and Disease Progression in Early Human Immunodeficiency Virus Type 1 Infection. *New England Journal of Medicine*, 337(18):1267–1274, October 1997. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199710303371803. URL http://www.nejm.org/doi/abs/10.1056/NEJM199710303371803.

[28] R. A. Koup, J. T. Safrit, Y. Cao, C. A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, and D. D. Ho. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *Journal of Virology*, 68(7):4650–4655, July 1994. ISSN 0022-538X.

[29] H W Sheppard and M S Ascher. The Natural History and Pathogenesis of HIV Infection. *Annual Review of Microbiology*, 46(1):533–564, October 1992. ISSN 0066-4227, 1545-3251. doi: 10.1146/annurev.mi.46.100192.002533. URL http://www.annualreviews.org/doi/abs/10.1146/annurev.mi.46.100192.002533.

[30] Gilbert R. Kaufmann, Chris Duncombe, John Zaunders, Philip Cunningham, and David Cooper. Primary HIV-1 Infection: A Review of Clinical Manifestations, Immunologic and Virologic Changes. *AIDS Patient Care and STDs*, 12(10):759–767, October 1998. ISSN 1087-2914, 1557-7449. doi: 10.1089/apc.1998.12.759. URL http://www.liebertonline.com/doi/abs/10.1089/apc.1998.12.759.

[31] Lauren E. Richey and Jason Halperin. Acute human immunodeficiency virus infection. *The American Journal of the Medical Sciences*, 345(2):136–142, February 2013. ISSN 1538-2990. doi: 10.1097/MAJ.0b013e31825d4b88.

[32] J. W. Mellors, C. R. Rinaldo, P. Gupta, R. M. White, J. A. Todd, and L. A. Kingsley. Prognosis in HIV-1 Infection Predicted by the Quantity of Virus in Plasma. *Science*, 272(5265):1167–1170, May 1996. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.272.5265.1167. URL http://www.sciencemag.org/cgi/doi/10.1126/science.272.5265.1167.

[33] Giuseppe Pantaleo, Cecilia Graziosi, James F. Demarest, Luca Butini, Maria Montroni, Cecil H. Fox, Jan M. Orenstein, Donald P. Kotler, and Anthony S. Fauci. HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature*, 362(6418):355–358, March 1993. ISSN 0028-0836. doi: 10.1038/362355a0. URL http://www.nature.com/doifinder/10.1038/362355a0.

[34] M Piatak, M. Saag, L. Yang, S. Clark, J. Kappes, K. Luk, B. Hahn, G. Shaw, and J. Lifson. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science*, 259(5102): 1749–1754, March 1993. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.8096089. URL http://www.sciencemag.org/cgi/doi/10.1126/science.8096089.

[35] J. Weber. The pathogenesis of HIV-1 infection. *British Medical Bulletin*, 58(1):61–72, September 2001. ISSN 14718391. doi: 10.1093/bmb/58.1.61. URL http://bmb.oupjournals.org/cgi/doi/10.1093/bmb/58.1.61.

[36] Michael S. Gottlieb, Robert Schroff, Howard M. Schanker, Joel D. Weisman, Peng Thim Fan, Robert A. Wolf, and Andrew Saxon. *Pneumocystis carinii* Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men: Evidence of a New Acquired Cellular Immunodeficiency. *New England Journal of Medicine*, 305(24):1425–1431, December 1981. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM198112103052401. URL http://www.nejm.org/doi/abs/10.1056/NEJM198112103052401.

[37] F Barre-Sinoussi, J. Chermann, F Rey, M. Nugeyre, S Chamaret, J Gruest, C Dauguet, C Axler-Blin, F Vezinet-Brun, C Rouzioux, W Rozenbaum, and L Montagnier. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, 220(4599): 868–871, May 1983. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.6189183. URL http://www.sciencemag.org/cgi/doi/10.1126/science.6189183.

[38] R. Gallo, S. Salahuddin, M Popovic, G. Shearer, M Kaplan, B. Haynes, T. Palker, R Redfield, J Oleske, B Safai, and al. et. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, 224(4648):500–503, May 1984. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.6200936. URL http://www.sciencemag.org/cgi/doi/10.1126/science.6200936.

[39] M Popovic, M. Sarngadharan, E Read, and R. Gallo. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*, 224(4648): 497–500, May 1984. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.6200935. URL http://www.sciencemag.org/cgi/doi/10.1126/science.6200935.

[40] Nathan Clumeck, Jean Sonnet, Henri Taelman, Françoise Mascart-Lemone, Marc De Bruyere, Philippe Vandeperre, Jean Dasnoy, Luc Marcelis, Monique Lamy, Claude Jonas, Luc Eyckmans, Henri Noel, Michel Vanhaeverbeek, and Jean-Paul Butzler. Acquired Immunodeficiency Syndrome in African Patients. *New England Journal of Medicine*, 310(8):492–497, February 1984. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM198402233100804. URL http://www.nejm.org/doi/abs/10.1056/NEJM198402233100804.

[41] J Brunet. ACQUIRED IMMUNODEFICIENCY SYNDROME IN FRANCE. *The Lancet*, 321 (8326):700–701, March 1983. ISSN 01406736. doi: 10.1016/S0140-6736(83)91985-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673683919852.

[42] J. Vandepitte, R. Verwilghen, and P. Zachee. AIDS AND CRYPTOCOCCOSIS (ZAIRE, 1977). *The Lancet*, 321(8330):925–926, April 1983. ISSN 01406736. doi: 10.1016/S0140-6736(83)91349-1. URL `http://linkinghub.elsevier.com/retrieve/pii/S0140673683913491`.

[43] D. Edwards, P.G. Harper, A.K. Pain, J. Welch, C. Barbatis, and C. Mallinson. KAPOSI'S SARCOMA ASSOCIATED WITH AIDS IN A WOMAN FROM UGANDA. *The Lancet*, 323(8377):631–632, March 1984. ISSN 01406736. doi: 10.1016/S0140-6736(84)91028-6. URL `http://linkinghub.elsevier.com/retrieve/pii/S0140673684910286`.

[44] I.C. Bygbjerg. AIDS IN A DANISH SURGEON (ZAIRE, 1976). *The Lancet*, 321(8330):925, April 1983. ISSN 01406736. doi: 10.1016/S0140-6736(83)91348-X. URL `http://linkinghub.elsevier.com/retrieve/pii/S014067368391348X`.

[45] A. Ellrodt, Ph. Le Bras, L. Palazzo, F. Brun-Vezinet, P. Segond, L. Montagnier, F. Barre-Sinoussi, M.T. Nugeyre, F. Rey, C. Rouzioux, R. Caquet, and J.C. Chermann. ISOLATION OF HUMAN T-LYMPHOTROPIC RETROVIRUS (LAV) FROM ZAIRIAN MARRIED COUPLE, ONE WITH AIDS, ONE WITH PRODROMES. *The Lancet*, 323(8391):1383–1385, June 1984. ISSN 01406736. doi: 10.1016/S0140-6736(84)91877-4. URL `http://linkinghub.elsevier.com/retrieve/pii/S0140673684918774`.

[46] P Sonigo, M Alizon, K Staskus, D Klatzmann, S Cole, O Danos, E Retzel, P Tiollais, A Haase, and S Wainhobson. Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. *Cell*, 42 (1):369–382, August 1985. ISSN 00928674. doi: 10.1016/S0092-8674(85)80132-X. URL `http://linkinghub.elsevier.com/retrieve/pii/S009286748580132X`.

[47] F Clavel, D Guetard, F Brun-Vezinet, S Chamaret, M. Rey, M. Santos-Ferreira, A. Laurent, C Dauguet, C Katlama, C Rouzioux, and e. al. Isolation of a new human retrovirus from West African patients with AIDS. *Science*, 233(4761):343–346, July 1986. ISSN 0036-8075, 1095-9203. doi: 10.1126/science. 2425430. URL `http://www.sciencemag.org/cgi/doi/10.1126/science.2425430`.

[48] Mireille Guyader, Michael Emerman, Pierre Sonigo, François Clavel, Luc Montagnier, and Marc Alizon. Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature*, 326 (6114):662–669, April 1987. ISSN 0028-0836. doi: 10.1038/326662a0. URL `http://www.nature.com/doifinder/10.1038/326662a0`.

[49] Lisa Chakrabarti, Mireille Guyader, Marc Alizon, Muthiah D. Daniel, Ronald C. Desrosiers, Pierre Tiollais, and Pierre Sonigo. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. *Nature*, 328(6130):543–547, August 1987. ISSN 0028-0836. doi: 10.1038/328543a0. URL `http://www.nature.com/doifinder/10.1038/328543a0`.

[50] Joris Hemelaar. The origin and diversity of the HIV-1 pandemic. *Trends in Molecular Medicine*, 18(3):182–192, March 2012. ISSN 14714914. doi: 10.1016/j.molmed.2011.12.001. URL `http://linkinghub.elsevier.com/retrieve/pii/S1471491411002103`.

[51] B. H. Hahn. AIDS as a Zoonosis: Scientific and Public Health Implications. *Science*, 287(5453):607–614, January 2000. ISSN 00368075, 10959203. doi: 10.1126/science.287.5453.607. URL http://www.sciencemag.org/cgi/doi/10.1126/science.287.5453.607.

[52] Philippe Lemey, Marco Salemi, and Anne-Mieke Vandamme. *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*. Cambridge University Press, Cambridge, UK ; New York, 2 edition, 2009. ISBN 978-0-521-73071-6.

[53] Avelin F. Aghokeng, Ahidjo Ayouba, Eitel Mpoudi-Ngole, Severin Loul, Florian Liegeois, Eric Delaporte, and Martine Peeters. Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. *Infection, Genetics and Evolution*, 10(3):386–396, April 2010. ISSN 15671348. doi: 10.1016/j.meegid.2009.04.014. URL http://linkinghub.elsevier.com/retrieve/pii/S1567134809000872.

[54] P. M. Sharp and B. H. Hahn. Origins of HIV and the AIDS Pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1):a006841–a006841, September 2011. ISSN 2157-1422. doi: 10.1101/cshperspect.a006841. URL http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a006841.

[55] P. M. Sharp and B. H. Hahn. The evolution of HIV-1 and the origin of AIDS. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552):2487–2494, August 2010. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2010.0031. URL http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2010.0031.

[56] Martine Peeters, Matthieu Jung, and Ahidjo Ayouba. The origin and molecular epidemiology of HIV. *Expert Review of Anti-Infective Therapy*, 11(9):885–896, September 2013. ISSN 1744-8336. doi: 10.1586/14787210.2013.825443.

[57] Martine Peeters, Valerie Courgnaud, Bernadette Abela, Philippe Auzel, Xavier Pourrut, Frederic Bibollet-Ruche, Severin Loul, Florian Liegeois, Cristelle Butel, Denis Koulagna, Eitel Mpoudi-Ngole, George M. Shaw, Beatrice H. Hahn, and Eric Delaporte. Risk to Human Health from a Plethora of Simian Immunodeficiency Viruses in Primate Bushmeat. *Emerging Infectious Diseases*, 8(5):451–457, May 2002. ISSN 1080-6040, 1080-6059. doi: 10.3201/eid0805.010522. URL http://wwwnc.cdc.gov/eid/article/8/5/01-0522_article.htm.

[58] T. Zhu, B. T. Korber, A. J. Nahmias, E. Hooper, P. M. Sharp, and D. D. Ho. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature*, 391(6667):594–597, February 1998. ISSN 0028-0836. doi: 10.1038/35400.

[59] Michael Worobey, Marlea Gemmel, Dirk E. Teuwen, Tamara Haselkorn, Kevin Kunstman, Michael Bunce, Jean-Jacques Muyembe, Jean-Marie M. Kabongo, Raphaël M. Kalengayi, Eric Van Marck, M. Thomas P. Gilbert, and Steven M. Wolinsky. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*, 455(7213):661–664, October 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature07390. URL http://www.nature.com/doifinder/10.1038/nature07390.

[60] B. Korber. Timing the Ancestor of the HIV-1 Pandemic Strains. *Science*, 288(5472):1789–1796, June 2000. ISSN 00368075, 10959203. doi: 10.1126/science.288.5472.1789. URL http://www.sciencemag.org/cgi/doi/10.1126/science.288.5472.1789.

[61] N. R. Faria, A. Rambaut, M. A. Suchard, G. Baele, T. Bedford, M. J. Ward, A. J. Tatem, J. D. Sousa, N. Arinaminpathy, J. Pepin, D. Posada, M. Peeters, O. G. Pybus, and P. Lemey. The early spread and epidemic ignition of HIV-1 in human populations. *Science*, 346(6205):56–61, October 2014. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1256739. URL http://www.sciencemag.org/cgi/doi/10.1126/science.1256739.

[62] Joel O. Wertheim and Michael Worobey. Dating the Age of the SIV Lineages That Gave Rise to HIV-1 and HIV-2. *PLoS Computational Biology*, 5(5):e1000377, May 2009. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000377. URL http://dx.plos.org/10.1371/journal.pcbi.1000377.

[63] W. H. Li, M. Tanimura, and P. M. Sharp. Rates and dates of divergence between AIDS virus nucleotide sequences. *Molecular Biology and Evolution*, 5(4):313–330, July 1988. ISSN 0737-4038.

[64] Joris Hemelaar, Eleanor Gouws, Peter D Ghys, and Saladin Osmanov. Global trends in molecular epidemiology of HIV-1 during 2000–2007:. *AIDS*, 25(5):679–689, March 2011. ISSN 0269-9370. doi: 10.1097/QAD.0b013e328342ff93. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201103130-00018.

[65] N. Vidal, M. Peeters, C. Mulanga-Kabeya, N. Nzilambi, D. Robertson, W. Ilunga, H. Sema, K. Tshimanga, B. Bongo, and E. Delaporte. Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *Journal of Virology*, 74(22):10498–10507, November 2000. ISSN 0022-538X.

[66] A. Rambaut, D. L. Robertson, O. G. Pybus, M. Peeters, and E. C. Holmes. Human immunodeficiency virus. Phylogeny and the origin of HIV-1. *Nature*, 410(6832):1047–1048, April 2001. ISSN 0028-0836. doi: 10.1038/35074179.

[67] Paul M. Sharp and Beatrice H. Hahn. AIDS: Prehistory of HIV-1. *Nature*, 455(7213):605–606, October 2008. ISSN 0028-0836, 1476-4687. doi: 10.1038/455605a. URL http://www.nature.com/doifinder/10.1038/455605a.

[68] P. A. Marx, P. G. Alcabes, and E. Drucker. Serial human passage of simian immunodeficiency virus by unsterile injections and the emergence of epidemic human immunodeficiency virus in Africa. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1410):911–920, June 2001. ISSN 0962-8436, 1471-2970. doi: 10.1098/rstb.2001.0867. URL http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2001.0867.

[69] Ernest Drucker, Phillip G Alcabes, and Preston A Marx. The injection century: massive unsterile injections and the emergence of human pathogens. *The Lancet*, 358(9297):1989–1992, December 2001. ISSN 01406736. doi: 10.1016/S0140-6736(01)06967-7. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673601069677.

[70] William H. Schneider. History of Blood Transfusion in Sub-Saharan Africa. *Transfusion Medicine Reviews*, 27(1):21–28, January 2013. ISSN 08877963. doi: 10.1016/j.tmrv.2012.08.001. URL http://linkinghub.elsevier.com/retrieve/pii/S0887796312000703.

[71] William H. Schneider and Ernest Drucker. Blood Transfusions in the Early Years of AIDS in Sub-Saharan Africa. *American Journal of Public Health*, 96(6):984–994, June 2006. ISSN 0090-0036, 1541-0048. doi: 10.2105/AJPH.2004.061630. URL http://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2004.061630.

[72] Tamara Giles-Vernick, Ch. Didier Gondola, Guillaume Lachenal, and William H. Schneider. SOCIAL HISTORY, BIOLOGY, AND THE EMERGENCE OF HIV IN COLONIAL AFRICA. *The Journal of African History*, 54(01):11–30, March 2013. ISSN 0021-8537, 1469-5138. doi: 10.1017/S0021853713000029. URL http://www.journals.cambridge.org/abstract_S0021853713000029.

[73] Amit Chitnis, Diana Rawls, and Jim Moore. Origin of HIV Type 1 in Colonial French Equatorial Africa? *AIDS Research and Human Retroviruses*, 16(1):5–8, January 2000. ISSN 0889-2229, 1931-8405. doi: 10.1089/088922200309548. URL http://www.liebertonline.com/doi/abs/10.1089/088922200309548.

[74] Jacques Pépin. The expansion of HIV-1 in colonial Leopoldville, 1950s: driven by STDs or STD control? *Sexually Transmitted Infections*, 88(4):307–312, June 2012. ISSN 1472-3263. doi: 10.1136/sextrans-2011-050277.

[75] M. T. P. Gilbert, A. Rambaut, G. Wlasiuk, T. J. Spira, A. E. Pitchenik, and M. Worobey. The emergence of HIV/AIDS in the Americas and beyond. *Proceedings of the National Academy of Sciences*, 104(47): 18566–18570, November 2007. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0705329104. URL http://www.pnas.org/cgi/doi/10.1073/pnas.0705329104.

[76] Peter Piot, Henri Taelman, Kapita Bila Minlangu, N. Mbendi, K. Ndangi, Kayembe Kalambayi, Chris Bridts, ThomasC. Quinn, FredM. Feinsod, Odio Wobin, P. Mazebo, Wim Stevens, Sheila Mitchell, and JosephB. Mccormick. ACQUIRED IMMUNODEFICIENCY SYNDROME IN A HETEROSEXUAL POPULATION IN ZAIRE. *The Lancet*, 324(8394):65–69, July 1984. ISSN 01406736. doi: 10.1016/S0140-6736(84)90241-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673684902411.

[77] R. C. Gallo. The AIDS virus. *Scientific American*, 256(1):46–56, January 1987. ISSN 0036-8733.

[78] K. E. Robbins, P. Lemey, O. G. Pybus, H. W. Jaffe, A. S. Youngpairoj, T. M. Brown, M. Salemi, A.-M. Vandamme, and M. L. Kalish. U.S. Human Immunodeficiency Virus Type 1 Epidemic: Date of Origin, Population History, and Characterization of Early Strains. *Journal of Virology*, 77(11):6359–6366, June 2003. ISSN 0022-538X. doi: 10.1128/JVI.77.11.6359-6366.2003. URL http://jvi.asm.org/cgi/doi/10.1128/JVI.77.11.6359-6366.2003.

[79] Christian Hoffmann and Stefan Esser. Kaposi's Sarcoma. In *HIV 2012/2013*. Medizin Fokus Verlag, Hamburg, Germany, 2012. ISBN 978-3-941727-11-3. URL https://hivbook.files.wordpress.com/2011/10/hivbook-2012.pdf.

[80] H. W. Jaffe, K. Choi, P. A. Thomas, H. W. Haverkos, D. M. Auerbach, M. E. Guinan, M. F. Rogers, T. J. Spira, W. W. Darrow, M. A. Kramer, S. M. Friedman, J. M. Monroe, A. E. Friedman-Kien, L. J. Laubenstein, M. Marmor, B. Safai, S. K. Dritz, S. J. Crispi, S. L. Fannin, J. P. Orkwis, A. Kelter, W. R. Rushing, S. B. Thacker, and J. W. Curran. National case-control study of Kaposi's sarcoma and Pneumocystis carinii pneumonia in homosexual men: Part 1. Epidemiologic results. *Annals of Internal Medicine*, 99 (2):145–151, August 1983. ISSN 0003-4819.

[81] Kevin M. De Cock, Harold W. Jaffe, and James W. Curran. The evolving epidemiology of HIV/AIDS:. *AIDS*, 26(10):1205–1213, June 2012. ISSN 0269-9370. doi: 10.1097/QAD.0b013e328354622a. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201206190-00009.

[82] David Klatzmann, Eric Champagne, Sophie Chamaret, Jacqueline Gruest, Denise Guetard, Thierry Hercend, Jean-Claude Gluckman, and Luc Montagnier. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature*, 312(5996):767–768, December 1984. ISSN 0028-0836. doi: 10.1038/312767a0. URL http://www.nature.com/doifinder/10.1038/312767a0.

[83] Angus G. Dalgleish, Peter C. L. Beverley, Paul R. Clapham, Dorothy H. Crawford, Melvyn F. Greaves, and Robin A. Weiss. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*, 312(5996):763–767, December 1984. ISSN 0028-0836. doi: 10.1038/312763a0. URL http://www.nature.com/doifinder/10.1038/312763a0.

