From DEPARTMENT OF LABORATORY MEDICINE
Karolinska Institutet, Stockholm, Sweden

ANTIRETROVIRAL TREATMENT OUTCOME AND MOLECULAR EPIDEMIOLOGY OF HIV-1 SUBTYPE C DRUG RESISTANCE IN ETHIOPIA

Nigus Fikrie Telele

Karolinska Institutet

Stockholm 2018

Published by Karolinska Institutet.
Printed by Printed by E-Print AB 2018

ISBN 978-91-7831-153-8

ANTIRETROVIRAL TREATMENT OUTCOME AND MOLECULAR EPIDEMIOLOGY OF HIV-1 SUBTYPE C DRUG RESISTANCE IN ETHIOPIA

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Nigus Fikrie Telele

Principal Supervisor:
Professor Anders Sönnerborg
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Co-supervisor:
Associate Professor Gaetano Marrone
Karolinska Institutet
Department of Public Health Sciences

Opponent:
Professor Rune Andersson
University of Gothenburg
Department of Biomedicine
Division of Infectious Diseases

Examination Board:
Professor Matti Ristola
University of Helsinki
Department of Infectious Diseases

Associate Professor Annelie Tjernlund
Karolinska Institutet
Department of Medicine Solna
Division of Infectious Diseases

Associate Professor Carl Johan Treutiger
Karolinska Institutet
Department of Medicine Huddinge
Center of Infectious Medicine

To my family!

ABSTRACT

Over the last decade, antiretroviral therapy (ART) has been rolled out in Ethiopia at a large scale. However, the outcome and aspects of drug resistance have been poorly studied at the national level. Also, the epidemic has been reported to be dominated by HIV-1 subtype C (HIV-1C), but this has not been verified at a nationwide level. In this thesis, I **aimed** at studying the outcome of ART and molecular epidemiology of HIV-1C drug resistance in Ethiopia using the first countrywide HIV-1 cohort. Data and plasma samples were collected from 874 adult and adolescent patients (age ≥ 14 years; 60% females), recruited from seven university hospitals during 2009 – 2011.

In **Paper I**, by on-treatment (OT) and intention-to-treat (ITT) analyses we determined the rate of treatment failure at month 6 and 12. Four multivariable logistic regression models were developed to identify baseline predictors of treatment outcome. OT analysis identified treatment failure in 8% and 7%, whereas with ITT analysis the figures were 23% and 34% at month 6 and 12, respectively. Hence, our study indicated early death and lost to follow up (LTFU) as the main risk factors for poor treatment outcome. In addition to the well-known baseline factors of ART outcomes, study site was identified as a strong predictor of failure, where regional hospitals had higher proportions of treatment failure and LTFU as compared to a national tertiary level hospital in the capital city.

In **Paper II**, we assessed surveillance drug resistance mutations (sDRM) in major and minor viral populations by population-based Sanger sequencing (PBSS) and next-generation sequencing (NGS). The short-term impact of sDRM on the outcome of ART was also assessed. NGS detected sDRM associated with reverse transcriptase inhibitors (RTIs) and protease inhibitors (PIs) more frequently than PBSS as well as major integrase strand transfer inhibitors (INSTIs) DRMs in minor viral variants. The baseline RTI-DRMs were associated with treatment failure at month 6 and 12.

In **Paper III**, by genotypic analysis of the V3-loop of *env*, we described the HIV-1C molecular epidemiology and co-receptor tropism trend. The epidemic has been still found to be monophylogenetically dominated by CCR5-tropic HIV-1C even in patients with advanced immunodeficiency, although a slight increasing temporal trend was observed in CXCR4-containing strains.

In **Paper IV**, genotypic resistance testing (GRT) of the HIV-1 *pol* gene described acquired DRMs at month 6 and 12 and near-full length genome (NFLG) analysis assessed amino acid changes in the *gag*, *pol*, *vif*, *vpr*, *tat*, *vpu*, and *nef* genes in paired baseline and month 6 samples of virologic failures. Broad major RTI- DRMs were detected in most failure patients. K65R, at a high rate, was identified only in TDF treated patients. The NFLG assay described all target regions of interest for HIVDR.

In conclusion, early death and LTFU are major contributors for ART failure in Ethiopia rather than detectable viremia; the universal test and treat strategy could possibly improve the treatment outcome. There is a geographical variation in ART outcome, which calls the need for provision of more support at regional hospitals. HIV-1C still dominates monophylogenetically the epidemic in Ethiopia. PBSS reports a broad DRM, but underestimates prevalence of the transmitted drug resistance (TDR). NGS identifies an even higher rate of DRMs including in the minor viral variants. Our NFLG assay covers all relevant target genes and is an attractive cost-efficient alternative for HIVDR surveillance.

LIST OF SCIENTIFIC PAPERS

I. **Telele NF**, Kalu AW, Marrone G, Gebre-Selassie S, Fekade D, Tegbaru B, Sönnerborg A. (2018). Baseline predictors of antiretroviral treatment failure and lost to follow up in a multicenter countrywide HIV-1 cohort study in Ethiopia. PLoS ONE 13(7): e0200505. https://doi.org/10.1371/journal.pone.0200505.

II. **Telele NF**, Kalu AW, Gebre-Selassie S, Fekade D, Abdurahman S, Marrone G, Neogi U, Tegbaru B, and Sönnerborg A. (2018). Pretreatment drug resistance in a large countrywide Ethiopian HIV-1C cohort: a comparison of Sanger and high-throughput sequencing. Sci Rep 8(1): 7556. doi:10.1038/s41598-018-25888-6.

III. Kalu AW, **Telele NF**, Gebreselasie S, Fekade D, Abdurahman S, Marrone G, and Söonnerborg A. (2017). Monophylogenetic HIV-1C epidemic in Ethiopia is dominated by CCR5-tropic viruses-an analysis of a prospective country-wide cohort. BMC Infect Dis 17(1): 37. doi:10.1186/s12879-016-2163-1.

IV. **Telele NF**, Kalu AW, Gebreselasie S, Fekade D, Marrone G, Neogi U, Grossman S, Tegbaru B, and Sönnerborg A. HIV drug resistance and genetic diversity among patients failing first-line therapy in the first large countrywide Ethiopian HIV cohort assessed by Sanger- and near-full length sequencing. Manuscript submitted.

CONTENTS

LIST OF ABBREVIATIONS

3TC	lamivudine
ABC	abacavir
ACM	Advanced Clinical Monitoring of ART in Ethiopia
ADR	acquired drug resistance
AIDS	acquired immunodeficiency syndrome
ART	antiretroviral therapy
ARV	antiretroviral
AS-PCR	allele-specific polymerase chain reaction
BMI	body mass index
cART	combined antiretroviral therapy
CRF	circulating recombinant form
d4T	stavudine
ddI	didanosine
DNA	deoxyribonucleic acid
DRM	drug resistance mutations
EFV	efavirenz
EHNRI	Ethiopian Health and Nutrition Research Institute
env	envelope
EPHI	Ethiopian Public Health Institute
ETR	etravirine
FDC	fixed-dose combinations
FPR	false positive rate
FTC	emtricitabine
G2P	Geno2Pheno
gp	glycoprotein
GRT	genotypic resistance test
HIV-1	human immunodeficiency virus type 1
HIV-1C	HIV-1 subtype C
HIVDR	human immunodeficiency virus drug resistance
IAS	International Antiviral Society–USA
IN	integrase
INSTI	integrase strand transfer inhibitor
ITT	intention-to-treat
kb	kilobases
LMIC	low- and middle-income countries

LTFU	lost to follow up
MTCT	mother to child transmission
NFLG	near-full length genome
NGS	next-generation sequencing
NNRTI	non-nucleoside reverse transcriptase inhibitors
NRTI	nucleoside reverse transcriptase inhibitors
NVP	nevirapine
OI	opportunistic infection
OT	on-treatment
PBSS	population-based Sanger sequencing
PCP	*Pneumocystis carinii (jiroveci) pneumonia*
PCR	polymerase chain reaction
PDR	pretreatment drug resistance
PI	protease inhibitor
PLHIV	people living with HIV
PR	protease
RLS	resource limited settings
RNA	ribonucleic acid
RPV	rilpivirine
RRS	resource rich settings
RT	reverse transcriptase
SIV	simian immunodeficiency virus
sSA	sub-Saharan Africa
TAM	thymidine analog mutation
TASH	Tikur Anbessa Specialised Hospital
TB	tuberculosis
TDR	transmitted drug resistance
URF	unique recombinant form
V3	third variable loop
vGWAS	viral wide genome association study
VL	viral load
WHO	World Health Organization
ZDV	Zidovudine

1 INTRODUCTION

Human immunodeficiency virus (HIV), an etiologic agent of acquired immunodeficiency syndrome (AIDS), is a chronic retroviral infection. Since AIDS was first reported in 1981 in the United States [1], the HIV epidemic has affected lives of millions of people, largely in sub-Saharan Africa (sSA). Like most other sSA, Ethiopia has experienced a severe HIV epidemic and remains to be one of the highly affected countries in the region [2]. Since its estimated introduction into the country in 1970s, the epidemic in Ethiopia has been monophylogenetically dominated by HIV-1 subtype C (HIV-1C), which was first identified in this country [3]. Currently, HIV-1C accounts for around half of the world HIV epidemic [4].

In response to the global HIV epidemic, antiretroviral therapy (ART) has become widely available worldwide including sSA countries, the hardest hit region by the epidemic. Since 2005, ART has been rolled out at a large scale in Ethiopia [5]. Subsequently, a significant decline in the epidemic has been observed over the years along with a decline in the national prevalence [6, 7]. However, a high number of people are still living with the virus throughout the country [8].

1.1 THE HIV PANDEMIC IN BRIEF

HIV has spread more rapidly than most other disease causing agents in recent human history and it still remains to be one of the threats to human health and development [9]. Few years after the first AIDS cases with occurrence of *Pneumocystis carinii (jiroveci) pneumonia* (PCP) among previously healthy homosexual Americans were reported [1], studies identified retrovirus [human T-lymphotropic virus type III/lymphadenopathy- associated virus (HTLV-III/LAV)] as the causative agent of AIDS [10, 11], which was later renamed as HIV [12].

During the last three and half decades the epidemic has claimed more than 35 million human lives in addition to the 36.7 million people living with HIV (PLHIV) as of 2016 [13]. Of the 36.7 million PLHIV worldwide, sSA accounted for nearly 70% (25.5 million). The UNAIDS reported that the eastern and southern Africa region alone harbor 19.4 million PLHIV. The report also showed that through 2010 to 2016, new HIV infections declined by 29% in this region, although the region accounted for 43% (790 000) of the 1.8 million new infections globally reported in the year 2016. During those years AIDS-related deaths also declined by 42% in this region, but still 420 000 of the one million total global AIDS-related deaths in the year 2016 were from this region.

1.2 ORIGIN AND SPREAD OF THE VIRUS

Studies have shown that the two types of HIV, HIV-1 and HIV-2, have evolved from simian immunodeficiency viruses (SIV) through multiple cross-species zoonotic transmissions [14]. The initial zoonotic cross-species transmissions believed to have occurred in west and central Africa and the virus might have entered into the human population at least after the early 1900s [15]. HIV-1 and HIV-2 evolved in humans via different pathways, where transmission

event involving SIVcpz from chimpanzees gave rise to HIV-1 in southeastern Cameroon [16] and another transmission event involving SIVsmm from sooty mangabey gave rise to HIV-2 [16, 17].

In addition to their evolutionary pathways, HIV-1 and HIV-2 have distinguished genomic organizations and both of them have already diversified into different subtypes [14, 16]. Although both types still co-circulate in West Africa, HIV-1 successfully spread throughout the world, whereas HIV-2 mainly remains there though limited number of cases are also reported from other parts of the world [18]. In addition, HIV-2 is less virulent characterised by much longer asymptomatic period, low plasma RNA load, slower CD4 cell decline, and cause lower mortality rate [19].

On the other hand, since its initial transmission to the human population in Africa during the early 20th Century, HIV-1 has evolved genetically and phenotypically due to its high genetic variabilities through mutations and recombination [20]. Phylogenetic analyses divide HIV-1 into four groups including group M (major), group O (outlier), group N (non-M, non-O), and group P (putative, a newly identified group) [21]. Group M, the pandemic, accounts for over 90% of the infections, whereas groups O and N are restricted in west and central Africa, and are responsible only for small proportion of infections. Group P was identified from individuals who were originally from Cameroon [22].

The HIV-1 M group comprises nine subtypes including A–D, F–H, J and K as well as several circulating recombinant forms (CRFs) and multiple unique recombinant forms (URFs) [23, 24]. In addition, some subtypes, such as subtype A and subtype F, are subdivided further. Generally, the genetic variation between subtypes and within subtypes range from 25% to 35% and 15% to 20%, respectively [22]. The subtypes A–D, subtype G, CRF 01_AE and CRF 02_AG account for over 90% of the global HIV epidemic [24]. Of them, HIV-1C represents for half of the global epidemic followed by subtype A (12%), subtype B (11%), CRF02_AG (8%), CRF01_AE (5%), subtype G (5%), and subtype D (2%) in descending order [4, 24].

The group M subtypes have rapidly expanded within the western and central Africa with subsequent establishment of multiple epidemics worldwide [22]. Generally, subtype distribution varies widely across the globe, but the greatest diversity occurred in sSA. Central Africa harbors an extremely high diversity, where all known HIV subtypes, many CRFs and unique strains are circulating, reflecting the age of the epidemic in the region [20, 24, 25].

The global expansion of the different subtypes has been partly linked to the advent of some specific modes and routes of transmission in different parts of the world. For instance, intravenous drug uses led to rapid spread of CRF01_AE in southeast Asia during 1980s [26] and subtype A in east Europe and Russia in 1990s [27]. In parallel, subtype B transmission believed to occur among homosexual men in North America and Europe at 1980s and it has been the predominant subtype in Europe, North America and Australia [23]. On the other hand, HIV-1C appeared to have slowly emerged [22, 27].

In addition, some studies have reported variation in the transmission efficiency of different subtypes. Indeed the difference in the epidemiology of the different HIV subtypes such as the rapid spread of HIV-1C in Africa and India could potentially be linked to the fact that this subtype has high viral load in the genital tract [28] as well as its efficient transmission in the heterosexual population than other subtypes [29].

The HIV-1C, which was first isolated in Ethiopia in the late 1980s by Prof Sönnerborg group [3], is highly predominant in the southern African epidemics and the Indian subcontinent [20, 23, 24, 30]. Phylogenetic analyses and historical inference also show the introduction of HIV-1C into the southern American region including Brazil from several African countries including Ethiopia [31]. However, unlike other sSA countries, even the neighboring East African countries, the HIV-1C has monophylogenetically dominated in Ethiopia since its introduction into the country in the 1970s, although the reason remains to be explained [32].

1.3 THE HIV EPIDEMIC IN ETHIOPIA

Like other sSA countries, Ethiopia has been experiencing a severe HIV epidemic [33]. Although Ethiopia still continues to be one of the most heavily affected countries, recent figures revealed that the incidence of HIV infection has significantly decreased over the years along with a decline in the national prevalence [6, 7, 34]. Yet, in 2017 about 740 000 PLHIV were estimated to live in the country, which corresponds to adult prevalence of 1.16% [8, 35]. HIV-1C has been claimed to dominate the epidemic in the country although no broad analysis of the whole country has been done. Within the Ethiopian HIV-1C, subclusters are leading the diversity into main C and C' (some literatures labeled respectively as C9 and C10) [4], which are co-circulating in the country [36, 37].

1.3.1 How was HIV introduced and spread in Ethiopia?

Although there is no firm evidence on when and how the HIV was introduced into Ethiopia, the first two HIV positive cases were identified through analysis of serum samples collected in 1984 [38] and subsequently the first clinically overt AIDS cases were diagnosed in 1986 [39]. The first study, which was conducted with the aim of defining when HIV was introduced into Ethiopia, stratified the samples into four groups. In groups 1 and 2, which included over 900 sera from 1982 and 1983, no samples were found to be HIV positive, whereas out of the 267 sera collected from 1984 (group 3), two samples were found to be positive and declared as the first to be diagnosed in the country followed by other 13 positive patients in group 4 in which 528 patients had been sampled during 1985 – 1987 at Tikur Anbessa Specialised Hospital (TASH) [38].

The virus is believed to have been introduced into the country at least a couple of years later than in most other sSA, but the feature of the epidemic generally resembles that of other eastern African countries [40]. The introduction of the virus into the country during the early stage of the epidemic could be considered as part of the transcontinental spread with movement of people in relation to refugees, military missions, entrance of non-governmental international organizations into the capital city, Addis Ababa [38, 40].

Although earlier studies failed to report from where the virus was introduced into Ethiopia, some reports with phylogenetic analysis claimed that the two HIV-1C clusters C and C' resembles to C9 from southern African (south Africa and Malawi) and C10 from other eastern African countries (Kenya, Djubuti, Burundi, and Uganda), respectively [4]. This could potentially demonstrate the fact that the two co-circulating subclusters of HIV-1C would have been introduced from other African countries through the transcontinental movement.

Again, like most other African countries, the infection is predominantly acquired through heterosexual contacts and perinatal transmission [41]. Early studies showed a high HIV prevalence among commercial sex workers followed by long truck drivers and soldiers, who are claimed to be instrumental for the most majority of the transmission in the country [40]. During the early stage of the epidemic the country was in a devastating civil war, which might have fueled the spread of the infection because there were a large displacement and movement of people, particularly militias, along the major transport routes, where commercial sex work is available. Mother to child transmission (MTCT) also played a role to a certain extent in expanding the epidemic in the country [2].

Even if the country has achieved promising progresses in preventing new HIV infections over the past decade with a decline in the national prevalence, still much remains to be done. For example, recent reports emphasised the need for sustainable efforts in preventing new infection among most-at-risk population (MARP) groups such as commercial sex workers, mobile population, seasonal farm workers, those working in constructions, as well as sero-discordant couples who could potentially spread the virus [6].

1.3.2 Factors influencing the HIV epidemic in Ethiopia

In general, the epidemiology of HIV is greatly influenced by a number of socio-economic, cultural, political, and geographical factors. In Ethiopia the spread and distribution of HIV has also been hugely affected by these factors [41]. The EDHS 2011 report showed nearly twice higher adult prevalence among women compared to men (1.9% vs 1.0%), which could be due to several biological, socio-economical and socio-cultural factors that could negatively affect the capacities of women in protecting themselves from infection [42].

In addition, less proportions of women have been shown to have sufficient knowledge towards HIV and AIDS than men [43-45], which could potentially lead women to disproportionally be affected by the virus. Traditionally arranged marriage practices with elder men increase risks of HIV infections for younger women through hindering their chances for open discussions on sexual relationship between the couples. Urban areas also have been shown to have seven folds higher prevalence as compared to rural areas in Ethiopia (4.2% vs 0.6%) [41] .

In the early stages of the epidemic a lack of awareness among health service providers led to the multiple uses of inadequately sterilized or unsterilized hypodermics and to blood transfusions without screening for HIV. Such practices are believed to have substantially

contributed to the spread of the virus in Ethiopia [38]. Like in other sSA, changes in tendencies towards sexual behaviors, especially towards extramarital relations and also multiple marriages contributed noticeably to the epidemic in the country. Further, economic problems particularly among youth group of the society together with gender inequities could lead females to start commercial sex work, which potentially predominates the transmission dynamics in urban areas [40].

HIV coinfection with tuberculosis (TB) is a major problem in sSA as the region hosts over three-fourths of the coinfection cases worldwide [46, 47]. Ethiopia is one of the 22 countries with high TB burden [47]. This has been associated with special diagnostic and therapeutic challenges among TB-HIV co-infected patients [48], which could have bilateral adverse effects in the epidemiology of the two public health important diseases. This could lead such patients to be diagnosed late with advanced diseases, even in the ART era. Studies also reported that most people diagnosed positive for HIV with severe diseases and low CD4 count (mostly < 200 cells/µl) [48-50]. Hence, the HIV epidemic is not only influenced by the virus, but also by other disease causing agents.

1.4 VIROLOGY OF HIV-1

HIV is a retrovirus belonging to genus Lentivirus in the Retroviridae family, Orthoretrovirinae subfamily [51]. It infects and replicates in a range of CD4+ host cells through reverse transcription with high rates of mutations that lead to a very heterogenous viral population [52].

1.4.1 HIV-1 structure, genomic organization, and genetic diversity

A mature HIV-1 measures about 100 nm in diameter and has a spherical shape. Its structure and genomic organization typically follow the retrovirus family that comprises a single stranded, positive sense ribonucleic acid (RNA) genome of approximately 9.7 kilobases (kb) [53]. The virus genome comprises nine genes, which encode 15 proteins as shown below in Figure 1.

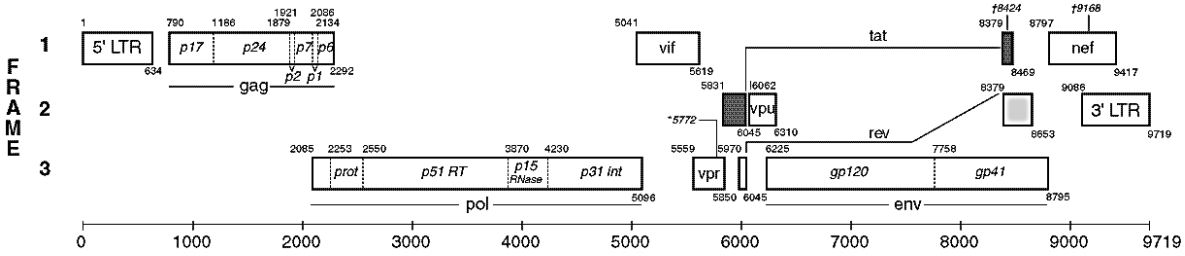

Figure 1. HIV-1 gene map. The rectangles represent open reading frames. The small number in the upper left corner of each rectangle indicates where the gene starts. Reference positions are based on HXB2 numbering system. Source: HIV Los Alamos Database, available at https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

The virus envelope is a lipid bilayer, which is derived from host cell membrane while the new virus particle buds [54, 55]. Each envelope subunit comprises noncovalently associated

glycoproteins (gp120 and gp41), which are used in the virus-cell attachment upon infection of the cell. Structurally, gp120 is divided into five loops of high genetic variability (V1–V5) interspersed with less variable/constant (C1–C5) domain. Hence, envelope is the most variable component of the virus. The V region's variability may result the envelope functionality, which can be demonstrated by the third variable loop (V3) region. Changes in amino acid within the V3 region could affect the virus co-receptor usage. The high genetic variability in its envelope confers a complex antigenic diversity for the HIV-1 [55].

Although all are required for a successful viral replication, the nine HIV genes can be classified into three functional groups: three major genes (*gag*, *pol* and *env*) code for structural proteins and the enzymes while the remaining six genes code for regulatory proteins (*tat* and *rev*) and accessory proteins (*vif*, *vpr*, *vpu* and *nef*) [53]. The *pol* gene encodes for three enzymes necessary for viral replication: reverse transcriptase (RT), integrase (IN) and protease (PR) [56].

