INVESTIGATING THE RISK FACTORS ASSOCIATED WITH LOW-LEVEL VIREMIA AND VIROLOGICAL FAILURE IN HIV-1 INFECTED CHILDREN UNDERGOING ANTIRETROVIRAL THERAPY

by

SHIVANI GUPTA

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Medical Microbiology & Infectious Diseases
University of Manitoba
Winnipeg, Manitoba

Abstract

The use of antiretroviral therapy (ART) for HIV-1 treatment effectively suppresses viral replication if managed appropriately, but sometimes patients experience incomplete viral suppression with manifestation of either persistent low-level viremia (LLV) or discernible virological failure (VF). In the present study, potential risk factors associated with LLV and VF were investigated in a cohort of HIV-1 infected Kenyan children receiving ART. Drug resistant (DR) variants in children with or without LLV were examined using a next-generation sequencing-based HIV DR typing protocol. The potential association between HIV DR mutations (DRMs) and LLV and/or VF was then examined. To measure the potential impacts from other clinical and epidemiological confounding factors, a database comprising of epidemiological and clinical information from this patient cohort was established and sanitized for ensured accountability. Statistically significant correlations between the examined factors and LLV or VF were determined using chi-square test, Kaplan-Meier survival analysis, and Cox proportional hazard models. Of 293 examined patients, 20% had LLV and 22% of the selected patients progressed to VF with no significant association observed between LLV and VF. ART adherence during therapy, baseline CD4 counts, DRMs at LLV, WHO clinical stage, gender, ART therapy stage ($1^{st}/2^{nd}$ line), ART drugs and co-morbidities were not significantly associated with LLV, whereas, the ART adherence, CD4 counts and co-infection with pneumonia was significantly associated with VF. This study highlights the factors predictive of VF, and the relevance of maintaining LLV in HIV-infected children.

Acknowledgements

This project would not have been possible without the help and guidance of many individuals. I value the skills and experience gained during my Master's, as this will go a long way. First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Gary Van Domselaar. Thank you for your understanding, guidance, and helpful advice throughout my project. I would also like to thank you for granting me the opportunity to work in your lab, providing me with facilities and/or resources to aid in my project and assisting me in surpassing obstacles to move forward with my project.

A special thanks to Dr. Binhua Liang for providing expertise to the success of this project. I am highly grateful to you for pushing me through the challenges that helped me to be competent in my field, and to Dr. Hezhao Ji for being a great mentor and for helping me enhance the lab skills necessary for the project. You have been very supportive and easy to approach every time I needed to discuss the project.

I am extremely thankful to my advisory committee members Dr. Blake Ball, Dr. Michael Domaratzki, and Dr. Elijah Songok for their valuable advice and encouragement over the course of my MSc. All of your comments and suggestions have been greatly appreciated.

Thanks to Dr. Michelle Liu for providing support to me in statistical analysis. Your knowledge and perspective really helped me rationale my analyses. I am also grateful to Joseph Kabogo and the members of KMRI and AVR Annex laboratory in Nairobi for being so welcoming and helpful during my stay there as well as for my project.

I would like to thank all the lab members at the Bioinformatics Core at Elgin and the HIV & Retrovirology lab at JC Wilt for camaraderie and insightfulness. I want to thank my fellow graduate students for being supportive and playing a pivotal role in keeping my head steady during times of disarray. I also have great appreciation for our collaborators at Nairobi for sharing their clinical samples and patient's data with us to help answer specific questions for this study. Thanks for MMSF and PHAC funding provided for this project.

The support staff of our Medical Microbiology department has also been very helpful with their assistance, especially Angela Nelson, Jude Zieske, Eva Lindsey, Sue Ramdahin and Sharon Tardi. Many Thanks to Dr. Adrienne Meyers for being informative about my trip to Nairobi as well as in student research seminars. I appreciate assistance from Dr. Matthew Gilmour for letting me stay in his Bacteriology Journal Club, even though I have a virology background.

I am indebted to my family, Dr. Rajesh Gupta, Poornima Gupta and Pratik Gupta, and close friends for being extremely motivating and continuously supporting me. I really appreciate everyone for celebrating my good days and for being tolerant of me during my bad days.

Dedication

To my parents; for their love and support throughout my educational endeavors. My dad helped me realize my interest in research and gave me the confidence in believing I can accomplish my graduate studies. My mom has always been there to lessen my stress. Thank you for teaching me patience and perseverance that was necessary to believe in myself and help me succeed in my career.

Table of Contents

List of Abbreviations

Abbreviations	Description
3TC	2'-deoxy-3'-thiacytidine (Lamivudine)
ABC	Abacavir
AIDS	Acquired Immune Deficiency Syndrome
ART	Antiretroviral Therapy
ASPCR	Allele-Specific Real-Time Polymerase Chain Reaction
AZT	Azidothymidine (Zidovudine)
CD4	Cluster of Differentiation 4
cDNA	Complementary DNA
CRF	Circulating Recombinant Form
d4T	2',3'-didehydro-2',3'-dideoxythymidine (Stavudine)
DC	Dendritic Cell
ddI	2',3'-dideoxyinosine (Didanosine)
DNA	Deoxyribose Nucleic Acid
DOA	Date Of Admission
DOB	Date Of Birth
DRM	Drug Resistant Mutations
EDTA	Ethylenediaminetetraacetic Acid
EFV	Efavirenz
Env	Envelope
EQAPOL	External Quality Assurance Program Oversight Laboratory
Gag	Group Specific Antigen
HIV	Human Immunodeficiency Virus
HyDRA	HIV Drug Resistant Analysis
IN	Integrase
KMRI	Kenya Medical Research Institute
LiPA	Line Probe Assay
LLV	Low Level Viremia
LPV	Lopinavir
LTP	Leo Toto Program
LTR	Long Terminal Repeat
$MgSO_4$	Magnesium Sulfate
mRNA	Messenger Ribonucleic Acid
MTC	Mother-to-Child Transmission
MVC	Maraviroc
N-PCR	Nested PCR
NDL	Nyumbani Diagnostics Laboratory
NGS	Next Generation Sequencing

NHP	Normal Human Plasma
NLHG	National Laboratory For HIV Genetics
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitors
NRTI	Nucleoside/Nucleotide Reverse Transcriptase Inhibitors
NVP	Nevirapine
OCR	Optical Character Recognition
OLA	Oligonucleotide Ligation Assay
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PI	Protease Inhibitors
pLLV	Persistent Low Level Viremia
Pol	Polymerase
PR	Protease
QC	Quality Control
RAL	Raltegravir
RNA	Ribonucleic Acid
RT	Reverse Transcriptase
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SGS	Single Genome Sequencing
SIV	Simian Immunodeficiency Virus
TAMs	Thymidine Analog Mutations
TB	Tuberculosis
TDF	Tenofovir Disoproxil Furmarate
TMA	Transcription Mediated Amplification
UDPS	Ultra-Deep Pyrosequencing
URF	Unusual Recombinant Form
UV	Ultraviolet Light
VF	Virologic Failure
VL	Viral Load
WHO	World Health Organization

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Origin & global impact of HIV-1

Human Immunodeficiency Virus (HIV) is a lentivirus that belongs to the *Retroviridae* family [1,2]. Lentiviruses are characterized by causing diseases with long incubation periods in a restricted host range. HIV virus infects CD4 positive cells, a key subset of human immune cells and impairs the host innate and adaptive immunity [3]. Without appropriate antiretroviral therapy (ART), HIV infection can progress to acquired immunodeficiency syndrome (AIDS) leaving the individual incapable of mounting effective immune responses against life-threatening opportunistic infections and cancers, which may eventually lead to death [4].

HIV can be classified into two separate groups known as HIV-1 and HIV-2. HIV-1 originated from a strain of simian immunodeficiency virus (SIV) found in chimpanzees while HIV-2 is more closely related to a strain of SIV from sooty mangabeys [5,6]. HIV was transmitted from non-human primates to human likely through contact with hunters and butchers infected bodily fluids or through the consumption of infected meat. HIV-2 is rare and less infectious than HIV-1 and is mainly found in a few countries in West Africa such as Mali, Mauritania, Nigeria and Sierra Leone [6].

HIV-1 is responsible for the majority of the global HIV epidemic among adults as well as children. In 2012, HIV/AIDS was the sixth leading cause of death worldwide and the leading cause of adult mortality in Africa [7]. By the end of 2015, there were an estimated 36.9 million people living with HIV/AIDS worldwide [8], including 2.6 million children under the age of 15 years [9]. On an average, 77% of all pregnant women living

with HIV receive ARTs to prevent transmission to their babies. In 2015, 17 million individuals were receiving ART globally, making it one of the largest increases of infected people accessing treatment over time [8]. Infants and children are at an exceptionally high risk of poor outcomes from infection as 52% of children die before two years of age in the absence of treatment [10]. Sub-Saharan Africa remains the most severely affected region, accounting for approximately 71% of all people living with HIV infection [8]. More than 90% of HIV-infected children in the world currently live in Sub-Saharan Africa [11]. Countries with a high prevalence of HIV in Sub-Saharan Africa have witnessed a decline in life expectancy at birth to just below 50 years of age [12]. Kenya is one of the countries most severely affected in Sub-Saharan Africa with 1.6 million Kenyan adults and children currently infected by HIV-1. Urban residents have a significantly higher risk of HIV infection (10%) than rural residents (6%). While new infections are stabilized at an average of 93,000 among adults and 11,000 among children in Kenya annually, females are overrepresented among all age groups except for the group of 45-49 [13]. Notably, HIV prevalence in young women is astonishingly higher than their male counterpart (i.e. 3.0% vs. 0.5% for the age group of 15-19 and 9.0% vs. 2.4% for the group of 20-24). This highlights young women as the main target population of HIV intervention strategies.

HIV-1 consists of at least nine genetically distinct subtypes known as A, B, C, D, F, G, H, J and K and accounts for over 90% of infection in humans worldwide. HIV-1 subtype B is most common in the Americas, Western Europe, Japan and Australia. The majority of HIV clinical research has been conducted in populations where subtype B predominates, although it represents only 12% of global HIV infections [14]. In contrast,

less research has been carried out on HIV-1 subtype C which is prevalent primarily in Sub-Saharan Africa and India accounting for ~50% of HIV infections across the world. Other commonly found subtypes in different regions of Africa are subtypes A (sub-subtype A1 being most common) and D infecting 38% and 5% of the population, respectively [15]. While subtypes G, H, J and K are also spread throughout Africa and Asia, these are not encountered as frequently as subtypes A, B, C or D. Further, novel strains can form through hybridization of different subtypes co-infecting the same individual and are preliminarily defined as unique recombinant forms (URFs). Co-infection with different but related HIV-1 strains gives rise to a novel strain that usually does not survive long, but those that are transmissible are known as circulating recombinant forms (CRFs). CRFs can vary from strong to negligible epidemic relevance, being more prominent in Asia and Western Africa [14,15]. It is noteworthy that the geographic distribution of varied HIV subtypes has been consistently changing over time due to international travel, immigration and migration.

1.2 HIV-1 transmission & pathogenesis

HIV/AIDS is a communicable disease transmitted predominantly by contact with infected bodily fluids [16]. It can also be transmitted vertically from mother to child during pregnancy, delivery and/or through breastfeeding [17]. The probability of HIV transmission is dependent on both host and viral factors, such as the stage of disease, the infectious dose of the infected bodily fluids, plasma viral load, contraceptive methods, lack of male circumcision, prolonged rupture of membranes or susceptibility of the uninfected partner [18,19] .

There is no clear predetermined linkage between a specific subtype and a unique mode of transmission. Therefore, different subtypes could have been influenced by a combination of different genetic, demographic, economic and social factors that separate the different risk groups for HIV-1. HIV is not an airborne virus and cannot be transmitted through casual contact, kissing, urine or even insect vectors [17]. The transmission is only possible if bodily fluids come in contact with a mucous membrane, damaged tissue or injected directly into the bloodstream [20]. Unprotected heterosexual intercourse is the main transmission route causing over 90% of HIV infections in adults [16]. In children, vertical transmission is the most common route for contracting the virus. Mother-to-child transmission (MTC) of HIV can occur from an infected mother to the fetus or an infant. The chance of MTC transmission via placenta during pregnancy or delivery is 15-25% with additional 10-32% risk from breastfeeding [8,17]. Replacement feeding is the most effective measure to decrease transmission through breast milk; however this is not always feasible in developing countries due to the associated cost [21]. Without any intervention, the MTC transmission rate varies between 15-45% per year [8]. Transmission routes via sharing of needles or syringe from an infected individual as well as contact with contaminated blood, blood products, and other infected body fluid exchange remains a concern mainly for hemophiliacs as well as intravenous drug users. However, safer practices in certain high-risk populations have shifted the focus of the epidemic [22].

It is estimated that 34 million children in Sub-Saharan Africa have lost one or both parents to AIDS [23]. Due to an unprecedented burden on social welfare services, these orphaned children are more likely to drop out of schools to earn for survival. This

places an enormous social burden on orphaned children. The expanding poverty and lack of preventive awareness increase the risk of these children to contract the virus themselves.

Additionally, large populations of HIV-infected individuals are either undiagnosed, remain untreated or enter treatment at a late stage of the disease. As of 2013, there are approximately 19 million HIV-infected but undiagnosed individuals worldwide, which represents a reservoir that can promote HIV transmission despite of expanded preventive programs and ART coverage [24]. Transmission is also enhanced by the fact that use of HIV prevention methods such as condom use, abstinence and safe sexual practices is inconsistent. These two factors contribute greatly to the burgeoning population of 2 million individuals that gets infected by HIV each year [18]. Thus, the development of novel and efficacious prevention technologies and interventions remain a public health priority.