[84] M. G. Sarngadharan, M. Popovic, L. Bruch, J. Schüpbach, and R. C. Gallo. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. *Science (New York, N.Y.)*, 224(4648):506–508, May 1984. ISSN 0036-8075.

[85] F. Brun-Vezinet, F. Barre-Sinoussi, A.G. Saimot, D. Christol, L. Montagnier, C. Rouzioux, D. Klatzmann, W. Rozenbaum, J.C. Gluckmann, and J.C. Chermann. DETECTION OF IgG ANTIBODIES TO LYMPHADENOPATHY-ASSOCIATED VIRUS IN PATIENTS WITH AIDS OR LYMPHADENOPATHY SYNDROME. *The Lancet*, 323(8389):1253–1256, June 1984. ISSN 01406736. doi: 10.1016/S0140-6736(84)92444-9. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673684924449.

[86] William M. Janda. Serologic tests for HTLV-III antibodies: Methods and interpretations. *Clinical Microbiology Newsletter*, 7(10):67–69, May 1985. ISSN 01964399. doi: 10.1016/S0196-4399(85)80088-X. URL http://linkinghub.elsevier.com/retrieve/pii/S019643998580088X.

[87] Laurence R. McCarthy. AIDS: The causative agent and new diagnostic tests. *Clinical Microbiology Newsletter*, 6(23):169–171, December 1984. ISSN 01964399. doi: 10.1016/S0196-4399(84)80012-4. URL http://linkinghub.elsevier.com/retrieve/pii/S0196439984800124.

[88] H. A. Perkins. Safety of the blood supply. *Journal of Clinical Apheresis*, 8(2):110–116, 1993. ISSN 0733-2459.

[89] Simon Wain-Hobson, Pierre Sonigo, Olivier Danos, Stewart Cole, and Marc Alizon. Nucleotide sequence of the AIDS virus, LAV. *Cell*, 40(1):9–17, January 1985. ISSN 00928674. doi: 10.1016/0092-8674(85)90303-4. URL http://linkinghub.elsevier.com/retrieve/pii/0092867485903034.

[90] Margaret A. Fischl, Douglas D. Richman, Michael H. Grieco, Michael S. Gottlieb, Paul A. Volberding, Oscar L. Laskin, John M. Leedom, Jerome E. Groopman, Donna Mildvan, Robert T. Schooley, George G. Jackson, David T. Durack, Dannie King, and The AZT Collaborative Working Group. The Efficacy of Azidothymidine (AZT) in the Treatment of Patients with AIDS and AIDS-Related Complex. *New England Journal of Medicine*, 317(4):185–191, July 1987. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM198707233170401. URL http://www.nejm.org/doi/abs/10.1056/NEJM198707233170401.

[91] Edward M. Connor, Rhoda S. Sperling, Richard Gelber, Pavel Kiselev, Gwendolyn Scott, Mary Jo O'Sullivan, Russell VanDyke, Mohammed Bey, William Shearer, Robert L. Jacobson, Eleanor Jimenez, Edward O'Neill, Brigitte Bazin, Jean-Francois Delfraissy, Mary Culnane, Robert Coombs, Mary Elkins, Jack Moye, Pamela Stratton, and James Balsley. Reduction of Maternal-Infant Transmission of Human Immunodeficiency Virus Type 1 with Zidovudine Treatment. *New England Journal of Medicine*, 331 (18):1173–1180, November 1994. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199411033311801. URL http://www.nejm.org/doi/abs/10.1056/NEJM199411033311801.

[92] Robert M. Grant, Javier R. Lama, Peter L. Anderson, Vanessa McMahan, Albert Y. Liu, Lorena Vargas, Pedro Goicochea, Martín Casapía, Juan Vicente Guanira-Carranza, Maria E. Ramirez-Cardich, Orlando Montoya-Herrera, Telmo Fernández, Valdilea G. Veloso, Susan P. Buchbinder, Suwat Chariyalertsak, Mauro Schechter, Linda-Gail Bekker, Kenneth H. Mayer, Esper Georges Kallás, K. Rivet Amico, Kathleen Mulligan, Lane R. Bushman, Robert J. Hance, Carmela Ganoza, Patricia Defechereux, Brian Postle, Furong Wang, J. Jeff McConnell, Jia-Hua Zheng, Jeanny Lee, James F. Rooney, Howard S. Jaffe, Ana I. Martinez, David N. Burns, and David V. Glidden. Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men. *New England Journal of Medicine*, 363(27): 2587–2599, December 2010. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa1011205. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa1011205.

[93] Myron S. Cohen, Ying Q. Chen, Marybeth McCauley, Theresa Gamble, Mina C. Hosseinipour, Nagalingeswaran Kumarasamy, James G. Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose H.S. Pilotto, Sheela V. Godbole, Sanjay Mehendale, Suwat Chariyalertsak, Breno R. Santos, Kenneth H. Mayer, Irving F. Hoffman, Susan H. Eshleman, Estelle Piwowar-Manning, Lei Wang, Joseph Makhema, Lisa A. Mills, Guy de Bruyn, Ian Sanne, Joseph Eron, Joel Gallant, Diane Havlir, Susan Swindells, Heather Ribaudo, Vanessa Elharrar, David Burns, Taha E. Taha, Karin Nielsen-Saines, David Celentano, Max Essex, and Thomas R. Fleming. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *New England Journal of Medicine*, 365(6):493–505, August 2011. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa1105243. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa1105243.

[94] Gero Hütter, Daniel Nowak, Maximilian Mossner, Susanne Ganepola, Arne Müßig, Kristina Allers, Thomas Schneider, Jörg Hofmann, Claudia Kücherer, Olga Blau, Igor W. Blau, Wolf K. Hofmann, and Eckhard Thiel. Long-Term Control of HIV by *CCR5* Delta32/Delta32 Stem-Cell Transplantation. *New England Journal of Medicine*, 360(7):692–698, February 2009. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa0802905. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa0802905.

[95] The Director-General. Global Strategy for the Prevention and Control of AIDS. Technical Report A42/11, World Health Organization, April 1989. URL http://apps.who.int/iris/bitstream/10665/171165/1/WHA42_11_eng.pdf.

[96] M M Braun, W L Heyward, and J W Curran. The Global Epidemiology of HIV Infection and Aids. *Annual Review of Microbiology*, 44(1):555–577, October 1990. ISSN 0066-4227, 1545-3251. doi: 10.1146/annurev.mi.44.100190.003011. URL http://www.annualreviews.org/doi/abs/10.1146/annurev.mi.44.100190.003011.

[97] World Health Organization. The World health report : 1999 : Making a difference. 1999. ISSN 1020-3311. URL http://www.who.int/iris/handle/10665/42167.

[98] H. D. Gayle and G. L. Hill. Global Impact of Human Immunodeficiency Virus and AIDS. *Clinical Microbiology Reviews*, 14(2):327–335, April 2001. ISSN 0893-8512. doi: 10.1128/CMR.14.2.327-335.2001. URL http://cmr.asm.org/cgi/doi/10.1128/CMR.14.2.327-335.2001.

[99] Frank J. Palella, Kathleen M. Delaney, Anne C. Moorman, Mark O. Loveless, Jack Fuhrer, Glen A. Satten, Diane J. Aschman, and Scott D. Holmberg. Declining Morbidity and Mortality among Patients with Advanced Human Immunodeficiency Virus Infection. *New England Journal of Medicine*, 338(13):853–860, March 1998. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199803263381301. URL http://www.nejm.org/doi/abs/10.1056/NEJM199803263381301.

[100] UNAIDS and WHO. *AIDS epidemic update: December 2007*. UNAIDS, Geneva, 2007. ISBN 978-92-9173-621-8.

[101] General Assembly of The United Nations. United Nations Millennium Declaration A/RES/55/2. Technical Report A /RES/55/2, September 2000. URL http://www.un.org/millennium/declaration/ares552e.pdf.

[102] General Assembly of The United Nations. Political Declaration on HIV and AIDS: Intensifying Our Efforts to Eliminate HIV and AIDS A/RES/65/277. Technical Report A / RES/65/277, United Nations, July 2011. URL www.unaids.org/sites/default/files/sub_landing/files/20110610_UN_A-RES-65-277_en.pdf.

[103] UNAIDS. *90-90-90 An ambitious treatment target to help end the AIDS epidemic*. Joint United Nations Programme on HIV/AIDS (UNAIDS), Geneva, October 2014. URL http://www.unaids.org/sites/default/files/media_asset/90-90-90_en_0.pdf.

[104] J. Roberts, K Bebenek, and T. Kunkel. The accuracy of reverse transcriptase from HIV-1. *Science*, 242 (4882):1171–1173, November 1988. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.2460925. URL http://www.sciencemag.org/cgi/doi/10.1126/science.2460925.

[105] L. M. Mansky and H. M. Temin. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of Virology*, 69(8):5087–5094, August 1995. ISSN 0022-538X.

[106] Redmond P. Smyth, Miles P. Davenport, and Johnson Mak. The origin of genetic diversity in HIV-1. *Virus Research*, 169(2):415–429, November 2012. ISSN 01681702. doi: 10.1016/j.virusres.2012.06.015. URL http://linkinghub.elsevier.com/retrieve/pii/S0168170212002122.

[107] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science*, 271(5255): 1582–1586, March 1996. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.271.5255.1582. URL http://www.sciencemag.org/cgi/doi/10.1126/science.271.5255.1582.

[108] Andrew Rambaut, David Posada, Keith A. Crandall, and Edward C. Holmes. The causes and consequences of HIV evolution. *Nature Reviews Genetics*, 5(1):52–61, January 2004. ISSN 1471-0056, 1471-0064. doi: 10.1038/nrg1246. URL http://www.nature.com/doifinder/10.1038/nrg1246.

[109] E. Domingo, J. Sheldon, and C. Perales. Viral Quasispecies Evolution. *Microbiology and Molecular Biology Reviews*, 76(2):159–216, June 2012. ISSN 1092-2172. doi: 10.1128/MMBR.05023-11. URL http://mmbr.asm.org/cgi/doi/10.1128/MMBR.05023-11.

[110] Adam S. Lauring and Raul Andino. Quasispecies Theory and the Behavior of RNA Viruses. *PLoS Pathogens*, 6(7):e1001005, July 2010. ISSN 1553-7374. doi: 10.1371/journal.ppat.1001005. URL http://dx.plos.org/10.1371/journal.ppat.1001005.

[111] Shiven B. Chabria, Shaili Gupta, and Michael J. Kozal. Deep Sequencing of HIV: Clinical and Research Applications. *Annual Review of Genomics and Human Genetics*, 15(1):295–325, August 2014. ISSN 1527-8204, 1545-293X. doi: 10.1146/annurev-genom-091212-153406. URL http://www.annualreviews.org/doi/abs/10.1146/annurev-genom-091212-153406.

[112] Arjen J. Stam, Monique Nijhuis, Walter M. van den Bergh, and Annemarie M. J. Wensing. Differential genotypic evolution of HIV-1 quasispecies in cerebrospinal fluid and plasma: a systematic review. *AIDS reviews*, 15(3):152–161, September 2013. ISSN 1698-6997.

[113] Bram Vrancken, Andrew Rambaut, Marc A. Suchard, Alexei Drummond, Guy Baele, Inge Derdelinckx, Eric Van Wijngaerden, Anne-Mieke Vandamme, Kristel Van Laethem, and Philippe Lemey. The genealogical population dynamics of HIV-1 in a large transmission chain: bridging within and among host evolutionary rates. *PLoS computational biology*, 10(4):e1003505, April 2014. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1003505.

[114] D. L. Robertson. HIV-1 Nomenclature Proposal. *Science*, 288(5463):55d–55, April 2000. ISSN 00368075, 10959203. doi: 10.1126/science.288.5463.55d. URL http://www.sciencemag.org/cgi/doi/10.1126/science.288.5463.55d.

[115] Jared M. Baeten, Bhavna Chohan, Ludo Lavreys, Vrasha Chohan, R. Scott McClelland, Laura Certain, Kishorchandra Mandaliya, Walter Jaoko, and Julie Overbaugh. HIV-1 Subtype D Infection Is Associated

with Faster Disease Progression than Subtype A in Spite of Similar Plasma HIV-1 Loads. *The Journal of Infectious Diseases*, 195(8):1177–1180, April 2007. ISSN 0022-1899, 1537-6613. doi: 10.1086/512682. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/512682.

[116] Pontiano Kaleebu, Neil French, Cedric Mahe, David Yirrell, Christine Watera, Fred Lyagoba, Jessica Nakiyingi, Alleluiah Rutebemberwa, Dilys Morgan, Jonathan Weber, Charles Gilks, and Jimmy Whitworth. Effect of Human Immunodeficiency Virus (HIV) Type 1 Envelope Subtypes A and D on Disease Progression in a Large Cohort of HIV-1–Positive Persons in Uganda. *The Journal of Infectious Diseases*, 185(9):1244–1250, May 2002. ISSN 0022-1899, 1537-6613. doi: 10.1086/340130. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/340130.

[117] Noah Kiwanuka, Oliver Laeyendecker, Merlin Robb, Godfrey Kigozi, Miguel Arroyo, Francine Mc-Cutchan, Leigh Anne Eller, Michael Eller, Fred Makumbi, Deborah Birx, Fred Wabwire-Mangen, David Serwadda, Nelson K. Sewankambo, Thomas C. Quinn, Maria Wawer, and Ronald Gray. Effect of Human Immunodeficiency Virus Type 1 (HIV-1) Subtype on Disease Progression in Persons from Rakai, Uganda, with Incident HIV-1 Infection. *The Journal of Infectious Diseases*, 197(5):707–713, March 2008. ISSN 0022-1899, 1537-6613. doi: 10.1086/527416. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/527416.

[118] A. Vasan, B. Renjifo, E. Hertzmark, B. Chaplin, G. Msamanga, M. Essex, W. Fawzi, and D. Hunter. Different Rates of Disease Progression of HIV Type 1 Infection in Tanzania Based on Infecting Subtype. *Clinical Infectious Diseases*, 42(6):843–852, March 2006. ISSN 1058-4838, 1537-6591. doi: 10.1086/499952. URL http://cid.oxfordjournals.org/lookup/doi/10.1086/499952.

[119] Philippa J Easterbrook, Mel Smith, Jane Mullen, Siobhan O'Shea, Ian Chrystie, Annemiek de Ruiter, Iain D Tatt, Anna Geretti, and Mark Zuckerman. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. *Journal of the International AIDS Society*, 13(1):4, 2010. ISSN 1758-2652. doi: 10.1186/1758-2652-13-4. URL http://archive.biomedcentral.com/1758-2652/13/4.

[120] Joris Hemelaar. Implications of HIV diversity for the HIV-1 pandemic. *Journal of Infection*, 66(5):391–400, May 2013. ISSN 01634453. doi: 10.1016/j.jinf.2012.10.026. URL http://linkinghub.elsevier.com/retrieve/pii/S016344531200312X.

[121] Thumbi Ndung'u and Robin A. Weiss. On HIV diversity:. *AIDS*, 26(10):1255–1260, June 2012. ISSN 0269-9370. doi: 10.1097/QAD.0b013e32835461b5. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201206190-00014.

[122] Eric G. Sandström and Joan C. Kaplan. Antiviral Therapy in AIDS: Clinical Pharmacological Properties and Therapeutic Experience to Date. *Drugs*, 34(3):372–390, September 1987. ISSN 0012-6667. doi: 10.2165/00003495-198734030-00004. URL http://link.springer.com/10.2165/00003495-198734030-00004.

[123] B. Oberg. Antiviral therapy. *Journal of Acquired Immune Deficiency Syndromes*, 1(3):257–266, 1988. ISSN 0894-9255.

[124] Paul A. Volberding, Stephen W. Lagakos, Matthew A. Koch, Carla Pettinelli, Maureen W. Myers, David K. Booth, Henry H. Balfour, Richard C. Reichman, John A. Bartlett, Martin S. Hirsch, Robert L. Murphy, W. David Hardy, Ruy Soeiro, Margaret A. Fischl, John G. Bartlett, Thomas C. Merigan, Newton E. Hyslop, Douglas D. Richman, Fred T. Valentine, Lawrence Corey, and the AIDS Clinical Trials Group of the National Institute of Allergy and Infectious Diseases*. Zidovudine in Asymptomatic Human Immunodeficiency Virus Infection: A Controlled Trial in Persons with Fewer Than 500 CD4-Positive Cells per Cubic Millimeter. *New England Journal of Medicine*, 322(14):941–949, April 1990. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199004053221401. URL http://www.nejm.org/doi/abs/10.1056/NEJM199004053221401.

[125] Margaret A. Fischl. The Safety and Efficacy of Zidovudine (AZT) in the Treatment of Subjects with Mildly Symptomatic Human Immunodeficiency Virus Type 1 (HIV) Infection: A Double-Blind, Placebo-Controlled Trial. *Annals of Internal Medicine*, 112(10):727, May 1990. ISSN 0003-4819. doi: 10.7326/0003-4819-112-10-727. URL http://annals.org/article.aspx?doi=10.7326/0003-4819-112-10-727.

[126] Jens D. Lundgren. Comparison of Long-term Prognosis of Patients With AIDS Treated and Not Treated With Zidovudine. *JAMA: The Journal of the American Medical Association*, 271(14):1088, April 1994. ISSN 0098-7484. doi: 10.1001/jama.1994.03510380044035. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1994.03510380044035.

[127] Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. Concorde Coordinating Committee. *Lancet (London, England)*, 343(8902):871–881, April 1994. ISSN 0140-6736.

[128] Paul A. Volberding, Stephen W. Lagakos, Janet M. Grimes, Daniel S. Stein, James Rooney, Tze-Chiang Meng, Margaret A. Fischl, Ann C. Collier, John P. Phair, Martin S. Hirsch, W. David Hardy, Henry H. Balfour, and Richard C. Reichman. A Comparison of Immediate with Deferred Zidovudine Therapy for Asymptomatic HIV-Infected Adults with CD4 Cell Counts of 500 or More per Cubic Millimeter. *New England Journal of Medicine*, 333(7):401–407, August 1995. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199508173330701. URL http://www.nejm.org/doi/abs/10.1056/NEJM199508173330701.

[129] State-of-the-art conference on azidothymidine therapy for early HIV infection. *The American Journal of Medicine*, 89(3):335–344, September 1990. ISSN 0002-9343.

[130] B. A. Larder, G. Darby, and D. D. Richman. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science (New York, N.Y.)*, 243(4899):1731–1734, March 1989. ISSN 0036-8075.

[131] B. A. Larder and S. D. Kemp. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science (New York, N.Y.)*, 246(4934):1155–1158, December 1989. ISSN 0036-8075.

[132] Margaret A. Fischl. Zalcitabine Compared with Zidovudine in Patients with Advanced HIV-1 Infection Who Received Previous Zidovudine Therapy. *Annals of Internal Medicine*, 118(10):762, May 1993. ISSN 0003-4819. doi: 10.7326/0003-4819-118-10-199305150-00002. URL http://annals.org/article.aspx?doi=10.7326/0003-4819-118-10-199305150-00002.

[133] Michelle I. Wilde and Heather D. Langtry. Zidovudine: An Update of its Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Efficacy. *Drugs*, 46(3):515–578, September 1993. ISSN 0012-6667. doi: 10.2165/00003495-199346030-00010. URL http://link.springer.com/10.2165/00003495-199346030-00010.

[134] Julie C. Adkins, David H. Peters, and Diana Faulds. Zalcitabine: An Update of Its Pharmacodynamic and Pharmacokinetic Properties and Clinical Efficacy in the Management of HIV Infection. *Drugs*, 53 (6):1054–1080, June 1997. ISSN 0012-6667. doi: 10.2165/00003495-199753060-00009. URL http://link.springer.com/10.2165/00003495-199753060-00009.

[135] Caroline M. Perry and Stuart Noble. Didanosine: An Updated Review of its Use in HIV Infection. *Drugs*, 58(6):1099–1135, 1999. ISSN 0012-6667. doi: 10.2165/00003495-199958060-00009. URL http://link.springer.com/10.2165/00003495-199958060-00009.

[136] Asa M. Margolis, Harry Heverling, Paul A. Pham, and Andrew Stolbach. A Review of the Toxicity of HIV Medications. *Journal of Medical Toxicology*, 10(1):26–39, March 2014. ISSN 1556-9039, 1937-6995. doi: 10.1007/s13181-013-0325-8. URL http://link.springer.com/10.1007/s13181-013-0325-8.

[137] J. S. Herman and P. J. Easterbrook. The metabolic toxicities of antiretroviral therapy. *International journal of STD & AIDS*, 12(9):555–562; quiz 563–564, September 2001. ISSN 0956-4624.