1.4.2 HIV-1 life-cycle and pathogenesis

Similar to other animal retroviruses, the life cycle of HIV-1 is a complex process that relies upon the transcription and translation machinery of its host cell. Seven steps occur in the HIV life cycle: binding, fusion, reverse transcription, integration, transcription, assembly, and budding [56, 57]. Binding occurs to the CD4 receptor and chemokine co-receptors (the virus co-receptor usage is briefly described below). The binding of gp120 to cell surface CD4 leads to the creation of a binding site for a chemokine receptor [56, 57]. This initiates conformational changes in the gp120-gp41 unit complex [57], which allows the virus to enter into the CD4 cell and release its genetic material into cytoplasm [56]. Then, viral RNA is retro-transcribed and the provirus integrates into the cellular genome; virus proteins synthesized, and finally the virus is assembled and budding occurs [56].

Once HIV enters host cells, it replicates very rapidly producing a viral burst that infects many CD4 cells and involves all lymphatic tissues. There is a rapid turnover of CD4 cells that ultimately leads to their destruction and to a change in lymphoid tissues that prevent immune responses [56]. As the number of CD4 cells decreases, the immune system becomes increasingly compromised. Following the acute HIV infection, most patients enter a prolonged, asymptomatic phase characterised by a progressive decline in the CD4 count. AIDS-defining opportunistic infections (OIs) and malignancies become increasingly common [56]. AIDS is defined by the occurrence of any of more than 20 OIs or HIV-related cancers. The most common manifestations of AIDS include TB, PCP, as well as debilitation weight loss, diarrhea, HIV dementia, Kaposi's sarcoma and lymphomas [58].

1.4.3 HIV-1 co-receptor tropism

As described above, HIV-1 uses its gp120 to bind to and infect $CD4^+$ cells. In order to facilitate its entry, the virus subsequently uses chemokine co-receptors additionally. There are three HIV-1 phenotypes depending on the virus ability to use different co-receptors, including R5-tropic (those using CCR5 chemokine receptor), X4-tropic (those using CXCR4

chemokine receptor), or dual-tropic (R5X4 variants using both CCR5 and CXCR4 chemokine receptors) strains [59]. The dual-tropic variants further classified as dual-R (more efficiently uses the CCR5 than CXCR4) or dual-X (more efficiently uses CXCR4 than CCR5) [60].

HIV-1 mainly targets primary T-cells and macrophages, but both CCR5 and CXCR4 co-receptors are expressed on all relevant target cells for the virus including dendritic cells. The co-receptor specificity is one of the critical factors to determine the virus ability in depleting the $CD4^+$ T-lymphocytes and pathogenesis, where a shift from CCR5 to CXCR4 usage is a common phenomenon at later stage of the disease [61]. In general, R5-tropic virus strains dominate the virus population during the primary infection [62]. Pure X4-tropic strains substantially contributes to the population at a more advanced stage of the disease [63]. Hence, the virus population usually contains a mixture of R5- and X4-tropic variants at later stage of the disease [64].

However, studies have revealed variations in the utilization of CCR5 and/or CXCR4 co-receptors across different HIV-1 subtypes, mainly later in the course of infection [65]. For example, HIV-1 subtype B emerges to use CXCR4 in approximately 40%–50% of infected individuals and subsequent shift in the virus population with increasing X4-tropic strains [63, 64]. On the other hand, only a limited number of X4-tropic strains of HIV-1C have been reported and R5-tropic strains dominate throughout the course of infection, including late stage at AIDS [32, 66].

In addition, although failing to clearly show how the tropism dynamics take place, studies have revealed a higher proportion of CXCR4 using strains among treatment experienced HIV-infected patients as compared to ART naïve individuals [67]. On top of a rapid CD4 cell count decline and enhanced pathogenesis, the emergence of X4-tropic strains have been shown to be associated with a poorer treatment response to ART [67, 68] and reduced survival time in untreated individuals [21, 69]. However, it is not clear whether X4-tropic strains are inherently responsible for the poor prognosis and inferior therapeutic response as the emergence of such variants is associated with advanced disease progression accompanied by rapid decline in CD4 cell count.

A large number of studies have revealed that the V3 region of gp120 is a critical determinant of co-receptor usage of HIV-1, although other regions have also been implicated [70]. Even though, most available data are generated from HIV-1 subtype B isolates, variations in amino acids within the V3 loop strongly influence the co-receptor usage of the virus [71]. This suggests that amino acid changes at certain specific positions in the V3 region play the most important roles in the interaction of gp120 with the co-receptors.

1.5 HIV TREATMENT AND MONITORING

Through extensive scientific research activities towards combating the global epidemic a number of potent antiretrovirals (ARVs) have been discovered for the treatment of HIV-1, which have turned the infection into a chronic and manageable disease [72].

1.5.1 Antiretroviral drugs

Early monotherapeutic treatment strategies, which were used up until the mid-1990s, had a number of challenges including incomplete virologic suppression, resulting in the emergence of multiple drug resistance mutations (DRM) [73]. Nevertheless, it was the introduction of combined antiretroviral therapy (cART) in 1996, which revolutionised the treatment of HIV-1 [74]. This strategy has brought effective virologic suppression, significant recovery of immune function, marked improvement in clinical symptoms, and remarkable extension of lifespan with improved quality of life.

During the last three decades, the US Food and Drug Administration (FDA) approved over 25 ARVs from six mechanistic classes or subclasses for the treatment of HIV. The six classes include: nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), integrase strand transfer inhibitors (INSTIs), a fusion inhibitor (FI), and an entry inhibitor- a CCR5 antagonist [74]. Three or more ARVs from at least two different drug classes are combined in the cART regimens. The three main classes recommended for the first-line ART regimen include NRTIs, NNRTIs, and PIs [72].

The cART are effective in dramatically suppressing the virus replication and reducing the plasma viral load (VL) to the level below the limit of detection of sensitive assays (50 copies/ml) [75]. The widespread use of the currently available ARVs has led to a significant decline in morbidity and mortality. As a result, AIDS-related deaths declined by over 48% from a peak of 1.9 million [1.7 million–2.2 million] in 2005 to 1.0 million [830 000–1 2 million] in 2016 [76]. The UNAIDS also sets an ambitious strategy called 90-90-90 (where, 90% of PLHIV knows their HIV status by 2020; 90% of people who know their HIV status accessing treatment; and 90% of people on treatment having suppressed viral loads, so they remain healthy) with targets to end the AIDS epidemic by 2030 [7]. About 20.9 million [18.4 million – 21.7 million] PLHIV were accessing ART in June 2017 [13]. About 11.7 million PLHIV in eastern and southern Africa were accessing ART with 60% regional coverage in 2016. Another 2.1 million PLHIV with about 35% coverage were accessing ART in western and central Africa during the same period [13].

1.5.2 HIV-1 drug resistance (HIVDR)

HIVDR is caused by mutations in the virus genetic structure that affects the activity of a particular single or combined drugs [77]. Emergence of HIVDR depends on the occurrence of genetic variation in the virus population plus the selection of drug resistance variants in the process [78]. The HIV-1 genetic variability is due to lack of ability by its RT enzyme to proofread and repair nucleotide sequences during the viral replication process [79]. This along with the high viral replication rate (at a rate of $10^{10} – 10^{11}$ viral particles per day) leads to continuous accumulation of a pool of genetically diverse viruses (quasispecies) [80]. The expansion of drug resistance variants is influenced by fitness of the virus; how long the

replication process has continued while the individual was receiving the ART; and the genetic barrier of the ARV agents [78, 81].

All the currently available ARVs, including newer classes, are at risk of being partially or fully inactive as a result of the emergence of HIVDR. Two type of HIVDR: pretreatment drug resistance (PDR), and acquired drug resistance (ADR) describe the way an infected individual got the drug resistance [82].Transmitted drug resistance (TDR), which is detected among treatment naïve individuals without history of ARV exposure, is described within PDR.

1.5.2.1 Pretreatment HIV-1 drug resistance

PDR is detected in ARV naïve individuals initiating ART for the first time or people with prior ARV exposure and initiating or reinitiating first-line ART [77]. It may have been transmitted at the time of infection, or it may be acquired as a result of prior ARV drug exposure such as in women exposed to ARVs for prevention of MTCT of HIV, in people who have received pre-exposure prophylaxis, or in individuals reinitiating first-line ART after a period of treatment interruption without documented virologic failure.

The prevalence and the extent of the problem widely vary with geographical locations and drug types. Along with the widespread use of ART and increasing number of people accessing the treatment, the prevalence of PDR has substantially increased in low- and middle- income countries (LMIC). Since 2001, the prevalence of NNRTI PDR has significantly increased globally [77]. Among 11 countries included in the WHO's survey of PDR conducted during 2014 – 2016, seven presented over 10% PDR prevalence in adults initiating ART. The WHO HIVDR report also showed PDR prevalence to NNRTIs of greater than the 10% threshold in six of the 11 countries. Among the LMICs, more rapid annual increments reported from eastern Africa (29%) and southern Africa (23%) compared to western and central Africa (17%), Latin America (15%), and Asia (11%). The survey also reported a higher PDR prevalence to NNRTIs among PLHIV initiating first-line ART with prior ARV exposure (21.6%) as compared to ART naïve (8.3%).

In addition, a systematic review from sSA showed significantly higher PDR NNRTI prevalence among children exposed to ARV drugs for prevention of MTCT (43%) as compared to unexposed once (13%) [83]. WHO recently recommended countries to make changes in their first-line regimen from NNRTI-based to non-NNRTI-based ART regimens, such as INSTIs if the population level of PDR NNRTI reaches the 10% threshold [84]. However, the risk of HIVDR associated to the newer class of second-generation INSTIs is not clearly known.

The prevalence of PDR to the three main ARV drug classes was reported increasing in resource rich settings (RRS), although it stabilized later as a result of the quality of ART services, where genotypic resistance test (GRT) is part of the routine care [85, 86]. Although the figures have declined overtime, the prevalence generally varies between 9% – 15% in the

US and European countries [86-89]. PLHIV in the RRS also have higher chance of accessing newer ARVs from different drug classes.

1.5.2.2 Acquired HIV Drug Resistance (ADR)

Despite the remarkable achievements observed over the last decades, emergence of HIVDR among individuals accessing ART remains to be an important issue. ADR develops when HIV mutations emerge due to viral replication in individuals receiving ARV drugs [84]. Suboptimal treatments and non-adherence to treatment are the main factors for emergence and accumulation of drug resistance HIV-1 variants. Through onward transmission of the ADR, ART naïve individuals could get drug resistance strains with subsequent effects on treatment outcome both at individual and population levels [82]. However, as a number of factors contribute for the emergence and onward transmission rate of the ADR, the prevalence and extent of the problem varies geographically.

The decline in number of patients exposed for monotherapy; the increasing use of more tolerable and potent cART; and routine use of viral load and GRT have largely contributed for the reduction of ADR in RRS [86]. For example, a study conducted in Sweden between 1997 and 2011 reported a rapid decrease in the prevalence of major HIVDR mutations, most rapidly between 2003 and 2007 [90].

In contrary to the RRS, in resource limited settings (RLS) the use of standardized ART regimens with limited treatment options; treatment interruptions due to drug stock outs; and poor treatment monitoring systems hugely influence the emergence and onward transmission of ADR [91, 92]. Reports from sSA cohorts have revealed higher rates of HIVDR among treated patients than what was expected [93, 94]. However, unlike in RRS, viral load and GRT are not used to monitor and guide the ART service in the RLS, where there could be potential misclassification of patients' treatment outcome [95, 96].

For example, a review from RLS on GRT among people who failed therapy after 12 months of ART reported a 60% drug resistance rate to any drug class (NRTI 55%, NNRTI 46%), but no drug resistance was identified from the remaining 40% treatment failure patients, which could be due to very poor adherence or treatment interruptions [97]. In the absence of GRT, individuals experiencing treatment failure, but without HIVDR, would have been unnecessarily switched to second-line ART regimens.

In a multicenter cohort study from six sSA countries, 70% of those successfully sequenced samples of virologic treatment failures harbored more than one DRM and 49% had dual-class (NRTI and NNRTI) resistance, with an average of 2.4 DRMs per sequence (range, 1–8) [98]. The most commonly observed DRMs include M184V, K103N, Y181C, and G190A. Thymidine analogue mutations (TAMs) reported to contribute to NRTI-associated DRMs in substantial proportion of the sequences. The study noted that K65R was frequently selected by tenofovir (TDF) (28%) or stavudine (d4T) (15%).

Another multicenter study on adult HIV treatment cohorts and clinical trials of HIVDR testing in Europe, Latin- and North- America, sSA, and Asia reported a high proportion of DRMs among virologic failure patients who were on a TDF-containing first-line regimen [92]. The prevalence of TDF-associated DRMs was highest among sSA patients (57%) compared to Europe (20%). Those individuals from sSA also had notable DRMs to other ARVs in their regimen, which could lead to virtually compromise all combination ART regimens in the region. TDF-resistant viruses had been reported to have substantial transmission potential. The study also noted that viral strains affected TDF-resistance in Europe, although it was not main driver of the resistance among viruses circulating in sSA. However, previous studies from Africa reported that HIV-1C rapidly develops K65R-related resistance to TDF [99-101].

1.5.2.3 Resistance to nucleoside analagoue reverse transcriptase inhibitors (NRTIs)

The first antiviral drug approved for the treatment of HIV was zidovudine (ZDV) from the NRTI group. The NRTI drug class has continued to be the core element of the cART [102]. The HIV-1 RT enzyme has been used as a major target for anti-HIV drug development such as ZDV, emtricitabine (FTC), lamivudine (3TC), TDF, d4T, abacavir (ABC), and didanosine (ddI) [72, 103]. NRTIs are prodrugs that require intracellular metabolic conversion into their pharmacologically active 5'-triphosphate form before they are incorporated into the proviral DNA at the 3'-end. The RT enzyme incorporates phosphorylated NRTIs via competing natural deoxynucleoside triphosphates (dNTPs) [104]. Then proviral DNA formation is terminated because of lack of a 3'-hydroxyl group or as a result of an altered sugar moiety, which prevents incorporation of incoming nucleotides [103-105].

HIV-1 develops resistance to NRTIs through two main mechanisms including NRTI discrimination and NRTI excision [72]. The first mechanism involves discriminatory mutations through which the RT enzyme avoids binding of the NRTI and retains its ability of recognizing the analogous natural dNTP substrate, which results in diminishing NRTI incorporation into the proviral DNA chain, where the virus carries on its proliferation [106]. A number of individual or groups of mutations that occur at or near the drug binding sites of the HIV-1 RT gene can promote resistance to most NRTIs through selective impairment of the enzyme's ability to incorporate analogues into the proviral DNA [107].

Discriminatory mutations including K65R, L74V, Q151M, and M184V/I weaken the enzyme's affinity towards NRTI, but without considerable affinity change towards the corresponding natural dNTP. For example, the K65R discriminatory mutation confers reduced polymerization rate for incorporation of all NRTIs and viruses with this mutation have decreased susceptibility for TDF, 3TC, FTC, ddI, and ABC, but no significant resistance for ZDV and d4T [72]. M184 is located at the heart of the catalytic site of the RT enzyme and M184V/I interfere with the proper positioning of lamivudine triphosphate within the catalytic site. Hence, these mutations confer the virus very high levels of resistance to 3TC and FTC [106, 107].

Multiple Q151M complex mutational patterns develop as a result of Q151M with other accompanying mutations A62V, V75I, F77L, and F116Y confer the virus high level resistance to most NRTIs (ZDV, d4T, ddI and ABC), but minimal susceptibility changes to 3TC, FTC and TDF. Other mutations at β3–β4 hairpin loop (positions 62 – 72 including T69N, K70E/G/Q mutations or T69 deletion) are also associated with the Q151M complex. L74V mutation also confers the virus resistance to ABC and ddI and it is associated to lower virologic response to TDF-based regimens [72].

Another major mechanism involves nucleotide excision that selectively removes NRTIs from 3' end of the chain terminated DNA primer in order to restore the DNA synthesis [106]. Removal of the nucleoside analogue from the terminated DNA chain is associated with a group of mutations commonly called TAMs [75, 107]. These mutations can promote resistance to almost all nucleoside and nucleotide analogues. There are two distinct TAM pathways: TAM1 pathway (M41L, L210W, and T215Y) and the TAM2 pathway (D67N, K70R, T215F, and K219E/Q) [75]. Once the RT acquires TAMs, further mutations such as T69 insertion complex could be developed with potential resistances to different NRTIs [72].

1.5.2.4 Resistance to non-nucleoside analagoue reverse transcriptase inhibitors (NNRTIs)

Unlike NRTIs, NNRTIs do not require intracellular metabolic conversion in order to employ their pharmacological activities [72]. An NNRTI binds to a noncatalytic allosteric pocket adjacent to the active site of the RT [107]. This noncompetitive binding induces conformational changes in the RT structure, which ultimately limit the enzymatic activity of RT and blocks the DNA polymerization process. Because of their minimum side effect and less toxicity, NNRTIs also play central roles in the treatment of HIV-1. The five approved NNRTIs include nevirapine (NVP), delavirdine (DLV), efavirenz (EFV), etravirine (ETR), and rilpivirine (RPV) [72].

Emergence of DRMs usually at the NNRTI-binding pocket reduces their binding affinity and could diminish or completely abolish their therapeutic efficacy. The earlier generation NNRTIs (NVP and EFV) have weak genetic barrier to drug resistance and DRMs quickly emerge. Only a single mutation at its binding site can induce a high level of resistance [72, 75]. Whereas, the later generation NNRTIs (such as ETR and RPV) have greater efficacy even against the drug resistant viruses and due to their conformational flexibility, they have a better genetic barrier towards the emergence of DRMs [72]. However, due to the narrowness of the binding site, NNRTI-associated DRMs are highly overlapping [108, 109].

The primary DRMs conferring resistance to NNRTIs are formed by amino acids with the following positions: L100, K101, K103, V106, T107, V108, V179, Y181, Y188, G190, F227, W229, L234, and Y318 of p66 and E138 of p51 [72]. K103N, Y181C and G190A are the three most common single mutations conferring high level resistance to all NNRTIs of the first-generation causing clinical failure [72, 75]. Although NVP resistance is often associated with Y181C mutation, it is also associated with Y188C, K103N, G190A, and V106A mutations [107, 108]. EFV has some common resistance mutations with other NNRTIs

including K103N, Y188L and G190A/S. Initial resistance to EFV is generally characterised by the K103N mutation, but Y188L mutation also involves [72, 107, 108, 110].

ETR and RPV have been developed in such a way that they have highly flexible structure with high genetic barrier, which allows them to efficiently bind even in the presence of regular NNRTI-associated DRMs [110-112]. Although it has a number of resistance associated mutations such as L100I, Y181C/I/V, K101P, and V179F, the K103N is not a major issue for ETR [112]. ETR, is therefore, recommended for use in patients who harbor the K103N mutant variants [108]. Due to the similar chemical structures, RPV has similar resistance mechanisms as ETR [72, 110].

1.5.2.5 Resistance to protease inhibitors (PIs)

The chemical structure of PIs resembles the viral peptides that are normally recognized and cleaved by the HIV-1 PR. PIs display a strong affinity for the active site of the PR and when they competitively bind to the active site of the enzyme, proteolytic cleavage of the gag and gag-pol is inactivated [75]. This in turn prevents the maturation of viral particles to infectious viruses. FDA approved nine PIs including atazanavir (ATV), darunavir (DRV), fosamprenavir (FPV), indinavir (IDV), lopinavir (LPV), nelfinavir (NFV), saquinavir (SQV), tipranavir (TPV), and ritonavir (RTV) for the treatment of HIV-1 [72, 108]. PIs are part of the first-line regimens in RRS, but their availability is limited in Africa, and they are reserved for second-line therapy [113].

HIV develops resistance to PIs by accumulating mutations that leads to reduced binding affinity to the drugs while retaining its binding affinity to natural substrates. PI-associated DRMs occur with amino acid substitutions within the active site as well as non-active distal sites [72, 107]. PI resistance occurs through a stepwise accumulation of mutations, which produce conformational changes in the PR binding site followed by secondary compensatory mutations that improve the enzymatic activity, and in some cases rescue the viral fitness [75]. The PI-associated DRMs include D30N, V32I, L33F, M46I/L, I47A/V, G48V, I50L/V, V82A/F/T/S/L, I84V, and L90M. Cross-resistance is common among PIs as they share relatively similar chemical structures [72].

1.5.2.6 Resistance to integrase strand transfer inhibitors (INSTIs)

The INSTIs interferes with catalytic activity of IN. They bind to the catalytic core domain (CCD, amino acids 50 – 212) of the enzyme and target the integration process at the strand transfer step [114]. They interact with two essential viral elements including the IN and the viral DNA. The three clinically approved INSTIs include raltegravir (RAL), elvitegravir (EVG), and dolutegravir (DTG), which bind to the catalytic core of IN [115]. INSTIs associated DRMs occur within the IN active site near the amino acid residues that coordinate the essential magnesium cofactors [114].

Resistance to RAL arises from three major independent pathways or sets of mutations as defined by primary/signature mutations at positions Y143, Q148, and N155 [75]. These

primary/signature mutations generally observed with specific secondary mutations and the mutational pathways are believed mutually exclusive, which do not appear in the same viral genome [114]. The Q148 mutational pathway includes Q148H/K/R accompanied by other subsequent mutations such as L74M, E92Q, T97A, G136R, E138A/K, G140A/S, and V151I. The N155H mutational pathway could be associated with L74M, T97A, E157Q, and G163K/R. The Y143 pathway consists of Y143C/H/R mutations with L74A/I, T97A, G163K/R accessory mutations [72].

The INSTIs share similar inhibition mechanism; hence cross-resistance becomes a common phenomenon among these drugs. For example, the two RAL major resistance mutational pathways at Q148 and N155 along with their accessory mutations have also been selected by EVG, which shows a similar resistance mutation profile as RAL [115]. Nevertheless, EVG does not structurally interact with the IN residue Y143 and the Y143 mutational pathway could not confer cross-resistance to this drug. On the other hand, specific mutations that confer resistance only to EVG appears at position T66 such as T66I/A/K. E92Q also confers high level resistance to EVG as compared to RAL [72]. In contrary, compared to RAL and EVG, DTG has lower genetic barrier to drug resistance, although R263K and G118R mutations confer certain level of resistance to this drug [115].

1.5.2.7 Resistance to entry inhibitors

HIV entry inhibitors can be classified into three groups: attachment inhibitors, co-receptor antagonists, and fusion inhibitors [116]. However, so far FDA approved only two drugs from two different classes of entry inhibitors: maraviroc (MVC) as a co-receptor antagonist, and enfuvirtide (ENF, T-20) as a fusion inhibitor [72]. MVC is a selective small molecule CCR5 antagonist, which targets the interaction between gp120 and CCR5 and inhibits the virus entry through altering the conformation of the receptor [117]. T-20 is a 36 amino acid long synthetic peptide fusion inhibitor targeting fusion activity of the HIV-1 gp41 [118].