The likelihood of HIV transmission is also tightly linked to the infectious dose of the infected partner. The strongest predictor of HIV sexual transmission is plasma viral load. There is a direct relationship between viral load in blood and sexual transmission of HIV with a cutoff at approximately 1,500 copies of viral RNA per milliliter of blood (copies/ml) [25]. As the viral load of a HIV-positive individual reduces, the chance of transmitting the virus also decreases. HIV can be transmitted during any stage of infection, but individuals in the acute infection stage that occurs within the first 2-4 weeks [20] as well as those suffering in late-stage of the disease have higher viral loads and therefore have a higher risk of transmitting the virus [20,26]. People who live with HIV and are taking anti-retroviral drugs typically have a low plasma viral load and are

less likely to transmit the virus [27]. Although, undetectable plasma viral load in the seminal fluid or genital secretions does not necessarily indicate a low viral load since host genetics and immune function also plays a role [28] .

In the event that HIV is successfully transmitted, the major pathology of the virus is through rapid dissemination to the lymphoid organs, thereby establishing chronic and persistent infection as a hallmark of HIV disease [29-31]. HIV infects and replicates in CD4 positive cells, primarily lymphocytes but macrophages and dendritic cells can also be infected [32]. Dendritic cells are thought to be one of the cell types near the mucosal surface that initiate the infection by binding to the HIV envelope glycoprotein in order to retain the infectious particle for days. Following local expansion, these cells facilitate the transportation of the virus from peripheral tissues to susceptible CD4 cells during the acute infection stage. This promotes the virus to replicate and establish the lymphatic tissue as the major viral reservoir of the body, further disseminating the massive destruction of CD4 cells and penetration of microbial translocation products in the systemic circulation [33]. In HIV-uninfected individuals, the CD4 T-cell count is approximately 500-1,500 cells/mm^3 but as HIV infection progresses to a chronic phase, the population of CD4 T-cell declines to less than 350 cells/mm^3 diminishing the host immunity [34]. Despite the vigorous cellular and humoral immune responses by host cells during primary HIV infection, the virus manages to escape an immune-mediated clearance [35]. Hence, the infection is never eliminated completely from the body once it has established in the lymph nodes. The depletion of immune cells leaves the patient susceptible to opportunistic infections and non-AIDS associated events that contribute to morbidity and mortality [36]. Since immunosuppressed individuals are more susceptible

to other communicable diseases including tuberculosis (TB), malaria and pneumonia, these co-infections are largely responsible for AIDS-related deaths. Currently, there is no cure for HIV/AIDS so HIV disease progression can only be managed through lifelong and extensive drug therapy. In the absence of a cure, HIV prevention remains the most effective means of stemming the global pandemic.

1.3 HIV-1 structure and replicative cycle

HIV is a spherical enveloped virus with a typical retroviral genomic organization. It encodes three main structural polyproteins as *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) genes in 5' to 3' order, flanked by 2 long terminal repeat (LTR) elements. The *gag* gene codes for viral capsid proteins and plays a pivotal role in assembly of virus-like particles and virion maturation. The gag polyprotein is post-transitionally cleaved into 4 proteins that form the internal structure of a virion. The core also contains the *pol* gene products, which is the RNA-dependent DNA polymerase or reverse transcriptase (RT), integrase (IN) and protease (PR) molecules. The reverse transcriptase enzyme transcribes the single-stranded viral genomic RNA into double-stranded DNA in order for the integrase enzyme to incorporate the HIV proviral genetic material into the chromosomal DNA of an infected cell. The protease enzyme cleaves the polyproteins into their component mature viral proteins so they can become part of new fully functioning HIV particles. The *env* gene codes for viral envelope-associated proteins that resides in the lipid layer which aids in the budding process and determines viral tropism [19].

The HIV-1 retrovirus life cycle follows a series of successive steps for viral replication via binding, fusion, viral entry, transcription, integration, replication,

processing, budding and maturation. The initial step of viral life cycle is the attachment in which HIV virion binds to the host target cell via interaction between viral envelope glycoproteins and receptor molecules on the surface of the target cell. The second step is fusion of viral envelope with the host cell membrane. After viral entry into the target cell, the virion uncoats and exposes the viral RT enzyme that allows transcription of viral RNA genome into a single stranded cDNA. The resulting viral cDNA is then translocated to the host nucleus where the viral IN enzyme inserts the viral DNA into the host genome, allowing replication upon host cell activation. Gene expression is highly influenced by the complex interplay between viral and host transcription factors [37]. The integrated DNA form of the virus is called a provirus, which could remain dormant (non-replicating) and lead to latent infection, or it can immediately begin producing more viral RNA. The viral mRNA undergoes protein synthesis, which utilizes the host translational machinery and viral protease, which cleaves the longer proteins into smaller functional components. The *env* gene encodes for the envelope glycoproteins and viral assembly occurs as RNA packaging starts near the host cell plasma membrane for budding. Newly formed immature virion buds out of the host cell and remains non-infectious until the gag-pol polyprotein undergoes autocatalysis to release the RT and IN gene products as well as PR. The protease cleaves the gag and gag-pol components into mature structural proteins [19,38]. This produces an intact HIV virion capable of carrying on the infection cycle.

A key factor for the appearance of resistant strains of HIV is the retroviral reverse transcriptase. This enzyme lacks the exonuclease-proofreading ability of cellular DNA polymerases and is therefore prone to nucleotide misincorporation during cDNA

synthesis. The rapid replication cycle generates approximately 10^{10}-10^{11} virions daily with an average of 1 mutation per genome per replication cycle [39-41]. Evolution of variable HIV sequences is an interaction between mutations and selection, where the high mutation rate during replication facilitates viral variability and selection favors more successful mutants to produce more offspring than less successful mutants. The distribution of mutants with their extensive genetic heterogeneity in virus populations has been defined as quasispecies. Viral quasispecies in patients are likely generated on a daily basis with this combination of replication, mutation and selection. Drug resistant mutations (DRMs) may be present in infected individuals prior to the initiation of therapy, which supports the need for combination therapy in HIV infected patients [42]. Generation of viral quasispecies due to high viral mutation rate along with the lack of understanding on HIV-1 pathogenesis and anti-HIV-1 immunity against the infection or disease progression has hindered the development of effective vaccines. Despite considerable efforts in the development of HIV vaccines, a safe and effective HIV-1 vaccine suitable for human use still remains an ideal. As a result, drug therapy is currently the only form of treatment available for HIV-1 infected patients.

1.4 Antiretroviral therapy

1.4.1 Drug classes

The introduction of combination therapy in 1985 reclassified HIV infection from being an inherently untreatable disease to a chronic, usually non-fatal one [43]. Antiretroviral therapy (ART) is the use of combination of drugs to treat HIV infection by interrupting the viral life cycle at different stages and protecting the host immune system.

The drugs do not eliminate the virus or cure the disease, rather they act to reduce the progression of the disease.

HIV drugs are grouped into different classes according to the stage of viral replicative cycle that the drug targets. Among the 5 classes of inhibitors, there are three main drug classes that comprise the first and second line treatment regimen. Nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs) include drugs that halt HIV DNA synthesis: lamivudine (3TC), stavudine (d4T), abacavir (ABC), azidothymidine (AZT), didanosine (ddI), tenofovir disoproxil furmarate (TDF). These nucleoside analogs mimics building blocks, binding to the active site of the polymerase domain in RT and inhibiting the synthesis of double stranded viral cDNA. Non-nucleoside RT inhibitors (NNRTIs) such as efavirenz (EFV) or nevirapine (NVP) bind to an RT hydrophobic pocket, modifying its structure and impairing the catalytic site which interferes with the enzyme's ability to convert HIV RNA into DNA. Protease inhibitors (PIs) such as lopinavir/ritonavir (Kaletra) mimic viral peptides and bind to the active site of the protease enzyme to prevent assembly and maturation of the virion in a late stage of the life cycle [44]. The other two classes of inhibitors interfere in the early steps of virus life cycle and are mostly used in salvage therapy that is used when a patient fails second line therapy. Integrase inhibitors such as raltegravir (RAL) block the integration of viral cDNA into the host genome. Entry inhibitors such as maraviroc (MVC) bind to envelope viral proteins and prevent the fusion of the viral envelope and host cell membrane. It also blocks virus attachment to host cell receptors to prevent viral entry into target cells.

1.4.2 Challenges of ART

Although the use of multiple antiretroviral drugs has dramatically reduced morbidity and mortality in developing nations, none of the current regimens are curative. Current regimens merely inhibit viral replication, block acute symptoms or slow down the disease progression. The level of HIV RNA in blood plasma is arguably the most important surrogate marker in the treatment of HIV infection. To measure the efficacy of ART regimen on an infected patient, HIV viral load (level of HIV-RNA in blood) and CD4 count (strength of immune system) are measured periodically. The goal of ART treatment is to reduce the plasma viral load to an undetectable level (<50 copies/ml) and sustain a healthy CD4 cell count [12,45]. Virologic suppression also decreases the ability of HIV to develop ART resistance. With an appropriate ART regimen, virologic suppression and CD4 increase should occur by six months after ART-initiation [45]. Despite these remarkable successes in the ART era, the cure is far from being achieved. Access and affordability of antiretroviral regimen is a challenge on its own, especially in resource-limited countries [45]. Even in countries not facing this problem, the life expectancy of people undergoing ART treatment remains lower than the HIV-uninfected population [46]. Another challenge is that viral rebound follows discontinuation of therapy; hence the therapy must be undertaken for a lifetime with proper adherence for ART to be effective [43].

Non-adherence is the most significant challenge to the successful management of HIV-infected individuals, especially children and adolescents which could affect their coping and adjustment to the demands of living with a chronic health condition. Adherence barriers include high cost of the drugs, difficulty swallowing drugs,

inconsistent parental supervision for HIV-infected children, self-sufficiency, lack of affordability, lack of knowledge in understanding disease and consequences, social stigma and lifestyle factors such as forgetfulness and worrying about disclosure. Adolescents who experience housing instability, lack of transportation to clinics, family impact of the condition and desire to gain autonomy may result in lowered adherence. Only 33% of sexually active adolescents report disclosing their HIV status to their partner to avoid the fear of rejection, raising the risk of transmission [45]. This physical and psychosocial evolution has prohibited many mothers infected with HIV from adhering to routine tests and ART compliance, further exposing their infants to risk of infection. High mortality rates in perinatal HIV-infected children suggest that those who survive into adolescence may be slow progressers [45]. Additionally, adolescence and older age has been related to poor adherence in resource-rich and resource-limited countries, possibly reflecting lack of host cell growth due to chronic pattern of poor adherence [45,47]. These results suggest that adherence counseling should be a key component of the clinical response to detectable low level viremia (LLV) with viral loads of 50-1000 copies/ml, and that choice of ART regimen may also impact LLV.

Moreover, chronic antiretroviral therapy is associated with cardiac or metabolic side effects and related disorders such as heart disease, diabetes and possibly coronary artery disease in adolescent and adult patients [48] . Opportunistic diseases largely caused by infectious agents can occur in the course of increasing disease progression, hastening the rate of host immunosuppression. Increased risk of anemia with certain ARTs (like AZT) has also enhanced co-morbidity conditions like malaria, pneumonia, malnutrition and tuberculosis. Younger children are at an increased risk of developing

HIV-associated chronic lung disease such as pneumonitis. Growth failure and reduced micronutrient status is also evident in paediatric HIV infection and generally distinguishes adolescents with perinatally acquired HIV infection from those infected sexually [49]. The higher prevalence of co-morbidities can affect drug metabolism and consequent drug-drug interactions from treatment may in turn have an impact on ART efficacy. Improvements in health infrastructures could improve care provision through treatment of co-morbidities and nutritional support.

ART drug dosage and regimen has resulted in toxic side effects and development of drug resistance where patients may fail to prevent disease progression even if they adhere to the regimen [45]. Drug resistant mutations have been described for ART drugs in clinical use. Some patients have DRMs prior to ART treatment initiation while others develop mutations after they have started ART regimen. Drug resistance affects up to 30-50% of all individuals on ART [50]. A major concern from increase in drug resistance is that it inevitably leads to a reliance on "last line" drug options that can be unaffordable, difficult to administer or have severe side effects. Several studies have shown correlation of resistance or reduced treatment options to morbidity and mortality in both resource-rich and resource-limited settings, especially affecting children [51].

HIV infection in younger children has been associated with rapid disease progression. The illness typically presents failure to thrive severe viral infections and higher chances of contracting AIDS-defining illnesses. On the contrary, older children with HIV infection are either asymptomatic or have multiple recurrent infections that are not recognized as an indication of underlying HIV infection [49]. The viral incubation time distribution was shown to be substantially shorter in adults compared with younger

individuals, likely enhancing the recurrent illness rate in children [52]. HIV-infected children are vulnerable to vaccine-preventable illnesses due to a lower response rate and declining immunity after early childhood immunizations. The amount of HIV RNA and CD4 counts were both independent predictors of antibody response to children who received hepatitis B vaccinations. The CD4 cell count in children is insufficient to determine the risk of diseases and duration of immunosuppression. As compared to adults, HIV-driven chronic immune activation in children can lead to premature ageing of the immune system and cause lymphatic memory cells or T-cells proliferative disorders, which promotes viral replication and presence of opportunistic illnesses in adulthood [49]. The growing population of perinatally infected children that are emerging into adolescence and adulthood require proper ART adherence to control viral replication, prevent immune system deterioration and to avert secondary transmission. The treatment options are increasing gradually but not at a pace required to effectively contain the HIV epidemic. Overcoming these daunting challenges is critical to ensure improved health outcome for HIV-infected children and adults to maximize treatment outcomes.