[138] Gail Skowron. Alternating and Intermittent Regimens of Zidovudine and Dideoxycytidine in Patients with AIDS or AIDS-Related Complex. *Annals of Internal Medicine*, 118(5):321, March 1993. ISSN 0003-4819. doi: 10.7326/0003-4819-118-5-199303010-00001. URL http://annals.org/article.aspx?doi=10.7326/0003-4819-118-5-199303010-00001.

[139] R Yarchoan. PHASE I STUDIES OF 2',3'-DIDEOXYCYTIDINE IN SEVERE HUMAN IMMUNODEFICIENCY VIRUS INFECTION AS A SINGLE AGENT AND ALTERNATING WITH ZIDOVUDINE (AZT). *The Lancet*, 331(8577):76–81, January 1988. ISSN 01406736. doi: 10.1016/S0140-6736(88)90283-8. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673688902838.

[140] V.S. Kitchen, C. Skinner, K. Ariyoshi, J.N. Weber, A.J. Pinching, E.A. Lane, I.B. Duncan, J. Burckhardt, H.U. Burger, and K. Bragman. Safety and activity of saquinavir in HIV infection. *The Lancet*, 345 (8955):952–955, April 1995. ISSN 01406736. doi: 10.1016/S0140-6736(95)90699-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673695906991.

[141] H. Jacobsen, M. Haenggi, M. Ott, I.B. Duncan, M. Andreoni, S. Vella, and J. Mous. Reduced sensitivity to saquinavir: an update on genotyping from phase I/II trials. *Antiviral Research*, 29(1):95–97, January

1996. ISSN 01663542. doi: 10.1016/0166-3542(95)00927-2. URL http://linkinghub.elsevier.com/retrieve/pii/0166354295009272.

[142] Jh Darbyshire. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *The Lancet*, 348(9023):283–291, August 1996. ISSN 01406736. doi: 10.1016/S0140-6736(96)05387-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673696053871.

[143] Scott M. Hammer, David A. Katzenstein, Michael D. Hughes, Holly Gundacker, Robert T. Schooley, Richard H. Haubrich, W. Keith Henry, Michael M. Lederman, John P. Phair, Manette Niu, Martin S. Hirsch, and Thomas C. Merigan. A Trial Comparing Nucleoside Monotherapy with Combination Therapy in HIV-Infected Adults with CD4 Cell Counts from 200 to 500 per Cubic Millimeter. *New England Journal of Medicine*, 335(15):1081–1090, October 1996. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199610103351501. URL http://www.nejm.org/doi/abs/10.1056/NEJM199610103351501.

[144] M. Tisdale, S. D. Kemp, N. R. Parry, and B. A. Larder. Rapid in vitro selection of human immunodeficiency virus type 1 resistant to 3'-thiacytidine inhibitors due to a mutation in the YMDD region of reverse transcriptase. *Proceedings of the National Academy of Sciences of the United States of America*, 90(12):5653–5656, June 1993. ISSN 0027-8424.

[145] Joseph J. Eron, Sharon L. Benoit, Joseph Jemsek, Rodger D. MacArthur, Jorge Santana, Joseph B. Quinn, Daniel R. Kuritzkes, Mary Ann Fallon, and Marc Rubin. Treatment with Lamivudine, Zidovudine, or Both in HIV-Positive Patients with 200 to 500 CD4+ Cells per Cubic Millimeter. *New England Journal of Medicine*, 333(25):1662–1669, December 1995. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199512213332502. URL http://www.nejm.org/doi/abs/10.1056/NEJM199512213332502.

[146] D. R. Kuritzkes, J. B. Quinn, S. L. Benoit, D. L. Shugarts, A. Griffin, M. Bakhtiari, D. Poticha, J. J. Eron, M. A. Fallon, and M. Rubin. Drug resistance and virologic response in NUCA 3001, a randomized trial of lamivudine (3tc) versus zidovudine (ZDV) versus ZDV plus 3tc in previously untreated patients. *AIDS (London, England)*, 10(9):975–981, August 1996. ISSN 0269-9370.

[147] C. Katlama, D. Ingrand, C. Loveday, N. Clumeck, J. Mallolas, S. Staszewski, M. Johnson, A. M. Hill, G. Pearce, and H. McDade. Safety and efficacy of lamivudine-zidovudine combination therapy in antiretroviral-naive patients. A randomized controlled comparison with zidovudine monotherapy. Lamivudine European HIV Working Group. *JAMA*, 276(2):118–125, July 1996. ISSN 0098-7484.

[148] J. A. Bartlett, S. L. Benoit, V. A. Johnson, J. B. Quinn, G. E. Sepulveda, W. C. Ehmann, C. Tsoukas, M. A. Fallon, P. L. Self, and M. Rubin. Lamivudine plus zidovudine compared with zalcitabine plus zidovudine in patients with HIV infection. A randomized, double-blind, placebo-controlled trial. North American HIV Working Party. *Annals of Internal Medicine*, 125(3):161–172, August 1996. ISSN 0003-4819.

[149] A. N. Phillips, J. Eron, J. Bartlett, D. R. Kuritzkes, V. A. Johnson, C. Gilbert, J. Johnson, A. Keller, and A. M. Hill. Correspondence between the effect of zidovudine plus lamivudine on plasma HIV level/CD4 lymphocyte count and the incidence of clinical disease in infected individuals. North American Lamivudine HIV Working Group. *AIDS (London, England)*, 11(2):169–175, February 1997. ISSN 0269-9370.

[150] Ann C. Collier, Robert W. Coombs, David A. Schoenfeld, Roland L. Bassett, Joseph Timpone, Alice Baruch, Michelle Jones, Karen Facey, Caroline Whitacre, Vincent J. McAuliffe, Harvey M. Friedman, Thomas C. Merigan, Richard C. Reichman, Carol Hooper, and Lawrence Corey. Treatment of Human Immunodeficiency Virus Infection with Saquinavir, Zidovudine, and Zalcitabine. *New England Journal of Medicine*, 334(16):1011–1018, April 1996. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199604183341602. URL http://www.nejm.org/doi/abs/10.1056/NEJM199604183341602.

[151] Jonathan M. Schapiro. The Effect of High-Dose Saquinavir on Viral Load and CD4+ T-Cell Counts in HIV-Infected Patients. *Annals of Internal Medicine*, 124(12):1039, June 1996. ISSN 0003-4819. doi: 10.7326/0003-4819-124-12-199606150-00003. URL http://annals.org/article.aspx?doi=10.7326/0003-4819-124-12-199606150-00003.

[152] S. H. Cheeseman, D. Havlir, M. M. McLaughlin, T. C. Greenough, J. L. Sullivan, D. Hall, S. E. Hattox, S. A. Spector, D. S. Stein, and M. Myers. Phase I/II evaluation of nevirapine alone and in combination with zidovudine for infection with human immunodeficiency virus. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology: Official Publication of the International Retrovirology Association*, 8(2):141–151, February 1995. ISSN 1077-9450.

[153] D. Havlir, M. M. McLaughlin, and D. D. Richman. A Pilot Study to Evaluate the Development of Resistance to Nevirapine in Asymptomatic Human Immunodeficiency Virus-Infected Patients with CD4 Cell Counts of >500/mm3 : AIDS Clinical Trials Group Protocol 208. *Journal of Infectious Diseases*, 172(5):1379–1383, November 1995. ISSN 0022-1899, 1537-6613. doi: 10.1093/infdis/172.5.1379. URL http://jid.oxfordjournals.org/lookup/doi/10.1093/infdis/172.5.1379.

[154] Richard T. D'Aquila. Nevirapine, Zidovudine, and Didanosine Compared with Zidovudine and Didanosine in Patients with HIV-1 Infection: A Randomized, Double-Blind, Placebo-Controlled Trial. *Annals of Internal Medicine*, 124(12):1019, June 1996. ISSN 0003-4819. doi: 10.7326/0003-4819-124-12-199606150-00001. URL http://annals.org/article.aspx?doi=10.7326/0003-4819-124-12-199606150-00001.

[155] Julio S. G. Montaner, Peter Reiss, David Cooper, Stefano Vella, Marianne Harris, Brian Conway, Mark A. Wainberg, D. Smith, Patrick Robinson, David Hall, Maureen Myers, Joep M. A. Lange, and for the INCAS Study Group. A Randomized, Double-blind Trial Comparing Combinations of Nevirapine, Didanosine, and Zidovudine for HIV-Infected Patients: The INCAS Trial. *JAMA*, 279(12):930, March 1998. ISSN 0098-7484. doi: 10.1001/jama.279.12.930. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.279.12.930.

[156] Sven A. Danner, Andrew Carr, John M. Leonard, Leah M. Lehman, Francesc Gudiol, Juan Gonzales, Antonio Raventos, Rafael Rubio, Emilio Bouza, Vicente Pintado, Antonio Gil Aguado, Juan Garcia de Lomas, Rafael Delgado, Jan C.C. Borleffs, Ann Hsu, Joaquin M. Valdes, Charles A.B. Boucher, and David A. Cooper. A Short-Term Study of the Safety, Pharmacokinetics, and Efficacy of Ritonavir, an Inhibitor of HIV-1 Protease. *New England Journal of Medicine*, 333(23):1528–1534, December 1995. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199512073332303. URL http://www.nejm.org/doi/abs/10.1056/NEJM199512073332303.

[157] Martin Markowitz, Michael Saag, William G. Powderly, Arlene M. Hurley, Ann Hsu, Joaquin M. Valdes, David Henry, Fred Sattler, Anthony La Marca, John M. Leonard, and David D. Ho. A Preliminary Study of Ritonavir, an Inhibitor of HIV-1 Protease, to Treat HIV-1 Infection. *New England Journal of Medicine*, 333(23):1534–1540, December 1995. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199512073332204. URL http://www.nejm.org/doi/abs/10.1056/NEJM199512073332204.

[158] D. Mathez, P. Bagnarelli, I. Gorin, C. Katlama, G. Pialoux, G. Saimot, P. Tubiana, P. De Truchis, J. P. Chauvin, R. Mills, R. Rode, M. Clementi, and J. Leibowitch. Reductions in viral load and increases in T lymphocyte numbers in treatment-naive patients with advanced HIV-1 infection treated with ritonavir, zidovudine and zalcitabine triple therapy. *Antiviral Therapy*, 2(3):175–183, July 1997. ISSN 1359-6535.

[159] D. W. Notermans, S. Jurriaans, F. de Wolf, N. A. Foudraine, J. J. de Jong, W. Cavert, C. M. Schuwirth, R. H. Kauffmann, P. L. Meenhorst, H. McDade, C. Goodwin, J. M. Leonard, J. Goudsmit, and S. A. Danner. Decrease of HIV-1 RNA levels in lymphoid tissue and peripheral blood during treatment with ritonavir, lamivudine and zidovudine. Ritonavir/3tc/ZDV Study Group. *AIDS (London, England)*, 12 (2):167–173, January 1998. ISSN 0269-9370.

[160] Roy M. Gulick, John W. Mellors, Diane Havlir, Joseph J. Eron, Charles Gonzalez, Deborah McMahon, Douglas D. Richman, Fred T. Valentine, Leslie Jonas, Anne Meibohm, Emilio A. Emini, Jeffrey A. Chodakewitz, Paul Deutsch, Daniel Holder, William A. Schleif, and Jon H. Condra. Treatment with Indinavir, Zidovudine, and Lamivudine in Adults with Human Immunodeficiency Virus Infection and Prior Antiretroviral Therapy. *New England Journal of Medicine*, 337(11):734–739, September 1997. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199709113371102. URL http://www.nejm.org/doi/abs/10.1056/NEJM199709113371102.

[161] Roy M. Gulick, John W. Mellors, Diane Havlir, Joseph J. Eron, Charles Gonzalez, Deborah McMahon, Leslie Jonas, Anne Meibohm, Daniel Holder, William A. Schleif, Jon H. Condra, Emilio A. Emini, Robin Isaacs, Jeffrey A. Chodakewitz, and Douglas D. Richman. Simultaneous vs Sequential Initiation of Therapy With Indinavir, Zidovudine, and Lamivudine for HIV-1 Infection: 100-Week Follow-up. *JAMA*, 280(1):35, July 1998. ISSN 0098-7484. doi: 10.1001/jama.280.1.35. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.280.1.35.

[162] Greg L. Plosker and Stuart Noble. Indinavir: A Review of its Use in the Management of HIV Infection. *Drugs*, 58(6):1165–1203, 1999. ISSN 0012-6667. doi: 10.2165/00003495-199958060-00011. URL http://link.springer.com/10.2165/00003495-199958060-00011.

[163] James S. Lewis, Colleen M. Terriff, Daniel R. Coulston, and Mark W. Garrison. Protease inhibitors: a therapeutic breakthrough for the treatment of patients with human immunodeficiency virus. *Clinical Therapeutics*, 19(2):187–214, March 1997. ISSN 01492918. doi: 10.1016/S0149-2918(97)80110-5. URL http://linkinghub.elsevier.com/retrieve/pii/S0149291897801105.

[164] D. J. Kempf, K. C. Marsh, G. Kumar, A. D. Rodrigues, J. F. Denissen, E. McDonald, M. J. Kukulka, A. Hsu, G. R. Granneman, P. A. Baroldi, E. Sun, D. Pizzuti, J. J. Plattner, D. W. Norbeck, and J. M. Leonard. Pharmacokinetic enhancement of inhibitors of the human immunodeficiency virus protease by coadministration with ritonavir. *Antimicrobial Agents and Chemotherapy*, 41(3):654–660, March 1997. ISSN 0066-4804.

[165] David D. Ho, Avidan U. Neumann, Alan S. Perelson, Wen Chen, John M. Leonard, and Martin Markowitz. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, 373 (6510):123–126, January 1995. ISSN 0028-0836. doi: 10.1038/373123a0. URL http://www.nature.com/doifinder/10.1038/373123a0.

[166] J. L. Martin, C. E. Brown, N. Matthews-Davis, and J. E. Reardon. Effects of antiviral nucleoside analogs on human DNA polymerases and mitochondrial DNA synthesis. *Antimicrobial Agents and Chemotherapy*, 38(12):2743–2749, December 1994. ISSN 0066-4804.

[167] S. E. Lim and W. C. Copeland. Differential Incorporation and Removal of Antiviral Deoxynucleotides by Human DNA Polymerase. *Journal of Biological Chemistry*, 276(26):23616–23623, June 2001. ISSN 0021-9258, 1083-351X. doi: 10.1074/jbc.M101114200. URL http://www.jbc.org/cgi/doi/10.1074/jbc.M101114200.

[168] Hélène C. F. Côté, Zabrina L. Brumme, Kevin J. P. Craib, Christopher S. Alexander, Brian Wynhoven, Lillian Ting, Hubert Wong, Marianne Harris, P. Richard Harrigan, Michael V. O'Shaughnessy, and Julio S. G. Montaner. Changes in mitochondrial DNA as a marker of nucleoside toxicity in HIV-infected patients. *The New England Journal of Medicine*, 346(11):811–820, March 2002. ISSN 1533-4406. doi: 10.1056/NEJMoa012035.

[169] William Lewis, Brian J. Day, and William C. Copeland. Mitochondrial toxicity of nrti antiviral drugs: an integrated cellular perspective. *Nature Reviews Drug Discovery*, 2(10):812–822, October 2003. ISSN 1474-1776, 1474-1784. doi: 10.1038/nrd1201. URL http://www.nature.com/doifinder/10.1038/nrd1201.

[170] M. G. Zaera, O. Miró, E. Pedrol, A. Soler, M. Picón, F. Cardellach, J. Casademont, and V. Nunes. Mitochondrial involvement in antiretroviral therapy-related lipodystrophy. *AIDS (London, England)*, 15 (13):1643–1651, September 2001. ISSN 0269-9370.

[171] C. M. Shikuma, N. Hu, C. Milne, F. Yost, C. Waslien, S. Shimizu, and B. Shiramizu. Mitochondrial DNA decrease in subcutaneous adipose tissue of HIV-infected individuals with peripheral lipoatrophy. *AIDS (London, England)*, 15(14):1801–1809, September 2001. ISSN 0269-9370.

[172] Andrea Cossarizza and Graeme Moyle. Antiretroviral nucleoside and nucleotide analogues and mitochondria. *AIDS (London, England)*, 18(2):137–151, January 2004. ISSN 0269-9370.

[173] K. Brinkman, H. J. ter Hofstede, D. M. Burger, J. A. Smeitink, and P. P. Koopmans. Adverse effects of reverse transcriptase inhibitors: mitochondrial toxicity as common pathway. *AIDS (London, England)*, 12(14):1735–1744, October 1998. ISSN 0269-9370.

[174] A. J White. Mitochondrial toxicity and HIV therapy. *Sexually Transmitted Infections*, 77(3):158–173, June 2001. ISSN 13684973. doi: 10.1136/sti.77.3.158. URL http://sti.bmj.com/cgi/doi/10.1136/sti.77.3.158.

[175] Paul E Sax, David Wohl, Michael T Yin, Frank Post, Edwin DeJesus, Michael Saag, Anton Pozniak, Melanie Thompson, Daniel Podzamczer, Jean Michel Molina, Shinichi Oka, Ellen Koenig, Benoit Trottier, Jaime Andrade-Villanueva, Gordon Crofoot, Joseph M Custodio, Andrew Plummer, Lijie Zhong, Huyen Cao, Hal Martin, Christian Callebaut, Andrew K Cheng, Marshall W Fordyce, and Scott McCallister. Tenofovir alafenamide versus tenofovir disoproxil fumarate, coformulated with elvitegravir, cobicistat, and emtricitabine, for initial treatment of HIV-1 infection: two randomised, double-blind, phase 3, non-inferiority trials. *The Lancet*, 385(9987):2606–2615, June 2015. ISSN 01406736. doi: 10.1016/S0140-6736(15)60616-X. URL http://linkinghub.elsevier.com/retrieve/pii/S014067361560616X.

[176] Darren Wong and Robert Grossberg. Tenofovir alafenamide: an effective option for HIV treatment with reduced risk. *Future Virology*, 10(9):1069–1075, September 2015. ISSN 1746-0794, 1746-0808. doi: 10.2217/fvl.15.72. URL http://www.futuremedicine.com/doi/10.2217/fvl.15.72.

[177] A. Carr, K. Samaras, S. Burton, M. Law, J. Freund, D. J. Chisholm, and D. A. Cooper. A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors. *AIDS (London, England)*, 12(7):F51–58, May 1998. ISSN 0269-9370.

[178] M. S. Hirsch. Chemotherapy of human immunodeficiency virus infections: current practice and future prospects. *The Journal of Infectious Diseases*, 161(5):845–857, May 1990. ISSN 0022-1899.

[179] C??dric Arvieux and Olivier Tribut. Amprenavir or Fosamprenavir plus Ritonavir in HIV Infection: Pharmacology, Efficacy and Tolerability Profile. *Drugs*, 65(5):633–659, 2005. ISSN 0012-6667. doi: 10.2165/00003495-200565050-00005. URL http://link.springer.com/10.2165/00003495-200565050-00005.

[180] Stuart Noble and Karen L. Goa. Amprenavir: A Review of its Clinical Potential in Patients with HIV Infection. *Drugs*, 60(6):1383–1410, December 2000. ISSN 0012-6667. doi: 10.2165/00003495-200060060-00012. URL http://link.springer.com/10.2165/00003495-200060060-00012.

[181] E. Jennifer Edelman, Kirsha S. Gordon, Janis Glover, Ian R. McNicholl, David A. Fiellin, and Amy C. Justice. The Next Therapeutic Challenge in HIV: Polypharmacy. *Drugs & Aging*, 30(8):613–628, August 2013. ISSN 1170-229X, 1179-1969. doi: 10.1007/s40266-013-0093-9. URL http://link.springer.com/10.1007/s40266-013-0093-9.

[182] Michelle D. Furler, Thomas R. Einarson, Sharon Walmsley, Margaret Millson, and Reina Bendayan. Polypharmacy in HIV: Impact of Data Source and Gender on Reported Drug Utilization. *AIDS Patient*

Care and STDs, 18(10):568–586, October 2004. ISSN 1087-2914, 1557-7449. doi: 10.1089/apc.2004.18. 568. URL http://www.liebertonline.com/doi/abs/10.1089/apc.2004.18.568.

[183] Benjamin Young. *Review:* Mixing New Cocktails: Drug Interactions in Antiretroviral Regimens. *AIDS Patient Care and STDs*, 19(5):286–297, May 2005. ISSN 1087-2914, 1557-7449. doi: 10.1089/apc.2005. 19.286. URL http://www.liebertonline.com/doi/abs/10.1089/apc.2005.19.286.