One of the resistance mechanisms employed by HIV-1 against MVC is tropism switching using CXCR4 co-receptor instead of CCR5 to enter into the cell [119]. However, studies have shown multiple mutations that occurred in V3 loop region of the *env* region allows the virus to use the MVC occupied CCR5 co-receptor as well [72]. Additional resistance mechanisms employed by the virus include increased affinity for the co-receptor, utilization of inhibitor-bound receptor for entry, and faster rate of entry [75]. In view of the fact that selection of CXCR4 tropic virus (which could lead to faster disease progression) due to CCR5 antagonist treatment; tropism switching has been a concern in the therapeutic administration of MVC [75, 108]. Resistance mutations to T-20 emerge through a stretch of amino acids, which is located at position 36 – 45 of gp41 [118].

1.6 MONITORING STRATEGIES

In sSA countries, when to start and monitor ART has been mainly based on a public health approach adapted from WHO. Clinical decision making and when available CD4 cell count,

rather than viral load testing, has been the main monitoring strategies in the region [113, 120]. However, the monitoring strategies have been continuously evolved [121].

1.6.1 When to start ART

The WHO recommendation about when to start ART has been continuously updated over time. For example, till 2013 in most RLS including Ethiopia, the decision to start ART was based on clinical and immunological criteria adapted from the WHO's 2006 guideline. In this guideline, the optimum time to commence ART was set to be before patients become unwell (present with their first OI). The guideline recommended not to initiate ART before patients' CD4 cell count drops below 350 copies/μl [5, 120]. Then, the 2013 WHO consolidated guideline recommended initiating ART for all HIV infected adults with a CD4 cell count at or below 500 cells/μl regardless of their WHO clinical stage, giving priority to those with severe or advanced diseases (WHO stage III or IV) or a CD4 count at or below 350 cells/μl [113]. Recently, a test and treat strategy has been recommended with final targets of universal treatment access [121, 122]. This recommendation has allowed the treatment of all PLHIV irrespective of their CD4 cell count [121].

Reports have revealed that immediate initiation of ART has impact in reducing both AIDS-related and serious non-AIDS-related events, as well as reduction in transmission of HIV [121]. However, there are challenges in retaining PLHIV in care particularly in LMIC, where, multiple factors related to the health care delivery systems could facilitate or hinder retention of patients [123]. There are also uncertainties in relation to the accelerated ART initiation, which may lead PLHIV to start the treatment before they are well prepared, with adverse consequences for adherence and treatment outcomes. Hence, starting the ART based on a person's informed decision including provision of information about the benefits of early treatment, but also the required lifelong commitment; the risk of delaying the treatment; and availabilities of adherence supports are recommended [121]. Once a patient has started the treatment, uninterrupted supply of ART with continuous monitoring are essential for sustainable viral suppression and optimal treatment outcome [123].

1.6.2 Monitoring response to ART and diagnosis of treatment failure

The currently available ART is not curative, but suppresses the virus replication. The treatment reduces the plasma viral load below the detection limit of sensitive assays (<50 copies/ml), preventing further loss of $CD4^+$ T-lymphocytes and deterring progression to AIDS [75]. This allows PLHIV accessing cART to have near-normal life expectancy and to decrease further transmission of HIV [124]. Hence, viral load (HIV-1 RNA level) and CD4 count are the two surrogate markers of ART outcomes and disease progressions, which have been used to manage and monitor the HIV-1 infection for decades [125].

Viral load is generally accepted as the gold standard for monitoring the ART response and is a marker of the treatment outcome [125] . It is the most important indicator of initial and sustained treatment response, which is recommended as the preferred monitoring approach to diagnose and confirm ART failure, but its use has mostly been limited in RRS. Although the

optimal threshold for defining virologic failure and for switching ART regimens has not been established, WHO recommended a threshold of 1000 copies/ml [121]. This recommendation in view of the fact that the risk of transmission of HIV to other individuals as well as the progression of an infection to AIDS is minimal when the HIV-1 RNA level is below 1000 copies/ml [126]. Below 1000 copies/ml, viral blips or intermittent low level viremia (50 – 1000 copies/ml) can occur during effective treatment, but have not been associated with an increased risk of treatment failure [127]. Most standard viral load platforms using plasma specimens have good diagnostic accuracy at this threshold level [121].

Compared to CD4 cell count, monitoring with viral load has been found to associate with satisfactory outcomes in early detecting treatment failure [128]. However, considering the large number of people accessing ART coupled with the deficiencies in the health care system in LMIC including sSA countries, viral load monitoring has not been the standard of care in such regions. Hence, when viral load is not routinely available, CD4 count and clinical monitoring are recommended to be used for the diagnosis of treatment failure. However, immunological and clinical criteria have poor sensitivity and specificity to detect treatment failure, particularly at higher CD4 cell counts, and more accurate immunological criteria are yet to be identified [121, 129].

1.6.3 Treatment outcome measurements

Once ART is initiated, the outcome can be measured in three ways: clinically, by disease progression and WHO clinical staging; immunologically, by measuring the temporal trends in CD4 cell counts; and virologically, by measuring the plasma HIV-1 RNA load [121]. Accordingly, treatment failure is defined clinically, virologically and immunologically, but slightly different definitions are used in children, and adult and adolescent populations. In adults and adolescents, the following definitions are given by WHO:-

Clinical failure: new or recurrent clinical event indicating severe immunodeficiency such as WHO stage IV diseases and certain WHO stage III conditions including pulmonary TB and severe bacterial infections after 6 months of effective treatment.

Immunological failure: CD4 cell count at or below 250 cells/µl following clinical failure; or persistent CD4 levels below100 cells/µl.

Virologic failure: a persistently detectable viral load exceeding 1000 copies/ml (that is, two consecutive viral load measurements within a 3 month interval with adherence support between measurements) after at least 6 months of starting a new ART regimen.

Retention in care and adherence to medication are the other most key indicators in monitoring the treatment outcome among patients accessing ART. For PLHIV who are accessing treatment, uninterrupted ART and continual monitoring are essential for sustained viral suppression and optimal treatment outcomes. Hence, treatment outcome is usually analysed in two ways: on-treatment (OT) analysis and intention-to-treat (ITT) analysis.

On-treatment (OT) analysis: accounts only for patients who continue the treatment (compliant patients). This type of analysis includes the number of patients with virological suppression or virological failures. However, OT analysis loses the benefits of initial randomization because it accounts only patients who continue the ART. In addition, because it is unlikely that compliance varies randomly among the comparison groups, bias is probable [130].

Intention-to-treat (ITT) analysis: compares patients in their randomly assigned treatment groups. ITT includes every subject who is randomized according to randomized treatment assignment. It ignores noncompliance, protocol deviations, withdrawal, and anything that happens after randomization. Hence, in addition to virologic failure, ITT analysis includes patients who have died or who were LTFU, and/or who stopped ART for different reason. However, a better application of the ITT approach is possible if complete outcome data are available for all randomized subjects [131].

1.7 GENOTYPIC ASSAYS FOR HIV-1 DRUG RESISTANCE

Genotypic assays are used to determine the presence of mutations that are known to confer decreased drug susceptibility to an organism [132]. The genotypic assays are based on analysis of the nucleotide sequence generated from genomic regions conferring phenotypic resistance to the organism. For various reasons, the genotypic assays have been commonly used for HIVDR testing [132, 133]. DRMs are identified through assessment of specific gene sequences in comparison to a reference (wild-type) virus, though the most commonly used reference sequence is the subtype B consensus sequence [78]. A number of genotypic approaches are available for detection of DRMs including population-based sequencing, whole-genome sequencing, allele-specific, and next-generation sequencing (NGS) [133-135].

1.7.1 Population-based sequencing

Population-based Sanger sequencing (PBSS), which is based on direct polymerase chain reaction (PCR) dideoxynucleotide sequencing, is the most conventional GRT method [136]. It has been used for HIVDR testing to guide ART and monitor DRMs at least in RRS [137, 138]. GRT can be done using either commercially available genotypic assays (ViroSeq and TruGene) [136] or *in-house* methods [139, 140]. However, unlike *in-house* methods, the commercial assays are more expensive and may not provide flexibilities [140]. These direct sequencing assays have been widely used for HIVDR testing in clinical settings as well as surveillance purposes [137, 141].

Although PBSS can provide detailed sequence information and detect all possible DRMs, the assay is not sensitive enough to detect low-level drug resistant variants that are present at levels below 20% of the circulating viral population [133]. Hence, more sensitive assays with features overcoming various limitations of the traditional PBSS have been developed for genotyping HIV including allele-specific PCR (AS-PCR) [134, 142, 143], and NGS [133, 144].

1.7.2 Allele-specific PCR (AS-PCR)

The drug resistant variants that are not detected by conventional assays have been shown to have influence on HIV treatment response. This calls for application of mutation-specific assays (such as AS-PCR) in defining the impact of drug resistant minority variants on therapeutic outcome [134]. Compared to the conventional PBSS method, minority-specific assays have high sensitivity and better capacity to detect low frequency mutations down to 0.01% [142, 145]. Currently, there are several allele-specific assays allowing specific investigation of DRMs with a higher analytical sensitivity for minority HIV variants [134, 143, 145-147].

Even if AS-PCR is currently one of the most sensitive assays being used in studying minority DRMs, it has limitations that it can only analyse one or a few mutations at a time [134]. Considering the fact that allele-specific assays are specific to mutations, their benefits depend on the mutation being studied. In view of such limitations, multiplex allele-specific assays that can simultaneously detect major HIVDR mutations at several loci have been developed [148, 149].

1.7.3 Next-generation sequencing (NGS)

The field of genomics is being transformed because of the recently evolved new deep sequencing techniques including NGS, which is also known as “second-generation”, “massive parallel” or “deep” sequencing [144, 150]. Several new high-throughput DNA sequencing instruments are made available with their own intrinsic performance metrics such as the number of reads obtained; the read lengths; the reading accuracies; the time needed to run the assay; and cost of the assay [133]. These technologies generate an enormous amount of information at relatively low cost in a short period of time than the conventional PBSS [96]. The technologies are continually evolving, opening new and more affordable opportunities for their implement in clinical laboratories [151, 152].

The use of deep sequencing technologies to evaluate the impact of minority variants and rare non-CCR5 variants promotes the assays to be important in the field of HIVDR [135, 144, 150]. For example, new assays to determine the specific tropism of viruses are essential for successful treatment with ARV classes targeting CCR5 or CXCR4 [144]. Hence, NGS assays with such increased sensitivity play an important role in HIVDR surveillance, to study transmission of drug resistant viruses, to determine the impact of drug resistant minority variants on therapeutic outcomes [133].

The commonly used NGS technologies include pyrosequencing by 454 Life Sciences/Roche (454/Roche); reversible terminator sequencing-by-synthesis (SBS) by Illumina; and single-molecule sequencing by Pacific Biosciences (PacBio) [153]. They differ in sequencing chemistry, speed, throughput, read length, error rates and error patterns. Each of the three technologies can be used to analyse viral genomes, but none enables sequencing regions larger than a few hundred bases with sufficiently high accuracy [154].

1.7.4 Whole-genome sequencing

Most of the currently available genotypic assays used in the clinic could provide information on DRMs, which is restricted to full-length PR and partial RT regions. As mentioned above, the currently available NGS technologies provide sequence reads of only a few hundred bases with sufficiently high accuracy, and subsequently virus diversity is typically assessed only by studying single nucleotide positions or mutations within a read length [155]. This led for the development of a growing number of *in-house* methods to address the need for simultaneous HIVDR testing against all ARV classes [135, 156].

Whole HIV-1 genome analysis also provides information about molecular epidemiology, population structures, and pathogenesis of the infection. This method could also contribute in the development of new effective preventive and treatment strategies including vaccine design [155]. In addition, the full genomes are also of interest for detection of quasispecies and tracking immune evasion through variation in $CD8^+$ T-lymphocyte epitopes with linkage elsewhere in the genome [154].

So far, genome-wide reconstructions were only achieved with simulated data *in silico*; thus, experimental amplifications and sequencing strategies, combined with appropriate bioinformatics methods for haplotype inference, are needed to phase multiple overlapping sequence reads, and to allow reconstruction of whole genomes. Due to its small genome, high intra-patient genetic diversity and clinical relevance, HIV-1 is an attractive model system for developing such a global haplotype inference methodology [154].

1.7.5 Bioinformatic tools

Because of the large number of mutations and mutational patterns that confer drug resistance it may be difficult to interpret results genarted from genotypic assays. To facilitate the use of the large amount of data generated from genotypic assays a range of bioinformatics tools and databases have been developed [157]. The bioinformatics tools and databases have been applied for predicting HIV subtypes, co-receptor tropism, response to combination of therapy, and drug resistance. These kinds of applications of the bioinformatics tools support clinical decision making based on the viral genomic data. With a number of available softwares, the development of computational methods that correlates specific genotypes with resistant phenotypes can allow effective therapy design against drug resistant HIV variants [157, 158]. Bioinformatic tools are also used in order to understand the mechanism of drug resistance development and to model evolution of the sequences under the selective pressure of combination therapies as well.

Generally, the methods used for the prediction of phenotypic drug resistance from genotypes can be subdivided into three groups including knowledge based, data based, and physics based approaches. The knowledge based approaches rely on extracting information on drug resistance from several sources, mostly from published literatures. Data based methods rely on applying machine learning techniques to matched genotype – phenotype pairs. Physics

based methods combine phenotypic drug resistance factors with computed inhibitor binding affinities, taking into account mutations in the genotype [159].

However, most of the genotype interpretation algorithms are knowledge based approaches, and clinical evaluation of the interpretations is not always straightforward. These interpretation algorithms apply a set of rules or scores for each mutation and drug. Hence, accurate predictions can be valuable for prescribing the most effective ARVs for infections with resistant HIV [160].

The commonly used interpretation algorithms include: Stanford HIVdb (https://hivdb.stanford.edu/), ANRS algorithm (http://www.hivfrenchresistance.org/), Geno2Pheno (https://www.geno2pheno.org/), HIV-grade (https://www.hiv-grade.de/cms/grade/homepage/), and REGA (http://www.bioafrica.net/HIVdrugresistancetools.html) [161]. There are also HIVDR databases, which are essential to store, represent, and analyse the data on drug resistance and to retrive these data when needed. The three common HIVDR databases include: EuResist database, Stanford database, and Los Alamos HIV Sequence Database. The Stanford database allows users to screen their own DNA sequences for resistance conferring mutations. The Los Alamos HIV Sequence Database (https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html) makes an easy access to all published genomic DNA, coding protein sequences of HIV with subtype and sequence annotations. The database also offers several tools and programs for analysis [162].

In addition to the prediction of HIVDR, bioinformatics tools have also been developed for prediction of the virus co-receptor tropism based on interpretation of V3 sequences of the *env* region [163]. The first genotypic algorithm designed to predict HIV-1 tropism took into account only the charge of amino acids at two key residues located within the V3 loop (amino acids 11 and 25). Although estimates using the "11/25 charge rule" are fairly satisfactory [67], other positions within the V3 region are also currently known to influence viral tropism as well, causing occasional disagreements. Hence, currently genotypic bioinformatics tools consider the entire V3 sequence [163]. In addition to the "11/25 charge rule", bioinformatics based algorithms including position specific scoring matrix (PSSM), Geno2Pheno, and PhenoSeq are used to predict the HIV-1 co-receptor tropism [164].

1.8 ANTIRETROVIRAL TREATMENT IN ETHIOPIA

In Ethiopia, the ART programme was officially started in 2003 as fee-based system [165] and free ART was launched in 2005 with support from the government, the U.S. President's Emergency Program for AIDS Relief (PEPFAR), and the Global Fund to Fight AIDS, Tuberculosis and Malaria [166]. Subsequently, the service has been decentralised and made available at increasingly large numbers of both health centers and hospitals throughout the country using the WHO public health approach. Integration of the service at lower-level health facilities and task shifting to low- and mid- level health workers have greatly enabled scaling up of ART in Ethiopia [167, 168]. By the end of 2016, over 420 000 PLHIV were accessing the treatment with an approximately 60% national coverage [35, 76]. Although a

high rate of ART coverage has been achieved in adults, the coverage remains low among children. Reports also revealed that a significant number of patients enrolled into the ART were dropped out from the service [6, 44].

In Ethiopia, the eligibility criteria for initiation of ART have been improved over time. For example, until 2012, the criterium for initiating ART was set at CD4 count at or below 200 cells/µl, or WHO stage IV; thenafter, it was improved to less than 350 cells/µl, or advanced diseases (WHO stage III or IV). Again in 2014, adaptation of the 2013 WHO consolidated guidelines improved the eligibility criteria to CD4 count at or below 500 cells/µl, regardless of WHO stages [113, 123]. According to the new national guidelines, initiation of ART is based on the recent test and treat eligibility criteria, i.e. ART should be initiated to all PLHIV, regardless of their WHO stage or CD4 cell count [35].

1.8.1 Antiretroviral regimens in Ethiopia

The first-line ART regimen in Ethiopia includes a triple therapy: two NRTIs and an NNRTI, reserving PIs for second-line, which is the center piece of the public health approach [35]. In order to reduce the pill burden and simplify the regimens the use fixed-dose combinations (FDCs) is recommended in the guidelines. Hence, a once-daily regimens comprising NRTI backbone (TDF + 3TC) and one NNRTI (EFV) is maintained as the preferred choice in adults, adolescents and children older than ten years. For children younger than three years a PI-based regimen is the preferred approach.

Thus, the new guideline recommend a once-daily FDC containing TDF + 3TC + EFV as prefered first-line regimen, whereas AZT + 3TC + EFV, AZT + 3TC + NVP, and TDF + 3TC + NVP as alternative regimens [35]. With the previous guidelines both NNRTIs were given equal preference [5], but the recent WHO guidelines recommended EFV as the preferred option [113, 121]. Regimens containing ABC, d4T and boosted PIs are recommended under special circumstances. Recently the use of the second-generation INSTI, DTG combined with two NRTIs is recommended as an alternative first line in RLS [121], but in Ethiopia it is recommended as a third-line regimen [35].

Second-line regimens containing a boosted PI plus two NRTIs are recommended as preferred strategy for adults, adolescents and also for children when NNRTI-containing regimens were used in first-line ART. In general, selection of a second-line regimen depends on the first-line regimen that the patient has been taking [35].

1.8.2 Monitoring strategies in Ethiopia

Ethiopia has been adapting and implementing series of WHO guidelines for monitoring of the ART outcome. For example, the first guideline on the use of ARV was issued in 2003 [165], which was later revised in 2005 and 2008 to facilitate a rapid scale-up of ART as per the WHO public health approach [5]. In 2014, the country adapted and implemented the 2013 WHO consolidated guideline of ART for treatment and prevention of HIV [113, 123], and recently adapted the 2016 WHO guidelines [121] in its new national guidelines [35].

Monitoring and evaluation tools are also routinely used to track key program performance and outcome indicators, which provided limited and primarily aggregated information about the effectiveness of ART [5, 35].

Clinical parameters and CD4 cell count have been used as the main criteria for monitoring of ART. As per the old guidelines, the criteria used to monitor treatment failure include: WHO stage IV disease; fall of CD4 cell count to pre-therapy baseline (or below); 50% fall in CD4 cell count from the on treatment peak value (if known); or persistent CD4 count below100 cells/µl [5, 120]. However, the performance of clinico-immunological criteria has been found to have poor sensitivity, specificity and reliability, which could result in inappropriate switching to second-line therapy and consequently it might lead to an increased number of resistance mutations [169, 170].

Although CD4 cell count measurement has been recommended every 6 months or as deemed necessary, because of limited availability of the facilities only at certain ART centres, the test has been performed through sample referral networking system [123]. The CD4 cell measurement facilities have not yet been completely available in every center throughout Ethiopia because of the high cost of the instrument and partly because of the large number of people accessing the treatment as well. Due to limited laboratory facilities, shortage of supplies and trained personnel, the clinical criteria have been used in a substantial proportion of patients for monitoring of the ART outcome in the country.

However, the new national guideline does not recommend CD4 cell count for monitoring purpose, rather viral load determination. CD4 cell count is recommended for patients on ART only when indicated. HIV-1 viral load measurement is recommended at 6- and 12- months after initiation of ART and every 12 months thereafter [35]. Aside from the routinely scheduled measurements, viral load test is also recommended whenever there is clinical or immunological indication of treatment failure.

1.8.3 Antiretroviral treatment outcome in Ethiopia

Despite the observed successes in scaling up of ART throughout Ethiopia, there are only few studies showing the treatment outcome in the country. The few available studies reported a high rate of early death and LTFU [44, 166]. Because of the fact that until recently viral load test had not been used in monitoring of ART outcome in Ethiopia, virologic treatment outcomes were reported only by very few sdudies conducted mainly for the purpose of drug resistance study. For example, a study from Jimma reported a virologic failure rate of 5.3% at month 6 after initiation of ART [171]. Whereas, a study from Gondar showed a virologic failure rate of up to 30% at a median time of 24 months of ART [172].

There are also other few studies with reports of LTFU and mortality after initation of first-line ART. A retrospective longitudinal analysis study on adults enrolled at 56 health facilities from 2006 – 2011 reported a total attration (LTFU and death) rate of 30% at month 36, most occurred in the first 6 months [173]. Another earlier retrospective longitudinal analysis study conducted at 55 health facilities reported that 9% and 12% patients were transferred out to

other facilities, 0.58% and 2.1% switched to second-line ART, 6% and 8% died, and 20% and 24% LTFU after 12 and 24 months of ART, respectively [174]. This study also stated that decentralisation of the ART service more to primary health facility levels than at centralised hospitals could improve retention in care in Ethiopia.

1.8.4 HIV-1 drug resistance in Ethiopia

In Ethiopia, there are only a few studies from limited geographical regions mainly from the capital city Addis Ababa and Northwest part of Ethiopia, where the prevalence of TDR have been reported [175-177]. These studies showed very low rate of TDR among treatment naïve individuals. However, so far no surveillance study has reported the extent of TDR at a national level. The reported prevalence among ART naïve individuals who were attending a public antenatal clinic in Addis Ababa was below 5% [175]. Another study conducted on samples from Addis Ababa using AS-PCR reported a 6.5% (6/92) TRD prevalence [134].

A study conducted among ART naïve HIV-1 infected patients who were seeking treatment at Gondar University Hospital, Northwest Ethiopia showed a 3.3% TDR prevalence [176]. In another study from the same geographic region reported a 5.6% TDR prevalence according to the Stanford University HIVDB drug resistance interpretation algorithms, but 13.1% as per the International Antiviral Society–USA (IAS) mutation lists [177]. However, a study from Jimma Hospital, Southern Ethiopia reported that half (6/12) of the baseline samples harbored resistance to NNRTIs, but all the 12 patients had virologic failure status at month 6 [171].

Again there are only a few studies with reports of ADR in Ethiopia. A study from Jimma Hospital reported an ADR rate of 81.8% (9/11) after 6 months of ART [171]. Two other studies from Gondar also reported ADR rates of 75% (6/8) at a median time of 30 months on ART [178] and 66% (15/21) after 8 years of treatment [172].

2 RATIONALE OF THE THESIS

Ethiopia remains one of the most heavily affected countries by the HIV-1 epidemic with about 740 000 PLHIV [35]. The treatment of HIV-1 has been scaled up for more than a decade using the WHO public health approach and a large number of PLHIV in the country have been receiving ART with an estimated coverage of 60% by the end of 2016 [35, 76]. Of course, the public health approach helped to decentralise and integrate the ART service to the low level health care system and indeed have made the treatment more accessible to the people in need. However, until recently viral load and GRT have not been part of the current standard monitoring strategies like in most other sSA.