1.4.3 Guidelines for treatment

Laboratory measurement of CD4 T-cells is a commonly used tool to establish a patient's degree of immunosuppression and their eligibility for treatment. Mild or advanced immunosuppression occurs when a patient has a CD4 T-cell count between 350-499 cell/mm^3 or 200-349 cell/mm^3 respectively, whereas a CD4 <200 cell/mm^3 is considered as severe immunosuppression [53]. The U.S. Center for Disease Control (CDC) classification system assesses the severity of HIV related immunosuppression by

CD4 T-cell counts to aid in clinical and therapeutic management of HIV-infected people. When there are insufficient resources to perform immunologic tests such as in certain resource poor settings, clinicians rely on clinical parameters when assessing a patient's disease status [54]. A World Health Organization (WHO) clinical staging and disease classification system developed in 1990 classifies HIV disease on the basis of clinical conditions or symptoms to guide the diagnosis, evaluation and management without requiring CD4 cell count measurement [4,53]. WHO clinical stages are categorized as stage I for asymptomatic infection, stage II as mildly symptomatic progressing to stage III with concerning symptoms emerging from HIV-infection, advancing in stage IV to HIV/AIDS. The WHO clinical staging system has shown to be a practical and accurate way to manage patients and compares favorably to laboratory markers including CD4 cell count [54].

The WHO guidelines recommend treatment for adults and adolescents with severe or advanced HIV clinical disease in stage 3 or 4, and individuals with CD4 count ≤350 cell/mm^3 regardless of clinical stage [10,45,53]. Patients co-infected with tuberculosis and pregnant mothers infected with HIV are recommended to be placed on ART, regardless of WHO stage and CD4 counts [55]. However, only 11% of eligible pregnant women currently receive the treatment in resource-limited countries [56] . HIV infected children less than five years of age should be initiated on ART, regardless of clinical stage or CD4 cell count. For children five years of age or older, CD4 cell count ≤500 cell/mm^3 or at WHO clinical stage 3 or 4 are recommended for treatment initiation [10].

According to WHO, an initial regimen has a combination of 1 NNRTI and 2 NRTI's, one of which is either AZT or TDF. A preferred option is TDF + 3TC + EFV but

if unavailable, TDF + 3TC + NVP or AZT + 3TC + EFV/NVP are recommended [10]. If a patient fails first line treatment, they are switched to the second line regimen that consists of 1 PI and 2 NRTI's. This usually includes antiretroviral drugs lopinavir (LPV) and either of AZT or TDF with 3TC [45,55]. AZT or TDF are commonly replaced by Abacavir (ABC) or Didanosine (ddI) in children less than 3 years of age. The last option of treatment is salvage therapy, where active drugs (drugs to which the HIV is not resistant) are combined with drugs of different classes (integrase or entry inhibitors) [20]. There are currently no effective guidelines for the management of patients failing second-line therapy [55]. ART guidelines can be used on patients across the globe, as it is reported that a patient's subtype does not influence a positive treatment response [57-59].

1.5 Low Level Viremia (LLV)

1.5.1 Defining low level viremia

Viral load (VL) is the amount of virus measured in the blood at a given point of time during the course of ART. The suppression of plasma HIV RNA viral load below the limit of quantification of <50 copies/ml with standard clinical assays is a widely accepted indicator of successful ART. Instead of achieving and maintaining virological suppression, a small portion of patients may first experience a rapid VL reduction episode but then retain a low but detectable VL for an extended period despite continued ART use. Low-level viremia (LLV) refers to a residual viremia with VL at 50-1,000 copies/ml whilst still receiving ART. When there are two consecutive LLV episodes over ≥ 1 month, it is known as persistent low-level viremia (pLLV) [60-62]. Among HIV-infected patients with sustained viral suppression, 25% on average will have a single episode of

LLV that is immediately preceded and followed by a viral load below the limit of detection; this is known as a blip [63-67]. Overall, about 10-40% of patients who have reached virologic suppression can have a rebound between 50 copies/ml to 1000 copies/ml, leading to either persistent LLV or blips [66,68] .

1.5.2 Kinetics of LLV

The exact mechanisms of accumulation of LLV remains unsolved but several studies have reported increased evolution of the HIV envelope and polymerase genes, slower decay of the latently infected pool and emergence of subpopulations of drug-resistant virus [69-71]. Maintaining virological suppression depends on a number of variables including regimen potency, adherence, and tolerability. Abrupt elevations in HIV RNA levels are observed when patients interrupt therapy or when the potency of a regimen is reduced [72]. Detection of LLV is likely to indicate significant levels of ongoing viral replication even after an undetectable VL has been achieved. It could be that some medications are less effective in suppressing the replication of HIV due to incomplete penetration into the target cells. The reduced rates of viral suppression can also be a result of higher baseline HIV RNA levels and/or the presence of baseline drug resistance and polymorphisms in receptor genes [73]. Alternatively, ongoing or intermittent bursts of inflammation may contribute to activation of HIV within reservoirs (lymphoid tissues) that escape into the bloodstream where it is detected. In some patients, transient episodes or blips may represent assay variation or laboratory errors in specimen processing, identification or reporting [74] .

Adherence and tolerability to an ART regimen is a key determinant of treatment success and avoiding an event of LLV, especially in children. The levels of ART adherence in resource-limited settings are comparable to those in resource-rich settings. ART adherence in children or adolescents is influenced by their social environment, the level of vigilance of their primary caregiver, lack of knowledge or awareness, pill burden, absence of a pediatric formulation, high viral load observed in some infants and imprecise pharmacokinetic data for children. Adolescents are especially prone to the fear of stigmatization and have an increased desire to assimilate with their HIV-negative peers [75]. Moreover, treatment fatigue is quite common in children and adolescents with prolonged ART usage. The reduction in morbidity during prolonged ART usage lead to a feeling of having been cured. This may result in decline or suspension of ART use, accompanied with viral load rebound to LLV, recurrence of opportunistic infections and introduction of drug resistant mutations [76] .

1.6 Implications of LLV

1.6.1 Drug Resistant Mutations (DRMs)

pLLV has been linked with the early occurrence of HIV DRMs as these are typically selected during periods of increasing HIV replication (viral rebound to >50 copies/ml). DRMs are usually found in the *pol* gene of HIV genome, mainly concentrating in protease and reverse transcriptase regions [77] . Pre-existing or minor DRMs at baseline can rapidly dominate the HIV quasispecies in patients when they start treatment [78] . Moreover, the drug-resistant variant in women during pregnancy can be transmitted to their children, which can have serious implications for the future of ART

success in those infants. Each drug can give rise to several new mutations. Some of the most common mutations from NRTI class have highest prevalence of M184V/I (53%) from 3TC/FTC, thymidine analog mutations (TAMs) such as T215F/Y (27%) from AZT and d4T and K65R (5%) from ABC/ddI/TDF. Major mutations from NNRTIs are K103N (36% prevalence), Y181C (20%) and G190A (15%). Major PI resistance mutations include L90M (28%) and I84V (11%) in ART-treated patients [79] .

Certain NNRTI-based regimen and PI regimens have been associated with higher rates of resistance. Differential kinetics between subtypes leads to distinct prevalence of DRM acquisition over time but the mechanisms are unknown. With respect to NNRTIs, a study in Africa on pregnant women and their infants facing mother-to-child transmission using NVP indicated the proportion of viral resistance, particularly K103N, Y181C and Y188C, was higher in subtype C (69%) than in subtype A (19%) and D (36%) [80,81] . Another study identified NVP mutations in 65-87% of women and infants between 6-36 weeks [82] . Persistence of DRMs has been documented in long-lived cellular reservoirs as well as in blood plasma spanning several months to years after treatment discontinuation in patients who never achieved viral suppression [83] . A study found that the vertically transmitted HIV-DR mutations K103N and Y181C could persist for up to 7 years, even in the absence of ART drug pressure [84] . In contrast to NNRTIs where a single point mutation can confer high level DR, substantial resistance to PIs requires an accumulation of multiple mutations [85]. Several studies assessing resistance levels in HIV infected population failing first and second-line regimens found that 34-99% children had NNRTI and NRTI drug resistant mutations, while 52% of perinatally infected adolescents had dual-class and 12% had triple-class ART resistance [86,87] .

Inadequate suppression of viremia compromises future treatment options and these concerns are magnified in low weight adolescents where appropriate lower dose formulations are not available, a common problem in resource-limited countries.

Patients are more likely to develop high-level resistance to their regimens during prolonged periods of non-adherence [88]. New DRMs have been associated with higher HIV-1 RNA levels during pLLV. During persistent LLV, new DRMs were detected in 37% cases with most common mutations being M184I/V, K103N and M230L [62] . Long-term persistence of DRMs at low viral load should be monitored closely, especially in patients restarting ART to avoid secondary transmission or in women to prevent mother-to-child transmission of HIV along with resistant mutations. Nevertheless, routine DRM testing for all patients is cost-prohibitive in resource-limited countries, eventually leading to less effective treatment in infants and children.

1.6.1.1 Amplification of LLV to detect DRMs

To optimize treatment decisions and for research purposes, a reliable and sensitive assay is essential for amplification of LLV specimens in order to understand PR and RT inhibitors and their relationship with DRMs. Conventional genotyping methods rely on efficient polymerase chain reaction (PCR) of HIV genes encoding targeted viral proteins. The challenge common to PCR amplification of any HIV gene is the high viral sequence diversity and limited template number when patients' VL is low due to either the early stage of infection or residual LLV [89] .

The standard methods fail to yield reliable amplification when pVL is <1000 copies/ml [90]. A three-step PCR strategy was used in a study to improve sensitivity of

the 2-step amplification procedure but could only amplify four out of eight samples with VL >80-1,000 copies/ml [90]. Another study used the ultrasensitive protocol with 1.5 ml plasma input volume, longer ultracentrifugation time and a three-step PCR. Their amplification was 61% successful with LLV >50 copies/ml, 37% LLV 20-50 copies/ml and 23% with VL <20 copies/ml [91]. Single genome sequencing (SGS) is a reliable method for amplification of the genome as it is non-labor intensive and has no PCR bias but is limited by primer design which constrains the HIV-1 subtype populations it can amplify [92]. Nucleic acid-amplification has been performed on plasma using an ultrasensitive isothermal Transcription Mediated Amplification (TMA) protocol. It is a highly specific and sensitive assay with 50% success rate at a detection limit of 3.6-14 copies/ml [93] .

Although the amplification success rate is worthy in some studies, other studies report that these methods have many limitations such as low consistency, lack of reproducibility due to absence of technical details, narrow HIV-1 subtype coverage or limited sensitivity on real-world LLV samples. The amplification variation and error rate tends to increase at the lower limits of detection of these assays. Moreover, most of these methods require a high plasma input volume of >1 ml [91-94]. Some procedures involve advanced instrumentation as a necessary component of the analysis (bead-based or automated extractions, ultracentrifuges) that might not be available in resource-limited settings, and also could be very costly to use.

1.6.1.2 Detecting DRMs in the LLV population

Most clinical DR genotyping tests rely on Sanger sequencing of the *pol* gene both for subtyping and for the detection of DRMs, but the interpretation is limited by the inability to detect minority variants below 20-30% [95-97] . Several assays have been developed to overcome the limitations of Sanger sequencing. Phenotypic assays have been modified for enhanced sensitivity for use in routine practice to assess DR by determining the effect on viral replication, which can be further characterized by DNA sequencing to potentially identify novel DRMs, but these do not quantify minority variants and have a long turnaround time as well as significant costs [98,99]. Population sequencing technologies can be hard to interpret, as they cannot determine whether mutations are from different variants in the population [40]. Point mutation assays such as allele-specific real-time PCR (ASPCR) [100], oligonucleotide ligation assays (OLA) [101,102] and line probe assays (LiPA) [103], detect known mutations with high specificity and sensitivity (<1%), are relatively inexpensive, but results show false-positive rates for DRMs present at the lower limit of detection. Additionally, these detect only a single mutation at a time and have a reduced ability to address genetic mutation linkages and polymorphisms similar to bulk quasispecies population sequencing [40,104]. Consequently, estimates from these assays may provide inaccurate quantification of low-abundance DR variants and, therefore, precaution needs to be taken in their interpretation. Whole genome sequencing is the dominant way to determine if all relevant mutations are present, thus providing a complete resistance assessment. Sequencing assays such as ultra-deep pyrosequencing (UDPS) and SGS have been the "gold standard" for low abundant DR variant detection as these are less susceptible to

polymorphisms. These can analyze the entire sequence context, and quantitatively detect any mutation with high sensitivity (~1% relative abundance) while providing both proportional and absolute number of sequencing reads harboring certain mutations. SGS enables a genetic linkage analysis of detected DR mutations within the same viral sequence and identification of complex mutations. Nevertheless, these assays have small read length, and tend to be labor-intensive and cost prohibitive [40,105] .

A common drawback for SGS and point mutation assays is that primers are designed for specific lab strains that might not perform well with different subtypes that circulate globally. Another issue for all genotyping technologies designed to detect low-level viral variants is that sensitivity is dependent on the number of viral templates provided for cDNA synthesis and amplification. Moreover, the sensitivity of variant detection may be lower than implied as the summation of enzymatic error rates (for extraction, PCR and sequencing) may be limiting for methods employing enzymes [104]. Extraction and amplification bias along with enzymatic and sequencing errors can lead to inaccurate results and draw concern for reliability of standard genotyping methods.

A lack of dependable sensitive assays and an emerging importance of detecting DRMs at low levels have motivated the development of diagnostic technologies employing large-scale genomics with Next Generation Sequencing (NGS) platforms like Roche 454 Life Science FLX, Ion Torrent and Illumina. These technologies share a common paradigm of massive, parallel, clonal analysis of DNA templates with high data throughput. These methods have high accuracy (~99%) and can capture low frequency mutations at levels of as low as 1% of the viral genome population but also have limitations that include short read length and sequencing errors in homopolymeric

regions (for semiconductor sequencing (Ion Torrent) and pyrosequencing (454) technologies) [106]. This makes it hard to reassemble the sequence population and estimate frequency of low-abundant mutations. Among the available NGS technologies, the Illumina platforms are preferred for HIV-related studies because of their accuracy (99.9%), fast turn-around time, and long read length of 600 base pairs (bp) when paired-end sequencing is performed.