[184] Hartmut B Krentz, Ian Cosman, Kathy Lee, Jinell Mah Ming, and M John Gill. Pill burden in HIV infection: 20 years of experience. *Antiviral Therapy*, 17(5):833–840, 2012. ISSN 13596535. doi: 10.3851/IMP2076. URL http://www.intmedpress.com/journals/avt/abstract.cfm?id=2076&pid=48.

[185] Annemarie M J Wensing, Noortje M van Maarseveen, and Monique Nijhuis. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. *Antiviral research*, 85(1):59–74, January 2010. ISSN 1872-9096. doi: 10.1016/j.antiviral.2009.10.003. URL http://www.ncbi.nlm.nih.gov/pubmed/19853627.

[186] D. Warnke, J. Barreto, and Z. Temesgen. Antiretroviral Drugs. *The Journal of Clinical Pharmacology*, 47(12):1570–1579, October 2007. ISSN 0091-2700. doi: 10.1177/0091270007308034. URL http://doi.wiley.com/10.1177/0091270007308034.

[187] Lianhong Xu, Hongtao Liu, Bernard P. Murray, Christian Callebaut, Melody S. Lee, Allen Hong, Robert G. Strickley, Luong K. Tsai, Kirsten M. Stray, Yujin Wang, Gerry R. Rhodes, and Manoj C. Desai. Cobicistat (GS-9350): A Potent and Selective Inhibitor of Human CYP3a as a Novel Pharmacoenhancer. *ACS Medicinal Chemistry Letters*, 1(5):209–213, August 2010. ISSN 1948-5875, 1948-5875. doi: 10.1021/ml1000257. URL http://pubs.acs.org/doi/abs/10.1021/ml1000257.

[188] Brian Conway. The Role of Adherence to Antiretroviral Therapy in the Management of HIV Infection:. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 45(Supplement 1):S14–S18, June 2007. ISSN 1525-4135. doi: 10.1097/QAI.0b013e3180600766. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00126334-200706011-00004.

[189] Sripal Bangalore, Gayathri Kamalakkannan, Sanobar Parkar, and Franz H. Messerli. Fixed-Dose Combinations Improve Medication Compliance: A Meta-Analysis. *The American Journal of Medicine*, 120(8):713–719, August 2007. ISSN 00029343. doi: 10.1016/j.amjmed.2006.08.033. URL http://linkinghub.elsevier.com/retrieve/pii/S000293430601151X.

[190] Vicki Oldfield, Gillian M Keating, and Greg Plosker. Enfuvirtide: A Review of its Use in the Management of HIV Infection: A Review of its Use in the Management of HIV Infection. *Drugs*, 65 (8):1139–1160, 2005. ISSN 0012-6667. doi: 10.2165/00003495-200565080-00007. URL http://link.springer.com/10.2165/00003495-200565080-00007.

[191] Caroline M. Perry. Maraviroc: A Review of its Use in the Management of CCR5-Tropic HIV-1 Infection. *Drugs*, 70(9):1189–1213, June 2010. ISSN 0012-6667. doi: 10.2165/11203940-000000000-00000. URL http://link.springer.com/10.2165/11203940-000000000-00000.

[192] Roy M. Gulick, Jacob Lalezari, James Goodrich, Nathan Clumeck, Edwin DeJesus, Andrzej Horban, Jeffrey Nadler, Bonaventura Clotet, Anders Karlsson, Michael Wohlfeiler, John B. Montana, Mary McHale, John Sullivan, Caroline Ridgway, Steve Felstead, Michael W. Dunne, Elna van der Ryst, and Howard Mayer. Maraviroc for Previously Treated Patients with R5 HIV-1 Infection. *New England Journal of Medicine*, 359(14):1429–1441, October 2008. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa0803152. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa0803152.

[193] David A. Cooper, Jayvant Heera, James Goodrich, Margaret Tawadrous, Michael Saag, Edwin DeJesus, Nathan Clumeck, Sharon Walmsley, Naitee Ting, Eoin Coakley, Jacqueline D. Reeves, Gustavo Reyes-Teran, Mike Westby, Elna Van Der Ryst, Prudence Ive, Lerato Mohapi, Horacio Mingrone, Andrzej Horban, Frances Hackman, John Sullivan, and Howard Mayer. Maraviroc versus Efavirenz, Both in Combination with Zidovudine-Lamivudine, for the Treatment of Antiretroviral-Naive Subjects with CCR5-Tropic HIV-1 Infection. *The Journal of Infectious Diseases*, 201(6):803–813, March 2010. ISSN 0022-1899, 1537-6613. doi: 10.1086/650697. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/650697.

[194] Hans-Jürgen Stellbrink, Eric Le Fevre, Andrew Carr, Michael S. Saag, Geoffrey Mukwaya, Silvia Nozza, Srinivas Rao Valluri, Manoli Vourvahis, Alex R. Rinehart, Lynn McFadyen, Carl Fichtenbaum, Andrew Clark, Charles Craig, Annie F. Fang, and Jayvant Heera. Once-daily maraviroc versus tenofovir/emtricitabine each combined with darunavir/ritonavir for initial HIV-1 treatment:. *AIDS*, 30(8):1229–1238, May 2016. ISSN 0269-9370. doi: 10.1097/QAD.0000000000001058. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201605150-00008.

[195] Jamie D. Croxtall and Susan J. Keam. Raltegravir: A Review of its Use in the Management of HIV Infection in Treatment-Experienced Patients. *Drugs*, 69(8):1059–1075, June 2009. ISSN 0012-6667. doi: 10.2165/00003495-200969080-00007. URL http://link.springer.com/10.2165/00003495-200969080-00007.

[196] Jamie D. Croxtall and Lesley J. Scott. Raltegravir: In Treatment-Naive Patients with HIV-1 Infection. *Drugs*, 70(5):631–642, March 2010. ISSN 0012-6667. doi: 10.2165/11204590-000000000-00000. URL http://link.springer.com/10.2165/11204590-000000000-00000.

[197] Jeffrey L Lennox, Edwin DeJesus, Adriano Lazzarin, Richard B Pollard, Jose Valdez Ramalho Madruga, Daniel S Berger, Jing Zhao, Xia Xu, Angela Williams-Diaz, Anthony J Rodgers, Richard JO Barnard, Michael D Miller, Mark J DiNubile, Bach-Yen Nguyen, Randi Leavitt, and Peter Sklar. Safety and efficacy of raltegravir-based versus efavirenz-based combination therapy in treatment-naive patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. *The Lancet*, 374(9692):796–806, September 2009. ISSN 01406736. doi: 10.1016/S0140-6736(09)60918-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673609609181.

[198] Roy T. Steigbigel, David A. Cooper, Princy N. Kumar, Joseph E. Eron, Mauro Schechter, Martin Markowitz, Mona R. Loutfy, Jeffrey L. Lennox, Jose M. Gatell, Jurgen K. Rockstroh, Christine Katlama, Patrick Yeni, Adriano Lazzarin, Bonaventura Clotet, Jing Zhao, Joshua Chen, Desmond M. Ryan,

Rand R. Rhodes, John A. Killar, Lucinda R. Gilde, Kim M. Strohmaier, Anne R. Meibohm, Michael D. Miller, Daria J. Hazuda, Michael L. Nessly, Mark J. DiNubile, Robin D. Isaacs, Bach-Yen Nguyen, and Hedy Teppler. Raltegravir with Optimized Background Therapy for Resistant HIV-1 Infection. *New England Journal of Medicine*, 359(4):339–354, July 2008. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa0708975. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa0708975.

[199] S. G. Deeks. Determinants of virological response to antiretroviral therapy: implications for long-term strategies. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 30 Suppl 2:S177–184, June 2000. ISSN 1058-4838. doi: 10.1086/313855.

[200] Caroline M. Perry. Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate Single-Tablet Regimen (Stribild®): A Review of Its Use in the Management of HIV-1 Infection in Adults. *Drugs*, 74 (1):75–97, January 2014. ISSN 0012-6667, 1179-1950. doi: 10.1007/s40265-013-0158-4. URL http://link.springer.com/10.1007/s40265-013-0158-4.

[201] Jean-Michel Molina, Anthony LaMarca, Jaime Andrade-Villanueva, Bonaventura Clotet, Nathan Clumeck, Ya-Pei Liu, Lijie Zhong, Nicolas Margot, Andrew K Cheng, and Steven L Chuck. Efficacy and safety of once daily elvitegravir versus twice daily raltegravir in treatment-experienced patients with HIV-1 receiving a ritonavir-boosted protease inhibitor: randomised, double-blind, phase 3, non-inferiority study. *The Lancet Infectious Diseases*, 12(1):27–35, January 2012. ISSN 14733099. doi: 10.1016/S1473-3099(11)70249-3. URL http://linkinghub.elsevier.com/retrieve/pii/S1473309911702493.

[202] Paul L. McCormack. Dolutegravir: A Review of Its Use in the Management of HIV-1 Infection in Adolescents and Adults. *Drugs*, 74(11):1241–1252, July 2014. ISSN 0012-6667, 1179-1950. doi: 10.1007/s40265-014-0256-y. URL http://link.springer.com/10.1007/s40265-014-0256-y.

[203] Sarah L. Greig and Emma D. Deeks. Abacavir/Dolutegravir/Lamivudine Single-Tablet Regimen: A Review of Its Use in HIV-1 Infection. *Drugs*, 75(5):503–514, April 2015. ISSN 0012-6667, 1179-1950. doi: 10.1007/s40265-015-0361-6. URL http://link.springer.com/10.1007/s40265-015-0361-6.

[204] Mark A. Wainberg and Ying-Shan Han. Will drug resistance against dolutegravir in initial therapy ever occur? *Frontiers in Pharmacology*, 6, April 2015. ISSN 1663-9812. doi: 10.3389/fphar.2015.00090. URL http://journal.frontiersin.org/article/10.3389/fphar.2015.00090/abstract.

[205] S. Hare, S. J. Smith, M. Metifiot, A. Jaxa-Chamiec, Y. Pommier, S. H. Hughes, and P. Cherepanov. Structural and Functional Analyses of the Second-Generation Integrase Strand Transfer Inhibitor Dolutegravir (S/GSK1349572). *Molecular Pharmacology*, 80(4):565–572, October 2011. ISSN 0026-895X. doi: 10.1124/mol.111.073189. URL http://molpharm.aspetjournals.org/cgi/doi/10.1124/mol.111.073189.

[206] Thibault Mesplède and Mark Wainberg. Is Resistance to Dolutegravir Possible When This Drug Is Used in First-Line Therapy? *Viruses*, 6(9):3377–3385, August 2014. ISSN 1999-4915. doi: 10.3390/v6093377. URL http://www.mdpi.com/1999-4915/6/9/3377/.

[207] Pedro Cahn, Anton L Pozniak, Horacio Mingrone, Andrey Shuldyakov, Carlos Brites, Jaime F Andrade-Villanueva, Gary Richmond, Carlos Beltran Buendia, Jan Fourie, Moti Ramgopal, Debbie Hagins, Franco Felizarta, Jose Madruga, Tania Reuter, Tamara Newman, Catherine B Small, John Lombaard, Beatriz Grinsztejn, David Dorey, Mark Underwood, Sandy Griffith, and Sherene Min. Dolutegravir versus raltegravir in antiretroviral-experienced, integrase-inhibitor-naive adults with HIV: week 48 results from the randomised, double-blind, non-inferiority SAILING study. *The Lancet*, 382(9893):700–708, August 2013. ISSN 01406736. doi: 10.1016/S0140-6736(13)61221-0. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673613612210.

[208] Bonaventura Clotet, Judith Feinberg, Jan van Lunzen, Marie-Aude Khuong-Josses, Andrea Antinori, Irina Dumitru, Vadim Pokrovskiy, Jan Fehr, Roberto Ortiz, Michael Saag, Julia Harris, Clare Brennan, Tamio Fujiwara, and Sherene Min. Once-daily dolutegravir versus darunavir plus ritonavir in antiretroviral-naive adults with HIV-1 infection (FLAMINGO): 48 week results from the randomised open-label phase 3b study. *The Lancet*, 383(9936):2222–2231, June 2014. ISSN 01406736. doi: 10.1016/S0140-6736(14)60084-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673614600842.

[209] Sharon L. Walmsley, Antonio Antela, Nathan Clumeck, Dan Duiculescu, Andrea Eberhard, Felix Gutiérrez, Laurent Hocqueloux, Franco Maggiolo, Uriel Sandkovsky, Catherine Granier, Keith Pappa, Brian Wynne, Sherene Min, and Garrett Nichols. Dolutegravir plus Abacavir–Lamivudine for the Treatment of HIV-1 Infection. *New England Journal of Medicine*, 369(19):1807–1818, November 2013. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa1215541. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa1215541.

[210] Dipen A. Patel, Sonya J. Snedecor, Wing Yu Tang, Lavanya Sudharshan, Jessica W. Lim, Robert Cuffe, Sonia Pulgar, Kim A. Gilchrist, Rodrigo Refoios Camejo, Jennifer Stephens, and Garrett Nichols. 48-Week Efficacy and Safety of Dolutegravir Relative to Commonly Used Third Agents in Treatment-Naive HIV-1–Infected Patients: A Systematic Review and Network Meta-Analysis. *PLoS ONE*, 9(9):e105653, September 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0105653. URL http://dx.plos.org/10.1371/journal.pone.0105653.

[211] David D. Ho. Time to Hit HIV, Early and Hard. *New England Journal of Medicine*, 333(7):450–451, August 1995. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJM199508173330710. URL http://www.nejm.org/doi/abs/10.1056/NEJM199508173330710.

[212] Charles C. J. Carpenter. Antiretroviral Therapy for HIV Infection in 1996: Recommendations of an International Panel. *JAMA*, 276(2):146, July 1996. ISSN 0098-7484. doi: 10.1001/jama.1996.03540020068031. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1996.03540020068031.

[213] Patrick G. Yeni, Scott M. Hammer, Charles C. J. Carpenter, David A. Cooper, Margaret A. Fischl, Jose M. Gatell, Brian G. Gazzard, Martin S. Hirsch, Donna M. Jacobsen, David A. Katzenstein, Julio S. G. Montaner, Douglas D. Richman, Michael S. Saag, Mauro Schechter, Robert T. Schooley, Melanie A. Thompson, Stefano Vella, and Paul A. Volberding. Antiretroviral treatment for adult HIV infection in 2002:

updated recommendations of the International AIDS Society-USA Panel. *JAMA*, 288(2):222–235, July 2002. ISSN 0098-7484.

[214] Timothy J. Wilkin and Roy M. Gulick. HIV/AIDS: When to Start Antiretroviral Therapy? *Clinical Infectious Diseases*, 47(12):1580–1586, December 2008. ISSN 1058-4838, 1537-6591. doi: 10.1086/593311. URL http://cid.oxfordjournals.org/lookup/doi/10.1086/593311.

[215] Ricardo A Franco and Michael S Saag. When to start antiretroviral therapy: as soon as possible. *BMC Medicine*, 11(1):147, 2013. ISSN 1741-7015. doi: 10.1186/1741-7015-11-147. URL http://www.biomedcentral.com/1741-7015/11/147.

[216] Tomas Cihlar and Adrian S. Ray. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. *Antiviral Research*, 85(1):39–58, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.09.014. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209004859.

[217] Stefan G. Sarafianos, Bruno Marchand, Kalyan Das, Daniel M. Himmel, Michael A. Parniak, Stephen H. Hughes, and Eddy Arnold. Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. *Journal of Molecular Biology*, 385(3):693–713, January 2009. ISSN 00222836. doi: 10.1016/j.jmb.2008.10.071. URL http://linkinghub.elsevier.com/retrieve/pii/S0022283608013442.

[218] Marie-Pierre de Béthune. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: A review of the last 20 years (1989–2009). *Antiviral Research*, 85(1):75–90, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.09.008. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209004616.

[219] Damian J. McColl and Xiaowu Chen. Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. *Antiviral Research*, 85(1):101–118, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.11.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209005336.

[220] Pinar Iyidogan and Karen Anderson. Current Perspectives on HIV-1 Antiretroviral Drug Resistance. *Viruses*, 6(10):4095–4139, October 2014. ISSN 1999-4915. doi: 10.3390/v6104095. URL http://www.mdpi.com/1999-4915/6/10/4095/.

[221] John C. Tilton and Robert W. Doms. Entry inhibitors in the treatment of HIV-1 infection. *Antiviral Research*, 85(1):91–100, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.07.022. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209003945.

[222] L. C. F. Mulder, A. Harari, and V. Simon. Cytidine deamination induced HIV-1 drug resistance. *Proceedings of the National Academy of Sciences*, 105(14):5501–5506, April 2008. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0710190105. URL http://www.pnas.org/cgi/doi/10.1073/pnas.0710190105.

[223] Sally Land, Kate Mcgavin, Chris Birch, and Ron Lucas. Reversion from zidovudine resistance to sensitivity on cessation of treatment. *The Lancet*, 338(8770):830–831, September 1991. ISSN 01406736. doi: 10.1016/0140-6736(91)90727-7. URL http://linkinghub.elsevier.com/retrieve/pii/0140673691907277.

[224] C. Delaugerre, M. A. Valantin, M. Mouroux, M. Bonmarchand, G. Carcelain, C. Duvivier, R. Tubiana, A. Simon, F. Bricaire, H. Agut, B. Autran, C. Katlama, and V. Calvez. Re-occurrence of HIV-1 drug mutations after treatment re-initiation following interruption in patients with multiple treatment failure. *AIDS (London, England)*, 15(16):2189–2191, November 2001. ISSN 0269-9370.

[225] Annemarie M Wensing, Vincent Calvez, Günthard F Huldrych, Victoria A Johnson, Roger Paredes, Deenan Pillay, Robert W Shafer, and Douglas D Richman. 2015 Update of the Drug Resistance Mutations in HIV-1. *Topics in antiviral medicine*, 23(4):132–146, November 2015. URL https://www.iasusa.org/sites/default/files/tam/october_november_2015.pdf.

[226] Luis Menéndez-Arias. Molecular basis of human immunodeficiency virus drug resistance: An update. *Antiviral Research*, 85(1):210–231, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.07.006. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209003854.

[227] Thibault Mesplède and Mark Wainberg. Resistance against Integrase Strand Transfer Inhibitors and Relevance to HIV Persistence. *Viruses*, 7(7):3703–3718, July 2015. ISSN 1999-4915. doi: 10.3390/v7072790. URL http://www.mdpi.com/1999-4915/7/7/2790/.

[228] Huldrych F. Günthard, Judith A. Aberg, Joseph J. Eron, Jennifer F. Hoy, Amalio Telenti, Constance A. Benson, David M. Burger, Pedro Cahn, Joel E. Gallant, Marshall J. Glesby, Peter Reiss, Michael S. Saag, David L. Thomas, Donna M. Jacobsen, and Paul A. Volberding. Antiretroviral Treatment of Adult HIV Infection: 2014 Recommendations of the International Antiviral Society–USA Panel. *JAMA*, 312(4):410, July 2014. ISSN 0098-7484. doi: 10.1001/jama.2014.8722. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2014.8722.

[229] European AIDS Clinical Society. *Guidelines. Version 8.0. October 2015.* European Aids Clinical Society, Brussels, Belgium, 8.0 edition, October 2015. URL http://www.eacsociety.org/files/2015_eacsguidelines_8.0-english_revised-20151104.pdf.

[230] Amit C Achhra and Mark A Boyd. Antiretroviral regimens sparing agents from the nucleoside(tide) reverse transcriptase inhibitor class: a review of the recent literature. *AIDS Research and Therapy*, 10 (1):33, 2013. ISSN 1742-6405. doi: 10.1186/1742-6405-10-33. URL http://www.aidsrestherapy.com/content/10/1/33.

[231] David A Margolis, Cynthia C Brinson, Graham H R Smith, Jerome de Vente, Debbie P Hagins, Joseph J Eron, Sandy K Griffith, Marty H St Clair, Marita C Stevens, Peter E Williams, Susan L Ford, Britt S Stancil, Melinda M Bomar, Krischan J Hudson, Kimberly Y Smith, and William R Spreen. Cabotegravir plus rilpivirine, once a day, after induction with cabotegravir plus nucleoside reverse transcriptase inhibitors in antiretroviral-naive adults with HIV-1 infection (LATTE): a randomised, phase 2b, dose-ranging trial. *The Lancet Infectious Diseases*, 15(10):1145–1155, October 2015.

ISSN 14733099. doi: 10.1016/S1473-3099(15)00152-8. URL http://linkinghub.elsevier.com/retrieve/pii/S1473309915001528.