The treatment outcome and the levels of HIV-1 drug resistance had been poorly described and reports are available only from a few geographic regions of the country involving a limited number of patients. Hence, data has been lacking at the national level. In addition, some fragmented reports also showed that a substantial proportion of patients on treatment were LTFU [44, 166, 173]. Yet, the factors for poor treatment outcome and discontinuation of ART by a substantial proportion of patients are poorly understood. Although the HIV-1C was first described in Ethiopia, the status of the epidemic has not been characterised with data and samples representative of the country, and hence description of the HIV epidemic remains to be determined at the national level.

Bearing in mind the use of ART for over a decade in Ethiopia with high rate of dropouts, but without a comprehensive data about the ART effectiveness and drug resistance at the national level, it has been essential to assess the treatment outcome and determine the level and patterns of HIVDR across the country. Identifying the baseline predictors of treatment failure and LTFU has also been shown to be an important issue that should be addressed in order to get information about the factors behind the poor treatment outcome. Therefore, the aims of my thesis were to assess the treatment outcome and to determine the level and patterns of HIVDR both among treatment naïve patients and virologic failures in Ethiopia. In addition, the thesis also aimed at identifying baseline predictors of treatment failure and LTFU. Further, molecular characterisation of the epidemic and co-receptor tropism of HIV-1C was also the aim of this thesis. To achieve these aims and bridge the prevailing information gaps, data and plasma samples were collected from a countrywide HIV-1 cohort, where patients from seven university affiliated hospitals in Ethiopia were enrolled.

3 OBJECTIVES OF THE THESIS

3.1 GENERAL OBJECTIVES

The main objectives of this thesis were to assess the outcome of antiretroviral therapy and to investigate the molecular epidemiology of HIV-1C drug resistance using a large countrywide HIV-1 cohort in Ethiopia.

3.2 SPECIFIC OBJECTIVES

Overall, the objectives were addressed through the following five specific objectives:

1) To identify baseline predictors of therapy outcome and lost to follow up after 6 and 12 months of first-line antiretroviral regimen (**Paper I**).

2) To assess the presence of surveillance drug resistance mutations and investigate minority mutations by population-based Sanger sequencing as compared to next-generation sequencing (**Paper II**).

3) To evaluate the short-term impact of baseline drug resistance mutations on antiretroviral therapy outcome (**Paper II**)

4) To describe the HIV-1C molecular epidemiology and tropism trend over the last two decades in Ethiopia and to assess the impact on antiretroviral therapy outcome (**Paper III**).

5) To investigate the level and patterns of acquired drug resistance mutations in HIV-1C *pol* gene and perform a viral genome wide association study in virologic failure patients (**Paper IV**)

4 MATERIALS AND METHODS

4.1 STUDY DESIGN AND PATIENT POPULATION

The studies were conducted on data and plasma specimens collected through the Advanced Clinical Monitoring of Antiretroviral Therapy in Ethiopia (ACM) cohort.

In this prospective multi-site adult and adolescent repository cohort of HIV-1 infected patients, 874 ART naïve study subjects (age ≥14, both sexes) were enrolled at seven university hospitals between October 2009 and December 2011. During the recruitment period, 112 to 134 participants were enrolled in each hospital ART clinic. Then, all patients started ART as per the national guidelines [5], and were followed up until 2013.

4.1.1 Establishment of the countrywide national HIV cohort

The ACM cohort was established through a partnership between ten local and two international institutions (Figure 2). The local participating institutions include: HAPCO (Federal HIV/AIDS Prevention and Control Office of Ethiopia), EHNRI (Ethiopian Health and Nutrition Research Institute, now EPHI- Ethiopian Public Health Institute), ESTA (Ethiopian Science and Technology Agency, now MOST- Ministry of Science and Technology), and seven university affiliated hospitals. The two international institutions were Centers for Disease Control and Prevention (CDC), and Johns Hopkins University (JHU).

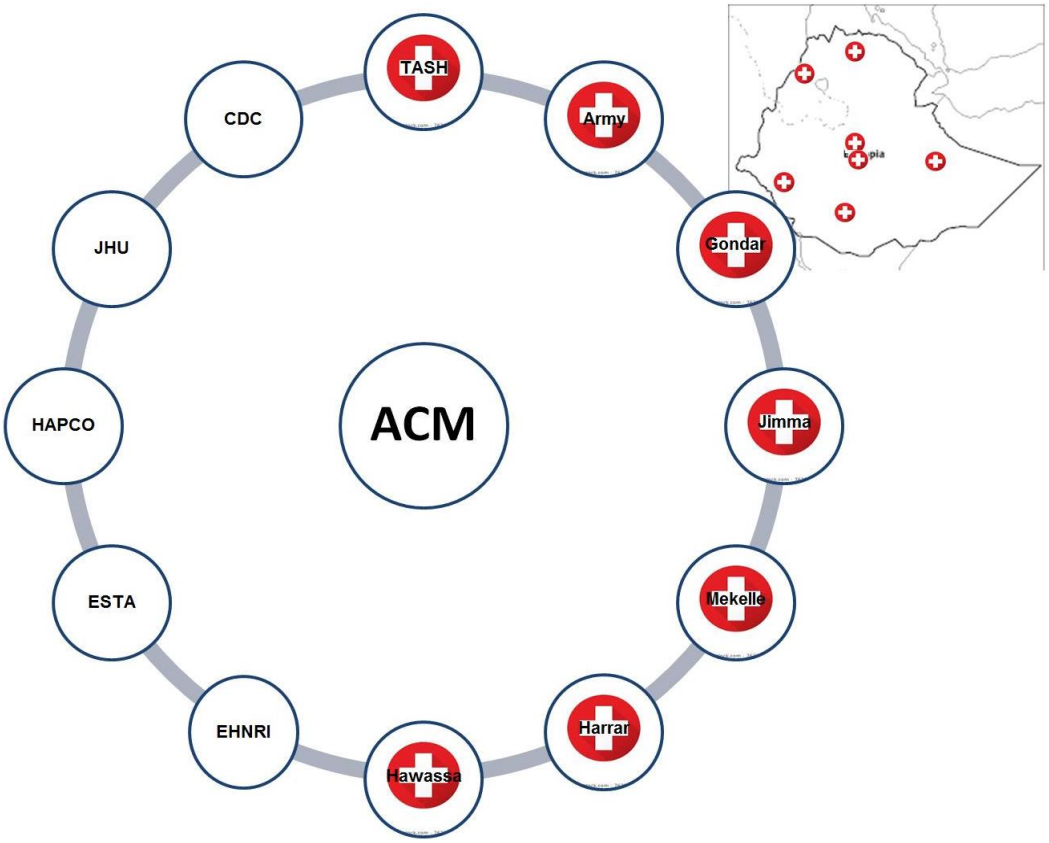

Figure 2. Advanced Clinical Monitoring of Antiretroviral Therapy in Ethiopia (ACM) project organizational structure. The map shows the geographical locations of the study sites (hospitals); : participating hospitals.

The ACM was one of the largest HIV cohorts in African settings with over 4000 patients who were enrolled from a well-defined study population at seven geographically distinct hospitals in Ethiopia. This cohort was also the first and largest clinical study on HIV infected patients aimed at assessing the longitudinal effectiveness of ART in Ethiopia. Clinical data were obtained from over 4000 patients and clinical and plasma samples from over 1000 patients including children.

4.1.2 Study sites in brief

The seven study sites include: TASH, Armed Force General Hospital, Gondar Referral and Teaching Hospital, Jimma University Specialised Hospital, Mekelle General Hospital, Hiwot Fana Referral Hospital, and Hawassa University Referral Hospital. The number of study participants enrolled at the respective sites were 119, 122, 134, 112, 126, 128, and 133. All of these hospitals are referral hospitals providing service to millions of individuals each. Geographically the study sites are well distributed throughout Ethiopia (a map showing the locations of the study sites depicted on Figure 3 in **Paper II**). Briefly, TASH is located in the capital city Addis Ababa, central Ethiopia; Gondar Hospital located in the northwest; Jimma in the west; Mekelle in the north; Harrar in the east; and Hawassa in the south parts of the country. In addition, Armed Force General Hospital is located in Addis Ababa providing service to mobile military staff and to their families.

4.2 DATA AND SPECIMEN COLLECTION AND STORAGE

Longitudinal data and plasma specimens were collected from each participant. Ten ml whole blood was collected at ART initiation and every 6 months thereafter and labeled with specific code and date of specimen collection. The blood specimens were then processed at the respective hospital laboratories and plasma specimens were transported to EHNRI (located in Addis Ababa), and centrally stored at -80°C after registered in a predesigned electronic repository database.

Quantification of HIV-1 RNA viral load was performed centrally at EHNRI using NucliSENS easyQ® HIV-1 Nucleic Acid Sequence-Based-Amplification (NASBA) assay (BioMérieux Diagnostics) with a lower detection limit of 150 HIV-1 RNA copies/ml. Viral load results for each participant at baseline, 6 and 12 months of ART initiation were stored in an ACM central electronic database at EHNRI. Other than VL, clinical and laboratory tests were performed at the respective hospitals as per the national guideline [5] and stored in the central electronic database after quality control checks.

For the purpose of this study, clinical and laboratory data at ART initiation and follow up time points (at month 6 and 12 after initiation of ART) were extracted from the central electronic database using a predesigned form. Patient data analysed in this thesis include: socio-demographic data (study site, gender, age, educational status, employment status, marital status, and spouse HIV sero-status); clinical data (WHO clinical stage, functional status, TB, OIs other than TB, BMI – body mass index, and adherence level); treatment

history (initial ART regimen, months on ART, history of treatment switch); and laboratory data (CD4 cell count, CD4/CD8 ratio, and VL).

In addition, 1 ml volume plasma (mostly in two 0.5 ml aliquots) of each study participant at each available visit (at ART initiation, month 6 and 12) was transported on dry ice from EHNRI, Ethiopia to Karolinska Institutet, Campus Flemingsberg, Sweden following the standard biosafety levels. At arrival all samples were checked for integrity, registered and immediately stored in −80 °C freezer.

4.3 TREATMENT OUTCOME MEASURES

The ART outcome was analysed using the two treatment analysis outcome measurements: on-treatment (OT) and intention-to-treat (ITT) analyses.

4.3.1 On-treatment (OT) analysis

In the OT analysis, only those patients who had VL at the follow up time points were included. Virologic treatment failure was defined using two threshold VL values: 150 copies/ml (the limit of detection of our assay) and 1000 copies/ml (as per the WHO virologic failure definition) [179]

4.3.2 Intention-to-treat (ITT) analysis

In the ITT analysis, treatment failure was defined as virologic failure as defined above (OT analysis), death or LTFU. In this analysis if a patient had no data at the follow up time points he or she was considered as LTFU.

4.4 GENOTYPIC HIV DRUG RESISTANCE TESTS

Genotypic resistance test was performed on plasma specimens using the following steps: nucleic acid extraction; cDNA synthesis; nested-PCR amplification of DNA; DNA purification; and population-based DNA sequencing.

4.4.1 RNA extraction

For studies included in **Paper II – Paper IV**, HIV-1 RNA was extracted using the QIAamp® viral RNA extraction mini-kit (Qiagen, Hilden, Germany) from 140 µl plasma specimen (if VL ≥10 000 copies/µl) as per the manufacturer's instruction. However, when the viral load was below 10 000 copies/ml, the RNA extraction was done from a 1 ml plasma sample through a high-speed centrifugation at 20 000 × g for 60 min in a refrigerated centrifuge at 4°C condition. Then, using a fine-tip pipet about 860 µl of the supernatant has been carefully removed leaving about 140 µl of plasma pellet. Quantity and purity of the extracted RNA was checked using NanoDrop (Thermo Scientific, DE, USA) and stored at stored at −80 °C until further use.

4.4.2 RT-PCR

For the purpose of **Paper II**, **Paper III**, and partly of **Paper IV** (PBSS assay) complementary DNA (cDNA) synthesis was performed by reverse transcriptase polymerase chain reaction (RT- PCR) using RevertAid H-minus reagents (Life technologies, Paisley, UK) following the manufacturer's instructions. For second part of **Paper IV** (NFLG assay), two strands of cDNA were synthesised including fragment-1 (Gag-Vpu, $F1_{Gag\text{-}Vpu}$); and fragment-2 (Tat-3 LTR, $F2_{Tat\text{-}3LTR}$) as described in Grossmann *et al*. [156].

4.4.3 Nested-PCR

PCR runs were performed with high fidelity KAPA HiFi HotStart ReadyMix (2×) (KAPA Biosystem, MA, USA) following the manufacturer's instructions. In relation to **Paper II** and partly of **Paper IV** (for the PBSS assay), nested-PCR was performed through a first-round PCR using JA203F-C (forward primer) and JA206R-C (reverse) primer pairs, followed by a second-round PCR, which was performed using JA204F-C (forward) and JA205R-C (reverse) primer pairs as described in Lindström and Albert "*in-house*" method [139]; and a product spanning PR and the first two-thirds of RT gene of the HIV-1 *pol*-gene (ref HXB2 position: 2135–3338) was amplified. In **Paper III**, the *env* gene encoding the V3-V4 region was amplified as described [180], using modified primers: ED5-1F and ED12-1R for the first-round PCR run; and ES7-2F and ES8-2R for the second-round PCR.

For the second part of **Paper IV** (the NFLG assay), an 8.8 kb genome was amplified into two fragments (fragment-1 and fragment-2) by nested- and semi-nested- PCR with seven primers as described in Grossmann *et al*. [156]. Briefly, the 5.5 kb fragment-1 ($F1_{Gag\text{-}Vpu}$) was amplified through a first round PCR, which was performed with 0682F and 6352R primers followed by a second round PCR run with 0776F and 6231R primers. The 3.7 kb fragment-2 ($F2_{Tat\text{-}3LTR}$) was amplified in a semi-nested-PCR using a first- and a second-round forward primers 5550F and 5831F, respectively and a common reverse primer 9555R. Details of the primers used in **Papers II – Paper IV** are given in Table 1.

4.4.4 DNA purification

A 5 µl nested-PCR product was analysed using a 5% ethidium bromide (EtBr)-stained agarose gel electrophoresis. Then, the amplified DNA products were purified using a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) as per the manufacturers' instructions. DNA yield and quality was checked using a NanoDrop spectrophotometer and/or agarose gel electrophoresis as required. When double bands were observed in the gel electrophoresis, an additional purification step was employed through QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) before proceeding to the genotypic analysis.

4.4.5 DNA sequencing

DNA sequencing was performed using three assays: PBSS, NGS, and near full-length genome (NFLG) as follows:

4.4.5.1 Population-based Sanger sequencing (PBSS)

As part of **Paper II** and **Paper IV**, the purified DNA products were sent to Eurofins MWG operon (MWG operon, Germany) for sequencing with four primers: JA204F-C, JA205R-C, PR2R, and RT07. Similarly, for **Paper III**, sequencing was performed using two primers: ES7-2F and ES8-2R.

4.4.5.2 Next-generation sequencing (NGS)

As part of **Paper II**, NGS was performed as described earlier [181]. Briefly, a fragment of gag-pol region was amplified, gel purified, and fragmented on the Coveris S200 followed by library preparation using NEBNext UltraTM DNA library Prep Kit. Forty-eight libraries were then pooled at equimolar (10 nM each) and run on Illumina HiSeq 2500.

4.4.5.3 Near full-length genome (NFLG)

As part of **Paper IV**, NFLG was performed as described earlier [156] using 23 sequences listed in the Table 1 below.

Table 1. List of primers used PCR amplifications and sequencing.

Primer ID	Sequence (5'→3')	HXB2 Position	Application	Paper
JA203F-C	GAAAGACTGTACTGAGAGACAGGC	2058→2081	Amp.	Paper II, Paper IV
JA204F-C	TTCAGAGCAGACCAGAGCCAACAG	2135→2158	Amp., seq.	Paper II, Paper IV
JA205R-C	TTTTCCCACTAACTTCTGTATATC	3338→3315	Amp., seq.	Paper II, Paper IV
JA206R-C	TTAATCCCTGGGTAAATCTGACTT	3373→3350	Amp.	Paper II, Paper IV
RT07	AAGCCAGGAATGGATGGCCCA	2586→2606	Seq.	Paper II, Paper IV
PR2R	GGATTTTCAGGCCCAATTTTTG	2713→2691	Seq.	Paper II, Paper IV
ED5-1F	ATGGGATCAAAGCCTAAAGCCATGTG	6556→6581	Amp.	Paper III
ED12-1R	AGTGCTTCCTGCTGCTCCCAAGAACCCAAG	7822→7792	Amp.	Paper III
ES7-ET	TTRTTAAATGGTAGTATAGC	7002→7021	Amp., seq.	Paper III, Paper IV
ES8-2R	CACTTCTCCAATTGTCCCTCA	7667→7647	Amp., seq.	Paper III
5550F	AGARGAYAGATGGAACAAGCCCCAG	5550→5574	Amp.	Paper III, Paper IV
9555R	TCTACCTAGAGAGACCCAGTACA	9555→9533	Amp., Seq.	Paper III, Paper IV
6433F	CYACCAACGCGTGTGTACCCACAGA	6433→6457	Amp., seq.	Paper III
8329R	CCCTGCCGGCCTCTATTYAYTATAGAAA	8356→8329	Amp.	Paper III
0682F	TCTCTCGACGCAGGACTCGGCTTGCTG	0682→0708	Amp.	Paper IV
0776F	CTAGAAGGAGAGAGAGATGGGTGCGAG	0776→0800	Amp., Seq.	Paper IV
6352R	GGTACCCCATAATAGACTGTRACCCACAA	6352→6324	Amp.	Paper IV
6231R	CTCTCATTGCCACTGTCTTCTGCTC	6231→6207	Amp., seq.	Paper IV
5861F	TGGAAGCATCCRGGAAGTCAGCCT	5861→5884	Amp., seq.	Paper IV
G1231F	TCACCTAGAACTTTRAATGCATGGG	1231→1255	Seq.	Paper IV
E30	GTGTACCCACAGACCCCAGCCCACAAG	6445→6471	Seq.	Paper IV
G100F	TAGAAGAAATGATGACAG	1817→1834	Seq.	Paper IV
3885R	CTGCTCCATCTACATAGAA	3885→3867	Seq.	Paper IV
RTR2	CACAGCTAGCTACTATTTCTTTTGC	4350→4326	Seq.	Paper IV
VIF1F	GGGTTTATTACAGGGACAGCAGAG	4900→4923	Seq.	Paper IV
Rev11	ATCATCACCTGCCATCTGTTTTCCAT	5066→5041	Seq.	Paper IV
E70	GGGATCAAAGCCTAAAGCCATGTGTAA	6559→6585	Seq.	Paper IV
E120	GTAGAAATTAATTGTACAAGACCC	7098→7121	Seq.	Paper IV
V3R	GAAAAATTCTCCTCYACAATTAAA	7373→7350	Seq.	Paper IV
E160	GTGGGAATAGGAGCTGTGTTCCTTGGG	7761→7787	Seq.	Paper IV
E75	GCGCCCATAGTGCTTCCTGCTGCTCCC	819→7793	Seq.	Paper IV
E230	AATATTCATAATGATAGTAGGAGG	8273→8296	Seq.	Paper IV
E15	CTCTCTCTCCACCTTCTTCTTC	8445→8424	Seq.	Paper IV
LTRF2	GCTTCTACGCGTAAGAAAAGGGGGGACTGGA	9059→9089	Seq.	Paper IV
9181R	GTGTGTAGTTCTGCCAATCAGGGAA	9181→9157	Seq.	Paper IV

Amp.: amplification; seq.: sequencing

4.4.6 Bioinformatics analysis

4.4.6.1 Sequence analysis

The sequences generated with PBSS assay were aligned, and manually edited using BioEdit software version 7.2.6.1 available at: http://www.mbio.ncsu.edu/bioedit/bioedit.html. For **Paper III**, the V3 loop sequence was generated using a gene cutter program available at: http://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html.

For the NGS analysis FASTQ file was demultiplexed and a consensus sequence was created for each sample followed by realignment again with the consensus sequence as input. The variant calling was performed at amino acid level. Only amino acid covering 5000X per position was considered quality passed.

For the NFLG assay, sequence analysis was performed in such a way that a multiple sequence alignment was generated with the reference genome in AliView version software 1.17.1 as described earlier [182] and analysed with an *in-house* Perl script that recognised the nucleotide changes from the reference sequence and created a corresponding number code as per HXB2 coordinates (HXB2 reference position 790 to 9417). The resulting matrix was plotted using the TraMineR package [183] in R ver. 3.1.2 [184] to obtain a diversity plot.

Maximum likelihood phylogenetic analysis was performed using Molecular Evolutionary Genetics Analysis software version 7 (MEGA 7).

4.4.6.2 HIV subtyping

HIV-1 subtyping was performed using three different online tools: Recombinant Identification Programme (RIP, http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html), REGA HIV-1 Subtyping Tool ver. 3.0 (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool), and COMET HIV-1 (http://comet.retrovirology.lu).

4.4.6.3 HIV-1 drug resistance test and identification of amino acid mutations

The mutations considered were for drug classes NRTI, NNRTI, and PIs. In addition, mutations associated with INSTIs drug class were considered in **Paper II** for the baseline samples analysed with the NGS assay.

In **Paper II**, for the baseline samples analysed by PBSS and NGS assays, surveillance drug resistance mutations (sDRM) were identified by the calibrated population resistance tool (http://cpr.stanford.edu/cpr.cgi) at Stanford HIVDR Database and using the WHO list of DRM for surveillance of TDR. In addition, for the INSTIs associated DRM, the Stanford drug resistance summaries for INSTIs (hivdb.stanford.edu).

In **Paper II** and **Paper IV**, acquired DRMs were identified by the Stanford HIVdb Programme (hivdb.stanford.edu). Genotypic drug resistance defined as the presence of one or more major amino acid substitution included in the 2017 IAS mutation list [185] and the Stanford algorithm was used to predict drug susceptibility.

For samples analysed with the NGS assay, mutations with frequencies of 1% or more were considered for DRM interpretation. In the NFLG assay, the amino acid changes at HIV-1 *gag*, *pol*, *vif*, *vpr*, *tat*, *vpu*, and *nef* genomes in month 6 samples as compared to their corresponding baseline were manually identified using AliView ver 1.17.1 and BioEdit ver 7.2.6.1 softwares.

4.4.6.4 Genotypic tropism testing (GTT)

In **Paper III**, GTT was employed to determine viral tropism. Co-receptor tropism was predicted from the V3 loop using Geno2Pheno clonal and clinical algorithms as described earlier [186, 187]. The Geno2Pheno with false positive rate (FPR) set to 10% and 5% were used to determine GTT. The clinical model includes clinical data with nadir CD4 count and viral load, to improve the GTT prediction among ART naïve patients. A virus was considered X4-tropic or R5-tropic, if all amino acid sequences were predicted. A virus that contained both type of amino acid sequences was assigned as mixed R5 and R4 (R5/X4-tropic). The Geno2Pheno interpretations include X4-tropic if all translations of a sequence displayed FPR ≤10%; R5-tropic if FPR >10%; or R5/X4-tropic if sequence translations from a given patient displayed FPR ≤10% and >10%. The Geno2PhenoFPR10% tool was used for prediction of the virus tropism from the V3 loop sequences of the paired samples generated by the NFLG assay in **Paper IV**.