1.6.2 Virologic Failure (VF)

Low-level viremia may lead to potential viral rebound and disease progression in HIV-1 infected individuals. Virologic failure (VF) occurs when an infected individual's plasma viral load rebounds to >1000 copies/ml after previously attaining at least one undetectable (<50 copies/ml) viral load during therapy [65-67,107-109]. Several factors previously linked to VF include ART adherence, DRMs, ART regimen characteristics, duration of ART, clinical stage at the start of therapy, co-morbidities, high baseline viral load, prior treatment failure, drug side effects and toxicities, suboptimal pharmacokinetics and drug-drug interactions [45,110] .

Some reports have suggested drug resistant mutations as a major factor causing VF [40,111,112]. DRMs arising from incomplete viral suppression, reducing the effectiveness of ART and causing viral rebound in infected individuals can further lead to VF. An analysis showed that low-frequency NNRTI mutations increased the risk of VF by two fold while the impact of PI DRMs on treatment response has been limited to a few studies that found no association [91,112]. DR variants in LLV present as a minority population are often missed by standard assays may lead to VF, hence maintaining

virologic suppression throughout the course of therapy is critical. Virologic suppression rates in adolescents range from 28-70%, which is much lower when compared to 90% for adults on a similar regimen, increasing the chances of VF in than in adults [45,88]. Studies have suggested the absolute number of DRMs may impact the time leading to virologic failure in a person on therapy [104,113].

A group in Uganda conducted a study on adults and children in resource-limited settings for 12 months and found a VF rate of 14% in adults and 26% in children [114]. Moreover, K103N mutation was found in 100% of those patients experienced VF. Another study with patients in a developed country found 34% of virologic failure rate in children aged 2-15 years and 71% of those with VF had at least one DRM, or 23% of the children in the study [11]. Overall, several findings have revealed HIV-infected children and adults have a VF rate between 16-42% and that chances of having VF is higher among ART-treated patients than treatment-naïve patients [11,109,114-116]. This is a significant amount of patients failing the treatment per year, which is a substantial global burden that could potentially be alleviated.

1.7 Gaps in knowledge & study rationale

1.7.1 Gaps in knowledge

The HIV-1 epidemic and the devastating consequences on human life presents highly variable genome sequence making HIV-1 one of the most challenging pathogens to combat for researchers, clinicians, vaccinologists and pharmaceutical professionals. Genetic variability is responsible for the virus establishing persistent infection and inducing disease after a long incubation period. Many studies support the clinical impact

of LLV and other risk factors on VF but in some cases, blips and low-level viremia represent viral production that is irrelevant for the long-term virological outcome. There is evidence of LLV (pLLV and/or blips) increasing the risk of VF [61-63,65,66,96,112,116,117] but some reports suggest otherwise [65,67,68,72,118,119]. Few studies have found no significant association between the presence of low-frequency DRMs in LLV or subsequent virological failure [90,91,95,120,121] while others reported an overt correlation [67,95,96,100,110,116,119,122]. A study done on 7 children and 11 adults concluded that LLV lacked DRMs and were not indicative of failure [90]. Hence, clinical implications of low-level viremia and DRMs at the time of VF remain inconclusive.

It has been demonstrated that low ART adherence and the nature of ART regimen increases the risk of LLV [67,108] but Castor et al. stated adherence did not modify the association between LLV and VF [123]. A study done by Vandenhende et al. found that time-period of ART initiation was associated with LLV and VF, but ART regimen was not. Geretti et al. related the effect of gender on VF where women have higher chances of failure but other studies do not support the association [110,124,125]. Although there are conflicting results for other risk factors as well, the impact of CD4 counts on LLV is well supported [126-129], while the effect of co-morbidities on VF has not shown significant results [124,125].

These conflicting results could be due to several limitations in previous studies. Lack of analysis on children population represents a major gap in knowledge. This can determine whether the clinical outcomes are similar to those of adults as studied previously and guide treatment of this age group for proper management. The

inconsistency in results among similar age groups studied arises from having differences in the case definition for the upper VL limit of persistent LLV, blips and VF. Studies reporting increased risk for VF had included participants regardless of whether they initially achieved virologic suppression [69,70,90]. Having VL <50 copies/ml before reaching failure is important to confirm the efficacy of regimen. Many studies did not address adherence in patients, which can be important to evaluate immune and clinical response. Having a small study population requires further investigation to confirm the findings as it provides low statistical power. It is largely unknown how frequently low-abundance HIV drug-resistant variants are present at levels under the limit of detection of conventional genotyping (<20% of quasispecies) in antiretroviral-experienced persons experiencing virologic failure. Emerging drug resistance along with other risk factors in patients with LLV and subsequent VF poses a major concern for HIV programs in settings when treatment options are limited.

1.7.2 Study rationale

Low-level viremia and various other risk factors have been linked directly or indirectly to virologic failure. This has been showcased mainly in the adult population but rarely any investigation has been done in children. The immune system in children is not fully developed for a strong response to the hard-hitting attack of the virus and is further weakened in children who lack proper nutrition and/or have co-infections. These individuals are more prone to acquiring drug resistant mutations from mother-to-child transmission, and therefore face additional challenges of ART and risk of LLV. Their ability to reach adulthood is challenged as they face limited treatment options along with

adherence problems at a very young age. An effective treatment for children is influenced by their care environment, peer influence, ARV pill burden, and even formulation (pill or syrup). They are prone to the fear of social stigmatization by HIV-negative peers. Treatment fatigue is common in children that can lead to poor outcomes, especially with the incidence of co-morbidities. Moreover, the mechanism behind virologic failure susceptibility is poorly understood in children. Simple statistical tests for data analysis may underappreciate the complexity of LLV and VF, thus additional analysis using survival curves, univariate and/or multivariate models should also be used for association studies. Additionally, a more sensitive method is required to amplify low-level viremia as well as to detect drug resistant mutations. Next generation sequencing coupled with statistical tests to validate mutations may more accurately determine and quantify DRMs that cannot be detected by traditional approaches, and thus provide new insights into the clinical data associated with LLV and VF. If LLV is associated with virologic failure due to drug-resistant populations, early therapy switches or intensification could potentially improve long-term rates of viral suppression. On the other hand, if intermittent viremia is not predictive of virologic failure, then patients may be able to maintain a regimen for longer periods and avoid premature/unnecessary interventions. Not only resource-limited countries but also resource–rich countries are working towards improved patient care and could have significance through emerging literature for efforts to eradicate HIV. A comprehensive analysis of the factors causing LLV and consequent virologic failure in children will have important implications for management of HIV.

1.8 Hypotheses and objectives

1.8.1 Hypotheses

1) Low-level viremia is associated with VF in HIV-1 infected children undergoing ART
2) Confounding factors including ART adherence, CD4 counts, WHO clinical stage, ART drugs and regimen and co-morbidities have an association with VF in HIV-infected children
3) The risk of reaching low-level viremia is multifactorial and does not exclude ART adherence, ART regimen, DRMs and co-morbidities in risk of LLV in HIV-1 infected children.

1.8.2 Objectives

1) Develop a protocol to amplify plasma samples of children with LLV
2) Detect and profile HIV DRMs in ART-experienced children during the period of LLV
3) Investigate and interpret the association between LLV and VF
4) Predict the association of potential risk factors such as ART adherence, ART regimen, co-morbidities, WHO clinical disease stage, gender, and duration of ART which are predictive of LLV and VF in the HIV-1 infected children

Chapter 2: Materials & Methods

2.1 Ethics

Scientific and ethical approval for this study was granted by the Research Ethics Board (HS19999 (B2016:078)) of the University of Manitoba in Winnipeg, Manitoba, Canada, and by the Scientific and Ethics Review Unit (KEMRI-SERU) of the Kenya Medical Research Institute (protocol #SSC-2500). Institutional approval for data collection and access to the medical records and plasma samples was granted by the Nyumbani Medical Board (NMB), which oversees the Leo Toto Program (LTP). The NMB granted the researchers consent on behalf of the caregivers and children. Patient confidentiality was ensured by substituting personally identifiable information with a unique alphanumerical code to each child in the study.

2.2 Study population

The Leo Toto program comprises 2,800 infants and adolescents infected with HIV-1. The cohort is operated by Nyumbani Children's Home in Nairobi, Kenya. This is a donor-funded community outreach program that hosts clinics in 8 resource-limited suburbs of Nairobi: Dagoretti, Dandora, Kangemi, Kariobangi, Kawangware, Kibera, Mukuru, and Zimmerman. The program provides to all infected children free ART drugs, CD4 T-cell count testing every 3 months, VL testing every 6 months, treatment for opportunistic infections, school fee assistance, food aid, and counseling. All the children were under 17 years of age and had received HIV-1 infection through mother-to-child transmission. This cohort only takes care of the children (not their caregivers) and all

children are on ART with first-line therapy, second-line therapy or salvage (third-line) therapy.

2.3 Data and sample collection

2.3.1 Data collection

Clinical, social and epidemiological data was collected for all patients from the 8 clinics. Upon admission to the clinic, each patient was given a unique client admission number to avoid disclosing identity. Data was collected upon admission for variables such as admission date to the clinic, birth date, gender, CD4 counts, WHO clinical stage and co-morbidities. Typically, CD4 counts were checked every six months to decide if the treatment should be started. Once the child was initiated on ART, information on their regimen, viral load, CD4 counts, therapy start date, switch/stop therapy date, therapy stage and ART adherence information was recorded periodically throughout the course of their treatment.

All the patients' data was initially recorded by a data officer present at each clinic in paper and/or electronically with no standard format among the clinics. The variables of interest were then recorded electronically for this study as listed above, except for the start date of ART, switch/stop date of ART, adherence, WHO clinical stage and co-morbidities information which were on a paper copy and were later added into the electronic database for each patient.

2.3.2 Plasma collection from children

The blood samples were collected and stored in a tube containing Ethylene-diamine-tetra-acetic acid (EDTA) to prevent coagulation. The plasma samples were prepared by spinning down non-coagulated blood at 2000× g for 5 minutes in a centrifuge at the Nyumbani Diagnostic Laboratory (NDL). In general, blood was collected every six months to measure CD4 counts and VL for all children admitted in the cohort, and the remainder of the blood collected was used for the study. For the proposed study, CD4 cell tests and viral load tests were periodically analyzed by the lab technologists at the NDL. CD4+ T-cell counts were determined using the BD FACSCalibur (BD, Franklin Lakes, New Jersey) and viral load was determined using Amplicor HIV-1 monitor kit version 1.5 (Roche Diagnostics, Alameda, California) according to manufacturer's instructions.

2.4 Amplification of RNA from LLV samples

2.4.1 EQAPOL samples

External Quality Assurance Program Oversight Laboratory (EQAPOL) HIV-spiked plasma samples with known viral load were used to develop a protocol for amplification of low-level viremia, for different subtypes. Eight HIV-1 subtypes A1, AE, AG, B, C, D, F2 and G were initially tested to ensure their quality. After familiarization with the standard automated protocol, four HIV-1 subtypes A1, B, C, D, mainly found in Africa were optimized for LLV protocol. The original stocks for these control samples were highly concentrated with viral load up to 6.87×10^5 copies/ml. Dilution factor was calculated for each sample and two-fold serial dilutions were prepared with Normal

Human Plasma (NHP) (SeraCare, USA) to dilute the sample to 50 copies/ml in order to mimic plasma samples with low viral load.

2.4.2 RNA extraction

2.4.2.1 Automated extraction

The first step of HIV genotyping was to isolate viral RNA from plasma samples. The automated, silica bead-based HIV RNA extraction using NucliSens easyMAG system (BioMerieux, Canada) was used to test the reliability of the procedure. This method is being used by the National HIV & Retrovirology laboratory at National Microbiology Laboratory (Winnipeg, MB) as part of the standard protocol for HIV DR typing in support of the Canadian HIV-1 Strain and Drug Resistance Surveillance Program.

Nucleic acid extraction was performed according to the manufacturer's instructions on EQAPOL control samples consisting of the eight aforementioned HIV-1 subtypes [130] . The input volume of 400 µl plasma samples was added to the appropriate wells in sample vessels containing lysis buffer and was incubated for 10 minutes. Silica beads were prepared by adding 100 µl of distilled water-beads solution to each well, followed by automatic magnetic separation. Nucleic acid was recovered using an elution volume of 110 µl and stored in -80 °C freezer. NHP and Accurun (SeraCare, USA) were used as negative and positive controls for this nucleic acid extraction step.

2.4.2.2 Manual extraction

The goal of optimizing the protocol was to amplify LLV plasma samples from children in Nairobi, so the QIAamp Viral RNA extraction (Qiagen) was chosen as it was readily available and could be easily shipped to Kenya without compromising the efficacy of reagents. Extraction was performed according to the QIAamp Viral RNA mini handbook instructions for "Purification of Viral RNA (Spin Protocol)" on HIV-1 subtypes A1, B, C and D [131]. The manual extraction was initially tested to compare its efficacy in comparison with the automated extraction method and was further optimized for downstream amplification.