[232] Pola de la Torre, Jomy George, and John D. Baxter. Nucleoside-Sparing Antiretroviral Regimens. *Current Infectious Disease Reports*, 16(7), July 2014. ISSN 1523-3847, 1534-3146. doi: 10.1007/s11908-014-0410-4. URL http://link.springer.com/10.1007/s11908-014-0410-4.

[233] Randall Tressler and Catherine Godfrey. NRTI Backbone in HIV Treatment: Will it Remain Relevant? *Drugs*, 72(16):2051–2062, November 2012. ISSN 0012-6667. doi: 10.2165/11640830-000000000-00000. URL http://link.springer.com/10.2165/11640830-000000000-00000.

[234] Ard van Sighem, Luuk Gras, Peter Reiss, Kees Brinkman, and Frank de Wolf. Life expectancy of recently diagnosed asymptomatic HIV-infected patients approaches that of uninfected individuals:. *AIDS*, 24(10):1527–1535, June 2010. ISSN 0269-9370. doi: 10.1097/QAD.0b013e32833a3946. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201006190-00015.

[235] Tammy M. Rickabaugh and Beth D. Jamieson. A challenge for the future: aging and HIV infection. *Immunologic Research*, 48(1-3):59–71, December 2010. ISSN 0257-277X, 1559-0755. doi: 10.1007/s12026-010-8167-9. URL http://link.springer.com/10.1007/s12026-010-8167-9.

[236] A. Calcagno, S. Nozza, C. Muss, B. M. Celesia, F. Carli, S. Piconi, G. V. De Socio, A. M. Cattelan, G. Orofino, D. Ripamonti, A. Riva, and G. Di Perri. Ageing with HIV: a multidisciplinary review. *Infection*, 43(5):509–522, October 2015. ISSN 0300-8126, 1439-0973. doi: 10.1007/s15010-015-0795-5. URL http://link.springer.com/10.1007/s15010-015-0795-5.

[237] Sean Slavin, Julian Elliott, Christopher Fairley, Martyn French, Jennifer Hoy, Matthew Law, and Sharon Lewin. HIV and aging: an overview of an emerging issue. *Sexual Health*, 8(4):449, 2011. ISSN 1448-5028. doi: 10.1071/SH11110. URL http://www.publish.csiro.au/?paper=SH11110.

[238] José R. Blanco, Ana M. Caro, Santiago Pérez-Cachafeiro, Félix Gutiérrez, José Antonio Iribarren, Juan González-García, Sara Ferrando-Martínez, Gema Navarro, and Santiago Moreno. HIV infection and aging. *AIDS reviews*, 12(4):218–230, December 2010. ISSN 1698-6997.

[239] Jared M. Baeten, Deborah Donnell, Patrick Ndase, Nelly R. Mugo, James D. Campbell, Jonathan Wangisi, Jordan W. Tappero, Elizabeth A. Bukusi, Craig R. Cohen, Elly Katabira, Allan Ronald, Elioda Tumwesigye, Edwin Were, Kenneth H. Fife, James Kiarie, Carey Farquhar, Grace John-Stewart, Aloysious Kakia, Josephine Odoyo, Akasiima Mucunguzi, Edith Nakku-Joloba, Rogers Twesigye, Kenneth Ngure, Cosmas Apaka, Harrison Tamooh, Fridah Gabona, Andrew Mujugira, Dana Panteleeff, Katherine K. Thomas, Lara Kidoguchi, Meighan Krows, Jennifer Revall, Susan Morrison, Harald Haugen, Mira Emmanuel-Ogier, Lisa Ondrejcek, Robert W. Coombs, Lisa Frenkel, Craig Hendrix, Namandjé N. Bumpus, David Bangsberg, Jessica E. Haberer, Wendy S. Stevens, Jairam R. Lingappa, and Connie Celum. Antiretroviral Prophylaxis for HIV Prevention in Heterosexual Men and Women. *New England Journal of Medicine*, 367(5):399–410, August 2012. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa1108524. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa1108524.

[240] Kachit Choopanya, Michael Martin, Pravan Suntharasamai, Udomsak Sangkum, Philip A Mock, Manoj Leethochawalit, Sithisat Chiamwongpaet, Praphan Kitisin, Pitinan Natrujirote, Somyot Kittimunkong, Rutt Chuachoowong, Roman J Gvetadze, Janet M McNicholl, Lynn A Paxton, Marcel E Curlin, Craig W Hendrix, and Suphak Vanichseni. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. *The Lancet*, 381(9883):2083–2090, June 2013. ISSN 01406736. doi: 10.1016/S0140-6736(13)61127-7. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673613611277.

[241] Douglas S. Krakower and Kenneth H. Mayer. Pre-Exposure Prophylaxis to Prevent HIV Infection: Current Status, Future Opportunities and Challenges. *Drugs*, 75(3):243–251, February 2015. ISSN 0012-6667, 1179-1950. doi: 10.1007/s40265-015-0355-4. URL http://link.springer.com/10.1007/s40265-015-0355-4.

[242] Douglas S. Krakower, Sachin Jain, and Kenneth H. Mayer. Antiretrovirals for Primary HIV Prevention: the Current Status of Pre- and Post-exposure Prophylaxis. *Current HIV/AIDS Reports*, 12(1):127–138, March 2015. ISSN 1548-3568, 1548-3576. doi: 10.1007/s11904-014-0253-5. URL http://link.springer.com/10.1007/s11904-014-0253-5.

[243] D. Wilson and N. Fraser. Who Pays and Why? Costs, Effectiveness, and Feasibility of HIV Treatment as Prevention. *Clinical Infectious Diseases*, 59(suppl 1):S28–S31, July 2014. ISSN 1058-4838, 1537-6591. doi: 10.1093/cid/ciu300. URL http://cid.oxfordjournals.org/lookup/doi/10.1093/cid/ciu300.

[244] David P. Wilson. HIV Treatment as Prevention: Natural Experiments Highlight Limits of Antiretroviral Treatment as HIV Prevention. *PLoS Medicine*, 9(7):e1001231, July 2012. ISSN 1549-1676. doi: 10.1371/journal.pmed.1001231. URL http://dx.plos.org/10.1371/journal.pmed.1001231.

[245] Myron S Cohen, M Kumi Smith, Kathryn E Muessig, Timothy B Hallett, Kimberly A Powers, and Angela D Kashuba. Antiretroviral treatment of HIV-1 prevents transmission of HIV-1: where do we go from here? *The Lancet*, 382(9903):1515–1524, November 2013. ISSN 01406736. doi: 10.1016/S0140-6736(13)61998-4. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673613619984.

[246] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. Springer New York, New York, NY, 2013. ISBN 978-1-4614-7137-0 978-1-4614-7138-7. URL http://link.springer.com/10.1007/978-1-4614-7138-7.

[247] Frequentist Inference. In Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, and Hiroki Yokota, editors, *Encyclopedia of Systems Biology*, pages 762–762. Springer New York, 2013. ISBN 978-1-4419-9862-0 978-1-4419-9863-7. URL http://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_100518. DOI: 10.1007/978-1-4419-9863-7_100518.

[248] Leonhard Held and Daniel Sabanés Bové. Elements of Frequentist Inference. In *Applied Statistical Inference*, pages 51–78. Springer Berlin Heidelberg, 2014. ISBN 978-3-642-37886-7 978-3-642-37887-4. URL http://link.springer.com/chapter/10.1007/978-3-642-37887-4_3. DOI: 10.1007/978-3-642-37887-4_3.

[249] Javier Rojo. Frequentist Inference. In Javier Rojo, editor, *Selected Works of E. L. Lehmann*, Selected Works in Probability and Statistics, pages 1075–1081. Springer US, 2012. ISBN 978-1-4614-1411-7 978-1-4614-1412-4. URL http://link.springer.com/chapter/10.1007/978-1-4614-1412-4_91. DOI: 10.1007/978-1-4614-1412-4_91.

[250] Eric-Jan Wagenmakers, Michael Lee, Tom Lodewyckx, and Geoffrey J. Iverson. Bayesian Versus Frequentist Inference. In Herbert Hoijtink, Irene Klugkist, and Paul A. Boelen, editors, *Bayesian Evaluation of Informative Hypotheses*, Statistics for Social and Behavioral Sciences, pages 181–207. Springer New York, 2008. ISBN 978-0-387-09611-7 978-0-387-09612-4. URL http://link.springer.com/chapter/10.1007/978-0-387-09612-4_9. DOI: 10.1007/978-0-387-09612-4_9.

[251] Vladimir N. Vapnik. Bounds on the Rate of Convergence of Learning Processes. In *The Nature of Statistical Learning Theory*, Statistics for Engineering and Information Science, pages 69–91. Springer New York, 2000. ISBN 978-1-4419-3160-3 978-1-4757-3264-1. URL http://link.springer.com/chapter/10.1007/978-1-4757-3264-1_4. DOI: 10.1007/978-1-4757-3264-1_4.

[252] Leonhard Held and Daniel Sabanés Bové. Bayesian Inference. In *Applied Statistical Inference*, pages 167–219. Springer Berlin Heidelberg, 2014. ISBN 978-3-642-37886-7 978-3-642-37887-4. URL http://link.springer.com/chapter/10.1007/978-3-642-37887-4_6. DOI: 10.1007/978-3-642-37887-4_6.

[253] Dr Roger Higdon. Bayesian Inference. In Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, and Hiroki Yokota, editors, *Encyclopedia of Systems Biology*, pages 71–73. Springer New York, 2013. ISBN 978-1-4419-9862-0 978-1-4419-9863-7. URL http://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_1179. DOI: 10.1007/978-1-4419-9863-7_1179.

[254] José M. Bernardo. Bayesian Statistics. In Miodrag Lovric, editor, *International Encyclopedia of Statistical Science*, pages 107–133. Springer Berlin Heidelberg, 2011. ISBN 978-3-642-04897-5 978-3-642-04898-2. URL http://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2_139. DOI: 10.1007/978-3-642-04898-2_139.

[255] Larry Wasserman. Bayesian Inference. In *All of Statistics*, Springer Texts in Statistics, pages 175–192. Springer New York, 2004. ISBN 978-1-4419-2322-6 978-0-387-21736-9. URL http://link.springer.com/chapter/10.1007/978-0-387-21736-9_11. DOI: 10.1007/978-0-387-21736-9_11.

[256] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2nd ed. edition, April 2011. ISBN 0-387-84857-6.

[257] Christopher J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, June 1998. ISSN 1384-5810, 1573-756X. doi: 10.1023/A: 1009715923555. URL http://link.springer.com/article/10.1023/A%3A1009715923555.

[258] H. W. Kuhn and A. W. Tucker. Nonlinear Programming. The Regents of the University of California, 1951. URL http://projecteuclid.org/euclid.bsmsp/1200500249.

[259] Vladimir N. Vapnik. Controlling the Generalization Ability of Learning Processes. In *The Nature of Statistical Learning Theory*, Statistics for Engineering and Information Science, pages 93–122. Springer New York, 2000. ISBN 978-1-4419-3160-3 978-1-4757-3264-1. URL http://link.springer.com/chapter/10.1007/978-1-4757-3264-1_5. DOI: 10.1007/978-1-4757-3264-1_5.

[260] Vladimir N. Vapnik. The Vicinal Risk Minimization Principle and the SVMs. In *The Nature of Statistical Learning Theory*, Statistics for Engineering and Information Science, pages 267–290. Springer New York, 2000. ISBN 978-1-4419-3160-3 978-1-4757-3264-1. URL http://link.springer.com/chapter/10.1007/978-1-4757-3264-1_9. DOI: 10.1007/978-1-4757-3264-1_9.

[261] Vladimir N. Vapnik. Methods of Pattern Recognition. In *The Nature of Statistical Learning Theory*, Statistics for Engineering and Information Science, pages 123–180. Springer New York, 2000. ISBN 978-1-4419-3160-3 978-1-4757-3264-1. URL http://link.springer.com/chapter/10.1007/978-1-4757-3264-1_6. DOI: 10.1007/978-1-4757-3264-1_6.

[262] Andreas Christmann and Ingo Steinwart. Kernels and Reproducing Kernel Hilbert Spaces. In *Support Vector Machines*, Information Science and Statistics, pages 110–163. Springer New York, 2008. ISBN 978-0-387-77241-7 978-0-387-77242-4. URL http://link.springer.com/chapter/10.1007/978-0-387-77242-4_4. DOI: 10.1007/978-0-387-77242-4_4.

[263] Alexander J. Smola and Bernhard Schölkopf. A Tutorial on Support Vector Regression. *NeuroCOLT2 Technical Report Series*, NC2-TR-1998-030, October 1998.

[264] David G. Kleinbaum and Mitchel Klein. Introduction to Survival Analysis. In *Survival Analysis*, Statistics for Biology and Health, pages 1–54. Springer New York, 2012. ISBN 978-1-4419-6645-2 978-1-4419-6646-9. URL http://link.springer.com/chapter/10.1007/978-1-4419-6646-9_1. DOI: 10.1007/978-1-4419-6646-9_1.

[265] Vanya Van Belle, Kristiaan Pelckmans, Sabine Van Huffel, and Johan A. K. Suykens. Support vector methods for survival analysis: a comparison between ranking and regression approaches. *Artificial Intelligence in Medicine*, 53(2):107–118, October 2011. ISSN 1873-2860. doi: 10.1016/j.artmed.2011.06.006.

[266] Correlation Coefficient. In *The Concise Encyclopedia of Statistics*, pages 115–119. Springer New York, New York, NY, 2008. ISBN 978-0-387-31742-7 978-0-387-32833-1. URL http://www.springerlink.com/index/10.1007/978-0-387-32833-1_83.

[267] Michael J. Pencina and Ralph B. D'Agostino. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in Medicine*, 23(13):2109–2123, July 2004. ISSN 0277-6715, 1097-0258. doi: 10.1002/sim.1802. URL http://doi.wiley.com/10.1002/sim.1802.

[268] F. E. Harrell, R. M. Califf, D. B. Pryor, K. L. Lee, and R. A. Rosati. Evaluating the yield of medical tests. *JAMA*, 247(18):2543–2546, May 1982. ISSN 0098-7484.

[269] Alejandro Pironti. *Probabilistic Disease-Gene Prioritization*. Master's Thesis, University of Saarland, Saarbrücken, Germany, September 2009.

[270] M. P. Wand and M. C. Jones. Comparison of Smoothing Parameterizations in Bivariate Kernel Density Estimation. *Journal of the American Statistical Association*, 88(422):520, June 1993. ISSN 01621459. doi: 10.2307/2290332. URL http://www.jstor.org/stable/2290332?origin=crossref.

[271] M. P. Wand and M. C. Jones. *Kernel Smoothing*. CRC Press, December 1994. ISBN 978-0-412-55270-0.

[272] Tarn Duong and Martin Hazelton. Plug-in bandwidth matrices for bivariate kernel density estimation. *Journal of Nonparametric Statistics*, 15(1):17–30, January 2003. ISSN 1048-5252, 1029-0311. doi: 10.1080/10485250306039. URL http://www.tandfonline.com/doi/abs/10.1080/10485250306039.

[273] Rob Hyndman, Xibin Zhang, and Maxwell King. Bandwidth Selection for Multivariate Kernel Density Estimation Using MCMC. Econometric Society 2004 Australasian Meetings 120, Econometric Society, August 2004. URL http://econpapers.repec.org/paper/ecmausm04/120.htm.

[274] Francisco J. Goerlich Gisbert. Weighted samples, kernel density estimators and convergence. *Empirical Economics*, 28(2):335–351, April 2003. ISSN 0377-7332, 1435-8921. doi: 10.1007/s001810200134. URL http://link.springer.com/article/10.1007/s001810200134.

[275] Nicolai Lohse, Niels Obel, Gitte Kronborg, Alex Laursen, Court Pedersen, Carsten S. Larsen, Birgit Kvinesdal, Henrik Toft Sørensen, and Jan Gerstoft. Declining risk of triple-class antiretroviral drug failure in Danish HIV-infected individuals. *AIDS (London, England)*, 19(8):815–822, May 2005. ISSN 0269-9370.

[276] Fiona C. Lampe, Jose M. Gatell, Schlomo Staszewski, Margaret A. Johnson, Christian Pradier, M. John Gill, Elisa de Lazzari, Brenda Dauer, Mike Youle, Eric Fontas, Hartmut B. Krentz, and Andrew N. Phillips. Changes over time in risk of initial virological failure of combination antiretroviral therapy: a multicohort analysis, 1996 to 2002. *Archives of Internal Medicine*, 166(5):521–528, March 2006. ISSN 0003-9926. doi: 10.1001/archinte.166.5.521.

[277] Valentina Svicher, Roberta D'Arrigo, Claudia Alteri, Massimo Andreoni, Gioacchino Angarano, Andrea Antinori, Guido Antonelli, Patrizia Bagnarelli, Fausto Baldanti, Ada Bertoli, Marco Borderi, Enzo Boeri, Isabella Bonn, Bianca Bruzzone, Anna Paola Callegaro, Roberta Cammarota, Filippo Canducci, Francesca Ceccherini-Silberstein, Massimo Clementi, Antonella D'Arminio Monforte, Andrea De Luca, Antonio Di Biagio, Simona Di Gianbenedetto, Giovanni Di Perri, Massimo Di Pietro, Lavinia Fabeni, Giovanni Fadda, Massimo Galli, William Gennari, Valeria Ghisetti, Andrea Giacometti, Andrea Gori, Francesco Leoncini, Franco Maggiolo, Renato Maserati, Francesco Mazzotta, Valeria Micheli, Genny Meini, Laura Monno, Cristina Mussini, Silvia Nozza, Stefania Paolucci, Saverio Parisi, Monica Pecorari, Daniele Pizzi, Tiziana Quirino, Maria Carla Re, Giuliano Rizzardini, Rosaria Santangelo, Alessandro Soria, Francesca Stazi, Gaetana Sterrantino, Ombretta Turriziani, Claudio Viscoli, Vincenzo Vullo,

Adriano Lazzarin, Carlo Federico Perno, and OSCAR Study Group. Performance of genotypic tropism testing in clinical practice using the enhanced sensitivity version of Trofile as reference assay: results from the OSCAR Study Group. *The new microbiologica*, 33(3):195–206, July 2010. ISSN 1121-7138.

[278] Simone E Langford, Jintanat Ananworanich, and David A Cooper. Predictors of disease progression in HIV infection: a review. *AIDS Research and Therapy*, 4(1):11, 2007. ISSN 17426405. doi: 10.1186/1742-6405-4-11. URL http://www.aidsrestherapy.com/content/4/1/11.

[279] Diane E. Bennett, Ricardo J. Camacho, Dan Otelea, Daniel R. Kuritzkes, Hervé Fleury, Mark Kiuchi, Walid Heneine, Rami Kantor, Michael R. Jordan, Jonathan M. Schapiro, Anne-Mieke Vandamme, Paul Sandstrom, Charles A. B. Boucher, David van de Vijver, Soo-Yon Rhee, Tommy F. Liu, Deenan Pillay, and Robert W. Shafer. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PloS One*, 4(3):e4724, 2009. ISSN 1932-6203. doi: 10.1371/journal.pone.0004724.

[280] Jens Lundgren, Jose M. Gatell, Hansjakob Furrer, and Jürgen Rockstroh. *Guidelines. Version 7.1. November 2014*. European Aids Clinical Society, 7.1 edition. URL http://www.eacsociety.org/files/guidelines-7.1-english.pdf.

[281] K. H. Mayer, G. J. Hanna, and R. T. D'Aquila. Clinical Use of Genotypic and Phenotypic Drug Resistance Testing to Monitor Antiretroviral Chemotherapy. *Clinical Infectious Diseases*, 32(5):774–782, March 2001. ISSN 1058-4838, 1537-6591. doi: 10.1086/319231. URL http://cid.oxfordjournals.org/lookup/doi/10.1086/319231.

[282] Kai Wang, Ram Samudrala, and John E. Mittler. Antivirogram or phenosense: a comparison of their reproducibility and an analysis of their correlation. *Antiviral Therapy*, 9(5):703–712, October 2004. ISSN 1359-6535.

[283] Michele W. Tang and Robert W. Shafer. HIV-1 antiretroviral resistance: scientific principles and clinical applications. *Drugs*, 72(9):e1–25, June 2012. ISSN 0012-6667. doi: 10.2165/11633630-000000000-00000.

[284] D. L. Mayers, F. E. McCutchan, E. E. Sanders-Buell, L. I. Merritt, S. Dilworth, A. K. Fowler, C. A. Marks, N. M. Ruiz, D. D. Richman, and C. R. Roberts. Characterization of HIV isolates arising after prolonged zidovudine therapy. *Journal of Acquired Immune Deficiency Syndromes*, 5(8):749–759, 1992. ISSN 0894-9255.