4.4.6.5 Statistical analysis

Overall socio-demographic, clinical, and laboratory parameters were summarized using descriptive statistics including mean, median, standard deviation, and percentiles for numerical variables, whereas frequencies and percentages for categorical variables. The variables analysed include: study site, gender, age, educational status, employment status, marital status, spouse HIV sero-status, functional status, WHO clinical stage, TB, OI, ART regimen, BMI, weight, CD4 cells, CD4/CD8 ratio, and VL. Differences between categorical variables were tested using Chi-square test or Fisher's Exact Test. Differences in numerical variables between two or more categories were tested with independent t-test, Mann-Whitney, Anova and Kruskal-Wallis. Multivariable logistic regression models also were developed in relation to **Paper I – Paper III**. Odds ratios (OR) with the 95% confidence intervals (CI) were presented. P-value <0.05 was considered statistically significant. Data analysis was performed using STATA software 14 (Stata Corp. College Station, Texas, USA).

In **Paper I**, a total of four multivariable logistic regression models were used to identify baseline predictors of treatment failure and LTFU at month 6 and 12. Study site, gender, age, educational status, employment status, marital status, spouse HIV sero-status, functional status, WHO clinical stage, TB, OI, ART regimen (NRTI or NNRTI), BMI, weight, CD4 cells, CD4/CD8 ratio, and VL were considered as potential predictors of treatment outcome. Both stepwise forward and backward methods in multivariable logistic regression models,

testing variable with p-value <0.1 in the bivariate analysis were used to select variables included in the best final models.

In **Paper II**, the impact of sDRM on treatment outcome during the follow up period was investigated with a multivariable logistic regression model adjusted for different confounding factors. In **Paper III**, multivariable logistic regression models also were used for analysis of virological responses to compare differences between R5-tropic and X4-tropic viruses infected patients, adjusting for different confounders. In **Paper IV**, treatment outcomes were compared between patients with different NRTI regimens and the prevalence and type of DRMs were compared between patients with TDF- or ZDV-based regimens by Chi-square or Fisher's exact test.

4.5 ETHICAL CONSIDERATIONS

Scientific and ethical approvals were obtained from the National Research Ethics Review Committee in Ethiopia (3.10|528|06) and the Institutional Review Board (IRB) of Ethiopian Health and Nutrition Research Institute (E.H.N.R.I 6.13|163 and E.H.N.R.I 6.13|164).

5 RESULTS

A total of 874 ART naïve individuals were enrolled into the first large countrywide HIV-1 cohort established at seven hospitals in Ethiopia, where 112 to 134 participants were recruited per site. The data and plasma specimens generated were analysed via studies included in the four different papers incorporated in this thesis. The thesis mainly studied predictors of the treatment outcome and LTFU (**Paper I**); prevalence and patterns of surveillance- and minority-DRM (**Paper II**); molecular characteristics of co-receptor tropism in HIV-1C (**Paper III**); and acquired DRM and viral genome wide amino acid changes in relation to the ART outcome (**Paper IV**). Summaries of the main findings of the studies addressed in each paper are given below.

5.1 BASELINE PREDICTORS OF ANTIRETROVIRAL TREATMENT FAILURE AND LOST TO FOLLOW UP IN A MULTICENTER COUNTRYWIDE HIV-1 COHORT STUDY IN ETHIOPIA (PAPER I)

Four multivariable logistic regression models were developed using the data collected from the 874 HIV-1 infected patients at baseline, month six and 12 after ART initiation. The details of socio-demographic characteristics, clinical and laboratory findings at baseline are presented in Table 1 (**Paper I**). Briefly, the majority of the participants (60%, 527/874) were females; median age 33 years (IQR: 28 – 39 years); one-fourth of the participants had TB co-infection; 38% had one or more OIs and/or AIDS-defining conditions; mean CD4 count was 144 cells/μl (SD: 87); three-fourth of the participants had below 200 cells/μl; and mean HIV-1 RNA load was 5.2 log10 copies/ml (SD: 0.8).

5.1.1 Treatment outcome

Among the 874 patients enrolled into the cohort, 743 (85%) and 690 (79%) reached their 6 and 12 months follow up period, respectively. On the other hand, 131 (15%) and 184 (21%) were LTFU, respectively. Confirmed death accounted for 62 (47%) and 75 (41%) of the LTFU at the respective time points. Among 676 and 459 patients with available VL data at month 6 and 12 after initiation of ART, the OT analysis identified 57 (8%) and 34 (7%) virologic failure patients (VL ≥1000 copies/ml), respectively. Hence, in the ITT analysis, the respective treatment failure rates were 23% and 34%. Details of the cohort status during the one year ART period with the subsequent treatment outcome is outlined below in Figure 3.

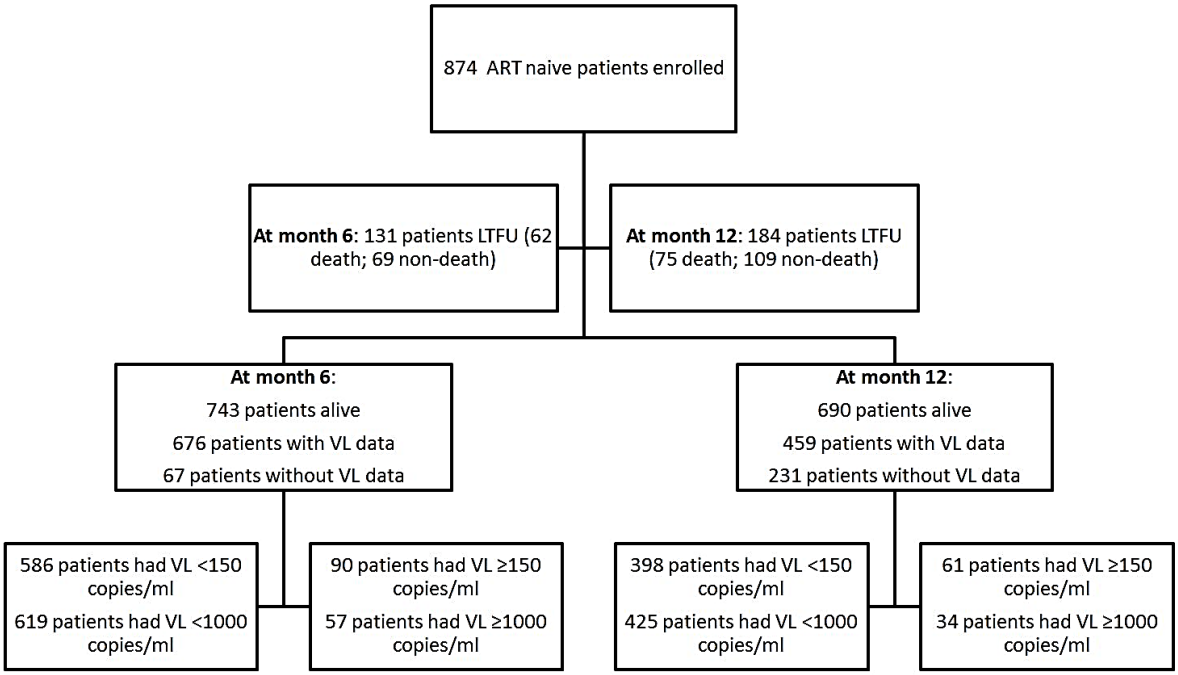

Figure 3. Advanced Clinical Monitoring of Antiretroviral Therapy in Ethiopia (ACM) cohort profile and treatment outcome during the one year treatment follow up period.

5.1.2 Baseline predictors of treatment failure

In a multivariable logistic regression model developed at month 6, the odds of treatment failure was predicted by gender, TB, OI, CD4 count, VL and study site. At month 12, by functional status, TB, CD4 count and study site. In brief, TB co-infected patients had nearly twice higher odds of treatment failure than TB negative participants both at month 6 (OR: 1.75; 95% CI: 1.23 to 2.50) and month 12 (OR: 1.81; 95% CI: 1.25 to 2.61). Patients who started ART with CD4 count below 50 cells/µl had nearly three times higher odds of treatment failures at month 6 (OR: 3.0; 95% CI: 1.77 to 5.07) and month 12 (OR: 2.54; 95% CI: 1.49 to 4.30) than those who started the treatment with over 200 cells/µl. The details of the baseline predictors included in the multivariable logistic regression models for treatment failure are presented in Table 2 (**Paper I**).

Study participants from some regional hospitals such as Harrar, Hawassa, and Mekelle had much higher odds of treatment failure at month 12 (OR: 3.89, 95% CI: 2.04 to 7.39; OR: 2.75, 95% CI: 1.53 to 4.94; OR: 2.01, 95% CI: 1.11 to 3.64, respectively) than those from the national tertiary level hospital TASH. The proportions of treatment failure from these hospitals were also found to be significantly higher: 41% from Harrar, 45% from Hawassa and 37% from Mekelle compared to 24% from TASH.

5.1.3 Baseline predictors of LTFU

Likewise, a multivariable analysis identified functional status, WHO stage, CD4 cell count and study sites as predictors of LTFU at month 6. At month 12, gender, functional status, WHO stage, TB, CD4 cell and study site were predictors of LTFU. Briefly, patients with functional disabilities had higher odds of LTFU (OR: 2.37; 95% CI: 1.54 to 3.63 for

ambulatory patients; OR: 2.22; 95% CI: 0.86 to 5.71 for bedridden patients at month 6 and OR: 2.20; 95% CI: 1.49 to 3.24; OR: 3.63; 95% CI: 1.61 to 8.18 at month 12) as compared to those who had working functional status at ART initiation. Patients with advanced WHO clinical stage had nearly twice higher odds of LTFU (OR: 1.79; 95% CI: 1.21 to 2.37 at month six and OR: 1.92; 95% CI: 1.36 to 2.72 at month 12) compared to those with early WHO clinical stage I/II. Patients who started ART with lower CD4 count such as below 50 cells/µl had higher odds of LTFU (OR: 2.32; 95% CI: 1.28 to 4.19 at month 6 and OR: 2.65; 95% CI: 1.57 to 4.46 at month 12) compared to those started with over 200 cells/µl. TB co-infected patients also had higher odds (OR: 1.97; 95% CI: 1.37 to 2.84) of LTFU at month 12 compared to TB negative participants. Details of the multivariable logistic regression models for the LTFU both at month 6 and 12 are presented in Table 3 (**Paper I**).

Again, participants from regional hospitals such as Harrar, Hawassa and Gondar had higher odds of LTFU (OR: 2.75, 95% CI: 1.30 to 5.81; OR: 2.39, 95% CI: 1.12 to 5.07; OR: 2.25, 95% CI: 1.06 to 4.80 at month 6; and OR: 2.20, 95% CI: 1.17 to 3.13; OR: 2.16, 95% CI: 1.15 to 4.06; OR: 1.83, 95% CI: 0.97 to 3.47 at month 12, respectively) compared to TASH. Higher proportions of LTFU were identified from Harrar (27%, and 51%), Hawassa (20%, and 38%), and Gondar (19%, and 29%) compared to TASH (10%, and 17%) at month 6 and 12, respectively. Comparative proportions of the treatment failure and LTFU across the study sites are presented on Figure 2 (**Paper I**).

5.2 PRETREATMENT DRUG RESISTANCE IN A LARGE COUNTRYWIDE ETHIOPIAN HIV-1C COHORT: A COMPARISON OF SANGER AND HIGH-THROUGHPUT SEQUENCING (PAPER II)

The prevalence and patterns of surveillance- and minority-DRM at ART initiation were determined using PBSS and NGS assays. The detected sDRM by the two assays were compared and the impact of TDR on first-line ART outcome was also evaluated. In addition, INSTI associated DRM were identified by NGS assay. The details of patients' baseline characteristics included in the two assays are presented in Table 1 (**Paper II**).

5.2.1 sDRM detected by PBSS

The PBSS assay was attempted on baseline plasma samples derived from 490 randomly selected patients after stratifying those with available samples by study sites (70 from each site). Of the 490 samples, the *pol*-gene was successfully sequenced for 461 (94%) and all sequences clustered within HIV-1C, except one CRF02_AG. Of the 461 sequences, sDRM was detected in18 (3.9%) including nine NRTI-, seven NNRTI-, and two PI- associated resistance mutations. No sample found with dual class sDRM, but three were found to have double mutations from the same ARV drug class. Among the baseline parameters tested for their association with the detected sDRM, only participants' age (OR: 0.93; 95% CI: 0.87–0.99) and viral load (OR: 2.67; 95% CI: 1.25–5.71) found to have significant correlation. The details of the sDRM detected by the PBSS assay and the study profile are presented in Table 2 and on Figure 1 (**Paper II**), respectively.

Considering the WHO virologic failure definition (VL >1000 copies/ml) and using the OT analysis, those patients with RTI- sDRM had higher odds of treatment failures after 6 and 12 months of ART than those without such mutations (OR: 9.0, 95% CI: 1.9–43.3 and OR: 7.4, 95% CI: 1.5–35.0 adjusted for tuberculosis co-infection, respectively). In the ITT analysis, those with RTI- sDRM had nearly four times higher odds of treatment failures at month 6 than those without the mutations (OR: 3.8, 95% CI: 1.4–10.5), but no significant association was found at month 12.

5.2.2 DRM detected by NGS

The NGS assay was performed on 109 selected baseline samples derived both from patients who had detectable viremia at month 6 and/or month 12 (n=71) and undetectable viremia (n=38). All of the 109 baseline samples were successfully sequenced and DRM (with mutation frequency >1%) was detected in 28 (25.7%) patients; 23 from the virologic failure patients, and five from the virologic suppressors. The NGS assay detected DRMs associated with one to three different classes of ARVs including NNRTIs alone in ten sequences; NRTIs in eight, PIs in three; INSTIs in two; NRTI and NNRTI in one; NRTI and PI in one; NRTI and INSTI in one; and three drug classes (NRTIs, PI and INSTIs) in two sequences. In addition to the RTI- and PI- sDRM, the NGS assay detected major INSTI-DRMs in minor viral variants from five patients including E138K (n=2), Q148R, Q148H, and T66I.

5.2.3 Comparative analysis of the sDRM detected by NGS vs PBSS

The NGS assay detected RTI- or PI- associated sDRM more frequently than PBSS (23.9%; 26/109 vs 6.4%; 7/109; p<0.0001). Again, when only the RTI-sDRM from virologic failures was separately considered, the NGS detected the mutations more frequently than the PBSS assay (28.2%, 20/71 vs 8.54%, 6/71; p=0.004). The odds of detection of RTI- and/or PI-associated sDRM by NGS was higher among patients with detectable viremia at month 6 and/or 12 as compared to those who had undetectable viremia (OR: 6.4; 95% CI: 1.6 – 26.4 adjusted for NRTI regimens and CD4 cells). Details of the DRM detected by NGS are presented in Table 3 (**Paper II**).

5.3 MONOPHYLOGENETIC HIV-1C EPIDEMIC IN ETHIOPIA IS DOMINATED BY CCR5-TROPIC VIRUSES-AN ANALYSIS OF A PROSPECTIVE COUNTRY-WIDE COHORT (PAPER III)

A randomly selected 420 baseline samples (60 from each site) was attempted for PBSS analysis of the V3 loop, and successful sequences were obtained from 352 (84%). The V3 loop was also successfully sequenced from 34 and 19 samples of virologic failure patients at month 6 and 12, respectively. Additionally, 387 historical sequences of the HIV-1C_{ET} V3 loop (archived from Ethiopia during 1984 – 2003) were obtained from the Los Alamos HIV Sequence Database for comparison with our sequences. Of the 352 successful baseline sequences, 350 (99.4%) clustered to HIV-1C and the reaming two assigned as subtype A1.

5.3.1 Genotypic tropism testing (GTT)

Of 352 V3 loop sequences generated from the baseline samples, the Geno2Pheno$_{FPR10\%}$ clinical model predicted R5 in 285 (81%); X4 in 60 (17%); dual R5/X4 in 7 (2%) sequences. The Geno2Pheno$_{FPR10\%}$ clonal model also predicted R5 in 291 (83%); X4 in 50 (14%); and dual R5/X4 in 11 (3%) sequences. Overall, the X4-tropic strains were present in 107 (30%) by any of the two GTT models, where only 21 (6%) were pure X4-tropic strains by both models. Of the parameters tested at baseline, CD4 count found to have significant association with the virus tropism predicted by the clinical model ($p < 0.001$), but not by the clonal model. When Geno2Pheno$_{FPR5\%}$ was used, X4-tropic strains were predicted in 47 (13.3%) sequences by any of the models. The virus tropism had significant association with CD4 count by both models ($p = 0.001$ and $p = 0.023$ by the clinical and clonal models, respectively). Details of the tropism predictions are presented in Table 2 and on Figure 2 (**Paper III**).

The Geno2Pheno$_{FPR10\%}$ comparative analysis of the historical sequences with ours revealed that the proportion of X4- and dual R5/X4-tropic strains increased in Ethiopia over time from 5.6% in 1984 – 1993, to 7.1% in 1994 – 2003, and to 17.3% in 2009 – 2011 ($p<0.001$).

5.3.2 Impact of baseline tropism on treatment outcome

The Geno2Pheno$_{FPR10\%}$ clonal model predicted R5 in 29 (85%) and X4 or dual X4/R5 in five (15%) of the 34 sequences from month 6 samples. A 7% (2/29) and a 40% (2/5) tropism switch were observed in the respective groups and the difference in the virus tropism switch was statistically significant ($p= 0.006$). At month 12, R5 was predicted in 15 (79%) and X4 or R5/X4 in four (21%) of the 19 V3 loop sequences. Only three of the later four X4 containing viruses switched from R5, but the difference was not statistically significant. In addition, the multivariable analysis could not find significant differences in treatment outcome between patients with X4-tropic and R5- tropic virus.

5.4 HIV DRUG RESISTANCE AND GENETIC DIVERSITY AMONG PATIENTS FAILING FIRST-LINE THERAPY IN THE FIRST LARGE COUNTRYWIDE ETHIOPIAN HIV COHORT ASSESSED BY SANGER- AND NEAR-FULL LENGTH SEQUENCING (PAPER IV)

Treatment outcomes and emergence of HIV DRMs and amino acid changes among virologic failure patients after 6 and 12 month of ART were described. OT and ITT analyses were performed on data collected from the 874 patients who had received first-line ART regimens including TDF, 467; ZDV, 362; d4T, 44; and ABC, one patient. However, there was no statistically significant difference in treatment outcome (both by OT and ITT analysis) across the NRTIs. Prevalence and patterns of acquired DRM and amino acid changes were described using PBSS and NFLG assays. The details of patient baseline characteristics included in the different studies and the treatment outcome across the NRTIs were presented in Table 1 and Table 2, respectively (**Paper IV**).

5.4.1 Acquired DRM detected by PBSS

Of the 57 month 6 samples attempted for PBSS assay, 47 were successfully sequenced from 28/34 (82.6%), 15/19 (78.9%), and 4/4 (100%) patients who failed TDF-, ZDV-, d4T-based first-line ART regimens, respectively. Major DRMs associated with RTIs were detected in 36 (76.6%) samples including 21 from TDF- and 12 from ZDV-failure patients. The DRM detected in the 21 TDF failures include 17 NRTI+NNRTI; and four NNRTI alone. Ten (35.7%) patients had K65R. The DRM among 12 ZDV failures include five NRTI+NNRTI; six NNRTI alone, and one NRTI only). No K65R was found from patients who received ZDV-based regimens.

Likewise, of the 33 month 12 samples attempted for PBSS analysis, 30 were successfully sequenced from 16/17 (94.1%), 10/12 (83.3%) and 4/4 (100%) patients who failed TDF-, ZDV-, d4T-based regimens, respectively. Major RTI-DRMs were found in 20 (66.7%) samples including 13 from TDF- and five from ZDV- failure patients. The DRM detected among the 13 TDF failures include 12 NRTI+NNRTI; and one NNRTI alone. Half of the TDF failure patients had K65R. The DRM found among ZDV failures include two NRTI+NNRTI; two NNRTI alone; and one NRTI alone. At month 6, no difference was found between TDF- and ZDV- failing patients with regard to the number of RTI-DRMs. However, at month 12 sequences from TDF-failure samples had more DRM associated with RTI than those from ZDV-failures ($p= 0.017$), and also when stratified by drug class including NNRTI-DRM ($p= 0.037$), and NRTI-DRM ($p= 0.040$). List of the DRM per patient and the details of DRM across NRTIs were given in Supplementary Table S1 and Table 3, respectively (**Paper IV**).

5.4.2 Amino acid changes identified by NFLG

Amino acid changes during the first six months of ART were analysed using NFLG assay, which generated successful sequences on HIV-1 *gag*, *pol*, *vif*, *vpr*, *tat*, *vpu*, and *nef* genomes from 32 (16 paired baseline and month 6) samples. In addition, the *env* (V3 loop) region was successfully sequenced in 13 baseline and 14 month 6 samples for analysis of co-receptor tropism. The LTR region was also successfully sequenced in all of the 32 samples for analyses of the NF-kB.

Analysis of the *pol* gene revealed that 11 (68.8%) patients had acquired one to five (median: three) major DRMs associated with RTIs after six months of ART including seven associated with NRTI+NNRTI; and four with NNRTI. Although none had known PI- or INSTI- DRMs, several other mutations were identified in the PR and IN regions. Four amino acid insertions (GLIP/GALN/GTLV/GALN) were also displayed at positions 48 – 55. The predicted sensitivities to N/NRTI were presented on Figure 2 (**Paper IV**).

Analysis of the *gag* gene showed that amino acid changes mostly clustered in the p6 and p17 regions. In the Gag-p6 region, frequent mutations and insertions were found. For example, PYKE insertion was found in nine (56.3%) of the 16 paired samples. In addition, duplicate tetra-peptide motif PTAP was identified in two samples at baseline, where the respective

month 6 samples had single motif (deletion of one of the double motifs). A triple PTAP tetra-peptide motif was also identified from paired samples of one patient. R4S mutation was identified in 14 (87.5%) patients at the C-terminal of p17. A number of mutations were also identified in the *vif* and *nef* genes. On the other hand, the *tat*, *vpu* and *vpr* genes showed very few amino acid changes. Detail representation of the amino acid changes in the different genomic regions are presented in Figure 3 (**Paper IV**).

The Geno2Pheno$_{FPR10\%}$ tool predicted R5 in 10 (77%) and 12 (86%) of the 13 and 14 baseline and month 6 V3 loop sequences, respectively. The two X4 and one R5/X4 dual tropic viruses found at baseline switched to R5 and X4, respectively. Of the 16 samples, 15 (94%) displayed three nuclear factor kappa B (NF-kB) each in the enhancer region of the LTR of HIV-1C. Whereas, one sample shows large insertions and four instead of two binding sites after 6 months of ART.