Initially, 200 µl plasma input with 60 µl elution volume was used to compare against EasyMag automated extraction. Later, the input and output volumes were optimized for the four subtypes with VL ranging from 100-5000 copies/ml. The plasma sample input volumes tested over different trials were 300, 400 or 500 µl with respective changes in buffer volumes, eluted in either 50 or 60 µl of Buffer AVE and stored at -80 °C freezer. The elution volume for Qiagen was almost half of the automated NucliSens method in order to keep an equal concentration ratio of eluted samples. A sample with high viral load that had been successfully extracted multiple times by easyMAG was kept as a positive control while NHP was used as a negative control.

2.4.3 Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR)

2.4.3.1 One-step RT-PCR

Following the viral RNA extraction, one-step RT-PCR was performed on *pol* gene to amplify PR and the first 233 amino acids of RT region on samples with different

viral loads among subtypes A1, B, C, D using Invitrogen SuperScript III One-step PCR kit (Invitrogen, Canada) with Platinum *Taq* High Fidelity Enzyme mix (Invitrogen, Canada). The cDNA was synthesized and PCR amplified in one step using this kit. The 50 µl reaction was prepared as follows: 4 µl RNase free water (Invitrogen, Canada), 25 µl SuperScript III 2× reaction buffer, 4 µl of 5 mM $MgSO_4$, 2 µl of 200 µM PR_2071_F1 primer, 2 µl of 200 µM RT_3370_R1 primer (Table 1), 1 µl SuperScript III enzyme mix, and either 10, 12, 14 µl RNA template to corresponding PCR tube. Along with control specimens carried forward from manual extraction, amplification was carried out on a GeneAmp PCR system 9700 Thermocycler (ABI) using the following settings: reverse transcriptase step for 30 min at 50 °C, 2 min at 94 °C to inactivate RT enzyme and activate Taq mixture, 40 cycles of (20 sec at 94 °C, 30 sec at 50 °C, 90 sec at 68 °C) followed by final extension of 5 min at 68 °C and then hold at 4 °C till the end of the reaction.

2.4.3.2 Two-step RT-PCR

The first step of cDNA synthesis from total RNA was performed using Invitrogen SuperScript VILO cDNA synthesis kit (Invitrogen, Canada), followed by PCR amplification using SuperScript III RT-PCR kit (Invitrogen, Canada). Master mix for the first-strand cDNA synthesis contained 4 µl 5× VILO reaction mix, 2 µl 10× SuperScript enzyme mix and different template volume, either 10 µl or 14 µl template RNA, in a 20 µl reaction. The thermocycler settings were as follows: incubate at 25 °C for 10 min and then at 42 °C for 90 min, terminate at 85 °C for 5 min followed by a hold at 4 °C.

The 50 µl master mix for the second step qPCR included: 25 µl SuperScript III 2× reaction buffer, 4 µl 5mM $MgSO_4$, 2 µl 5 µM PR_2071_F1 primer, 2 µl 5 µM RT_3370_R1 primer (Table 1), 1 µl SuperScript III enzyme mix, and 16 µl template RNA. The Thermocycler conditions were the same as mentioned in one-step RT-PCR with the omission of RT step.

2.4.4 Nested PCR (N-PCR)

2.4.4.1 One-amplicon nested PCR

To further amplify the target genes, Phusion Hot Start II High Fidelity DNA-polymerase (Thermo Fisher Scientific, Canada) kit was used for one round of nested PCR with primers listed in Table 1. The master mix for 50 µl reactions were prepared as follows: 24.5 µl RNase free water (Invitrogen, Canada), 10 µl of 5× HF buffer, 1 µl of 10 mM dNTP mix, 5 µl of 200 µM PR_2243_F2 primer, 5 µl of 200 µM RT_3326_R2 primer (Table 1), 0.5 µl phusion enzyme (1 unit), and template volume of either 2, 4 6 or 8 µl DNA. Amplification was carried out on a GeneAmp PCR system 9700 thermocycler (ABI) with cycling conditions as follows: 30 sec at 98 °C to activate the enzyme, 35 cycles of (10 sec at 98 °C, 20 sec at 62 °C, 40 sec at 72 °C) followed by final extension of 10 min at 72 °C and then hold at 4 °C.

2.4.4.2 Two-amplicon nested PCR

For enhanced sensitivity in a cost effective manner, the same kit and protocol was used as 1-amplicon nested PCR, but with modified primers and minor changes in cycling parameters. For this process, two set of overlapping primers with ligated adaptors (~600

bp total) were designed to amplify the PR and partial RT genes. The first primer set binds to sequences outside the target DNA as expected in a standard PCR along with some binding to other areas of the template. The second primer set binds specifically to the sequences in the target DNA that are within the portion amplified by the first set (that is, the primers are nested). Thus, the second set of primers will bind and amplify target DNA within the products of the first reaction.

The forward overhang adaptor sequence was TCGTCGGCAGCGTCAGATGTG TATAAGAGACAG and the reverse overhang adaptor sequence was GTCTCGTGGGC TCGGAGATGTGTATAAGAGACAG for each primer used in the two-amplicon nested PCR procedure (Table 1). For every sample, two master mixes were prepared separately for PR and RT region using the same protocol as mentioned for 1-amplicon nested PCR procedure. The cycling parameters were as follows: 30 sec at 98 °C to activate the enzyme, 40 cycles of (10 sec at 98 °C, 30 sec at 58 °C, 20 sec at 72 °C) followed by final extension of 10 min at 72 °C. and holding temperature of 4 °C. Each master mix was then run on a separate gel to detect DNA amplification of the PR and RT genes, with expected bands around 609 bp and 645 bp, respectively.

Table 1: **Primers used in RT-PCR and nested-PCR procedures**. Detailed information of primers used in RT-PCR and Nested-PCR steps. The RT-PCR forward primer (F1) is from Richard Harrigan Lab (University of British Columbia, Canada) and Nested PCR forward primer (F2) is from Zhou et al [132]. Reverse primers for both PCR's are from Yang et al [133] .

	Primer Name	Location at HXB2 loci	Sequence 5' – 3'
RT-PCR	PR_2071_F1 (5µM)	2071-2095	GAR AGA CAG GCT AAT TTT TTA GGG A
	RT_3370_R1 (5µM)	3348-3370	ATC CCT GCA TAA ATC TGA CTT GC
Nested PCR [1-amplicon]	PR_2243_F2 (5µM)	2243-2266	CTT TAR CTT CCC TCA RAT CAC TCT
	RT_3326_R2 (5µM)	3304-3326	CTT CTG TAT GTC ATT GAC AGT CC
Nested PCR [2-amplicon]	PR_F1 (5µM)	2243-2266	CTT TAR CTT CCC TCA RAT CAC TCT
	PR_R1 (5µM)	2761-2785	TCT CTG AAA TCT ACT AAT TTY CTC C
	RT_F2 (5µM)	2692-2713	CAA AAA TTG GGC CTG AAA ATC C
	RT_R2 (5µM)	3249-3270	CTG TCC ATT TRT CAG GAT GRA

2.4.5 Gel electrophoresis

2.4.5.1 Automated capillary electrophoresis

QIAxcel (Qiagen) is an automated capillary electrophoresis system with high resolution, accuracy and high detection sensitivity compared to the conventional agarose gel electrophoresis [134]. For each DNA sample, 10 µl were loaded onto the machine with 5 kb DNA ladder, according to the manufacturer's instructions. The results were then read from QIAxcel ScreenGel software that is designed for use with QIAxcel system.

2.4.5.2 Manual agarose gel electrophoresis

Agarose gel electrophoresis was used to compare the detection sensitivity with the automated method as well as for the detection of DNA present in LLV patients in Nairobi. Briefly, reactions were resolved on a 1% agarose gel in 1× TBE running buffer along with a 1 kb plus DNA ladder (Invitrogen). 10 µl of GelGreen Nucleic Acid Stain (10,000× in water, Biotium) was added as a less toxic alternative to ethidium bromide. To load each sample onto the gel, 2 µl of 10× Blue Juice (Invitrogen) gel loading buffer was added in 8 µl of the DNA sample. The samples were then run for one hour at 100 V and photographed under the ultraviolet (UV) light on a transilluminator.

2.4.6 Validation of the optimized protocol

2.4.6.1 Validation with EQAPOL samples

The optimized protocol was validated using EQAPOL HIV-spiked plasma samples in triplicates (Table 2) with known viral load. These samples covered major

subtypes in Africa including A1, B, C and D. Prior to HIV RNA extraction, serial dilution was performed on all samples to obtain a VL range of 10,000, 5,000, 2,500, 1,000, 500, 200, 100 and 50 copies/ml. All samples were prepared in triplicates and then subjected to RNA extraction with 400 µl plasma eluted in 60 µl. One-step RT PCR with 12 µl template and one-amplicon nested PCR with 4 µl DNA were used throughout the workflow for amplification.

Table 2: Details of control samples from EQAPOL. To validate the consistency of the developed protocol on major subtypes prevalent in Kenya, EQAPOL samples were tested in triplicates. The ID for each sample is shown with their respective Gen Bank numbers that were used for each subtype.

EQAPOL ID	GenBank	Subtype
DEMA108RU003.01	KF716491	A1
DEMA106ES002.01	JX140651	
DEMA111KE002.01	KF716474	
DEMB11US015.01	KC473835	B
DEMB10VE001.01	JX140659	
DEMB11US004.01	KC473832	
DEMC06ES003.01	KC473844	C
DEMC09ZA114.01	JX140669	
DEMC07BR003.01	JX140663	
DEMD11UG003.01	KF716480	D
DEMD08UG001.01	KC596072	
DEMD10UG004.01	KF716480	

2.4.6.2 Validation with clinical samples

The LLV plasma samples were collected from HIV-infected children at LTP a year prior to the experiments. The optimized protocol was tested on 49 children plasma samples with low-level viremia having viral loads between 50 to 1000 copies/ml from Leo Toto cohort in Nairobi, Kenya. The volumes of initial plasma sample were between 50 µl to 500 µl thus phosphate-buffered saline (PBS) solution was added to increase the volume to approximately 500 µl for manual Qiagen extraction. The samples were further processed with one-step RT-PCR and one-amplicon nested PCR (Figure 1), then visualized on 1% agarose gel electrophoresis with 100 V for 30 minutes. For positive controls, positively extracted EQAPOL samples from EasyMag protocol were used, while NHP was used as a negative control. The nucleic acid concentration was measured on Nano Drop with MilliQ water used as a blank.

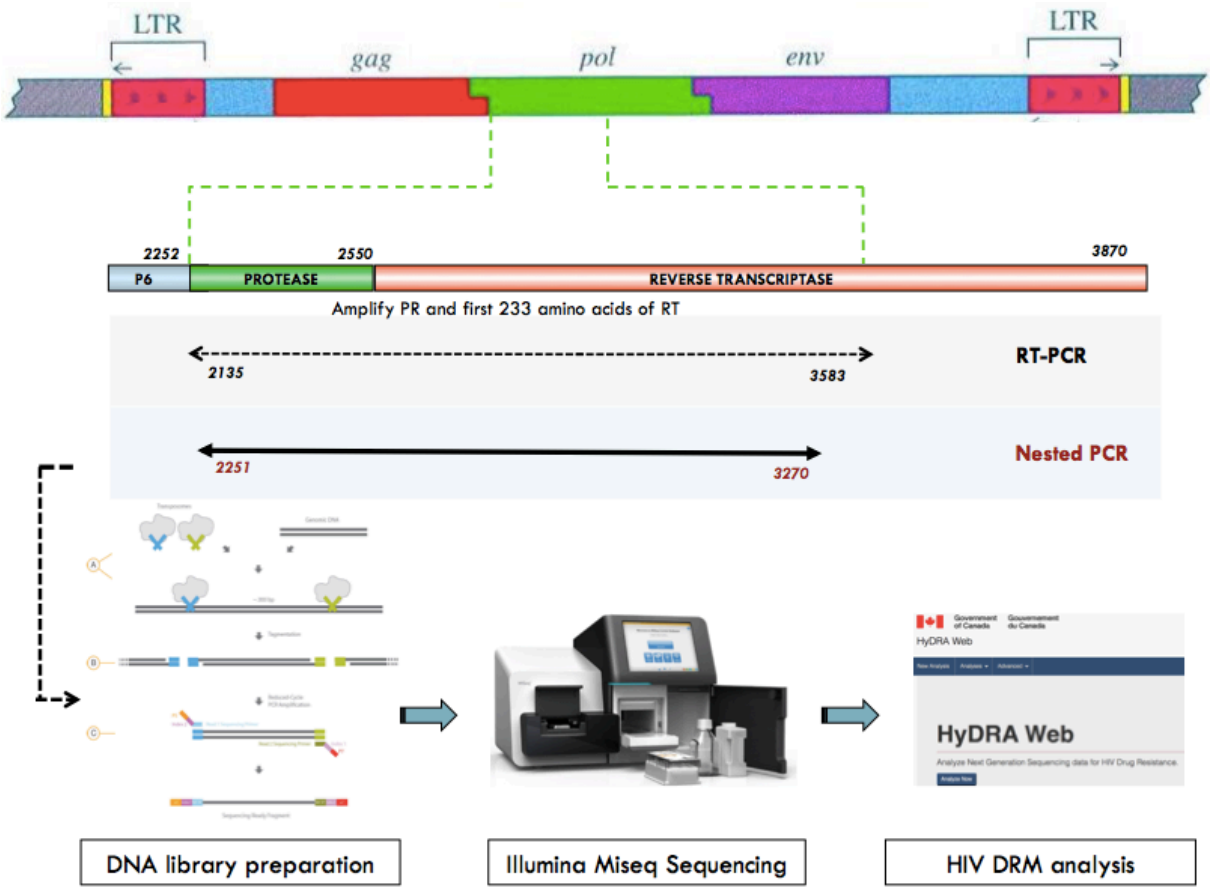


Figure 1: Schema of low-level viremia sample processing. Total RNA was extracted, followed by two-step amplification of protease and first 233 amino acids of reverse transcriptase region of *pol* gene. RT-PCR was performed followed by nested PCR generating an amplicon covering the region of interest. Purified and quantified amplicons undergo library preparation using Illumina Nextera XT DNA Library kit and then are clonally amplified and sequenced using Illumina Miseq platform. Using proprietary HyDRA pipeline from National Microbiology Laboratory, PHAC, all the quality controlled reads were used for HIV drug resistant mutation (DRM) identification.