[285] A. J. Japour, D. L. Mayers, V. A. Johnson, D. R. Kuritzkes, L. A. Beckett, J. M. Arduino, J. Lane, R. J. Black, P. S. Reichelderfer, and R. T. D'Aquila. Standardized peripheral blood mononuclear cell culture assay for determination of drug susceptibilities of clinical human immunodeficiency virus type 1 isolates. The RV-43 Study Group, the AIDS Clinical Trials Group Virology Committee Resistance Working Group. *Antimicrobial Agents and Chemotherapy*, 37(5):1095–1101, May 1993. ISSN 0066-4804.

[286] P. Kellam and B. A. Larder. Recombinant virus assay: a rapid, phenotypic assay for assessment of drug susceptibility of human immunodeficiency virus type 1 isolates. *Antimicrobial Agents and Chemotherapy*, 38(1):23–30, January 1994. ISSN 0066-4804.

[287] K. Hertogs, M. P. de Béthune, V. Miller, T. Ivens, P. Schel, A. Van Cauwenberge, C. Van Den Eynde, V. Van Gerwen, H. Azijn, M. Van Houtte, F. Peeters, S. Staszewski, M. Conant, S. Bloor, S. Kemp, B. Larder, and R. Pauwels. A rapid method for simultaneous detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. *Antimicrobial Agents and Chemotherapy*, 42(2): 269–276, February 1998. ISSN 0066-4804.

[288] C. J. Petropoulos, N. T. Parkin, K. L. Limoli, Y. S. Lie, T. Wrin, W. Huang, H. Tian, D. Smith, G. A. Winslow, D. J. Capon, and J. M. Whitcomb. A Novel Phenotypic Drug Susceptibility Assay for Human Immunodeficiency Virus Type 1. *Antimicrobial Agents and Chemotherapy*, 44(4):920–928, April 2000. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.44.4.920-928.2000. URL http://aac.asm.org/cgi/doi/10.1128/AAC.44.4.920-928.2000.

[289] Shoukat H. Qari, Richard Respess, Hillard Weinstock, Elise M. Beltrami, Kurt Hertogs, Brendan A. Larder, Christos J. Petropoulos, Nicholas Hellmann, and Walid Heneine. Comparative analysis of two commercial phenotypic assays for drug susceptibility testing of human immunodeficiency virus type 1. *Journal of Clinical Microbiology*, 40(1):31–35, January 2002. ISSN 0095-1137.

[290] Martin Schutten. Resistance assays. In Anna Maria Geretti, editor, *Antiretroviral Resistance in Clinical Practice*. Mediscript, London, 2006. ISBN 978-0-9551669-0-7. URL http://www.ncbi.nlm.nih.gov/books/NBK2252/.

[291] P. Simmonds, L. Q. Zhang, F. McOmish, P. Balfe, C. A. Ludlam, and A. J. Brown. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 env sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. *Journal of Virology*, 65(11):6266–6276, November 1991. ISSN 0022-538X.

[292] K. L. Kroodsma, M. J. Kozal, K. A. Hamed, M. A. Winters, and T. C. Merigan. Detection of drug resistance mutations in the human immunodeficiency virus type 1 (HIV-1) pol gene: differences in semen and blood HIV-1 RNA and proviral DNA. *The Journal of Infectious Diseases*, 170(5):1292–1295, November 1994. ISSN 0022-1899.

[293] Terry L. Riss, Richard A. Moravec, Andrew L. Niles, Helene A. Benink, Tracy J. Worzella, and Lisa Minor. Cell Viability Assays. In G. Sitta Sittampalam, Nathan P. Coussens, Henrike Nelson, Michelle Arkin, Douglas Auld, Chris Austin, Bruce Bejcek, Marcie Glicksman, James Inglese, Philip W. Iversen, Zhuyin Li, James McGee, Owen McManus, Lisa Minor, Andrew Napper, John M. Peltier, Terry Riss, O. Joseph Trask, and Jeff Weidner, editors, *Assay Guidance Manual*. Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004. URL http://www.ncbi.nlm.nih.gov/books/NBK144065/.

[294] Robert W Shafer. Rationale and uses of a public HIV drug-resistance database. *The Journal of infectious diseases*, 194 Suppl 1:S51–58, September 2006. ISSN 0022-1899. doi: 10.1086/505356.

[295] J. G. Garcia-Lerma, H. MacInnes, D. Bennett, H. Weinstock, and W. Heneine. Transmitted Human Immunodeficiency Virus Type 1 Carrying the D67n or K219q/E Mutation Evolves Rapidly to Zidovudine Resistance In Vitro and Shows a High Replicative Fitness in the Presence of Zidovudine. *Journal*

of Virology, 78(14):7545–7552, July 2004. ISSN 0022-538X. doi: 10.1128/JVI.78.14.7545-7552.2004. URL http://jvi.asm.org/cgi/doi/10.1128/JVI.78.14.7545-7552.2004.

[296] Victoria A Johnson, Vincent Calvez, Huldrych F Gunthard, Roger Paredes, Deenan Pillay, Robert W Shafer, Annemarie M Wensing, and Douglas D Richman. Update of the drug resistance mutations in HIV-1: March 2013. *Topics in antiviral medicine*, 21(1):6–14, March 2013. ISSN 2161-5853.

[297] M. Obermeier, A. Pironti, T. Berg, P. Braun, M. Däumer, J. Eberle, R. Ehret, R. Kaiser, N. Kleinkauf, K. Korn, C. Kücherer, H. Müller, C. Noah, M. Stürmer, A. Thielen, E. Wolf, and H. Walter. HIV-GRADE: A publicly available, rules-based drug resistance interpretation algorithm integrating bioinformatic knowledge. *Intervirology*, 55(2):102–107, 2012. ISSN 0300-5526. doi: 10.1159/000331999.

[298] Tommy F Liu and Robert W Shafer. Web resources for HIV type 1 genotypic-resistance test interpretation. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, 42(11):1608–1618, June 2006. ISSN 1537-6591. doi: 10.1086/503914. URL http://www.ncbi.nlm.nih.gov/pubmed/16652319.

[299] N. Beerenwinkel, M. Däumer, M. Oette, K. Korn, D. Hoffmann, R. Kaiser, T. Lengauer, J. Selbig, and H. Walter. Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes. *Nucleic Acids Research*, 31(13):3850–3855, 2003. ISSN 0305-1048. doi: 10.1093/nar/gkg575.

[300] Maurizio Zazzi, Francesca Incardona, Michal Rosen-Zvi, Mattia Prosperi, Thomas Lengauer, Andre Altmann, Anders Sonnerborg, Tamar Lavee, Eugen Schülter, and Rolf Kaiser. Predicting response to antiretroviral treatment by machine learning: the EuResist project. *Intervirology*, 55(2):123–127, 2012. ISSN 1423-0100. doi: 10.1159/000332008.

[301] Robert W Shafer, Soo-Yon Rhee, Deenan Pillay, Veronica Miller, Paul Sandstrom, Jonathan M Schapiro, Daniel R Kuritzkes, and Diane Bennett. HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *AIDS (London, England)*, 21(2):215–223, January 2007. ISSN 0269-9370. doi: 10.1097/QAD.0b013e328011e691. URL http://www.ncbi.nlm.nih.gov/pubmed/17197813.

[302] M Zazzi, R Kaiser, A Sönnerborg, D Struck, A Altmann, M Prosperi, M Rosen-Zvi, A Petroczi, Y Peres, E Schülter, C A Boucher, F Brun-Vezinet, P R Harrigan, L Morris, M Obermeier, C-F Perno, P Phanuphak, D Pillay, R W Shafer, A-M Vandamme, K van Laethem, A M J Wensing, T Lengauer, and F Incardona. Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study). *HIV medicine*, 12(4):211–218, April 2011. ISSN 1468-1293. doi: 10.1111/j.1468-1293.2010.00871.x.

[303] Andrea-Clemencia Pineda-Peña, Nuno Rodrigues Faria, Stijn Imbrechts, Pieter Libin, Ana Barroso Abecasis, Koen Deforche, Arley Gómez-López, Ricardo J Camacho, Tulio de Oliveira, and Anne-Mieke Vandamme. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*, 19:337–348, October 2013. ISSN 1567-7257. doi: 10.1016/j.meegid.2013.04.032.

[304] Daniel Struck, Glenn Lawyer, Anne-Marie Ternes, Jean-Claude Schmit, and Danielle Perez Bercoff. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Research*, 42(18):e144, October 2014. ISSN 1362-4962. doi: 10.1093/nar/gku739.

[305] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):1–27, April 2011. ISSN 21576904. doi: 10.1145/1961189.1961199. URL http://dl.acm.org/citation.cfm?doid=1961189.1961199.

[306] R Core Team. R: A Language and Environment for Statistical Computing, 2014. URL http://www.R-project.org.

[307] Tobias Sing, Oliver Sander, Niko Beerenwinkel, and Thomas Lengauer. ROCR: visualizing classifier performance in R. *Bioinformatics (Oxford, England)*, 21(20):3940–3941, October 2005. ISSN 1367-4803. doi: 10.1093/bioinformatics/bti623. URL http://www.ncbi.nlm.nih.gov/pubmed/16096348.

[308] Frank Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80, December 1945. ISSN 00994987. doi: 10.2307/3001968. URL http://www.jstor.org/stable/10.2307/3001968?origin=crossref.

[309] John Platt. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. In *Advances in Large Margin Classifiers*, pages 61–74. MIT Press.

[310] Yang Xiang, Sylvain Gubian, Brian Suomela, and Julia Hoeng. Generalized Simulated Annealing for Global Optimization: The GenSA Package. *The R Journal*, 5(1):13–28, June 2013.

[311] Andre Altmann, Martin P Däumer, Alejandro Pironti, Juliane Perner, Joachim Büch, Carlo Torti, Mattia Prosperi, Rolf Kaiser, and Thomas Lengauer. Application of a probabilistic Optimization Method for Deriving Clinical Cutoffs for (Predicted) Durg Resistance Phenotypes. In *Reviews in Antiviral Therapy and Infectious Diseases*, Sorrento, Italy, March 2010.

[312] Tarn Duong. ks: Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data in R. *Journal of Statistical Software*, 21(7):1–16, October 2007. ISSN 1548-7660. URL http://www.jstatsoft.org/v21/i07.

[313] Timur V. Elzhov, Katharine M. Mullen, Andrej-Nikolai Spiess, and Ben Bolker. minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds, November 2015. URL https://cran.r-project.org/web/packages/minpack.lm/index.html.

[314] J G Garcia-Lerma, S Nidtha, K Blumoff, H Weinstock, and W Heneine. Increased ability for selection of zidovudine resistance in a distinct class of wild-type HIV-1 from drug-naive persons. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24):13907–13912, November 2001. ISSN 0027-8424. doi: 10.1073/pnas.241300698. URL http://www.ncbi.nlm.nih.gov/pubmed/11698656.

[315] H. Vermeiren, E. Van Craenenbroeck, P. Alen, L. Bacheler, G. Picchio, and P. Lecocq. Prediction of HIV-1 drug susceptibility phenotype from the viral genotype using linear regression modeling. *Journal of Virological Methods*, 145(1):47–55, October 2007. ISSN 0166-0934. doi: 10.1016/j.jviromet.2007.05. 009. URL http://www.sciencedirect.com/science/article/pii/S0166093407001735.

[316] A. Altmann, T. Sing, H. Vermeiren, B. Winters, E. Van Craenenbroeck, K. Van der Borght, S.-Y. Rhee, R.W. Shafer, E. Schülter, R. Kaiser, Y. Peres, A. Sönnerborg, W.J. Fessel, F. Incardona, M. Zazzi, L. Bacheler, H. Van Vlijmen, and T. Lengauer. Advantages of predicted phenotypes and statistical learning models in inferring virological response to antiretroviral therapy from HIV genotype. *Antiviral Therapy*, 14(2):273–283, 2009. ISSN 1359-6535.

[317] Philippe Flandre, Colombe Chappey, Anne Genevieve Marcelin, Kirk Ryan, Jen-Fue Maa, Mike Bates, Daniel Seekins, Marie Charlotte Bernard, Vincent Calvez, and Jean Michel Molina. Phenotypic susceptibility to didanosine is associated with antiviral activity in treatment-experienced patients with HIV-1 infection. *The Journal of Infectious Diseases*, 195(3):392–398, February 2007. ISSN 0022-1899. doi: 10.1086/510754.

[318] Andrew R. Zolopa. Incorporating Drug-Resistance Measurements into the Clinical Management of HIV-1 Infection. *The Journal of Infectious Diseases*, 194(s1):S59–S64, September 2006. ISSN 0022-1899, 1537-6613. doi: 10.1086/505360. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/505360.

[319] Michael D. Miller, Nicolas Margot, Biao Lu, Lijie Zhong, Shan-Shan Chen, Andrew Cheng, and Michael Wulfsohn. Genotypic and phenotypic predictors of the magnitude of response to tenofovir disoproxil fumarate treatment in antiretroviral-experienced patients. *The Journal of Infectious Diseases*, 189(5): 837–846, March 2004. ISSN 0022-1899. doi: 10.1086/381784.

[320] Françoise Brun-Vézinet, Dominique Costagliola, Mounir Ait Khaled, Vincent Calvez, François Clavel, Bonaventura Clotet, Richard Haubrich, Dale Kempf, Marty King, Daniel Kuritzkes, Randall Lanier, Michael Miller, Veronica Miller, Andrews Phillips, Deenan Pillay, Jonathan Schapiro, Janna Scott, Robert Shafer, Maurizio Zazzi, Andrew Zolopa, and Victor DeGruttola. Clinically validated genotype analysis: guiding principles and statistical concerns. *Antiviral Therapy*, 9(4):465–478, August 2004. ISSN 1359-6535.

[321] Anna Maria Geretti. Clinical implications of HIV drug resistance to nucleoside and nucleotide reverse transcriptase inhibitors. *AIDS reviews*, 8(4):210–220, December 2006. ISSN 1139-6121.

[322] Bart Winters, Julio Montaner, P Richard Harrigan, Brian Gazzard, Anton Pozniak, Michael D Miller, Sean Emery, Frank van Leth, Patrick Robinson, John D Baxter, Marie Perez-Elias, Delivette Castor, Scott Hammer, Alex Rinehart, Hans Vermeiren, Elke Van Craenenbroeck, and Lee Bacheler. Determination of Clinically Relevant Cutoffs for HIV-1 Phenotypic Resistance Estimates Through a Combined Analysis of Clinical Trial and Cohort Data:. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 48(1):26–34, May 2008. ISSN 1525-4135. doi: 10.1097/QAI.0b013e31816d9bf4. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00126334-200805010-00004.

[323] Johan Vingerhoets, Lotke Tambuyzer, Hilde Azijn, Annemie Hoogstoel, Steven Nijs, Monika Peeters, Marie-Pierre de Béthune, Goedele De Smedt, Brian Woodfall, and Gastón Picchio. Resistance profile of etravirine: combined analysis of baseline genotypic and phenotypic data from the randomized, controlled Phase III clinical studies. *AIDS (London, England)*, 24(4):503–514, February 2010. ISSN 1473-5571. doi: 10.1097/QAD.0b013e32833677ac.

[324] Carlo-Federico Perno and Ada Bertoli. Clinical cut-offs in the interpretation of phenotypic resistance. In Anna Maria Geretti, editor, *Antiretroviral Resistance in Clinical Practice*. Mediscript, London, 2006. ISBN 978-0-9551669-0-7. URL http://www.ncbi.nlm.nih.gov/books/NBK2254/.

[325] P. R. Harrigan, J. S. Montaner, S. A. Wegner, W. Verbiest, V. Miller, R. Wood, and B. A. Larder. Worldwide variation in HIV-1 phenotypic susceptibility in untreated individuals: biologically relevant values for resistance testing. *AIDS (London, England)*, 15(13):1671–1677, September 2001. ISSN 0269-9370.

[326] N. T. Parkin, N. S. Hellmann, J. M. Whitcomb, L. Kiss, C. Chappey, and C. J. Petropoulos. Natural Variation of Drug Susceptibility in Wild-Type Human Immunodeficiency Virus Type 1. *Antimicrobial Agents and Chemotherapy*, 48(2):437–443, February 2004. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.48.2.437-443.2004. URL http://aac.asm.org/cgi/doi/10.1128/AAC.48.2.437-443.2004.

[327] Nancy S. Shulman, Michael D. Hughes, Mark A. Winters, Robert W. Shafer, Andrew R. Zolopa, Nicholas S. Hellmann, Michael Bates, Jeannette M. Whitcomb, and David A. Katzenstein. Subtle decreases in stavudine phenotypic susceptibility predict poor virologic response to stavudine monotherapy in zidovudine-experienced patients. *Journal of Acquired Immune Deficiency Syndromes (1999)*, 31(2): 121–127, October 2002. ISSN 1525-4135.

[328] Bart Winters, Elke Van Craenenbroeck, Koen Van der Borght, Pierre Lecocq, Jorge Villacian, and Lee Bacheler. Clinical cut-offs for HIV-1 phenotypic resistance estimates: update based on recent pivotal clinical trial data and a revised approach to viral mixtures. *Journal of Virological Methods*, 162(1-2): 101–108, December 2009. ISSN 1879-0984. doi: 10.1016/j.jviromet.2009.07.023.

[329] B Foley, T Leitner, C Apetrei, B Hahn, I Mizrachi, J Mullins, A Rambaut, S Wolinsky, and B Korber. *HIV Sequence Compendium 2013*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, 2013. URL http://www.hiv.lanl.gov/.

[330] Soo-Yon Rhee, Jose Blanco, Tommy F Liu, Iñaki Pere, Rolf Kaiser, Maurizio Zazzi, Francesca Incardona, William Towner, Josep Gatell, Andrea De Luca, W Fessel, and Robert W Shafer. Standardized representation, visualization and searchable repository of antiretroviral treatment-change episodes. *AIDS Research and Therapy*, 9(1):13, 2012. ISSN 1742-6405. doi: 10.1186/1742-6405-9-13. URL http://www.aidsrestherapy.com/content/9/1/13.

[331] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical optimization*. Academic Press, London ; New York, 1981. ISBN 978-0-12-283952-8.

[332] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 57(1):289–300, 1995.

[333] M. A. Wainberg, B. G. Brenner, and D. Turner. Changing Patterns in the Selection of Viral Mutations among Patients Receiving Nucleoside and Nucleotide Drug Combinations Directed against Human Immunodeficiency Virus Type 1 Reverse Transcriptase. *Antimicrobial Agents and Chemotherapy*, 49 (5):1671–1678, May 2005. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.49.5.1671-1678.2005. URL http://aac.asm.org/cgi/doi/10.1128/AAC.49.5.1671-1678.2005.

[334] R. K. Zeldin. Pharmacological and therapeutic properties of ritonavir-boosted protease inhibitor therapy in HIV-infected patients. *Journal of Antimicrobial Chemotherapy*, 53(1):4–9, December 2003. ISSN 1460-2091. doi: 10.1093/jac/dkh029. URL http://www.jac.oupjournals.org/cgi/doi/10.1093/jac/dkh029.

[335] Caroline Fenton and Caroline M Perry. Darunavir: In the Treatment of HIV-1 Infection. *Drugs*, 67(18): 2791–2801, 2007. ISSN 0012-6667. doi: 10.2165/00003495-200767180-00010. URL http://link.springer.com/10.2165/00003495-200767180-00010.

[336] E. L. Murphy, A. C. Collier, L. A. Kalish, S. F. Assmann, M. F. Para, T. P. Flanigan, P. N. Kumar, L. Mintz, F. R. Wallach, G. J. Nemo, and Viral Activation Transfusion Study Investigators. Highly active antiretroviral therapy decreases mortality and morbidity in patients with advanced HIV disease. *Annals of Internal Medicine*, 135(1):17–26, July 2001. ISSN 0003-4819.

[337] Joep MA Lange and Jintanat Ananworanich. The discovery and development of antiretroviral agents. *Antiviral Therapy*, 19(Suppl 3):5–14, 2014. ISSN 13596535. doi: 10.3851/IMP2896. URL http://www.intmedpress.com/journals/avt/abstract.cfm?id=2896&pid=88.

[338] Caroline A Sabin. Do people with HIV infection have a normal life expectancy in the era of combination antiretroviral therapy? *BMC Medicine*, 11(1):251, 2013. ISSN 1741-7015. doi: 10.1186/1741-7015-11-251. URL http://www.biomedcentral.com/1741-7015/11/251.