6 DISCUSSIONS

For more than a decade, ART has been widely and rapidly scaled up in Ethiopia through the WHO public health approach and the treatment coverage is reported to have reached 60% among PLHIV [76]. However, the overall health care system is inadequate to meet the complex monitoring standard required by the ART service such as viral load and GRT. Besides, a high rate of LTFU has been reported [44, 166]. All this contributes to a high rate of late diagnosis of treatment failure patients with subsequent emergence and accumulation of drug resistance as reported also from other sSA [91]. However, data is not earlier available from Ethiopia, at least at the national level. In addition, even if HIV-1C was discovered in Ethiopia [3], the distribution of subtypes has not been characterised at a countrywide level.

Previous studies have had a limited geographical coverage, representing mainly the capital Addis Ababa and Northwest part of the country. Thus a clear picture of the epidemic in the country had not earlier been provided. Therefore, in view of filling the abovementioned research gaps, I have performed the studies included in this thesis using a countrywide HIV-1 cohort. A total of 874 ART naïve individuals were enrolled in this HIV-1 cohort established at seven hospitals in Ethiopia, where 112 to 134 participants were recruited per site. The cohort had a good geographical distribution, which should represent the national HIV epidemic profile for the first time. In addition, among the 874 patients enrolled, 743 (85%) and 690 (79%) reached their 6 and 12 months follow up time point, respectively. The GTR study conducted certainly met the WHO recommendation mentioned in its "2012 *Global Strategy for the Surveillance and Monitoring of HIV Drug Resistance* [188]," to conduct cross-sectional surveys of acquired HIVDR at nationally representative university affiliated hospital ART clinics. Therefore, the findings described in my thesis most likely reflect the status of the ART outcome including LTFU and death; drug resistance both PDR and ADR; and the HIV-1C epidemic at the national level. The main findings from each paper are discussed below:-

In **Paper I**, we found that the virologic failure rate as determined by OT analysis (VL >1000 copies/ml) was below 10% both at month 6 and 12. This was comparable with earlier reports from other sSA studies [189, 190]. The result was in harmony with the last 90 in the UNAIDS 90-90-90 ambitious strategy [122]. However, in the ITT analysis (VL >1000 copies/ml plus LTFU including death), the treatment failure rate was substantially higher, 23% and 34%, at month six and 12, respectively. Although this was also comparable with reports from other sSA [191-193], it was substantially lower than the retention rate targeted by the UNAIDS 90-90-90 strategy. The study identified early mortality and LTFU as the main reasons for poor ART outcome in Ethiopia, as reported from other sSA countries [194]. The early mortality and the low CD4 T-cell counts at diagnosis also clearly indicate that the proportion of PLHIV who know their status is very much lower than the first 90. This suggest a need for continued supportive efforts in order to reduce early mortality and LTFU in the ART programme but also to diagnose patients at a much earlier phase of infection.

We developed four multivariable models, which identified male sex, functional disabilities, advanced WHO clinical stages, TB and other OIs, low CD4 cell count, and high VL as baseline predictors associated with increased odds of viral failure and/or LTFU at month 6 and 12. Of course, this has already been described in previous studies done in other LMIC. However, the design of this multisite countrywide HIV-1 cohort allowed us to identify study site as one of the strongest predictors of ART outcome in Ethiopia. Regional differences in the prevalence of sexually transmitted infections and variations with regard to knowledge towards HIV prevention have previously been reported in Ethiopia [195]. In the present study, we have for the first time identified significant regional differences in ART outcome in the country. This may be attributed to socio-demographic and socio-cultural variations along with cultural and religious practices within different parts of the country, although the design of the cohort did not allow us to study this. In view of the ART outcome differences observed across the hospitals, we recommend provision of additional support to strengthen capacity of the regional health care system.

Due to deficiencies in the health care systems, the rapid scale-up of ART in LMIC has raised concerns for enhanced accumulation of PDR and its onward transmission to newly infected patients [196]. The PDR rate has recently been reported to increase in the African region. However, there are only few studies showing the pattern and prevalence of PDR in Ethiopia and these studies have had limited geographical coverage, lacking continuousness and the national prevalence has remained to be determined. Therefore, using the countrywide HIV-1 cohort, prevalence and patterns of PDR were studied in **Paper II** by employing both PBSS and NGS assays. The respective assays detected 3.9% and 25.7% DRM rates, among 461 and 109 baseline samples of ART naïve patients. The prevalence detected by the PBSS assay was low, especially when considering the recently observed increased rates of PDR in the East African regions [77]. It was also below the WHO threshold for surveillance of TDR. However, it should be noted that the samples analysed in this study were collected during 2009 – 2011 and the current TDR prevalence rate may have been changed.

Further, considering the fact that majority of the participants had low CD4 cell counts, they can be classified as late presenters. Consequently, the TDR rate might have been underestimated by the PBSS assay since some drug resistant variants frequently disappear from the major viral population, especially during the period when ART is not initiated [77]. Hence, we also determined the baseline DRM by the robust NGS assay with capacity of detecting mutations also in the minor population (>1% frequency), where the PDR prevalence was found to be 25.7% in our study. Although the samples were not randomly selected and there could be a sample selection bias, our baseline DRMs prevalence detected by the NGS assay was supported by the recently observed increasing trends in East Africa [77].

Since the rollout of ART a temporal increasing trend of TDR has been observed across the few available scattered reports from Ethiopia including <5% from Addis Ababa in 2005 [175]; 5.6% from Gondar, Northwest Ethiopia in 2008 [177]; and 7.2% again from Gondar in

2010 [197]; all studies were done with Sanger-based assays. Effects of PDR on the virologic outcome and emergence of DRM among treated patients have been previously reported from sSA [98]. The observed increasing trend could compromise the effectiveness of the few available first-line ART regimen options in the region thereby also limit the clinical benefits from ART. Hence, genotypic resistance screening would be essential at ART initiation. Nevertheless, subsequent studies on samples collected from newly infected individuals are recommended to ascertain the current magnitude of TDR in Ethiopia.

In our comparative analysis NGS detected a higher rate of sDRM than PBSS. Additional baseline DRMs were detected from 17 patients selected for the NGS, which was consistent with an earlier report of a high detection rate of NNRTI- TDR by an allele-specific PCR (AS-PCR) (6.5%) in Addis Ababa, performed by our group [134]. The discrepancy between our PBSS and NGS assays suggests that NGS could facilitate detections of sDRMs in LMIC by the same or lower costs if high-throughput assays are used and could therefore be an alternative method for screening surveillance of TDR/PDR [198].

With the NGS, we also identified major DRMs associated with INSTIs in five ART naïve individuals, including T66I, E138K, Q148R, and Q148H. These DRMs alone or in combinations associated to different levels of RAL and/or EVG resistances [72]. Although INSTIs are planned to be integrated in the ART regimen in some African countries, they are not part of the treatment strategy in Ethiopia so far neither is an approval of its use in the country documented. However, it is not unlikely that the second-generation INSTI, such as DTG and BIC, will become valid alternatives also in Ethiopia since the cost is now becoming dramatically lower, as compared to HIC.

In addition, with maximum likelihood phylogenetic analysis, we confirmed that there was no clustering across the samples harboring the INSTI associated DRMs (Figure 2). Patients whose baseline samples harboured the INSTI-DRM were from five different geographically distinct hospitals, where possible transmission of the INSTI-resistant viral variants was less likely. Thus, wild type HIV-1C strains might harbor low abundance of INSTI-DRMs in Ethiopia. However, potential clinical impacts of the detected DRMs still remain to be evaluated because some studies reported that low abundance of INSTI-DRMs had no effect [199, 200]. These circumstances warrant further studies on INSTI-DRMs in the minor virus population in Africa and elsewhere, where some countries plan to deploy the use of DTG in the near future.

In **Paper III**, we assessed the HIV-1 subtype distribution across the different parts of Ethiopia through sequencing and analyzing the *env* gene. Of the 352 successful sequences, 350 (99.4%) clustered to HIV-1C, which was consistent with previous reports. Since its first identification in Ethiopia in 1986 [3], HIV-1C has consistently been reported as the predominant subtype in the country. For example, studies conducted in the mid- and late-1990s mostly among commercial sex workers and pregnant women in Addis Ababa using the *env* gene identified HIV-1C in 98.5% to 100% [36, 201, 202]. Later studies from Gondar,

Northwest Ethiopia, using *gag-pol* and *pol* sequences also identified HIV-1C in 97.8% [176] to 98.7% [177] of the infections.

Although our study and earlier reports have shown that HIV-1C have monophylogenetically dominated (>99%) the epidemic in Ethiopia since its estimated introduction into the country in 1970 [203], this situation contradicts with other African countries' reports, even with the neighboring East African, where different HIV-1 subtypes are co-circulating [20]. Studies from neighboring countries showed that HIV-1C has been co-circulating with other subtypes mostly with subtype A and D. For example, in Djibouti, HIV-1C accounted only for 65.7% of the subtypes [204]; in northern Kenya, 27% [205]; in western Kenya, 6% [206]; and in Sudan, 30% [207]. HIV-1 subtype A and D together accounted only for <1% of the epidemic in Ethiopia, which are likely to have been introduced from the abovementioned neighboring countries.

In this study, we also assessed the distribution of co-receptor tropism at baseline and its impact on virologic- and immunologic-treatment outcome after 6 and 12 months of ART. Tropism switch after 6 and 12 months of first-line ART was also determined. The $\text{Geno2Pheno}_{\text{FPR10\%}}$ clinical model predicted R5-tropic viruses in 81% of the baseline sequences generated from the V3 loop; X4 in 17%; dual R5/X4 in 2%. The respective predictions by the $\text{Geno2Pheno}_{\text{FPR10\%}}$ clonal model were 83%, 14%; and 3%. Overall, both GTT tools predicted that the epidemic in Ethiopia is dominated by CCR5-using HIV-1C strains, which is in line with earlier studies involving samples from Ethiopia [208, 209]. However, comparative analysis using historical sequences and the sequences generated in our study showed a slight temporal increment in proportion of X4- and dual R5/X4-tropic HIV-1C strains in Ethiopia. This finding is in agreement with other studies done in Botswana [210], South Africa [211], and India [212], where increasing trends in X4-tropic HIV-1C strains have been documented during the last decade.

Although studies have shown an effect of the baseline co-receptor tropism on disease progression and treatment outcome [213, 214], the multivariable analysis in our study could not find such effect. Even if we found too few X4-tropic strains at baseline, a bidirectional tropism switch was identified both among month 6 and 12 samples of virologic failures, but a switch from X4 to R5 was more common. However, whether this is a common feature of HIV-1C is not yet fully understood.

In **Paper IV**, we described treatment outcomes in relation to the emergence of HIV drug resistance and amino acid changes among virologic failure patients. As discussed in **Paper I**, findings from this countrywide HIV-1 cohort confirmed that early death and LTFU are the main reasons for poor treatment outcome in Ethiopia. However, a wide range of DRMs to the first-line ART regimens was commonly detected among virologic failures, who had also continued the treatment. Up to seven (median three) and eight (median four) HIV-1 DRMs were identified per sequence from 76.6% and 66.7% of the virologic failures at month 6 and 12, respectively. Majority (66.7% and 64.0%, respectively) of the detected DRMs per sequence conferred resistance to dual drug classes (NRTIs and NNRTIs).

Although the viral *pol* sequences from earlier studies in Ethiopia are few and the follow up time points and viral load cut-off differ from ours, the DRMs rates observed in our study were lower than reports from Jimma (81.8%, 9/11) at month six [171] and from Gondar (75%, 6/8) at a median time of 30 months on ART [178]. We have also reported a wider pattern of NRTI-DRMs as compared to the studies from Jimma and Gondar, where only limited types of NRTI associated DRMs were reported. However, the rate of DRMs and mutation patterns identified in our study were generally in agreement with a report from a multicenter cohort in six sSA countries, where 70% of the sequence from virologic failures (after 12 month of ART) harbored DRMs and 49% had dual class (NRTI and NNRTI) resistance, with an average of 2.4 DRMs per sequence (range: 1 – 8) [98]. However, the average number of DRMs and the proportion of dual class drug resistance observed in our study were higher than the sSA multicenter cohort.

Neither a difference in virologic failure rate nor in DRM rate was observed across the NRTI-based regimens. The most frequently observed NRTI-DRMs were K65R, M184V/I, and TAM, but the K65R mutation was only found among TDF-treated patients. In a recent report from Gondar a low frequency of K65R was reported [172]. The DRM patterns we observed were relatively in agreement with the multicenter sSA countries cohort, where M184V, K65R, and TAMs were the most commonly observed NRTI-DRMs [98]; and a Ugandan study, where M184V, K65R, and TAMs were most frequently observed [193]. In contrary, a study from Kenya did not report the K65R mutation, but none of their patients took TDF-based regimens [206].

K65 is one of the notable residues due to its pivotal position in the deoxynucleotide triphosphate (dNTP)-binding pocket; and its involvement in nucleoside analogue resistance and polymerase fidelity [215]. Mutations at the K65 position of HIV-1 RT reduces susceptibility of the virus to different NRTIs including TDF, 3TC, FTC, and ABC [216]. Studies have also shown that K65R mutation is preferentially selected by HIV-1C in *ex vivo* and *in vitro* analyses [217, 218]. As per the WHO recommendations, TDF has replaced thymidine analogues including ZDV and d4T in first-line ART regimens. The majority of our study participants received TDF-based first-line regimens. As described in the background section (section 1.5.2.2), studies have shown that TDF–resistant viruses have a substantial transmission potential [92, 99-101], and subsequently an increasing prevalence of resistance against this drug is an important concern.

Overall, our findings showed that the main reason or consequence of virologic failure was the emergence of drug resistance. The most important factors influencing the emergence of HIVDR in sSA include limited treatment options; treatment interruptions due to drug stock outs; and poor treatment monitoring strategies, where viral load and GRT are not used to monitor the ART outcomes [91, 92, 97]. Our study also pointed out that further expansion of TDF in HIV-1C dominated sSA region including Ethiopia, where viral load monitoring is not routinely used (at least presently) may result in an extensive spread of the K65R mutation. Together with the high potential transmission capacity of K65R strains, the mutant viruses

may jeopardize future therapeutic and prophylactic use of TDF and tenofovir alafenamide [219].

A possible way forward to improve treatment outcome could be the use of long-acting drugs, such as the INSTIs cabotegravir, which inhibits HIV-1C at least as efficient as HIV-1B [220], and the NNRTI rilpivirine. However, both approaches should be further evaluated by several reasons of which one reason might be the presence of major INSTI-DRMs in minor viral variants, as identified among five ART naïve patient samples in our cohort [96]. Also rilpivirine treatment can be questioned due to the high viral load in the majority of our patients at diagnosis and the less binding efficacy of the drug to the HIV-1C reverse transcriptase [96, 221].

In addition to analyses of the *pol* region by PBSS, we also employed NFLG to describe all relevant target regions of HIV-1 for drug resistance. The NFLG analysis was done in 32 (16 paired baseline and month 6) samples of virologic failure patients. Although the point mutations we identified are not known to influence the response to PIs, the NFLG assay identified *gag* mutations in the majority of virologic failures, especially in the p6 and p17 regions. Gag is essential for virus particle formation and its C-terminal p6 domain harbors short peptide motifs that facilitate virus release from the plasma membrane and mediate incorporation of the viral Vpr protein. Hence, although the sites and functional relevance of Gag-p6 is not completely known, changes in the C-terminal Gag-p6 domain plays important role in the release of viral particles [222]. HIV-1C also has unique amino acid frequencies in this region.

We found a PYKE tetra-peptide in the ALIX-binding motif of Gag-p6 in nine paired samples (both at baseline and month 6) which remained unaffected by the treatment in all patient samples. This tetra-peptide motif was observed also among half of treatment naïve Ethiopian patients in our earlier study, but the status was not known among ART experienced patients [223]. However, this study identified the PYKE motif only in few sequences from South African and Indian ART naïve HIV-1C infected patients (1% and 3%, respectively). Therefore, it is important to study the clinical relevance of the PYKE motif in terms of viral fitness and susceptibility to ART, especially to PI drugs with a larger number of HIV-1C viruses.

On the other hand, deletion of the PTAP motif in the Gag-p6 was identified in two samples (at month 6), which initially had double motifs in their corresponding baseline samples. Subtype specific differences have earlier been observed in Gag-p6 with regard to this tetra-peptide motif [224, 225]. A study also showed a difference in ART outcome in relation to duplication of the PTPA motifs including HIV-1C [225].The duplication probably restores the ALIX mediated virus release pathway, which is lacking in HIV-1C as PTAP motif is thought to be a key player in viral budding [224]. However, the disappearance of the duplicated PTPA tetra-peptide motifs in most of our patient paired samples contradicted with what has been reported in a study which also included HIV-1C [225]. In addition, to the best of our knowledge, so far no study has shown a triple PTPA motif. Therefore, we recommend

further studies about the significance of this tetra-peptide motif on treatment outcome and its clinical relevance.

It can also be noted that resistance to DTG has been described outside the integrase, mainly in *nef* and LTR overlapping region *in vitro,* and in patients failing monotherapy with DTG [226]. In contrast to the findings in the *gag* gene, no common pattern of amino acid changes was seen in the other genes, with the exception of known RTI-DRM in the *pol* gene. However, in the NFLG sequences we identified three NF-kB binding regions in the LTR which is unique for HIV-1C as earlier described by our group [227].

In summary, this thesis presents for the first time major findings on the HIV-1C epidemic and the antiretroviral treatment response including the molecular epidemiology of drug resistance at a countrywide level in Ethiopia. Our findings from the on-treatment vs intention-to-treat analyses pointed out that early LTFU and mortality are outstanding problem in the treatment of HIV rather than detectable viremia while being on treatment. The genotypic resistance tests by PBSS, NGS, and NFLG detected broad DRMs associated with different ARV drug classes including RTIs, PIs, and INSTIs. Our findings also showed that R5-tropic HIV-1C monophylogenetically dominate the epidemic, although a temporal slight increasing trend in X4-tropism usage observed. Hence, this thesis provides a range of fundamental information from the molecular epidemiology of HIV-1C epidemic to its therapeutic responses in Ethiopia that could be used by researchers, clinicians, policy makers, and other stakeholders. In addition, the thesis is based on one of the largest countrywide HIV-1 cohort in Africa; the findings have most likely implications in other HIV-1C dominated countries.

7 CONCLUSIONS AND FUTURE PERSPECTIVES

This thesis provides an account of the HIV-1C epidemic and ART outcome in Ethiopia using the first large countrywide HIV-1 cohort, which most likely reflects the national profile during the study period. The following specific conclusions and recommendations can been drawn from the studies included in the thesis:

- The on-treatment virologic suppression rate was above the WHO-suggested targets for HIV drug resistance prevention, but a high rate of treatment failure was identified by the intention-to-treat analysis, reflecting that early death and LTFU are major reasons for poor ART outcome in Ethiopia. It is possible that the recently recommended universal test and treat strategy could positively contribute in reducing the observed LTFU and death rate because the majority of the participants in this study started ART with low CD4 cell counts and advanced WHO clinical stages.
- In addition to well-known determinants such as advanced WHO clinical stages, low CD4 cell count, TB and other OIs , the study site was found to be a strong baseline predictor of treatment failure. Compared to national tertiary hospital located in the capital city, the regional hospitals had a higher proportion of treatment failure, which highlights the need for provision of more regional support .
- Our comparative study on sDRM showed that NGS detected a higher rate of TDR as compared to PBSS, which underestimates the prevalence of TDR. The NGS also identified major INSTI resistance mutations in minor viral variants. However, their clinical impact should be confirmed. In addition, our findings highlighted the potential value of using high-throughput NGS in surveillance studies of PDR in LMIC.
- Broad NRTI- and NNRTI-DRMs patterns were identified among three quarters of the virologic failure patients who continued the treatment, but there was no significant difference with regard to treatment outcome or level of drug resistance between patients who were treated with TDF- or ZDV-based first-line ART regimens. This shows the need for application of viral load testing and GRT monitoring of ART in LMIC. Our NLFG assay amplified efficiently the HIV-1 genes with known or potential relevance for drug resistance and is thereby an attractive alternative for surveillance studies in LMIC.
- CCR5-tropic HIV-1C dominates monophylogenetically the epidemic in all regions of Ethiopia despite many years since its introduction, even among patients with very advanced disease, though a slight temporal increasing trend of CXCR4-tropic and dual-tropic strains might have been occurred since the epidemic started.

8 ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge **Karolinska Institute**, which provided me all the necessary knowledge and skill while doing my PhD study.

I am highly grateful to my main supervisor, **Prof. Anders Sönnerborg**, who hugely impacted my academic path right from the time he accepted me to be part of his big research group. Ever since I started working with him, he devoted his precious time and energy in fixing outstanding issues through recurrent discussions, which made my way possible during the entire PhD journey. I would not have completed the PhD without his tireless follow up of my activities with prompt communications and huge supports. In parallel to the scholarly guidance, he made me comprehend the power of critical thinking both in scientific and nonscientific issues to be an independent and a productive researcher. I don't have words to him except thank you.

I am sincerely thankful to my co-supervisor, **Assoc. Prof. Gaetano Marrone**, for his huge contribution in biostatistics. In addition, he taught me how to independently organise research data and perform relevant statistical analyses.

I am highly grateful to my mentor, **Prof. Lars Lindquist**, who has been positively and wisely involved in solving the outstanding issues happened in between. Without his wise engagement and advice this journey wouldn't have reached to this point.

I am happy to have this opportunity in order to express my gratefulness to my PhD fellow and a friend, **Amare Worku**, who have been working and sharing with me the day-to-day research activities. Together we have been walking all the PhD way and passing all the challenges. We have been sharing many great days with unforgettable memories over the past years.

I acknowledge my **coauthors** in all scientific papers for their scholarly inputs.

I am grateful to **Gudran Rafi**, **Marita Wallenberg Lundgren**, and **Arja Kramsu** for their unreserved administrative help during my study period.

I am also thankful to all members of the **Division of Clinical Microbiology, Laboratory Medicine, Karolinska Institutet**.

I would like also to take this opportunity to extend my gratefulness to all my former colleagues in **Sönnerborg group**: **Drs Jenny Svärd**, **Babilonia Barqasho**, **Halime Ekici**, **Jessica Nyström**, **Amanda Häggblom**, **Piotr Nowak**, **Irene Bontell**, **Eva Agneskog**, **Kajsa Noyan**, and the new members as well for their companion during my laboratory activities, scientific group meetings, and social activities. Thank you **Lydia Scharf** for being a nice long term office mate.

I am grateful to all **participating institutions in the ACM project** (advanced clinical monitoring of antiretroviral therapy in Ethiopia) which offered me the chance to be affiliated

with the project through their representatives in the ACM steering committee. I also grateful to **all study participants** who have given their voluntarily consents through the ACM project.

I am also extremely grateful to all my friend and colleagues at the **Department of Microbiology, Immunology, and Parasitology**, School of Medicine, Addis Ababa University for their kind support throughout my study period.

I am very much grateful to **Dr. Samir Abdurahman** for his companionable approach and brotherly good care since my arrival at Arlanda Airport. Ever since I met him, I felt at easy discussing any scientific and nonscientific issue. He and his beloved family (**Joddy** and **Nabbe**) have been helpful and their contribution has been critical for my family life in Sweden. I don't have words to them too except thank you for everything you have done to me.