2.5 Detection of drug resistant mutations in LLV population

The amplified clinical samples from children with LLV in Nairobi were genotyped for examining the DRMs in the target HIV-1 genes. Prior to Illumina MiSeq sequencing, the library preparation was done based on the Nextera XT protocol as described by the manufacturer. The library preparation kit uses an engineered transposome to simultaneously fragment and tagment DNA. Limited-cycle PCR reaction uses the interest-specific primers in ligated adapter sequences to amplify the DNA samples and also adds index sequences on both ends of the DNA to enable dual-indexed sequencing. The libraries are quantified using a fluorometric quantification method, normalized and then pooled in equimolar concentrations for MiSeq sequencing. Small sample size after tagmentation step or low sample concentration during DNA normalization step can cause fluctuations in quality metrics for preparing DNA libraries for sequencing [135] .

Data generated from MiSeq sequencing is examined for minor variant analysis. The Bioinformatics Core and the National Laboratory for HIV Genetics (NLHG) at Public Health Agency of Canada (Winnipeg, MB) have jointly developed a web-based HIV-1 drug resistant variant analysis (HyDRA) pipeline for this purpose. HyDRA is an annotated reference-based bioinformatics pipeline for NGS-based HIV DR typing (Figure 2). It reliably detects mutations at frequencies as low as 1% in HIV-1 quasispecies using fully validated NGS read quality control, reference mapping and variant calling strategies. An annotated HXB2 sequence (GenBank Accession number: K03455) is used for reference mapping. The main output is a report on HIV-1 DRMs by referring to the Stanford SDRM 2009 list for the major HIV DRMs and the general Stanford HIV DR

database for minor and other reportable HIV mutation [79] . Along with the names and loci of these mutations, it reports the frequencies of all identified mutations and the sequencing coverage at the specific loci.

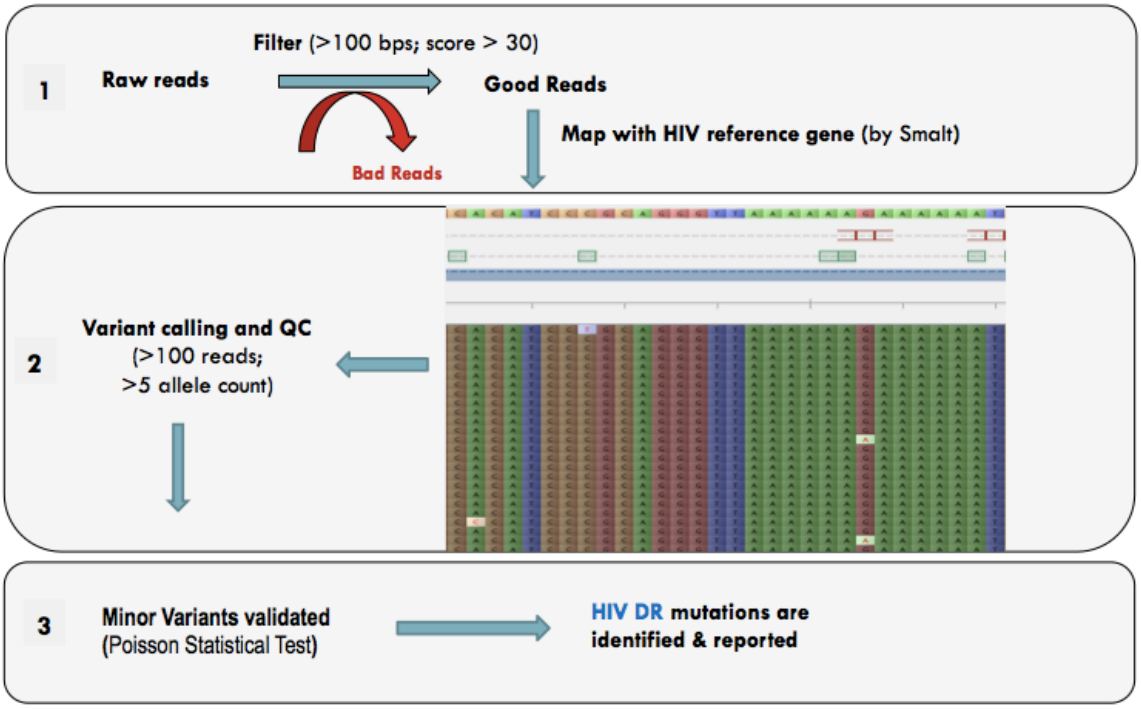


Figure 2: Workflow overview for Illumina MiSeq based HIV drug resistant analysis (HyDRA). The input raw reads are first subjected to quality control (QC). Low quality reads are filtered out via QC. Secondly, pairwise alignment of the good reads is done by Smalt, a tool that maps the reads to a reference gene. Lastly, HIV drug resistant mutations are identified followed by confidence testing with Poisson distribution model. Output report provides information on HIV drug resistant mutations, coverage and alignments.

Default parameters were used to analyze the data (Table A1) for HIV DRMs identification. The read length cutoff is 100 nucleotides so any reads that fall short of this length is filtered out as it is generally more error-prone and difficult to align. The quality control reads are mapped against HXB2 reference HIV genome using Smalt [136] for mapping reads to the reference, variants in mapped regions exceeding 1000x coverage are identified and their probability of being a true variant versus a sequencing error is estimated using a cumulative Poisson distribution-based calculation. The high probability variants are searched against a DRM database (Stanford HIV DR database and SDRM 2009). Finally, the high probability DRMs with >1% abundance are reported in the form of a CSV table including information on the gene location of the mutations (PR, RT, IN), the classification and surveillance status of the mutations, the position and identity of wild type and mutant amino acids, the variant frequency and depth of coverage at those loci.

2.6 Statistical analysis of the risk factors

2.6.1 Database construction and organization

2.6.1.1 Data sanitizing

Clinical, epidemiological and social data was initially recorded by data officers assigned to each of the eight participating clinics in a paper format and/or electronically with no standard format. The electronic data was collected by Mr. Joseph Mbugua Kabogo, a graduate student under the supervision of Dr. Elijah Maritim Songok at the Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi. A data dictionary defining each variable was made to understand social, clinical and

epidemiological variables (~140 total) in the Leo Toto Program database. The variables were assigned as dependent or independent variables, categorical or continuous, and existing or derived variables (such as calculating age from DOB). Each variable was then divided into categories from the most to the least importance according to their relevance to the study. A flow diagram showing association between variables that have a direct and indirect impact on virologic failure was then prepared from the categories (Figure A1). A few participants from different clinics had the same admission number, hence a unique key was made to identify any patient in the cohort with the clinic name in addition to the client admission number.

The data was merged in a standardized format from all the eight clinics after the assembly of significant variables (27 existing variables) required for this study. These variables included: client admission number, clinic, gender, date of birth (DOB), date of admission (DOA), viral load, CD4 count, ART regimen (AZT, d4T, ETC, NVP, EFV, DDL, ABC, Kaletra, TDF) and therapy stage (i.e. 1^{st} line, 2^{nd} line or salvage). A few variables were derived from the database such as age at the start of therapy, duration of therapy, number of blood tests, status of LLV or VF, time to start ART and time to reach VF or LLV (Table A2).

Patient forms in the hard copy were accessed later in order to accumulate information on some of the important variables that were initially missing from the electronic database. The information was then manually transferred from the scanned sheets (~2500 files) to the electronic standardized database for each patient's adherence, WHO stage at the start of therapy, ART start date, ART switch/stop date and co-morbidities (pneumonia, TB, malaria, malnutrition) information. Use of optical character

recognition (OCR) testing or any coding would not have worked on the scanned files as each file had its own format and no proper uniformity was maintained in recording the data among different clinics. Moreover, OCR is not ideal for the typical volume of the data used in this study.

The database was sanitized prior to downstream analysis. A major problem encountered throughout the database was duplicates for each entry. Generally, each variable during periodic blood tests was generally recorded more than once for all the patients. Many of these entries had typing errors for variables such as viral load, CD4 count or blood test dates. All the duplicates were removed and the typing errors were fixed in order to keep the database clean and to avoid misinterpretation during data analysis. The typing errors were confirmed by checking the duplicates and from prior knowledge of clinic regulations (for example: duration between the blood tests). The date format was different in most of the date-related variables including DOB, DOA, ART start and stop dates so it was uniformly changed to a mm/dd/yyyy format.

2.6.1.2 Patient selection

All patients lacking information on ART start dates were excluded from the final database. The start dates are important to analyze the duration of therapy and to allow appropriate judgment of whether the failure was due to drug regimen, drug resistance or adherence issues. Patients without ART regimen and viral load information were also excluded, as it would have obstructed tracing the mutations back to the drug classes in LLV samples. Some patients had an odd start date; meaning patients did not actually start the therapy that matched their ART start date as stated. In fact, some patients had their

first blood test date as late as eight years after the recorded start date. To accommodate patients in a similar therapy stage during their treatment, patients who did not start their therapy within two years of the stated date were also excluded. Patients were also excluded if ART was interrupted/stopped for more than 15 months during the course of their treatment [66,125,137] .

For a patient to be included in the final database, they must be receiving ART and should have a start date. They should have received ART for at least six months overall, as it takes about 3-6 months for ART to have an effect on the host immune system. The data however was not analyzed starting from the sixth month but was analyzed from baseline to record the overall VL pattern. Each patient must have more than two blood test time points to allow for analysis of their VL pattern. If a patient had only two time points then it should span over ≥ 3 months. Many patients did not follow the standard six-month follow-up routine, and instead had routine tests in less amount of time (<3 months), which does not provide long time interval to accurately analyze the clinical impact of the therapy. Therefore, the time span of three months was accounted for and patients were removed if they had multiple follow-ups in less than three months, or only had one-time point as it can lead to gaps in understanding the effect of the regimen on VL and DRMs.

2.6.1.3 Definitions

Each patient was stratified into categories based on their VL pattern. Previous literature and WHO guidelines were followed while defining variables with very few exceptions made according to the data available.

For this study, the viral load cut-off for persistent low-level viremia was defined as two consecutive viral loads between 50-1000 copies/ml over one or more months [61,62,107] . The duration of having the viral load typically varies among papers and also highly depends on the type of data available. Due to inconsistency in time periods between blood tests among patients, duration of more than one month was chosen to accommodate patients with variable time intervals. A patient is considered to have a blip when there is only one episode of VL between 50-1000 copies/ml, preceded and followed by an undetectable (<50 copies/ml) viral load [64-67,118]. When a patient presents viral load suppression before and after the blip, the efficacy of regimen is ensured along with the confirmation of a "true blip". An exception was made for patients to be included in the blips category if they had their VL between 50-1000 copies/ml within a year after starting ART, but was not preceded by an undetectable VL. This is due to a possibility of missing a blood test or data entry. All the patients that passed the inclusion/exclusion criteria but did not belong to either pLLV or blips categories were selected as a control group. A true control group would have been patients who do not receive ART but this information was not available as every patient in this cohort received ART. Many researchers have included patients who maintain suppression throughout their ART as their control group, but this introduces bias into downstream analysis and does not represent the true population.

All the patients with an LLV time-point during the course of their therapy were observed for their virological outcome. A patient who had two consecutive VL >1000 copies/ml, after previously attaining at least one undetectable VL during their therapy, was deduced to have a virologic failure (VF) [61,66]. In the case of VF, the first event

was recorded as the time to reach VF for analyses; otherwise the patient was followed till the last blood test available in the database if they did not have VF. Having a suppression episode during the therapy validates the failure and confirms it is not due to high baseline VL or regimen. Defining VF with a lack of consecutive VL events is apparent in the literature reviews [65,108,109,117] but it is important to avoid "false VF" or viral rebounds. A patient was considered to have viral rebound when they maintain a suppressed VL throughout the therapy but had a single VL >1000 copies/ml [138]. These most commonly occur due to poor adherence, and since the database lacks longitudinal adherence information, it was necessary to consider these patients as mere rebounds and not as virological failure cases.

2.6.2 Risk factors of low-level viremia & virologic failure

Apart from assessing DRMs as a risk factor of LLV, other relevant factors were also investigated for their association with LLV and VF. A total of 293 patients were selected for downstream statistical analyses to investigate whether children with HIV-1 infection having LLV has an association with VF. Potential risk factors including adherence, WHO clinical stage, CD4 counts, gender, therapy stage, ART drugs with recommended regimen, and co-morbidities that can correlate to VF and LLV were assessed. All the patients were stratified into pLLV, blips or control categories with VF or no failure outcomes. The impact of these social or clinical risk factors on pLLV was also analyzed. Standard parametric tests such as chi-square (χ^2) and standard survival analysis with Kaplan-Meier curves were used for univariate analysis of time-dependent data for some variables. Cox proportional hazards regression model was used to calculate

the hazard ratio (HR) with 95% confidence interval for studying the dependence of survival time on predictor variables. The false discovery rate (FDR) approach was used to adjust the proportion of false positives among the multiple factors tested for co-morbidities and ART drugs.