[339] C. Delaugerre, J. Ghosn, J.-M. Lacombe, G. Pialoux, L. Cuzin, O. Launay, A. Menard, P. de Truchis, D. Costagliola, for the FHDH-ANRS CO4, S. Abgrall, F. Barin, E. Billaud, F. Boue, L. Boyer, A. Cabie, F. Caby, D. Costagliola, L. Cotte, P. De Truchis, X. Duval, C. Duvivier, P. Enel, J. Gasnault, C. Gaud, J. Gilquin, S. Grabar, M. A. Khuong, O. Launay, J. Le Bail, A. Mahamat, M. Mary-Krause, S. Matheron, J. L. Meynard, J. Pavie, L. Piroth, I. Poizot-Martin, C. Pradier, J. Reynes, E. Rouveix, E. Salat, A. Simon, P. Tattevin, H. Tissot-Dupont, J. P. Viard, N. Viget, C. Bronnec, D. Martin, D. Costagliola, S. Abgrall, S. Grabar, M. Guiguet, S. Lang, L. Lievre, M. Mary-Krause, H. Selinger-Leneman, J. M. Lacombe, and V. Potard. Significant Reduction in HIV Virologic Failure During a 15-Year Period in a Setting With Free Healthcare Access. *Clinical Infectious Diseases*, 60(3):463–472, February 2015. ISSN 1058-4838, 1537-6591. doi: 10.1093/cid/ciu834. URL http://cid.oxfordjournals.org/lookup/doi/10.1093/cid/ciu834.

[340] V. Pirrone, N. Thakkar, J. M. Jacobson, B. Wigdahl, and F. C. Krebs. Combinatorial Approaches to the Prevention and Treatment of HIV-1 Infection. *Antimicrobial Agents and Chemotherapy*, 55(5):1831–1842, May 2011. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.00976-10. URL http://aac.asm.org/cgi/doi/10.1128/AAC.00976-10.

[341] Maurizio Zazzi, Francesca Incardona, Michal Rosen-Zvi, Mattia Prosperi, Thomas Lengauer, Andre Altmann, Anders Sonnerborg, Tamar Lavee, Eugen Schülter, and Rolf Kaiser. Predicting response to antiretroviral treatment by machine learning: the EuResist project. *Intervirology*, 55(2):123–127, 2012. ISSN 1423-0100. doi: 10.1159/000332008.

[342] Asa M. Margolis, Harry Heverling, Paul A. Pham, and Andrew Stolbach. A Review of the Toxicity of HIV Medications. *Journal of Medical Toxicology*, 10(1):26–39, March 2014. ISSN 1556-9039, 1937-6995. doi: 10.1007/s13181-013-0325-8. URL http://link.springer.com/10.1007/s13181-013-0325-8.

[343] Pinar Iyidogan and Karen Anderson. Current Perspectives on HIV-1 Antiretroviral Drug Resistance. *Viruses*, 6(10):4095–4139, October 2014. ISSN 1999-4915. doi: 10.3390/v6104095. URL http://www.mdpi.com/1999-4915/6/10/4095/.

[344] E. L. Asahchop, M. A. Wainberg, R. D. Sloan, and C. L. Tremblay. Antiviral Drug Resistance and the Need for Development of New HIV-1 Reverse Transcriptase Inhibitors. *Antimicrobial Agents and Chemotherapy*, 56(10):5000–5008, October 2012. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.00591-12. URL http://aac.asm.org/cgi/doi/10.1128/AAC.00591-12.

[345] Jean-Jacques Parienti, David R. Bangsberg, Renaud Verdon, and Edward M. Gardner. Better Adherence with Once-Daily Antiretroviral Regimens: A Meta-Analysis. *Clinical Infectious Diseases*, 48(4):484–488, February 2009. ISSN 1058-4838, 1537-6591. doi: 10.1086/596482. URL http://cid.oxfordjournals.org/lookup/doi/10.1086/596482.

[346] L Bansi, C Sabin, V Delpech, T Hill, M Fisher, J Walsh, T Chadborn, P Easterbrook, R Gilson, M Johnson, K Porter, J Anderson, M Gompels, C Leen, J Ainsworth, C Orkin, M Nelson, B Rice, A Phillips, and for the UK Collaborative HIV Cohort (CHIC) Study and the Health Protection Agency. Trends over calendar time in antiretroviral treatment success and failure in HIV clinic populations. *HIV Medicine*, February 2010. ISSN 14642662, 14681293. doi: 10.1111/j.1468-1293.2009.00809.x. URL http://doi.wiley.com/10.1111/j.1468-1293.2009.00809.x.

[347] José Valdez Madruga, Pedro Cahn, Beatriz Grinsztejn, Richard Haubrich, Jacob Lalezari, Anthony Mills, Gilles Pialoux, Timothy Wilkin, Monika Peeters, Johan Vingerhoets, Goedele de Smedt, Lorant Leopold, Roberta Trefiglio, and Brian Woodfall. Efficacy and safety of TMC125 (etravirine) in treatment-experienced HIV-1-infected patients in DUET-1: 24-week results from a randomised, double-blind, placebo-controlled trial. *The Lancet*, 370(9581):29–38, July 2007. ISSN 01406736. doi: 10.1016/S0140-6736(07)61047-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673607610472.

[348] José Valdez Madruga, Daniel Berger, Marilyn McMurchie, Fredy Suter, Denes Banhegyi, Kiat Ruxrungtham, Dorece Norris, Eric Lefebvre, Marie-Pierre de Béthune, Frank Tomaka, Martine De Pauw, Tony Vangeneugden, and Sabrina Spinosa-Guzman. Efficacy and safety of darunavir-ritonavir compared with that of lopinavir-ritonavir at 48 weeks in treatment-experienced, HIV-infected patients in TITAN: a randomised controlled phase III trial. *The Lancet*, 370(9581):49–58, July 2007. ISSN 01406736. doi: 10.1016/S0140-6736(07)61049-6. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673607610496.

[349] Jean-Michel Molina, Anthony LaMarca, Jaime Andrade-Villanueva, Bonaventura Clotet, Nathan Clumeck, Ya-Pei Liu, Lijie Zhong, Nicolas Margot, Andrew K Cheng, and Steven L Chuck. Efficacy and safety of once daily elvitegravir versus twice daily raltegravir in treatment-experienced patients with HIV-1 receiving a ritonavir-boosted protease inhibitor: randomised, double-blind, phase 3, non-inferiority study. *The Lancet Infectious Diseases*, 12(1):27–35, January 2012. ISSN 14733099. doi: 10.1016/S1473-3099(11)70249-3. URL http://linkinghub.elsevier.com/retrieve/pii/S1473309911702493.

[350] Adriano Lazzarin, Thomas Campbell, Bonaventura Clotet, Margaret Johnson, Christine Katlama, Arend Moll, William Towner, Benoit Trottier, Monika Peeters, Johan Vingerhoets, Goedele de Smedt, Benny Baeten, Greet Beets, Rekha Sinha, and Brian Woodfall. Efficacy and safety of TMC125 (etravirine) in treatment-experienced HIV-1-infected patients in DUET-2: 24-week results from a randomised, double-blind, placebo-controlled trial. *The Lancet*, 370(9581):39–48, July 2007. ISSN 01406736. doi: 10.1016/S0140-6736(07)61048-4. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673607610484.

[351] Luke C. Swenson, Jeong Eun Min, Conan K. Woods, Eric Cai, Jonathan Z. Li, Julio S.G. Montaner, P. Richard Harrigan, and Alejandro Gonzalez-Serna. HIV drug resistance detected during low-level viraemia is associated with subsequent virologic failure:. *AIDS*, 28(8):1125–1134, May 2014. ISSN 0269-9370. doi: 10.1097/QAD.0000000000000203. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201405150-00005.

[352] M.-A. Vandenhende. Impact of low-level viremia on clinical and virological outcomes in treated HIV-1-infected patients; The antiretroviral therapy cohort collaboration (ART-CC). *AIDS*, 29(3):373–383, 2015. ISSN 0269-9370. doi: 10.1097/QAD.0000000000000544.

[353] Marc Wirden, Eve Todesco, Marc-Antoine Valantin, Sidonie Lambert-Niclot, Anne Simon, Ruxandra Calin, Roland Tubiana, Gilles Peytavin, Christine Katlama, Vincent Calvez, and Anne-Genevieve Marcelin. Low-level HIV-1 viraemia in patients on HAART: risk factors and management in clinical practice. *Journal of Antimicrobial Chemotherapy*, 70(8):2347–2353, August 2015. ISSN 0305-7453, 1460-2091. doi: 10.1093/jac/dkv099. URL http://jac.oxfordjournals.org/content/70/8/2347.

[354] Joanna Kryst, Paweł Kawalec, and Andrzej Pilc. Efavirenz-Based Regimens in Antiretroviral-Naive HIV-Infected Patients: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *PLOS ONE*, 10(5):e0124279, May 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0124279. URL http://dx.plos.org/10.1371/journal.pone.0124279.

[355] Bonaventura Clotet, Judith Feinberg, Jan van Lunzen, Marie-Aude Khuong-Josses, Andrea Antinori, Irina Dumitru, Vadim Pokrovskiy, Jan Fehr, Roberto Ortiz, Michael Saag, Julia Harris, Clare Brennan, Tamio Fujiwara, and Sherene Min. Once-daily dolutegravir versus darunavir plus ritonavir in antiretroviral-naive adults with HIV-1 infection (FLAMINGO): 48 week results from the randomised open-label phase 3b study. *The Lancet*, 383(9936):2222–2231, June 2014. ISSN 01406736. doi: 10.1016/S0140-6736(14)60084-2. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673614600842.

[356] Calvin J. Cohen, Jean-Michel Molina, Isabel Cassetti, Ploenchan Chetchotisakd, Adriano Lazzarin, Chloe Orkin, Frank Rhame, Hans-Jürgen Stellbrink, Taisheng Li, Herta Crauwels, Laurence Rimsky, Simon Vanveggel, Peter Williams, and Katia Boven. Week 96 efficacy and safety of rilpivirine in treatment-naive, HIV-1 patients in two Phase III randomized trials:. *AIDS*, 27(6):939–950, March 2013. ISSN 0269-9370. doi: 10.1097/QAD.0b013e32835cee6e. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201303270-00010.

[357] S. Staszewski, P. Keiser, J. Montaner, F. Raffi, J. Gathe, V. Brotas, C. Hicks, S. M. Hammer, D. Cooper, M. Johnson, S. Tortell, A. Cutrell, D. Thorborn, R. Isaacs, S. Hetherington, H. Steel, W. Spreen, and CNAAB3005 International Study Team. Abacavir-lamivudine-zidovudine vs indinavir-lamivudine-zidovudine in antiretroviral-naive HIV-infected adults: A randomized equivalence trial. *JAMA*, 285 (9):1155–1163, March 2001. ISSN 0098-7484.

[358] Karin J Metzner, Pia Rauch, Viktor von Wyl, Christine Leemann, Christina Grube, Herbert Kuster, Jürg Böni, Rainer Weber, and Huldrych F Günthard. Efficient suppression of minority drug-resistant HIV type 1 (HIV-1) variants present at primary HIV-1 infection by ritonavir-boosted protease inhibitor-containing antiretroviral therapy. *The Journal of infectious diseases*, 201(7):1063–1071, April 2010. ISSN 1537-6613. doi: 10.1086/651136. URL http://www.ncbi.nlm.nih.gov/pubmed/20196655.

[359] Annemarie M.J. Wensing, Noortje M. van Maarseveen, and Monique Nijhuis. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. *Antiviral Research*, 85(1):59–74, January 2010. ISSN 01663542. doi: 10.1016/j.antiviral.2009.10.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0166354209004902.

[360] Kristel Van Laethem, Paul De Munter, Yoeri Schrooten, Rene Verbesselt, Marc Van Ranst, Eric Van Wijngaerden, and Anne-Mieke Vandamme. No response to first-line tenofovir+lamivudine+efavirenz despite optimization according to baseline resistance testing: Impact of resistant minority variants on efficacy of low genetic barrier drugs. *Journal of Clinical Virology*, 39(1):43–47, May 2007. ISSN 13866532. doi: 10.1016/j.jcv.2007.02.003. URL http://linkinghub.elsevier.com/retrieve/pii/S1386653207000686.

[361] Melanie Balduin, Mark Oette, Martin P Däumer, Daniel Hoffmann, Herbert J Pfister, and Rolf Kaiser. Prevalence of minor variants of HIV strains at reverse transcriptase position 103 in therapy-naïve patients and their impact on the virological failure. *Journal of clinical virology: the official publication of the Pan American Society for Clinical Virology*, 45(1):34–38, May 2009. ISSN 1873-5967. doi: 10.1016/j.jcv.2009.03.002. URL http://www.ncbi.nlm.nih.gov/pubmed/19375978.

[362] Michela Violin, Alessandro Cozzi-Lepri, Rossella Velleca, Antonella Vincenti, Salvatore D'Elia, Francesco Chiodo, Florio Ghinelli, Ada Bertoli, Antonella d'Arminio Monforte, Carlo Federico Perno, Mauro Moroni, and Claudia Balotta. Risk of failure in patients with 215 HIV-1 revertants starting their first thymidine analog-containing highly active antiretroviral therapy:. *AIDS*, 18(2):227–235, January 2004. ISSN 0269-9370. doi: 10.1097/00002030-200401230-00012. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-200401230-00012.

[363] M Pingen, Me van der Ende, Am Wensing, A el Barzouhi, Bb Simen, M Schutten, and Cab Boucher. Deep sequencing does not reveal additional transmitted mutations in patients diagnosed with HIV-1 variants with single nucleoside reverse transcriptase inhibitor resistance mutations: Clinical relevance of transmitted singletons. *HIV Medicine*, 14(3):176–181, March 2013. ISSN 14642662. doi: 10.1111/j.1468-1293.2012.01037.x. URL http://doi.wiley.com/10.1111/j.1468-1293.2012.01037.x.

[364] Y. Mitsuya, V. Varghese, C. Wang, T. F. Liu, S. P. Holmes, P. Jayakumar, B. Gharizadeh, M. Ronaghi, D. Klein, W. J. Fessel, and R. W. Shafer. Minority Human Immunodeficiency Virus Type 1 Variants in Antiretroviral-Naive Persons with Reverse Transcriptase Codon 215 Revertant Mutations. *Journal of Virology*, 82(21):10747–10755, November 2008. ISSN 0022-538X. doi: 10.1128/JVI.01827-07. URL http://jvi.asm.org/cgi/doi/10.1128/JVI.01827-07.

[365] Sara Gianella and Douglas D. Richman. Minority Variants of Drug-Resistant HIV. *The Journal of Infectious Diseases*, 202(5):657–666, September 2010. ISSN 0022-1899, 1537-6613. doi: 10.1086/655397. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/655397.

[366] Bernard Masquelier. Minority drug-resistant HIV-1 variants: transmission and response to antiretroviral therapy. *HIV Therapy*, 3(1):55–61, January 2009. ISSN 1758-4310. doi: 10.2217/17584310.3.1.55. URL http://www.futuremedicine.com/doi/abs/10.2217/17584310.3.1.55.

[367] Anne-Mieke Vandamme, Ricardo J. Camacho, Francesca Ceccherini-Silberstein, Andreu de Luca, Lucia Palmisano, Dimitrios Paraskevis, Roger Paredes, Mario Poljak, Jean-Claude Schmit, Vincent Soriano, Hauke Walter, Anders Sönnerborg, and European HIV Drug Resistance Guidelines Panel. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. *AIDS reviews*, 13(2): 77–108, June 2011. ISSN 1698-6997.

[368] Mattia C. F. Prosperi, Simona Di Giambenedetto, Iuri Fanti, Genny Meini, Bianca Bruzzone, Annapaola Callegaro, Giovanni Penco, Patrizia Bagnarelli, Valeria Micheli, Elisabetta Paolini, Antonio Di Biagio, Valeria Ghisetti, Massimo Di Pietro, Maurizio Zazzi, Andrea De Luca, and ARCA cohort. A prognostic model for estimating the time to virologic failure in HIV-1 infected patients undergoing a new combination antiretroviral therapy regimen. *BMC medical informatics and decision making*, 11:40, 2011. ISSN 1472-6947. doi: 10.1186/1472-6947-11-40.

[369] Mattia C. F. Prosperi and Andrea De Luca. Computational models for prediction of response to antiretroviral therapies. *AIDS reviews*, 14(2):145–153, June 2012. ISSN 1698-6997.

[370] Andrew D. Revell, Dechao Wang, Mark A. Boyd, Sean Emery, Anton L. Pozniak, Frank De Wolf, Richard Harrigan, Julio S. G. Montaner, Clifford Lane, Brendan A. Larder, and RDI Study Group. The development of an expert system to predict virological response to HIV therapy as part of an online treatment support tool. *AIDS (London, England)*, 25(15):1855–1863, September 2011. ISSN 1473-5571. doi: 10.1097/QAD.0b013e328349a9c2.

[371] Andrew D Revell. Modelling Treatment Response Could Reduce Virological Failure in Different Patient Populations. *Journal of AIDS & Clinical Research*, 01(S5), 2013. ISSN 21556113. doi: 10.4172/2155-6113.S5-008. URL http://www.omicsonline.org/2155-6113/2155-6113-S5-008.digital/2155-6113-S5-008.html.

[372] BA Larder, D Wang, AD Revell, S Emery, F DeWolf, M Nelson, MJ Pérez-Elías, PR Harrigan, and JS Montaner. The Development of Computational Models That Accurately Predict Virological Response to HIV Therapy to Power an Online Treatment Selection Tool. Dubrovnik, Croatia, 2010.

[373] A. Altmann, M. Däumer, Y. Peres, E. Schülter, J. Büch, S.Y. Rhee, A. Sönnerborg, W.J. Fessel, R.W. Shafer, M. Zazzi, R. Kaiser, and T. Lengauer. Predicting the response to combination antiretroviral therapy: retrospective validation of geno2pheno-theo on a large clinical database. *Journal of Infectious Diseases*, 199(7):999–1006, 2009. ISSN 0022-1899. doi: 10.1086/597305.

[374] J. Bogojeska and T. Lengauer. Hierarchical Bayes model for predicting effectiveness of HIV combination therapies. *Statistical applications in genetics and molecular biology*, 11(3), 2012. ISSN 1544-6115.

[375] M Zazzi, R Kaiser, A Sönnerborg, D Struck, A Altmann, M Prosperi, M Rosen-Zvi, A Petroczi, Y Peres, E Schülter, Ca Boucher, F Brun-Vezinet, Pr Harrigan, L Morris, M Obermeier, C-F Perno, P Phanuphak, D Pillay, Rw Shafer, A-M Vandamme, K van Laethem, Amj Wensing, T Lengauer, and F Incardona. Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study): EuResist vs. experts (EVE) study. *HIV Medicine*, 12(4):211–218, April 2011. ISSN 14642662. doi: 10.1111/j.1468-1293.2010.00871.x. URL http://doi.wiley.com/10.1111/j.1468-1293.2010.00871.x.

[376] Ad Revell, Ma Boyd, D Wang, S Emery, B Gazzard, P Reiss, Ai van Sighem, Js Montaner, Hc Lane, and Ba Larder. A comparison of computational models with and without genotyping for prediction of response to second-line HIV therapy: Modelling response to second-line HIV therapy. *HIV Medicine*, 15(7):442–448, August 2014. ISSN 14642662. doi: 10.1111/hiv.12156. URL http://doi.wiley.com/10.1111/hiv.12156.

[377] A. D. Revell, D. Wang, R. Wood, C. Morrow, H. Tempelman, R. Hamers, G. Alvarez-Uria, A. Streinu-Cercel, L. Ene, A. Wensing, P. Reiss, A. I. van Sighem, M. Nelson, S. Emery, J. S. G. Montaner, H. C. Lane, B. A. Larder, on behalf of the RDI Study Group, P. Reiss, A. van Sighem, J. Montaner, R. Harrigan, T. Rinke de Wit, R. Hamers, K. Sigaloff, B. Agan, V. Marconi, S. Wegner, W. Sugiura, M. Zazzi, A. Streinu-Cercel, G. Alvarez-Uria, J. Gatell, E. Lazzari, B. Gazzard, M. Nelson, A. Pozniak, S. Mandalia, L. Ruiz, B. Clotet, S. Staszewski, C. Torti, C. Lane, J. Metcalf, M.-J. Perez-Elias, A. Carr, R. Norris, K. Hesse, E. Vlahakis, H. Tempelman, R. Barth, C. Morrow, R. Wood, L. Ene, G. Dragovic, S. Emery, D. Cooper, C. Torti, J. Baxter, L. Monno, C. Torti, J. Gatell, B. Clotet, G. Picchio, M.-P. deBethune,

266

and M.-J. Perez-Elias. An update to the HIV-TRePS system: the development of new computational models that do not require a genotype to predict HIV treatment outcomes. *Journal of Antimicrobial Chemotherapy*, 69(4):1104–1110, April 2014. ISSN 0305-7453, 1460-2091. doi: 10.1093/jac/dkt447. URL http://www.jac.oxfordjournals.org/cgi/doi/10.1093/jac/dkt447.