I am sincerely thankful to **Dr. Björn Eriksson** and his beloved wife **Viviann** for inviting me and my family a couple of times to Sävsjo, where we have got memories of the Swedish culture in the countryside with kind hospitality. Ever since I met **Björn** in Ethiopia, I learned a lot from his huge life expertise.

I am deeply indebted to my great and unconditional friends **Drs Abiy Habtewold** and **Solomon Mequanente**, **Adane Mihret**, **Tamrat Abebe**, and **Tadesse Kebede**. Wherever they go and apart physically, they are near to me, standing always by my side. Their brotherly encouragements, supports, and companions have made me stronger and boosted up my confidence, which helped me a lot to pass through recurrent challenges I faced throughout the long journey.

Even though there were lots of ups and downs, all my beloved family members have been always by my side and made me felt at ease bearing the burdens. My family has been my walking stick and driving force while I was being stuck with outstanding issues at certain stages. I am very much and deeply indebted to my wife **Emebet Negash**, my children **Musie**, **Aron** and **Michael**, my mother **Zewude Abuhay**, my father **Fikrie Telele**, my mother-in-law **Zimam Belay**, my brothers **Birhanu Fikrie**, **Erashachin Fikrie**, **Temesgen Birhanu**, my sisters **Fetlework Tezera**, **Selamawit Yohannes**, **Hanna Birhanu**, and **Ferewoyini Negash** for their unconditional loves, unreserved encouragements, moral supports and prayers thorough the long journey.

Finally, I do not know how to put into words my wholeheartedly thankfulness to God the Almighty for giving me all the opportunities and strengths of making this all possible. You have been infinitively good to me and your infinitive love, priceless care and true forgiveness helped me to daily progress and made me pass through all the difficulties to accomplish the long journey; THANK YOU **GOD THE ALMIGHTY** and **Holy Virgin MARY**, **Mother of JESUS**.

9 REFERENCES

1. Gottlieb, M.S., et al., *Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency.* N Engl J Med, 1981. **305**(24): p. 1425-31.
2. HAPCO, *Ethiopian Strategic Plan for Intensifying Multi-Sectoral HIV/AIDS Response (2004 - 2008). National HIV/AIDS Prevention and Control Office, Federal Ministry of Health. Addis Ababa, Ethiopia December 2004.* 2004.
3. Ayehunie, S., et al., *New subtype of HIV-1 in Ethiopia.* Lancet, 1990. **336**(8720): p. 942.
4. Thomson, M.M. and A. Fernandez-Garcia, *Phylogenetic structure in African HIV-1 subtype C revealed by selective sequential pruning.* Virology, 2011. **415**(1): p. 30-8.
5. HAPCO, *Guidelines for Management of Opportunistic Infections and Anti-retroviral Treatment in Adolescents and Adults in Ethiopia, March 2008. Available at: http://apps.who.int/medicinedocs/en/d/Js19146en/, accessed 31 May 2016.* 2008.
6. HAPCO, *Country progress report on the HIV response. Federal HIV/AIDS Prevention and Control Office (FHAPCO), March 2014.* 2014.
7. UNAIDS, *How AIDS changed everything. MDG 6: 15 years, 15 lessons of hope from the AIDS response.* 2015.
8. EPHI, *HIV Related Estimates and Projections for Ethiopia–2017. Ethiopian Public Health Institute. March 2017, Addis Ababa.* 2017.
9. Lamptey, P., J. Johnson, and M. Khan, *"The Global Challenge of HIV and AIDS," Population Bulletin 61, no. 1 (Washington, DC: Population Reference Bureau, 2006).* 2006.
10. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.
11. Gallo, R.C., et al., *Frequent Detection and Isolation of Cytopathic Retroviruses (Htlv-Iii) from Patients with Aids and at Risk for Aids.* Science, 1984. **224**(4648): p. 500-503.
12. Case, K., *Nomenclature - Human-Immunodeficiency-Virus.* Annals of Internal Medicine, 1986. **105**(1): p. 133-133.
13. UNAIDS, *Fact sheet - Latest statistics on the status of the AIDS epidemic.* 2017.
14. Watanabe, M.E., *Origins of HIV: The interrelationship between nonhuman primates and the virus.* Bioscience, 2004. **54**(9): p. 810-814.
15. Rambaut, A., et al., *The causes and consequences of HIV evolution.* Nat Rev Genet, 2004. **5**(1): p. 52-61.
16. Sharp, P.M. and B.H. Hahn, *Origins of HIV and the AIDS Pandemic.* Cold Spring Harbor Perspectives in Medicine, 2011. **1**(1).
17. Hirsch, V.M., et al., *An African Primate Lentivirus (Sivsm) Closely Related to Hiv-2.* Nature, 1989. **339**(6223): p. 389-392.
18. de Silva, T.I., M. Cotten, and S.L. Rowland-Jones, *HIV-2: the forgotten AIDS virus.* Trends in Microbiology, 2008. **16**(12): p. 588-595.

19. Zheng, N.N., et al., *Comparison of human immunodeficiency virus (HIV)-specific T-cell responses in HIV-1- and HIV-2-infected individuals in Senegal.* Journal of Virology, 2004. **78**(24): p. 13934-13942.

20. Lihana, R.W., et al., *Update on HIV-1 Diversity in Africa: A Decade in Review.* Aids Reviews, 2012. **14**(2): p. 83-100.

21. Robertson, D.L., et al., *HIV-1 nomenclature proposal.* Science, 2000. **288**(5463): p. 55-6.

22. Santoro, M.M. and C.F. Perno, *HIV-1 Genetic Variability and Clinical Implications.* ISRN Microbiol, 2013. **2013**: p. 481314.

23. Hemelaar, J., et al., *Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004.* AIDS, 2006. **20**(16): p. W13-23.

24. Hemelaar, J., et al., *Global trends in molecular epidemiology of HIV-1 during 2000-2007.* AIDS, 2011. **25**(5): p. 679-89.

25. Tebit, D.M., et al., *HIV-1 group O genotypes and phenotypes: relationship to fitness and susceptibility to antiretroviral drugs.* AIDS Res Hum Retroviruses, 2016.

26. Lau, K.A., B. Wang, and N.K. Saksena, *Emerging trends of HIV epidemiology in Asia.* Aids Reviews, 2007. **9**(4): p. 218-229.

27. Arien, K.K., G. Vanham, and E.J. Arts, *Is HIV-1 evolving to a less virulent form in humans?* Nature Reviews Microbiology, 2007. **5**(2): p. 141-151.

28. John-Stewart, G.C., et al., *Subtype C is associated with increased vaginal shedding of HIV-1.* Journal of Infectious Diseases, 2005. **192**(3): p. 492-496.

29. Campbell, T.B., *Are all HIV type 1 strains created equal?* Clinical Infectious Diseases, 2006. **42**(6): p. 853-854.

30. Hemelaar, J., *The origin and diversity of the HIV-1 pandemic.* Trends Mol Med, 2012. **18**(3): p. 182-92.

31. Veras, N.M., et al., *High-resolution phylogenetics and phylogeography of human immunodeficiency virus type 1 subtype C epidemic in South America.* J Gen Virol, 2011. **92**(Pt 7): p. 1698-709.

32. Kalu, A.W., et al., *Monophylogenetic HIV-1C epidemic in Ethiopia is dominated by CCR5-tropic viruses-an analysis of a prospective country-wide cohort.* BMC Infect Dis, 2017. **17**(1): p. 37.

33. Kloos, H., D. Haile Mariam, and B. Lindtjorn, *The AIDS Epidemic in a Low-Income Country: Ethiopia.* Human Ecology Review, 2007. **Vol. 14**(No. 1, 2007): p. 39-55.

34. WHO, *Ethiopia | HIV/AIDS. World Health Organization Regional Office for Africa. WHO Country Office-Ethiopia.* 2015.

35. Federal Minstry of Health, *Ethiopia National Guidelines for Comprehensive HIV Prevention, Care, and Treatment. Available at: https://aidsfree.usaid.gov/resources/ethiopia-national-guidelines-comprehensive-hiv-prevention-care-and-treatment-0, accessed 05 June 2018. .* 2017.

36. Abebe, A., et al., *HIV type 1 subtype C in Addis Ababa, Ethiopia.* AIDS Res Hum Retroviruses, 1997. **13**(12): p. 1071-5.

37. Pollakis, G., et al., *Recombination of HIV type 1C (C '/C '') in Ethiopia: Possible link of EthHIV-1C ' to subtype C sequences from the high-prevalence epidemics in India and southern Africa.* Aids Research and Human Retroviruses, 2003. **19**(11): p. 999-1008.

38. Tsega, E., et al., *Serological survey of human immunodeficiency virus infection in Ethiopia.* Ethiop Med J, 1988. **26**(4): p. 179-84.

39. Lester, F.T., S. Ayehunie, and D. Zewdie, *Acquired immunodeficiency syndrome: seven cases in an Addis Ababa hospital.* Ethiop Med J, 1988. **26**(3): p. 139-45.

40. Kloos, H., *HIV/AIDS in Ethiopia: The Epidemic and Social, Economic, and Demographic Impacts (2001). International Conference on African Development Archives (available at: http://scholarworks.wmich.edu/africancenter_icad_archive/25, accessed 10 June 2017).* 2001.

41. EDHS, *Central Statistical Agency [Ethiopia] and ICF International. 2012. Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.* 2011.

42. UNAIDS, *The Gap Report.* 2016.

43. CSA, *Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.* Central Statistical Agency (Ethiopia), 2012.

44. HAPCO, *Country progress report on HIV/AIDS response, Federal Democratic Republic of Ethiopia, Federal HIV/AIDS Prevention and Control Office, Addis Ababa, Ethiopia, April 2012.* Federal HAPCO, 2012.

45. Shiferaw, Y., et al., *Assessment of knowledge, attitude and risk behaviors towards HIV/AIDS and other sexual transmitted infection among preparatory students of Gondar town, north west Ethiopia.* BMC Res Notes, 2011. **4**: p. 505.

46. Kwan, C.K. and J.D. Ernst, *HIV and tuberculosis: a deadly human syndemic.* Clin Microbiol Rev, 2011. **24**(2): p. 351-76.

47. World Health Organization., *Global tuberculosis report 2015*. 20th ed ed. 2015, Geneva: World Health Organization. 192 p.

48. Wondimeneh, Y., D. Muluye, and Y. Belyhun, *Prevalence of pulmonary tuberculosis and immunological profile of HIV co-infected patients in Northwest Ethiopia.* BMC Res Notes, 2012. **5**: p. 331.

49. Beyene, M.B. and H.B. Beyene, *Predictors of Late HIV Diagnosis among Adult People Living with HIV/AIDS Who Undertake an Initial CD4 T Cell Evaluation, Northern Ethiopia: A Case-Control Study.* PLoS One, 2015. **10**(10): p. e0140004.

50. Amogne, W., et al., *Efficacy and Safety of Antiretroviral Therapy Initiated One Week after Tuberculosis Therapy in Patients with CD4 Counts < 200 Cells/muL: TB-HAART Study, a Randomized Clinical Trial.* PLoS One, 2015. **10**(5): p. e0122587.

51. Chiu, I.M., et al., *Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses.* Nature, 1985. **317**(6035): p. 366-8.

52. Cowley, S., *The biology of HIV infection.* Leprosy Review, 2001. **72**(2): p. 212-220.

53. Swanson, C.M. and M.H. Malim, *SnapShot: HIV-1 proteins.* Cell, 2008. **133**(4): p. 742, 742 e1.

54. Freed, E.O., *HIV-1 gag proteins: diverse functions in the virus life cycle.* Virology, 1998. **251**(1): p. 1-15.

55. Turner, B.G. and M.F. Summers, *Structural biology of HIV.* J Mol Biol, 1999. **285**(1): p. 1-32.

56. Sierra, S., B. Kupfer, and R. Kaiser, *Basics of the virology of HIV-1 and its replication.* J Clin Virol, 2005. **34**(4): p. 233-44.

57. Clapham, P.R. and A. McKnight, *Cell surface receptors, virus entry and tropism of primate lentiviruses.* J Gen Virol, 2002. **83**(Pt 8): p. 1809-29.

58. Jamison, D.T., *Investing in Health,* in *Disease Control Priorities in Developing Countries*, D.T. Jamison, et al., Editors. 2006: Washington (DC).

59. Berger, E.A., et al., *A new classification for HIV-1.* Nature, 1998. **391**(6664): p. 240-240.

60. Huang, W., et al., *Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations.* J Virol, 2007. **81**(15): p. 7885-93.

61. Malkevich, N., et al., *Human immunodeficiency virus type 1 (HIV-1) non-B subtypes are similar to HIV-1 subtype B in that coreceptor specificity is a determinant of cytopathicity in human lymphoid tissue infected ex vivo.* Journal of Virology, 2001. **75**(21): p. 10520-10522.

62. Michael, N.L., et al., *The role of viral phenotype and CCR-5 gene defects in HIV-1 transmission and disease progression.* Nature Medicine, 1997. **3**(3): p. 338-340.

63. Berger, E.A., P.M. Murphy, and J.M. Farber, *Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease.* Annual Review of Immunology, 1999. **17**: p. 657-700.

64. Connor, R.I., et al., *Change in coreceptor use correlates with disease progression in HIV-1-infected individuals.* Journal of Experimental Medicine, 1997. **185**(4): p. 621-628.

65. Zhang, L., et al., *HIV-1 subtype and second-receptor use.* Nature, 1996. **383**(6603): p. 768.

66. Abebe, A., et al., *HIV-1 subtype C syncytium- and non-syncytium-inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS.* Aids, 1999. **13**(11): p. 1305-1311.

67. Brumme, Z.L., et al., *Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals.* Journal of Infectious Diseases, 2005. **192**(3): p. 466-474.

68. Moore, J.P., et al., *The CCR5 and CXCR4 coreceptors - Central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection.* Aids Research and Human Retroviruses, 2004. **20**(1): p. 111-126.

69. Schuitemaker, H., et al., *Biological Phenotype of Human-Immunodeficiency-Virus Type-1 Clones at Different Stages of Infection - Progression of Disease Is Associated with a Shift from Monocytotropic to T-Cell-Tropic Virus Populations.* Journal of Virology, 1992. **66**(3): p. 1354-1360.

70. Hartley, O., et al., *V3: HIV's switch-hitter.* Aids Research and Human Retroviruses, 2005. **21**(2): p. 171-189.

71. Hoffman, N.G., et al., *Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop.* Journal of Virology, 2002. **76**(8): p. 3852-3864.

72. Iyidogan, P. and K.S. Anderson, *Current perspectives on HIV-1 antiretroviral drug resistance.* Viruses, 2014. **6**(10): p. 4095-139.

73. Tseng, A., J. Seet, and E.J. Phillips, *The evolution of three decades of antiretroviral therapy: challenges, triumphs and the promise of the future.* British Journal of Clinical Pharmacology, 2015. **79**(2): p. 182-194.

74. Zhang, X.Q., *Anti-retroviral drugs: current state and development in the next decade.* Acta Pharmaceutica Sinica B, 2018. **8**(2): p. 131-136.

75. Arts, E.J. and D.J. Hazuda, *HIV-1 Antiretroviral Drug Therapy.* Cold Spring Harb Perspect Med, 2012. **2**(4): p. a007161.

76. UNAIDS, *UNAIDS DATA 2017. Joint United Nations Programme on HIV/AIDS (UNAIDS). Available at: http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf, accessed 05 May 2018.* 2017.

77. WHO, *HIV drug resistance report 2017 (available at http://www.who.int/hiv/pub/drugresistance/hivdr-report-2017/en/, accessed 25 May 2018).* 2017.

78. Shafer, R.W., *Genotypic testing for human immunodeficiency virus type 1 drug resistance.* Clin Microbiol Rev, 2002. **15**(2): p. 247-77.

79. Mansky, L.M., *Retrovirus mutation rates and their role in genetic variation.* Journal of General Virology, 1998. **79**: p. 1337-1345.

80. Coffin, J.M., *HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy.* Science, 1995. **267**(5197): p. 483-9.

81. Mackie, N., *Resistance to non-nucleoside reverse transcriptase inhibitors*, in *Antiretroviral Resistance in Clinical Practice*, A.M. Geretti, Editor. 2006: London.

82. De Luca, A. and M. Zazzi, *Interplay Between Transmitted and Acquired HIV Type 1 Drug Resistance: Reasons for a Disconnect.* Journal of Infectious Diseases, 2015. **212**(1): p. 5-7.

83. Boerma, R.S., et al., *Alarming increase in pretreatment HIV drug resistance in children living in sub-Saharan Africa: a systematic review and meta-analysis.* Journal of Antimicrobial Chemotherapy, 2017. **72**(2): p. 365-371.

84. WHO, *Guidelines on the public health response to pre-treatment HIV drug resistance. Geneva: World health Organization; 2017 (available at http://who.int/hiv/pub/guidelines/hivdr-guidelines-2017/, accessed 24 May 2018).* 2017.

85. Wainberg, M.A., G.J. Zaharatos, and B.G. Brenner, *Development of antiretroviral drug resistance.* N Engl J Med, 2011. **365**(7): p. 637-46.

86. Schmidt, D., et al., *Estimating trends in the proportion of transmitted and acquired HIV drug resistance in a long term observational cohort in Germany.* PLoS One, 2014. **9**(8): p. e104474.

87. Pingen, M., et al., *Persistence of frequently transmitted drug-resistant HIV-1 variants can be explained by high viral replication capacity.* Retrovirology, 2014. **11**.

88. Wittkop, L., et al., *Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study.* Lancet Infect Dis, 2011. **11**(5): p. 363-71.

89. Karlsson, A., et al., *Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010.* PLoS One, 2012. **7**(3): p. e33484.

90. Bontell, I., et al., *Trends in antiretroviral therapy and prevalence of HIV drug resistance mutations in Sweden 1997-2011.* PLoS One, 2013. **8**(3): p. e59337.

91. Hamers, R.L., et al., *Emerging HIV-1 drug resistance after roll-out of antiretroviral therapy in sub-Saharan Africa.* Curr Opin HIV AIDS, 2013. **8**(1): p. 19-26.

92. TenoRes Study, G., *Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection: a multicentre retrospective cohort study.* Lancet Infect Dis, 2016.

93. Hamers, R.L., et al., *Effect of pretreatment HIV-1 drug resistance on immunological, virological, and drug-resistance outcomes of first-line antiretroviral treatment in sub-Saharan Africa: a multicentre cohort study.* Lancet Infect Dis, 2012. **12**(4): p. 307-17.

94. THE LANCET INFECTIOUS, D., *Investing wisely in HIV/AIDS.* Lancet Infect Dis, 2012. **12**(1): p. 1.

95. El-Khatib, Z., et al., *Drug resistance patterns and virus re-suppression among HIV-1 subtype C infected patients receiving non-nucleoside reverse transcriptase inhibitors in South Africa.* J AIDS Clin Res, 2011. **2**(117).

96. Telele, N.F., et al., *Pretreatment drug resistance in a large countrywide Ethiopian HIV-1C cohort: a comparison of Sanger and high-throughput sequencing.* Sci Rep, 2018. **8**(1): p. 7556.

97. WHO, *WHO HIV Drug Resistance Report 2012.* 2012.

98. Hamers, R.L., et al., *Patterns of HIV-1 drug resistance after first-line antiretroviral therapy (ART) failure in 6 sub-Saharan African countries: implications for second-line ART strategies.* Clin Infect Dis, 2012. **54**(11): p. 1660-9.

99. Sunpath, H., et al., *High rate of K65R for antiretroviral therapy-naive patients with subtype C HIV infection failing a tenofovir-containing first-line regimen.* AIDS, 2012. **26**(13): p. 1679-84.

100. Invernizzi, C.F., et al., *Signature Nucleotide Polymorphisms at Positions 64 and 65 in Reverse Transcriptase Favor the Selection of the K65R Resistance Mutation in HIV-1 Subtype C.* Journal of Infectious Diseases, 2009. **200**(8): p. 1202-1206.

101. Hawkins, C.A., et al., *Clinical and genotypic findings in HIV-infected patients with the K65R mutation failing first-line antiretroviral therapy in Nigeria.* J Acquir Immune Defic Syndr, 2009. **52**(2): p. 228-34.

102. Mehellou, Y. and E. De Clercq, *Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go?* J Med Chem, 2010. **53**(2): p. 521-38.

103. Menendez-Arias, L., *Mechanisms of resistance to nucleoside analogue inhibitors of HIV-1 reverse transcriptase.* Virus Res, 2008. **134**(1-2): p. 124-46.

104. Sarafianos, S.G., et al., *Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition.* J Mol Biol, 2009. **385**(3): p. 693-713.

105. Vijayan, R.S., E. Arnold, and K. Das, *Molecular dynamics study of HIV-1 RT-DNA-nevirapine complexes explains NNRTI inhibition and resistance by connection mutations.* Proteins, 2014. **82**(5): p. 815-29.

106. Marcelin, A.G., *Resistance to nucleoside reverse transcriptase inhibitors*, in *Antiretroviral Resistance in Clinical Practice*, A.M. Geretti, Editor. 2006: London.

107. Clavel, F. and A.J. Hance, *HIV drug resistance.* N Engl J Med, 2004. **350**(10): p. 1023-35.

108. Lu, X.F. and Z.W. Chen, *The development of anti-HIV-1 drugs.* Yao Xue Xue Bao, 2010. **45**(2): p. 165-76.

109. Soriano, V. and C. de Mendoza, *Genetic mechanisms of resistance to NRTI and NNRTI.* HIV Clin Trials, 2002. **3**(3): p. 237-48.

110. Usach, I., V. Melis, and J.E. Peris, *Non-nucleoside reverse transcriptase inhibitors: a review on pharmacokinetics, pharmacodynamics, safety and tolerability.* J Int AIDS Soc, 2013. **16**: p. 1-14.

111. Schauer, G., S. Leuba, and N. Sluis-Cremer, *Biophysical Insights into the Inhibitory Mechanism of Non-Nucleoside HIV-1 Reverse Transcriptase Inhibitors.* Biomolecules, 2013. **3**(4): p. 889-904.

112. Seminari, E., A. Castagna, and A. Lazzarin, *Etravirine for the treatment of HIV infection.* Expert Rev Anti Infect Ther, 2008. **6**(4): p. 427-33.

113. WHO, *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. Geneva: World Health Organization; 2013 (available at: http://www.who.int/hiv/pub/arv/arv-2016/en/, accessed 20 May 2017).* 2013.

114. Metifiot, M., C. Marchand, and Y. Pommier, *HIV integrase inhibitors: 20-year landmark and challenges.* Adv Pharmacol, 2013. **67**: p. 75-105.

115. Blanco, J.L., et al., *HIV-1 integrase inhibitor resistance and its clinical implications.* J Infect Dis, 2011. **203**(9): p. 1204-14.

116. Lobritz, M.A., A.N. Ratcliff, and E.J. Arts, *HIV-1 Entry, Inhibitors, and Resistance.* Viruses, 2010. **2**(5): p. 1069-105.