HIV-infected children on ART were counseled for their HIV status and adherence to the treatment was monitored via random pill count. Optimal adherence to ART was defined as the child or adolescent ingesting 95-100% of the pills, as per random pill counts by health practitioners at the LTP Centre, and suboptimal adherence to ART indicated less than 95% of the pills were ingested. Data provided on adherence information was further categorized into good (optimal) and poor (suboptimal and poor) groups to study the effect of adherence on LLV and VF. The four WHO clinical stages were also divided into mild (stage 1 and 2) and severe categories (stage 3 and 4). To evaluate the risk of an immunologically HIV-1 infected and uninfected individual leading to LLV and VF, children were divided into groups according to their CD4 cell counts of <500 cells/mm^3, 500-1200 cells/mm^3 and 1200 cells/mm^3. All nine ART drugs were investigated individually as well as with the recommended regimen for first and second-line treatment.

Chapter 3: Results

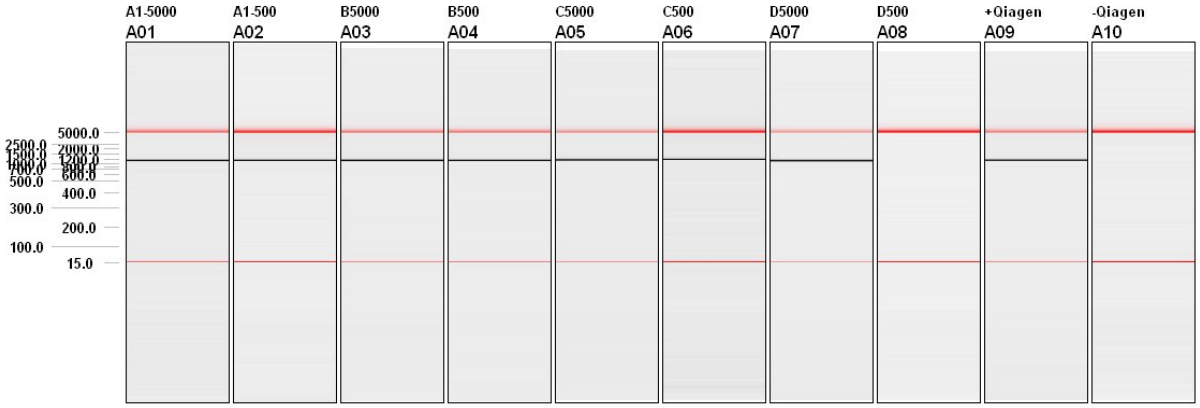
3.1 Developing the LLV amplification protocol

3.1.1 Automated versus manual extraction

Genome amplification of a target HIV gene fragment is a widely used method, but using it for samples with low viral load in a resource limited setting is not well established with no detailed protocols available for such samples. The objective of such protocol development was to amplify low copy number of HIV RNA obtained from HIV-infected children with equipments available at KMRI lab, Nairobi, Kenya, where the lab work was to be performed. The lab protocol routinely used for HIV drug resistance surveillance testing in the National HIV & Retrovirology Laboratories at JC Wilt, Infectious Diseases Research Centre, NML, was used initially for optimization in order to enhance its sensitivity over different HIV-1 subtypes prominent in Kenya. Due to limited resource available for extraction in Nairobi, a manual Qiagen RNA extraction method was also initially compared for its efficacy with the automated NucliSENS easyMag extraction protocol (Figure 3) before proceeding with further optimization. The two RNA extraction methods were found to work equally well on EQAPOL control samples. Subsequently, different input (200, 300, 400, 500 μl) and output (50 or 60 μl) volumes were tested from which 400 μl of maximized plasma input with 60 μl of elution volume was found to provide the most reliable results for LLV samples. Accurun (SeraCare, USA), whole cell or whole organism external control, was used as a positive control and NHP (SeraCare, USA) was used as a negative control. By comparing results from different trials among lab members, it was later discovered that Accurun is not a strong

control and has shown to not work well in general. Thus, previously extracted RNA from high viral loads was used as positive controls from downstream procedures.

A



B

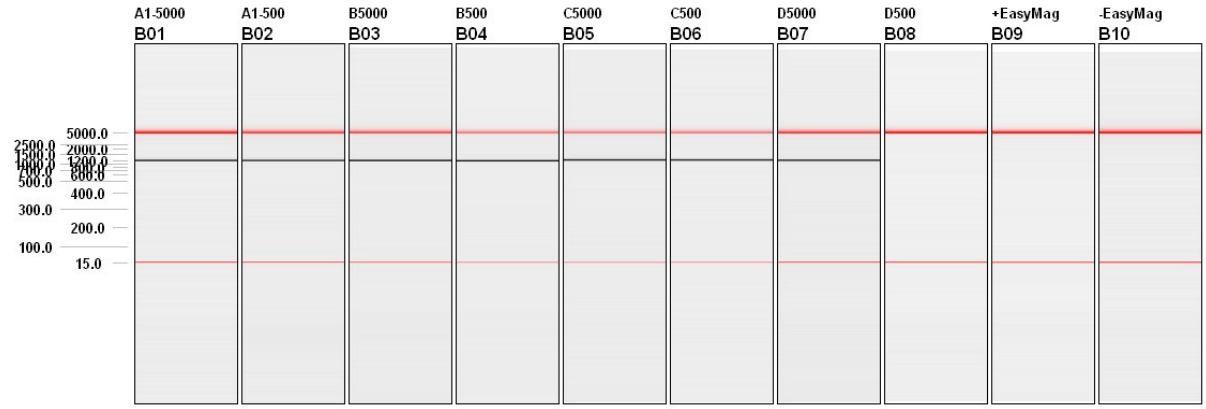


Figure 3: Comparison between manual and automated HIV RNA extraction methods. This image shows the Qiaxcel (automated gel electrophoresis) results of samples amplified using (A) the standard protocol with Qiagen manual extraction and (B) NucliSENS easyMag automated extraction. 1 kb DNA ladder is shown on the left with an expected band at 1100 bp. The red lines represent alignment marker for each well. Subtypes A1, B, C and D were diluted to 5000 copies/ml and 500 copies/ml to measure sensitivity at a high and a low range viral load, as indicated above each well. The columns A09 and B09 represents Accurun (positive control) while columns A10 and B10 represents NHP (negative control) for each method used.

3.1.2 One-step versus two-step RT-PCR

HIV is an RNA virus, thus the first step in preparing a sequence library involves reverse transcription of the HIV RNA genome to complimentary DNA (cDNA), which could then be PCR amplified. This can be carried by one-step RT-PCR, which combines the reverse transcription (cDNA preparation) and the initial PCR amplification of a specific target gene fragment in one reaction, or two-step PCR in which the cDNA preparation and subsequent PCR amplification are carried out separately. One-step method is convenient as both the RT and PCR reactions occur in a single tube making it less labor intensive. On the other hand, two-step protocol is more flexible as one RT reaction can provide templates for multiple PCR reactions. One-step RT-PCR is less expensive as only one kit is needed for the whole procedure whereas two-step RT-PCR can be more sensitive. Both methods were evaluated for their ability to amplify low copy number HIV genomic template.

Two-step RT-PCR did not amplify any of the four subtypes at any viral load dilution from 1000 copies/ml down to 50 copies/ml. Only the positive controls, which were previously extracted samples with consistent amplification among multiple trials, were found to have amplified. Generally, the one-step approach worked fairly well except for subtype D that fluctuated during every individual trial. Different RNA templates volumes of 10.0 μl, 12.0 μl and 14.0 μl were tested to determine the best concentration and positive amplification, from which 12.0 μl template input produced the best results.

3.1.3 One-amplicon versus two-amplicon N-PCR

The aim of nested PCR is to reduce product contamination due to the amplification of unintended primer binding sites and amplify the target gene(s). Typically, a larger outer fragment is unambiguously amplified with RT-PCR from which a smaller inner fragment is subsequently amplified via nested PCR to yield target amplification with minimal contamination. This can be performed with one-amplicon, where the entire target region is covered with a single PCR amplicon, or by two-amplicon approach, where the target region to be amplified is covered by two smaller overlapping amplicons. Reagent preparation for one-amplicon N-PCR was fast and easy as only one master mix is needed, which also reduces pipetting errors. Two-amplicon N-PCR requires preparation of two separate master-mixes to amplify the protease and reverse transcriptase region and thus uses more reagents along with two separate sets of primers. While one-amplicon nested PCR was less labor-intensive, it is more expensive as compared to the two-amplicon nested PCR approach. Having adaptor sequences ligated with the primers in the two-amplicon approach saves time and expense during library preparation for Nextera-based MiSeq sequencing as it omits the segmentation step and allows amplification by PCR in subsequent steps. The intent here was to compare the sensitivity of the two methods with the standard protocol for the more sensitive approach.

The two-amplicon approach did not work as well as the one-amplicon approach at low viral loads. For approximately 60% of samples, either the PR or the RT gene was amplified from using two-amplicon protocol, but both the genomic regions rarely amplified for all the subtypes (Table 3). In summary, one-amplicon N-PCR approach with 4 μl DNA template proved to be the most sensitive and reliable method.

Table 3: Comparison of one-amplicon and two-amplicon nested PCR amplification. The top row shows the HIV-1 subtype with its respective diluted viral load for each approach. The columns on the left show whether it is a one-amplicon or a two-amplicon approach, where two separate gels are run for each genomic region, protease and reverse transcriptase, amplified for the same sample. "+" represents the sample had amplified while "-" represents no band was seen after PCR amplification.

Nested PCR		A1-5000	A1-500	B-5000	B-500	C-5000	C-500	D-5000	D-500
One amplicon		+	+	+	+	+	+	+	-
Two amplicon	Protease	+	+	-	-	+	+	+	-
	Reverse-transcriptase	+	-	+	+	+	-	+	-

3.1.4 Standard versus optimized protocol

The protocol from National HIV & Retrovirology lab at JC Wilt National HIV & Retrovirology Laboratory at National Microbiology Laboratory (Winnipeg, MB) was modified and further optimized with the prime objective of amplifying low-level viremia plasma samples from a cohort of children in Nairobi. The amplification and sequencing was to be performed in Nairobi, thus, kits/methods available for use in Nairobi were chosen. Qiagen manual extraction was followed by one-step RT-PCR and one-amplicon nested PCR were the most accessible and effective methods. Each step of the protocol (extraction, RT-PCR, nested PCR) was further optimized to allow amplification at low viral loads. The aim was to allow amplification of samples with the lower limit of 50 copies/ml in EQAPOL HIV-1 subtypes A1, B, C and D.

Initially, the standard protocol included 200 µl plasma with 60 µl eluent using Qiagen manual extraction, 10 µl RNA template and 2 µl DNA template with 5 µM primer concentrations. After experimenting with different approaches for PCR's (1 amplicon vs. 2 amplicon), different enzyme and primer concentrations as well as plasma or RNA/DNA input volumes; the finalized protocol included 400 µl plasma input with 60 µl eluent using Qiagen manual extraction kit, 12 µl RNA template and 4 µl DNA template with 200 µM primer concentrations. The average volume of plasma samples available from the children cohort in Nairobi was 500 µl, thus the protocol was optimized to be able to test Nairobi-based LLV samples.

3.1.5 Validation results

3.1.5.1 EQAPOL samples

With the optimized protocol, 400 µl of plasma was used for RNA extraction and eluted in 60 µl volume. With an assumption of 80% extraction efficiency, the HIV RNA input is estimated at ~3 copies per RT-PCR reaction if the initial plasma VL is at 50 copies/ml. Even with such low viral template input, the protocol reliably amplifies HIV-1 subtypes A1 and B at VL as low as 50 copies/ml, although the consistency decreased below 100 copies/ml. It also performs well consistently on subtype C when the VL is ≥500 copies/ml, but with lower consistency when VL is at 100 copies/ml. The sensitivity of this protocol is determined at 1000 copies/ml for subtype D, but fluctuated for VL below this concentration (Table 4).

Table 4: EQAPOL sample amplification using optimized protocol. The efficiency of the developed PCR-based protocol to amplify HIV-1 subtypes A1, B, C and D was validated at various viral loads with the lower limit of 50 copies/ml. First row (in bold numbers) represents the diluted viral load from 5000 to 50 copies/ml. The second row indicates the theoretical number of input copies per RT-PCR assuming an 80% RNA extraction efficiency.

Subtypes	Sensitivity by VL (copies/mL) *						
	5000	**2500**	**1000**	**500**	**200**	**100**	**50**
	320	160	64	32	12.8	6.4	3.2
A1	+++	+++	+++	+++	+++	+++	+--
B	+++	+++	+++	+++	+++	+++	--+
C	+++	+++	+++	+++	+--	-+-	---
D	+++	+++	+++	---	+-+	---	-+-

* *Each +/- sign represents an individual run of the same sample type, run in triplicates.*

3.1.5.2 Children clinical samples

49 children plasma samples were available due to an initial loss of biological material from ~100 longitudinal samples, from these 36 children LLV plasma samples and 13 baseline samples were subjected to RNA extraction. The optimized protocol successfully amplified 56% LLV samples (20/36) and 85% baseline samples (11/13), yielding a 63% success rate (31/49). All the negative and positive controls performed as expected.

To amplify the samples initially failing PCR, several troubleshooting attempts were made that included changing primer or dNTP concentration, reducing $MgSO_4$ concentration to half, adjusting PCR cycling parameters as well as different agarose concentration for gel electrophoresis. None of the attempts were successful in amplifying additional plasma samples with the same mRNA target. Even the positive controls that had earlier amplified were also found to be failing with these changes.

3.2 Sequencing analysis of DRMs in LLV population

The purpose of this study was to see if the patients with LLV have a higher occurrence of DRMs and if the presence of DRMs in samples with LLV has an association with VF. All the 31 PCR amplicons successfully obtained from the LLV samples were further processed using Illumina Nextera XT DNA library kit and eventually sequenced using Illumina Miseq platform. Among them, 13 were either lacking biological material necessary to confirm resistance by sequencing (n=7), or had failed the laboratory's quality control for library preparation (n=6), permitting sequencing on 58% (18/31) samples. By the time the sequencing machine was prepped for the

sequencing run (~4 months), 9/18 samples that had previously passed quality control had reduced concentration rendering them unfit for sequencing. Hence, 9 samples were processed through MiSeq sequencing that was comprised of 3 baseline samples and 6 LLV samples (Figure 4). The range of VL among these 6 LLV samples was between 169 copies/ml and 926 copies/ml.