[378] Mattia C. F. Prosperi, Michal Rosen-Zvi, André Altmann, Maurizio Zazzi, Simona Di Giambenedetto, Rolf Kaiser, Eugen Schülter, Daniel Struck, Peter Sloot, David A. van de Vijver, Anne-Mieke Vandamme, Anders Sönnerborg, and for the EuResist and Virolab study groups. Antiretroviral Therapy Optimisation without Genotype Resistance Testing: A Perspective on Treatment History Based Models. *PLoS ONE*, 5(10):e13753, October 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0013753. URL http://dx.plos.org/10.1371/journal.pone.0013753.

[379] R. A. Fisher. On the Interpretation of &#967; 2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, 85(1):87, January 1922. ISSN 09528385. doi: 10.2307/2340521. URL http://www.jstor.org/stable/2340521?origin=crossref.

[380] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, June 2006. ISSN 01678655. doi: 10.1016/j.patrec.2005.10.010. URL http://linkinghub.elsevier.com/retrieve/pii/S016786550500303X.

[381] André Altmann. *Bioinformatical Approaches to Ranking of anti-HIV Combination Therapies and Planning of Treatment Schedules*. Doctoral Dissertation, University of Saarland, Saarbrücken, Germany, 2010.

[382] M. J. Mugavero, S. Napravnik, S. R. Cole, J. J. Eron, B. Lau, H. M. Crane, M. M. Kitahata, J. H. Willig, R. D. Moore, S. G. Deeks, M. S. Saag, and on behalf of the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS) Cohort Study. Viremia Copy-Years Predicts Mortality Among Treatment-Naive HIV-Infected Patients Initiating Antiretroviral Therapy. *Clinical Infectious Diseases*, 53(9):927–935, November 2011. ISSN 1058-4838, 1537-6591. doi: 10.1093/cid/cir526. URL http://cid.oxfordjournals.org/lookup/doi/10.1093/cid/cir526.

[383] Vincent C. Marconi, Greg Grandits, Jason F. Okulicz, Glenn Wortmann, Anuradha Ganesan, Nancy Crum-Cianflone, Michael Polis, Michael Landrum, Matthew J. Dolan, Sunil K. Ahuja, Brian Agan, Hemant Kulkarni, and Infectious Disease Clinical Research Program (IDCRP) HIV Working Group. Cumulative viral load and virologic decay patterns after antiretroviral therapy in HIV-infected subjects influence CD4 recovery and AIDS. *PLoS One*, 6(5):e17956, 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0017956.

[384] Emma D. Deeks and Caroline M. Perry. Efavirenz/Emtricitabine/Tenofovir Disoproxil Fumarate Single-Tablet Regimen (Atripla®). *Drugs*, 70(17):2315–2338, November 2012. ISSN 0012-6667, 1179-1950. doi: 10.2165/11203800-000000000-00000. URL http://link.springer.com/article/10.2165/11203800-000000000-00000.

[385] Julie C. Adkins and Stuart Noble. Efavirenz. *Drugs*, 56(6):1055–1064, September 2012. ISSN 0012-6667, 1179-1950. doi: 10.2165/00003495-199856060-00014. URL http://link.springer.com/article/10.2165/00003495-199856060-00014.

[386] Joel E. Gallant, Edwin DeJesus, José R. Arribas, Anton L. Pozniak, Brian Gazzard, Rafael E. Campo, Biao Lu, Damian McColl, Steven Chuck, Jeffrey Enejosa, John J. Toole, and Andrew K. Cheng. Tenofovir DF, Emtricitabine, and Efavirenz vs. Zidovudine, Lamivudine, and Efavirenz for HIV. *New England Journal of Medicine*, 354(3):251–260, January 2006. ISSN 0028-4793. doi: 10.1056/NEJMoa051871. URL http://dx.doi.org/10.1056/NEJMoa051871.

[387] Toni M. Dando and Lesley J. Scott. Abacavir plus Lamivudine. *Drugs*, 65(2):285–302, September 2012. ISSN 0012-6667, 1179-1950. doi: 10.2165/00003495-200565020-00010. URL http://link.springer.com/article/10.2165/00003495-200565020-00010.

[388] Monique Nijhuis, Rob Schuurman, Dorien de Jong, Remko van Leeuwen, Joep Lange, Sven Danner, Wilco Keulen, Tom de Groot, and Charles A. B. Boucher. Lamivudine-Resistant Human Immunodeficiency Virus Type 1 Variants (184v) Require Multiple Amino Acid Changes to Become Co-Resistant to Zidovudine In Vivo. *The Journal of Infectious Diseases*, 176(2):398–405, August 1997. ISSN 0022-1899, 1537-6613. doi: 10.1086/514056. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/514056.

[389] Caroline M. Perry and Julia A. Balfour. Didanosine: An Update on its Antiviral Activity, Pharmacokinetic Properties and Therapeutic Efficacy in the Management of HIV Disease. *Drugs*, 52(6):928–962, December 1996. ISSN 0012-6667. doi: 10.2165/00003495-199652060-00014. URL http://link.springer.com/10.2165/00003495-199652060-00014.

[390] Katherine F Croom and Susan J Keam. Tipranavir: A Ritonavir-Boosted Protease Inhibitor. *Drugs*, 65(12):1669–1677, 2005. ISSN 0012-6667. doi: 10.2165/00003495-200565120-00005. URL http://link.springer.com/10.2165/00003495-200565120-00005.

[391] Jean-Michel Molina, Pedro Cahn, Beatriz Grinsztejn, Adriano Lazzarin, Anthony Mills, Michael Saag, Khuanchai Supparatpinyo, Sharon Walmsley, Herta Crauwels, Laurence T Rimsky, Simon Vanveggel, and Katia Boven. Rilpivirine versus efavirenz with tenofovir and emtricitabine in treatment-naive adults infected with HIV-1 (ECHO): a phase 3 randomised double-blind active-controlled trial. *The Lancet*, 378(9787):238–246, July 2011. ISSN 01406736. doi: 10.1016/S0140-6736(11)60936-7. URL http://linkinghub.elsevier.com/retrieve/pii/S0140673611609367.

[392] Alejandro Pironti, Saleta Sierra, Rolf Kaiser, Thomas Lengauer, and Nico Pfeifer. Effects of sequence alterations on results from genotypic tropism testing. *Journal of Clinical Virology*, February 2015. ISSN 13866532. doi: 10.1016/j.jcv.2015.02.006. URL http://linkinghub.elsevier.com/retrieve/pii/S1386653215000530.

[393] Saleta Sierra, Rolf Kaiser, Nadine Lübke, Alexander Thielen, Eugen Schuelter, Eva Heger, Martin Däumer, Stefan Reuter, Stefan Esser, Gerd Fätkenheuer, Herbert Pfister, Mark Oette, and Thomas Lengauer. Prediction of HIV-1 coreceptor usage (tropism) by sequence analysis using a genotypic approach. *Journal of visualized experiments: JoVE*, (58), 2011. ISSN 1940-087X. doi: 10.3791/3264.

[394] N. Sichtig, S. Sierra, R. Kaiser, M. Daumer, S. Reuter, E. Schulter, A. Altmann, G. Fatkenheuer, U. Dittmer, H. Pfister, and S. Esser. Evolution of raltegravir resistance during therapy. *Journal of*

Antimicrobial Chemotherapy, 64(1):25–32, July 2009. ISSN 0305-7453, 1460-2091. doi: 10.1093/jac/dkp153. URL http://www.jac.oxfordjournals.org/cgi/doi/10.1093/jac/dkp153.

[395] T. Leitner, E. Halapi, G. Scarlatti, P. Rossi, J. Albert, E. M. Fenyö, and M. Uhlén. Analysis of heterogeneous viral populations by direct DNA sequencing. *BioTechniques*, 15(1):120–127, July 1993. ISSN 0736-6205.

[396] B. A. Larder, A. Kohli, P. Kellam, S. D. Kemp, M. Kronick, and R. D. Henfrey. Quantitative detection of HIV-1 drug resistance mutations by automated DNA sequencing. *Nature*, 365(6447):671–673, October 1993. ISSN 0028-0836. doi: 10.1038/365671a0. URL http://www.nature.com/doifinder/10.1038/365671a0.

[397] Takahiro Kanagawa. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*, 96(4):317–323, 2003. ISSN 1389-1723. doi: 10.1016/S1389-1723(03)90130-7.

[398] DNA Sequencing by Capillary Electrophoresis Chemistry Guide (PN 4305080) - cms_041003.pdf. URL https://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_041003.pdf.

[399] Lore Vinken, Sarah Megens, Yoeri Schrooten, Anne-Mieke Vandamme, and Kristel Van Laethem. Evaluation of the automatic editing tool RECall for HIV-1 pol and V3 loop sequences. *Journal of Virological Methods*, 193(1):135–139, October 2013. ISSN 1879-0984. doi: 10.1016/j.jviromet.2013.05.017.

[400] Richard M. Gibson, Christine L. Schmotzer, and Miguel E. Quiñones-Mateu. Next-Generation Sequencing to Help Monitor Patients Infected with HIV: Ready for Clinical Use? *Current Infectious Disease Reports*, 16(4):401, April 2014. ISSN 1523-3847. doi: 10.1007/s11908-014-0401-5.

[401] B Ewing and P Green. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome research*, 8(3):186–194, March 1998. ISSN 1088-9051.

[402] Erasmus Smit. Antiviral resistance testing. *Current Opinion in Infectious Diseases*, 27(6):566–572, December 2014. ISSN 1473-6527. doi: 10.1097/QCO.0000000000000108.

[403] Bram Vrancken, Sébastian Lequime, Kristof Theys, and Philippe Lemey. Covering all bases in HIV research: unveiling a hidden world of viral evolution. *AIDS reviews*, 12(2):89–102, June 2010. ISSN 1698-6997.

[404] Luisa Barzon, Enrico Lavezzo, Valentina Militello, Stefano Toppo, and Giorgio Palù. Applications of Next-Generation Sequencing Technologies to Diagnostic Virology. *International Journal of Molecular Sciences*, 12(11):7861–7884, November 2011. doi: 10.3390/ijms12117861. URL http://www.mdpi.com/1422-0067/12/11/7861.

[405] Mattia CF Prosperi, Laura Bracciale, Massimiliano Fabbiani, Simona Di Giambenedetto, Francesca Razzolini, Genny Meini, Manuela Colafigli, Angela Marzocchetti, Roberto Cauda, Maurizio Zazzi, and Andrea De Luca. Comparative determination of HIV-1 co-receptor tropism by Enhanced Sensitivity Trofile, gp120 V3-loop RNA and DNA genotyping. *Retrovirology*, 7(1):56, 2010. ISSN 1742-4690. doi:

10.1186/1742-4690-7-56. URL http://www.retrovirology.com/content/7/1/56.

[406] P. Recordon-Pinson, C. Soulie, P. Flandre, D. Descamps, M. Lazrek, C. Charpentier, B. Montes, M.-A. Trabaud, J. Cottalorda, V. Schneider, L. Morand-Joubert, C. Tamalet, D. Desbois, M. Mace, V. Ferre, A. Vabret, A. Ruffault, C. Pallier, S. Raymond, J. Izopet, J. Reynes, A.-G. Marcelin, B. Masquelier, , and the ANRS AC11 Resistance Study Group. Evaluation of the Genotypic Prediction of HIV-1 Coreceptor Use versus a Phenotypic Assay and Correlation with the Virological Response to Maraviroc: the ANRS GenoTropism Study. *Antimicrobial Agents and Chemotherapy*, 54(8):3335–3340, June 2010. ISSN 0066-4804, 1098-6596. doi: 10.1128/AAC.00148-10. URL http://aac.asm.org/cgi/doi/10.1128/AAC.00148-10.

[407] Eva Poveda, Roger Paredes, Santiago Moreno, José Alcamí, Juan Córdoba, Rafael Delgado, Félix Gutiérrez, Josep M Llibre, Miguel García Deltoro, José Hernández-Quero, Federico Pulido, José Antonio Iribarren, and Federico García. Update on clinical and methodological recommendations for genotypic determination of HIV tropism to guide the usage of CCR5 antagonists. *AIDS reviews*, 14(3):208–217, September 2012. ISSN 1698-6997.

[408] Trofile® DNA (HIV Tropism Test). URL http://www.monogrambio.com/hiv/tropism/trofile-dna/.

[409] T. Lengauer, O. Sander, S. Sierra, A. Thielen, and R. Kaiser. Bioinformatics prediction of HIV coreceptor usage. *Nature Biotechnology*, 25(12):1407–1410, 2007. ISSN 1087-0156. doi: 10.1038/nbt1371.

[410] J Symons, L Vandekerckhove, R Paredes, C Verhofstede, R Bellido, E Demecheleer, P M van Ham, S F L van Lelyveld, A J Stam, D van Versendaal, M Nijhuis, and A M J Wensing. Impact of triplicate testing on HIV genotypic tropism prediction in routine clinical practice. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 18(6):606–612, June 2012. ISSN 1469-0691. doi: 10.1111/j.1469-0691.2011.03631.x.

[411] Rachel A McGovern, Alexander Thielen, Simon Portsmouth, Theresa Mo, Winnie Dong, Conan K Woods, Xiaoyin Zhong, Chanson J Brumme, Douglass Chapman, Marilyn Lewis, Ian James, Jayvant Heera, Hernan Valdez, and P Richard Harrigan. Population-based sequencing of the V3-loop can predict the virological response to maraviroc in treatment-naive patients of the MERIT trial. *Journal of acquired immune deficiency syndromes (1999)*, 61(3):279–286, November 2012. ISSN 1944-7884. doi: 10.1097/QAI.0b013e31826249cf.

[412] Rachel A McGovern, Alexander Thielen, Theresa Mo, Winnie Dong, Conan K Woods, Douglass Chapman, Marilyn Lewis, Ian James, Jayvant Heera, Hernan Valdez, and P Richard Harrigan. Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies. *AIDS (London, England)*, 24(16):2517–2525, October 2010. ISSN 1473-5571. doi: 10.1097/QAD.0b013e32833e6cfb.

[413] Lpr Vandekerckhove, Amj Wensing, R Kaiser, F Brun-Vezinet, B Clotet, A De Luca, S Dressler, F Garcia, Am Geretti, T Klimkait, K Korn, B Masquelier, Cf Perno, J Schapiro, V Soriano, A Sönnerborg, àM Vandamme, C Verhofstede, H Walter, M Zazzi, and Ca Boucher. Consensus statement of the European

guidelines on clinical management of HIV-1 tropism testing. *Journal of the International AIDS Society*, 13(Suppl 4):O7, 2010. ISSN 1758-2652. doi: 10.1186/1758-2652-13-S4-O7. URL http://archive.biomedcentral.com/10.1186/1758-2652/13/S4/07.

[414] L P R Vandekerckhove, A M J Wensing, R Kaiser, F Brun-Vézinet, B Clotet, A De Luca, S Dressler, F Garcia, A M Geretti, T Klimkait, K Korn, B Masquelier, C F Perno, J M Schapiro, V Soriano, A Sönnerborg, A-M Vandamme, C Verhofstede, H Walter, M Zazzi, C A B Boucher, and European Consensus Group on clinical management of tropism testing. European guidelines on the clinical management of HIV-1 tropism testing. *The Lancet infectious diseases*, 11(5):394–407, May 2011. ISSN 1474-4457. doi: 10.1016/S1473-3099(10)70319-4.

[415] Hauke Walter, J Eberle, and Hans Müller. Empfehlung zur Bestimmung des HIV-1-Korezeptor-Gebrauchs (DAIG Recommendations for the HIV-1 Tropism Testing), October 2013. URL http://www.daignet.de/site-content/hiv-therapie/leitlinien-1/Empfehlungen%20zur%20Bestimmung%20des%20HIV_1_Korezeptor_Gebrauchs.pdf.

[416] Stéphanie Raymond, Patricia Recordon-Pinson, Adrien Saliou, Pierre Delobel, Florence Nicot, Diane Descamps, Anne-Geneviève Marcelin, Philippe Flandre, Vincent Calvez, Bernard Masquelier, Jacques Izopet, and ANRS AC11 Resistance Study Group. Improved V3 genotyping with duplicate PCR amplification for determining HIV-1 tropism. *The Journal of Antimicrobial Chemotherapy*, 66(9):1972–1975, September 2011. ISSN 1460-2091. doi: 10.1093/jac/dkr224.

[417] Rachel A McGovern, P Richard Harrigan, and Luke C Swenson. Genotypic inference of HIV-1 tropism using population-based sequencing of V3. *Journal of visualized experiments: JoVE*, (46), 2010. ISSN 1940-087X. doi: 10.3791/2531.

[418] Conan K Woods, Chanson J Brumme, Tommy F Liu, Celia K S Chui, Anna L Chu, Brian Wynhoven, Tom A Hall, Christina Trevino, Robert W Shafer, and P Richard Harrigan. Automating HIV drug resistance genotyping with RECall, a freely accessible sequence analysis tool. *Journal of clinical microbiology*, 50(6):1936–1942, June 2012. ISSN 1098-660X. doi: 10.1128/JCM.06689-11.

[419] A. D. Johnson. An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics*, 26(10):1386–1389, May 2010. ISSN 1367-4803, 1460-2059. doi: 10.1093/bioinformatics/btq098. URL http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btq098.

[420] Sanger Sequencing Instruments and Accessories. URL https://www.thermofisher.com/de/de/home/life-science/sequencing/sanger-sequencing/sanger-sequencing-technology-accessories.html.

[421] G Myers. A dataset generator for whole genome shotgun sequencing. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, pages 202–210, 1999. ISSN 1553-0833.

[422] Sherry McLaughlin, Luke C. Swenson, Shengbo Hu, Paul Hughes, P. Richard Harrigan, Robert W. Coombs, and Lisa M. Frenkel. Development of a Novel Codon-Specific Polymerase Chain Reaction for

the Detection of CXCR4-Utilizing HIV Type 1 Subtype B. *AIDS Research and Human Retroviruses*, 29(5):814–825, May 2013. ISSN 0889-2229, 1931-8405. doi: 10.1089/aid.2012.0024. URL http://online.liebertpub.com/doi/abs/10.1089/aid.2012.0024.

[423] Kristel Van Laethem, Kristof Theys, and Anne-Mieke Vandamme. HIV-1 genotypic drug resistance testing: digging deep, reaching wide? *Current Opinion in Virology*, 14:16–23, October 2015. ISSN 1879-6265. doi: 10.1016/j.coviro.2015.06.001.

[424] Jonathan Z. Li and Daniel R. Kuritzkes. Clinical implications of HIV-1 minority variants. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 56(11):1667–1674, June 2013. ISSN 1537-6591. doi: 10.1093/cid/cit125.

[425] Hsin-Yun Sun, Wang-Huei Sheng, Mao-Song Tsai, Kuan-Yeh Lee, Sui-Yuan Chang, and Chien-Ching Hung. Hepatitis B virus coinfection in human immunodeficiency virus-infected patients: A review. *World Journal of Gastroenterology : WJG*, 20(40):14598–14614, October 2014. ISSN 1007-9327. doi: 10.3748/wjg.v20.i40.14598. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4209527/.

[426] Vincent Soriano, Pablo Labarga, Carmen de Mendoza, José M. Peña, José V. Fernández-Montero, Laura Benítez, Isabella Esposito, and Pablo Barreiro. Emerging challenges in managing hepatitis B in HIV patients. *Current HIV/AIDS reports*, 12(3):344–352, September 2015. ISSN 1548-3576. doi: 10.1007/s11904-015-0275-7.

[427] Mattias Mandorfer, Philipp Schwabl, Sebastian Steiner, Thomas Reiberger, and Markus Peck-Radosavljevic. Advances in the management of HIV/HCV coinfection. *Hepatology International*, 10(3):424–435, May 2016. ISSN 1936-0541. doi: 10.1007/s12072-015-9691-4.

[428] Elizabeth Phillips and Pavlos. Individualization of antiretroviral therapy. *Pharmacogenomics and Personalized Medicine*, page 1, December 2011. ISSN 1178-7066. doi: 10.2147/PGPM.S15303. URL http://www.dovepress.com/individualization-of-antiretroviral-therapy-peer-reviewed-article-PGPM.