117. De Feo, C.J. and C.D. Weiss, *Escape from Human Immunodeficiency Virus Type 1 (HIV-1) Entry Inhibitors.* Viruses-Basel, 2012. **4**(12): p. 3859-3911.

118. Covens, K., et al., *The rare HIV-1 gp41 mutations 43T and 50V elevate enfuvirtide resistance levels of common enfuvirtide resistance mutations that did not impact susceptibility to sifuvirtide.* Antiviral Research, 2010. **86**(3): p. 253-260.

119. Moore, J.P. and D.R. Kuritzkes, *A piece de resistance: how HIV-1 escapes small molecule CCR5 inhibitors.* Curr Opin HIV AIDS, 2009. **4**(2): p. 118-24.

120. Gilks, C., M. Vitoria, and World Health Organization. Dept. of HIV/AIDS., *Antiretroviral therapy for HIV infection in adults and adolescents: recommendations*

for a public health approach. 2006 rev. ed. 2006, Geneva: World Health Organization. 128 p.

121. WHO, *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection Recommendations for a public health approach - Second edition, 2016. Available at: http://www.who.int/hiv/pub/arv/arv-2016/en/, accessed 10 une 2018.* 2016.

122. UNAIDS, *90-90-90 An ambitious treatment target to help end the AIDS epidemic. Available at: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf, accessed 20 May 2018.* 2014.

123. Federal Minstry of Health, *Ethiopia national guidelines for comprehensive HIV prevention, care, and treatment. Availabel at: https://aidsfree.usaid.gov/sites/default/files/ethiopia_natl_gl_2014.pdf, accessed on 05 June 2018.* 2014.

124. Martin, A.R. and R.F. Siliciano, *Progress Toward HIV Eradication: Case Reports, Current Efforts, and the Challenges Associated with Cure.* Annu Rev Med, 2016. **67**: p. 215-28.

125. DHHS, *Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents Living with HIV. Laboratory Testing; Plasma HIV-1 RNA (Viral Load) and CD4 Count Monitoring*

Last Updated: May 1, 2014; Last Reviewed: May 1, 2014. U.S. Department of Health & Human Services. Available at: https://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-arv/458/plasma-hiv-1-rna--viral-load--and-cd4-count-monitoring, accessed 10 May 2018. 2018.

126. Loutfy, M.R., et al., *Systematic Review of HIV Transmission between Heterosexual Serodiscordant Couples where the HIV-Positive Partner Is Fully Suppressed on Antiretroviral Therapy.* Plos One, 2013. **8**(2).

127. Havlir, D.V., et al., *Prevalence and predictive value of intermittent viremia with combination HIV therapy.* Jama-Journal of the American Medical Association, 2001. **286**(2): p. 171-179.

128. Keiser, O., et al., *Outcomes of antiretroviral treatment in programmes with and without routine viral load monitoring in Southern Africa.* AIDS, 2011. **25**(14): p. 1761-9.

129. Rutherford, G.W., et al., *Predicting treatment failure in adults and children on antiretroviral therapy: a systematic review of the performance characteristics of the 2010 WHO immunologic and clinical criteria for virologic failure.* AIDS, 2014. **28 Suppl 2**: p. S161-9.

130. Chene, G., et al., *Intention-to-treat vs. on-treatment analyses of clinical trial data: experience from a study of pyrimethamine in the primary prophylaxis of toxoplasmosis in HIV-infected patients. ANRS 005/ACTG 154 Trial Group.* Control Clin Trials, 1998. **19**(3): p. 233-48.

131. Gupta, S.K., *Intention-to-treat concept: A review.* Perspect Clin Res, 2011. **2**(3): p. 109-12.

132. Hanna, G.J. and R.T. D'Aquila, *Clinical use of genotypic and phenotypic drug resistance testing to monitor antiretroviral chemotherapy.* Clinical Infectious Diseases, 2001. **32**(5): p. 774-782.

133. Gibson, R.M., C.L. Schmotzer, and M.E. Quinones-Mateu, *Next-Generation Sequencing to Help Monitor Patients Infected with HIV: Ready for Clinical Use?* Curr Infect Dis Rep, 2014. **16**(4): p. 401.

134. Ekici, H., et al., *Minority drug-resistant HIV-1 variants in treatment naive East-African and Caucasian patients detected by allele-specific real-time PCR.* PLoS One, 2014. **9**(10): p. e111042.

135. Gall, A., et al., *Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes.* J Clin Microbiol, 2012. **50**(12): p. 3838-44.

136. Grant, R.M., et al., *Accuracy of the TRUGENE HIV-1 genotyping kit.* J Clin Microbiol, 2003. **41**(4): p. 1586-93.

137. Vandamme, A.M., et al., *European Recommendations for the Clinical Use of HIV Drug Resistance Testing: 2011 Update.* Aids Reviews, 2011. **13**(2): p. 77-108.

138. Vandamme, A.M., et al., *Updated European recommendations for the clinical use of HIV drug resistance testing.* Antiviral Therapy, 2004. **9**(6): p. 829-848.

139. Lindstrom, A. and J. Albert, *A simple and sensitive 'in-house' method for determining genotypic drug resistance in HIV-1.* J Virol Methods, 2003. **107**(1): p. 45-51.

140. Wallis, C.L., et al., *Affordable in-house antiretroviral drug resistance assay with good performance in non-subtype B HIV-1.* J Virol Methods, 2010. **163**(2): p. 505-8.

141. Bertognolio, S., et al., *World Health Organization/HIVResNet drug resistance laboratory strategy.* Antiviral Therapy, 2008. **13**: p. 49-57.

142. Hauser, A., et al., *Detection and Quantification of Minor Human Immunodeficiency Virus Type 1 Variants Harboring K103N and Y181C Resistance Mutations in Subtype A and D Isolates by Allele-Specific Real-Time PCR.* Antimicrobial Agents and Chemotherapy, 2009. **53**(7): p. 2965-2973.

143. Guo, D.X., et al., *Allele-specific real-time PCR testing for minor HIV-1 drug resistance mutations: assay preparation and application to reveal dynamic of mutations in vivo.* Chin Med J (Engl), 2010. **123**(23): p. 3389-95.

144. Swenson, L.C., M. Daumer, and R. Paredes, *Next-generation sequencing to assess HIV tropism.* Curr Opin HIV AIDS, 2012. **7**(5): p. 478-85.

145. Nishizawa, M., et al., *Highly-sensitive allele-specific PCR testing identifies a greater prevalence of transmitted HIV drug resistance in Japan.* PLoS One, 2013. **8**(12): p. e83150.

146. Rowley, C.F., et al., *Improvement in allele-specific PCR assay with the use of polymorphism-specific primers for the analysis of minor variant drug resistance in HIV-1 subtype C.* J Virol Methods, 2008. **149**(1): p. 69-75.

147. Nishizawa, M., et al., *Longitudinal Detection and Persistence of Minority Drug-Resistant Populations and Their Effect on Salvage Therapy.* Plos One, 2015. **10**(9).

148. Zhang, G.Q., et al., *Simultaneous Detection of Major Drug Resistance Mutations in the Protease and Reverse Transcriptase Genes for HIV-1 Subtype C by Use of a Multiplex Allele-Specific Assay.* Journal of Clinical Microbiology, 2013. **51**(11): p. 3666-3674.

149. Ellis, G.M., et al., *Simultaneous and sensitive detection of human immunodeficiency virus type 1 (HIV) drug resistant genotypes by multiplex oligonucleotide ligation assay.* Journal of Virological Methods, 2013. **192**(1-2): p. 39-43.

150. Zhang, J. and H. Xing, *[Application of Next-generation Sequencing Techniques in the Dynamics of HIV-1 Quasispecies]*. Bing Du Xue Bao, 2015. **31**(5): p. 573-8.

151. Zhang, J., et al., *Utility of next-generation sequencing technologies for the efficient genetic resolution of haematological disorders.* Clin Genet, 2016. **89**(2): p. 163-72.

152. Metzker, M.L., *Sequencing technologies - the next generation.* Nat Rev Genet, 2010. **11**(1): p. 31-46.

153. Metzker, M.L., *Applications of Next-Generation Sequencing Sequencing Technologies - the Next Generation.* Nature Reviews Genetics, 2010. **11**(1): p. 31-46.

154. Giallonardo, F.D., et al., *Full-length haplotype reconstruction to infer the structure of heterogeneous virus populations.* Nucleic Acids Res, 2014. **42**(14): p. e115.

155. Henn, M.R., et al., *Whole Genome Deep Sequencing of HIV-1 Reveals the Impact of Early Minor Variants Upon Immune Recognition During Acute Infection.* Plos Pathogens, 2012. **8**(3).

156. Grossmann, S., P. Nowak, and U. Neogi, *Subtype-independent near full-length HIV-1 genome sequencing and assembly to be used in large molecular epidemiological studies and clinical management.* Journal of the International Aids Society, 2015. **18**.

157. Beerenwinkel, N., et al., *Computational methods for the design of effective therapies against drug resistant HIV strains.* Bioinformatics, 2005. **21**(21): p. 3943-50.

158. Zazzi, M., et al., *Comparative evaluation of three computerized algorithms for prediction of antiretroviral susceptibility from HIV type 1 genotype.* Journal of Antimicrobial Chemotherapy, 2004. **53**(2): p. 356-360.

159. Cordes, F., R. Kaiser, and J. Selbig, *Bioinformatics approach to predicting HIV drug resistance.* Expert Review of Molecular Diagnostics, 2006. **6**(2): p. 207-215.

160. Vercauteren, J. and A.M. Vandamme, *Algorithms for the interpretation of HIV-1 genotypic drug resistance information.* Antiviral Res, 2006. **71**(2-3): p. 335-42.

161. Yu, X., I.T. Weber, and R.W. Harrison, *Prediction of HIV drug resistance from genotype with encoded three-dimensional protein structure.* BMC Genomics, 2014. **15 Suppl 5**: p. S1.

162. Kuiken, C., B. Korber, and R.W. Shafer, *HIV sequence databases.* AIDS Rev, 2003. **5**(1): p. 52-61.

163. Garrido, C., et al., *Evaluation of eight different bioinformatics tools to predict viral tropism in different human immunodeficiency virus type 1 subtypes.* J Clin Microbiol, 2008. **46**(3): p. 887-91.

164. Kalu, A.W., et al., *Prediction of coreceptor usage by five bioinformatics tools in a large Ethiopian HIV-1 subtype C cohort.* PLoS One, 2017. **12**(8): p. e0182384.

165. MOH, *Guidelines for use of antiretroviral drugs in Ethiopia, Ministry of Health Disease Prevention and Control Department in collaboration with HIV/ AIDS Prevention and Control office (HAPCO) and Drug Administration and Control Authority (DACA), February 2003 Addis Ababa, Ethiopia.* 2003.

166. HAPCO, *Report on progress towards implementation of the UN Declaration of Commitment on HIV/AIDS, Federal Democratic Republic of Ethiopia, Federal HIV/AIDS Prevention and Control Office, Addis Ababa, Ethiopia, March 2010.* 2010.

167. Assefa, Y., et al., *Scaling up antiretroviral treatment and improving patient retention in care: lessons from Ethiopia, 2005-2013.* Global Health, 2014. **10**: p. 43.

168. Assefa, Y., et al., *Effectiveness and acceptability of delivery of antiretroviral treatment in health centres by health officers and nurses in Ethiopia.* J Health Serv Res Policy, 2012. **17**(1): p. 24-9.

169. Sigaloff, K.C., et al., *Unnecessary antiretroviral treatment switches and accumulation of HIV resistance mutations; two arguments for viral load monitoring in Africa.* J Acquir Immune Defic Syndr, 2011. **58**(1): p. 23-31.

170. Svard, J., et al., *Drug resistance testing through remote genotyping and predicted treatment options in human immunodeficiency virus type 1 infected Tanzanian subjects failing first or second line antiretroviral therapy.* Plos One, 2017. **12**(6).

171. Abdissa, A., et al., *Drug resistance in HIV patients with virological failure or slow virological response to antiretroviral therapy in Ethiopia.* BMC Infect Dis, 2014. **14**: p. 181.

172. Mulu, A., M. Maier, and U.G. Liebert, *Upward trends of acquired drug resistances in Ethiopian HIV-1C isolates: A decade longitudinal study.* PLoS One, 2017. **12**(10): p. e0186619.

173. Melaku, Z., et al., *Characteristics and outcomes of adult Ethiopian patients enrolled in HIV care and treatment: a multi-clinic observational study.* Bmc Public Health, 2015. **15**.

174. Assefa, Y., et al., *Outcomes of antiretroviral treatment program in Ethiopia: retention of patients in care is a major challenge and varies across health facilities.* BMC Health Serv Res, 2011. **11**: p. 81.

175. Abegaz, W.E., et al., *Threshold survey evaluating transmitted HIV drug resistance among public antenatal clinic clients in Addis Ababa, Ethiopia.* Antivir Ther, 2008. **13 Suppl 2**: p. 89-94.

176. Kassu, A., et al., *Molecular epidemiology of HIV type 1 in treatment-naive patients in north Ethiopia.* AIDS Res Hum Retroviruses, 2007. **23**(4): p. 564-8.

177. Mulu, A., et al., *Clade homogeneity and Pol gene polymorphisms in chronically HIV-1 infected antiretroviral treatment naive patients after the roll out of ART in Ethiopia.* BMC Infect Dis, 2014. **14**: p. 158.

178. Mulu, A., M. Maier, and U.G. Liebert, *Low Incidence of HIV-1C Acquired Drug Resistance 10 Years after Roll-Out of Antiretroviral Therapy in Ethiopia: A Prospective Cohort Study.* PLoS One, 2015. **10**(10): p. e0141318.

179. Bennett, D.E., et al., *The World Health Organization's global strategy for prevention and assessment of HIV drug resistance.* Antivir Ther, 2008. **13 Suppl 2**: p. 1-13.

180. Delwart, E.L., et al., *Genetic subtyping of human immunodeficiency virus using a heteroduplex mobility assay.* PCR Methods Appl, 1995. **4**(5): p. S202-16.

181. Aralaguppe, S.G., et al., *Multiplexed next-generation sequencing and de novo assembly to obtain near full-length HIV-1 genome from plasma virus.* J Virol Methods, 2016. **236**: p. 98-104.

182. Larsson, A., *AliView: a fast and lightweight alignment viewer and editor for large datasets.* Bioinformatics, 2014. **30**(22): p. 3276-8.

183. Gabadinho, A., et al., *Analyzing and visualizing state sequences in R with TraMineR.* Journal of Statistical Software, 2011. **40**(4): p. 1-37.

184. R Development Core Team, *R: A language and environment for statistical computing.* 2014, R Foundation for Statistical Computing: Vienna, Austria.

185. Wensing, A.M., et al., *2017 Update of the Drug Resistance Mutations in HIV-1.* Top Antivir Med, 2017. **24**(4): p. 132-133.

186. Lengauer, T., et al., *Bioinformatics prediction of HIV coreceptor usage.* Nat Biotechnol, 2007. **25**(12): p. 1407-10.

187. Sing, T., et al., *Predicting HIV coreceptor usage on the basis of genetic and clinical covariates.* Antivir Ther, 2007. **12**(7): p. 1097-106.

188. WHO, *WORLD HEALTH ORGANIZATION GLOBAL STRATEGY FOR THE SURVEILLANCE AND MONITORING OF HIV DRUG RESISTANCE, HIV/AIDS Programme.* 2012.

189. Petersen, M., et al., *Association of Implementation of a Universal Testing and Treatment Intervention With HIV Diagnosis, Receipt of Antiretroviral Therapy, and Viral Suppression in East Africa.* JAMA, 2017. **317**(21): p. 2196-2206.

190. Iwuji, C.C., et al., *Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial.* Lancet HIV, 2018. **5**(3): p. e116-e125.

191. Mberi, M.N., et al., *Determinants of loss to follow-up in patients on antiretroviral treatment, South Africa, 2004-2012: a cohort study.* BMC Health Serv Res, 2015. **15**: p. 259.

192. Gupta, A., et al., *Early mortality in adults initiating antiretroviral therapy (ART) in low- and middle-income countries (LMIC): a systematic review and meta-analysis.* PLoS One, 2011. **6**(12): p. e28691.

193. Kaleebu, P., et al., *Virological Response and Antiretroviral Drug Resistance Emerging during Antiretroviral Therapy at Three Treatment Centers in Uganda.* PLoS One, 2015. **10**(12): p. e0145536.

194. Haas, A.D., et al., *Retention and mortality on antiretroviral therapy in sub-Saharan Africa: collaborative analyses of HIV treatment programmes.* Journal of the International Aids Society, 2018. **21**.

195. EDHS, *Central Statistical Agency (CSA) [Ethiopia] and ICF. 2016. Ethiopia Demographic and Health Survey 2016. Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF.* . 2016.

196. Hamers, R.L., et al., *The status of HIV-1 resistance to antiretroviral drugs in sub-Saharan Africa.* Antivir Ther, 2008. **13**(5): p. 625-39.

197. Huruy, K., et al., *Immune restoration disease and changes in CD4+ T-cell count in HIV- infected patients during highly active antiretroviral therapy at Zewditu memorial hospital, Addis Ababa, Ethiopia.* AIDS Res Ther, 2010. **7**: p. 46.

198. Inzaule, S.C., et al., *The evolving landscape of HIV drug resistance diagnostics for expanding testing in resource-limited settings.* AIDS Rev, 2017. **19**(2).

199. Charpentier, C., et al., *High frequency of integrase Q148R minority variants in HIV-infected patients naive of integrase inhibitors.* AIDS, 2010. **24**(6): p. 867-73.

200. Liu, J., et al., *Analysis of low-frequency mutations associated with drug resistance to raltegravir before antiretroviral treatment.* Antimicrob Agents Chemother, 2011. **55**(3): p. 1114-9.

201. Abebe, A., et al., *Identification of a genetic subcluster of HIV type 1 subtype C (C') widespread in Ethiopia.* AIDS Res Hum Retroviruses, 2000. **16**(17): p. 1909-14.

202. Hussein, M., et al., *HIV-1 subtype C in commerical sex workers in Addis Ababa, Ethiopia.* J Acquir Immune Defic Syndr, 2000. **23**(2): p. 120-7.

203. Delatorre, E.O. and G. Bello, *Phylodynamics of HIV-1 subtype C epidemic in east Africa.* PLoS One, 2012. **7**(7): p. e41904.

204. Abar, A.E., et al., *[Human immunodeficiency virus type 1 subtypes in Djibouti]*. Arch Inst Pasteur Tunis, 2012. **89**(1-4): p. 33-7.

205. Khamadi, S.A., et al., *HIV type 1 genetic diversity in Moyale, Mandera, and Turkana based on env-C2-V3 sequences.* AIDS Res Hum Retroviruses, 2008. **24**(12): p. 1561-4.

206. Kantor, R., et al., *HIV diversity and drug resistance from plasma and non-plasma analytes in a large treatment programme in western Kenya.* J Int AIDS Soc, 2014. **17**: p. 19262.

207. Hierholzer, M., et al., *HIV type 1 strains from East and West Africa are intermixed in Sudan.* AIDS Res Hum Retroviruses, 2002. **18**(15): p. 1163-6.

208. Bjorndal, A., et al., *Phenotypic characteristics of human immunodeficiency virus type 1 subtype C isolates of Ethiopian AIDS patients.* AIDS Res Hum Retroviruses, 1999. **15**(7): p. 647-53.

209. Kalinkovich, A., et al., *Increased CCR5 and CXCR4 expression in Ethiopians living in Israel: environmental and constitutive factors.* Clin Immunol, 2001. **100**(1): p. 107-17.

210. Sollerkvist, L.P., et al., *Increased CXCR4 Use of HIV-1 Subtype C Identified by Population Sequencing in Patients Failing Antiretroviral Treatment Compared with Treatment-Naive Patients in Botswana.* Aids Research and Human Retroviruses, 2014. **30**(5): p. 436-445.

211. Connell, B.J., et al., *Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa.* Aids, 2008. **22**(7): p. 896-899.

212. Gupta, S., et al., *HIV-1 Coreceptor Tropism in India: Increasing Proportion of X4-Tropism in Subtype C Strains Over Two Decades.* Jaids-Journal of Acquired Immune Deficiency Syndromes, 2014. **65**(4): p. 397-404.

213. Brumme, Z.L., et al., *Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy.* AIDS, 2004. **18**(4): p. F1-9.

214. Gijsbers, E.F., et al., *The presence of CXCR4-using HIV-1 prior to start of antiretroviral therapy is an independent predictor of delayed viral suppression.* PLoS One, 2013. **8**(10): p. e76255.

215. Garforth, S.J., C. Lwatula, and V.R. Prasad, *The lysine 65 residue in HIV-1 reverse transcriptase function and in nucleoside analog drug resistance.* Viruses, 2014. **6**(10): p. 4080-94.

216. Ross, L.L., et al., *A rare HIV reverse transcriptase mutation, K65N, confers reduced susceptibility to tenofovir, lamivudine and didanosine.* Aids, 2006. **20**(5): p. 787-789.

217. Brenner, B.G., et al., *HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture.* AIDS, 2006. **20**(9): p. F9-13.

218. Coutsinos, D., et al., *A template-dependent dislocation mechanism potentiates K65R reverse transcriptase mutation development in subtype C variants of HIV-1.* PLoS One, 2011. **6**(5): p. e20208.

219. van Tienen, C., et al., *Letter to the editor: Pre-exposure prophylaxis for HIV in Europe: The need for resistance surveillance.* Euro Surveill, 2017. **22**(11).

220. Neogi, U., et al., *Ex-vivo antiretroviral potency of newer integrase strand transfer inhibitors cabotegravir and bictegravir in HIV type 1 non-B subtypes.* AIDS, 2018. **32**(4): p. 469-476.

221. Neogi, U., et al., *Factors influencing the efficacy of rilpivirine in HIV-1 subtype C in low- and middle-income countries.* J Antimicrob Chemother, 2016. **71**(2): p. 367-71.

222. Radestock, B., et al., *Comprehensive Mutational Analysis Reveals p6(Gag) Phosphorylation To Be Dispensable for HIV-1 Morphogenesis and Replication.* Journal of Virology, 2013. **87**(2): p. 724-734.

223. Neogi, U., et al., *Novel tetra-peptide insertion in Gag-p6 ALIX-binding motif in HIV-1 subtype C associated with protease inhibitor failure in Indian patients.* AIDS, 2014. **28**(15): p. 2319-22.

224. Sharma, S., et al., *The PTAP sequence duplication in HIV-1 subtype C Gag p6 in drug-naive subjects of India and South Africa.* BMC Infect Dis, 2017. **17**(1): p. 95.

225. Martins, A.N., et al., *Accumulation of P(T/S)AP late domain duplications in HIV type 1 subtypes B, C, and F derived from individuals failing ARV therapy and ARV drug-naive patients.* AIDS Res Hum Retroviruses, 2011. **27**(6): p. 687-92.

226. Wijting, I.E.A., et al., *HIV-1 resistance dynamics in patients failing dolutegravir maintenance monotherapy.* J Infect Dis, 2018.

227. Johansson, B., K. Sherefa, and A. Sonnerborg, *Multiple enhancer motifs in HIV type 1 strains from Ethiopia.* AIDS Res Hum Retroviruses, 1995. **11**(6): p. 761-4.