All but one sample had enough query data to pass the quality control setup by HyDRA pipeline to detect DRMs, leaving 5 LLV and 3 baseline samples with complete data. As expected, children were infected by the virus resistant to two or more drug classes among pre-treated individuals. Drug resistant mutations were identified in 1/3 (33%) baseline samples and 1/5 (20%) LLV samples (Figure 4). This included one of the two patients with persistent LLV suggesting an association of increased DRMs in pLLV, but the small sample size does not allow confirmation of this observation (Table 5).

Most of the mutations did not associate with the ART regimen of the patient. The drug regimen for the patient having pLLV was 3TC/ABC/Kaletra, providing evidence that the mutation D67E is associated with the drug ABC. The ART regimen for the patients having mutation in the baseline specimen was on 3TC/NVP/ABC regimen, indicating that the mutation D67G may emerge in the presence of the drug ABC. Other mutations did not match the regimen of each patient indicating it is likely not caused by the regimen the patient was on at the time of sample collection and is possibly from baseline (transmitted) resistance. Moreover, both patients with DRMs did not proceed to virologic failure during the course of their therapy, implying DRMs do not increase the risk of VF in patients with pLLV. Nonetheless, the findings presented here rely on a very low sample size and thus requires further investigation to confirm the conclusions.

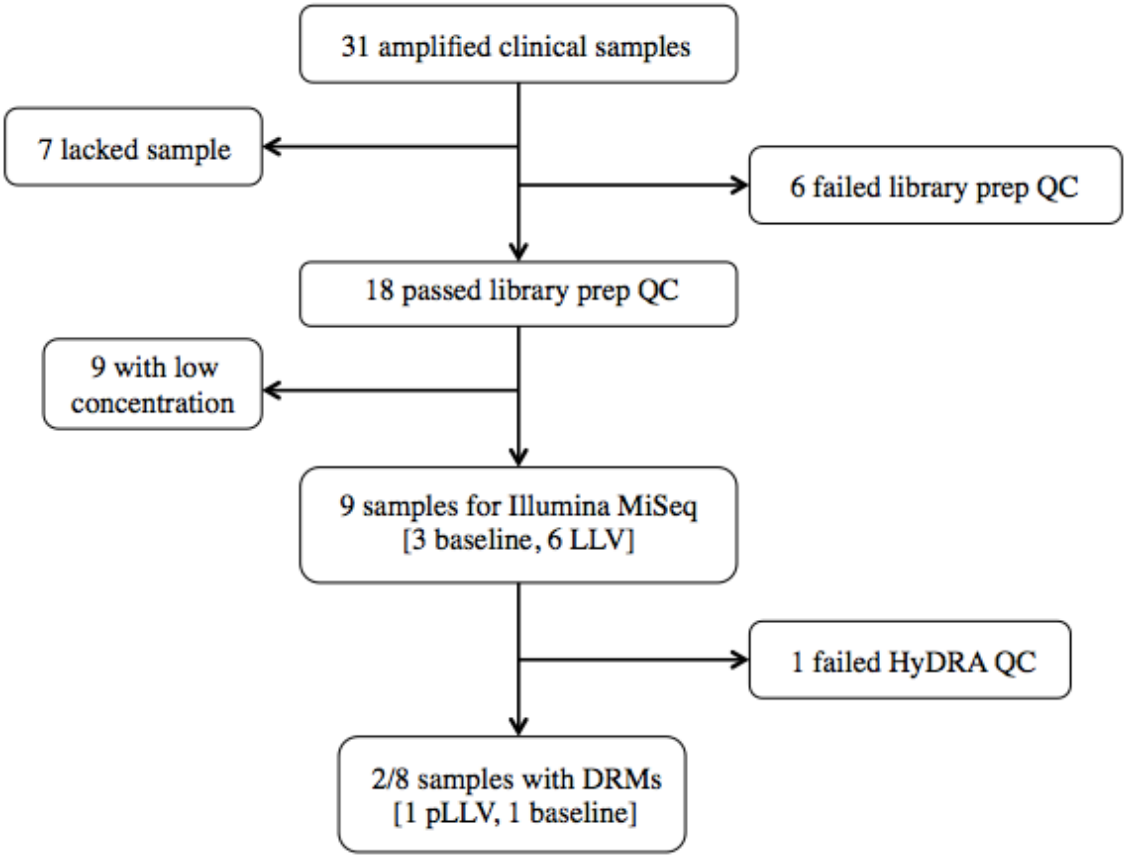


Figure 4: NGS-based HIV-1 DR typing.

Table 5: Drug resistant mutations detected by Illumina MiSeq sequencing. Prevalence of mutations in protease and reverse transcriptase genomic regions associated with drug resistance from different drug classes in the paediatric HIV-1 infected population. Drug resistant mutations were identified in children with persistent low level viremia and baseline sample specimens using HyDRA pipeline.

Sample Type	Gene	Classification	Mutation	Frequency (%)
pLLV	RT	NRTI	D67E	2.20
	PR	Other	L89M	96.30
Baseline	RT	NRTI	D67G	2.23
	PR	PI Major	L76V	4.43
	PR	PI Major	N88S	1.04
	PR	PI Major	I47V	1.07
	PR	PI Minor	F53L	3.49
	PR	PI Minor	M46V	1.13
	PR	PI Major	I54V	1.21
	PR	Other	I85V	1.26
	PR	Other	L89M	94.97

3.3 Statistical analysis of the risk factors for LLV and VF

Approximately 2800 children with HIV-1 infection were enrolled in the LTP cohort receiving ART. Overall, 293 children met the inclusion criteria for this study with the date of HIV diagnosis varying from 1991-2011. For this analysis, the database was sanitized in multiple stages and patients were filtered from exclusion criteria, performed both manually and via SAS software. The majority of the patients were removed due to incomplete clinical information in the database, leaving 654 patients. 17 patients were excluded due to lack of two blood-test (blood drawn from patients to test for VL and CD4 counts) time-points in ≥3 months, 158 were excluded because they did not initiate ART within the first-two years of the date stated in the records, 132 had interruption in their therapy for more than 15 months, and 54 participants were excluded due to an ART treatment duration lasting fewer than six months (Figure 5). These 293 remaining children were further stratified based on the definitions (pLLV, blips, control group) or outcomes (VF, adherence, co-morbidities, other risk factors) for analysis.

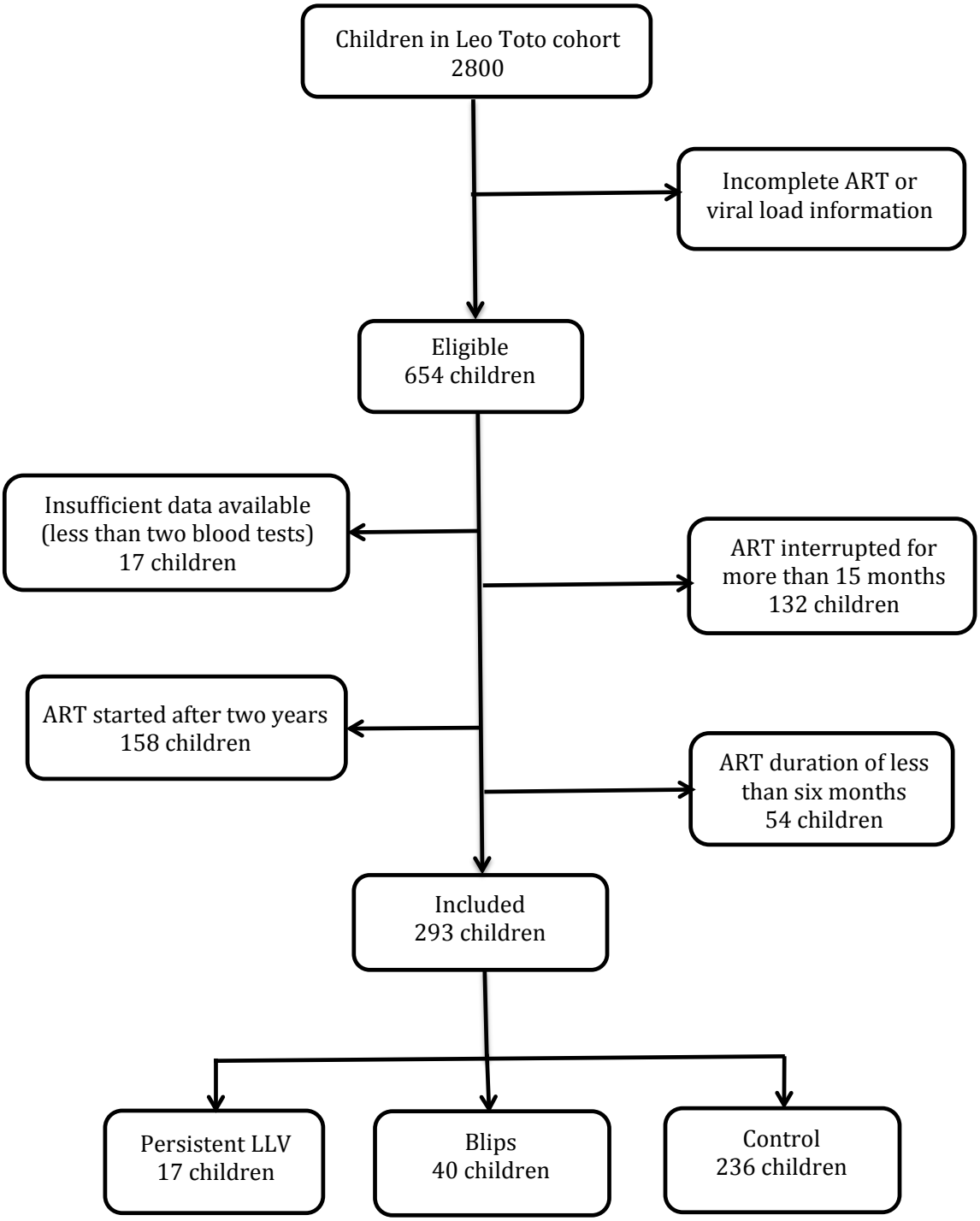


Figure 5: Flow chart for patient selection.

The study population with the required criteria belongs to the clinics enrolled in LTP except for Zimmerman and Mukuru clinics. As the route of infection for these children is via vertical transmission, the majority of the children initiate therapy in late years of their life instead of at birth. Although the HIV-infected pregnant mothers should have been on ART to avoid this transmission, some refused to take antiretrovirals while others insisted on breastfeeding. Many had refused to enroll their children in an ART program due to cultural, religious or financial complications.

Amongst the 293 selected patients, there was no gender bias in the cohort as the ratio of males and females was similar. 263 children were receiving first-line ART regimen with only three reported cases of second line regimen. A few patients (n=2) had stopped ART due to unknown reasons. Most of the clinical and epidemiological features of the children were recorded upon enrollment to the program (Table 6). A summary of the key clinical characteristics recorded during the course of their therapy on an average of 6 months interval is presented in Table 7. The mean age of the study participants at the time of admission to the clinic was 6.40 years, making this one of the youngest pediatric cohorts in Nairobi. Nearly two-thirds presented WHO clinical stages II and III of HIV/AIDS (Table 6). Children were enrolled in the program with baseline viral load (before therapy) and were started on therapy as the symptoms developed, exhibiting 1017 cells/mm^3 as the mean CD4 T-cell count. To further stratify the LLV categories by low and high VL, 6 pLLV patients and 15 blips patients had their VL between 50-500 copies/ml, and 11 patients with pLLV and 25 patients with blip exhibited VL between 500-1000 copies/ml. This represents 40 (13.65%) patients with blips and 17 (5.8%) patients with pLLV, providing a total of 57 (~20%) people with LLV.

Table 6: Baseline features of HIV-infected children. The clinical or epidemiological features for selected study population were recorded at the time of enrollment in the Leo Toto Program.

Features	**Patients (n=293)**
Clinic	
Dagoretti	26
Dandora	1
Kangemi	45
Kariobangi	1
Kawangware	97
Kibera	123
Sex	
Male	152
Female	141
Age	
0-9 years	219
10-17 years	74
Therapy Stage	
1st line	263
2nd line	3
Unknown	27
Adherence	
Optimal	76
Suboptimal	9
Poor	40
Unknown	168
WHO stage	
Clinical stage I	1
Clinical stage II	89
Clinical stage III	74
Clinical stage IV	3
Unknown	126
Pneumonia	
Present	106
Absent	145
Unknown	42
Tuberculosis	
Present	67
Absent	43
Unknown	183
Malaria	
Present	43
Absent	177
Unknown	73
Malnutrition	
Present	49
Absent	166
Unknown	78

Table 7: Characteristics of the study population

Variable	**Mean (min-max) (years)**
Age	6.40 (0-17)
Clinical follow-up	0.50 (0.17-1.24)
Data collection for this study	2010 (2007-2013)
Persistent low level viremia duration	0.89 (0.17-2.76)
Duration of antiretroviral therapy	3.01 (1.00-5.36)
Time to reach persistent low level viremia	1.63 (0.04-4.48)
Time to reach blips	1.77 (0.00-5.62)
Time to reach virological failure	2.21 (0.21-4.86)

For an in-depth understanding of an association of LLV with VF, the two LLV categories (pLLV and blips) were separately analyzed. Among the 293 children, 21.8% (n=64) of the patients had experienced virologic failure. 5.8% had pLLV and 13.7% had blips adding to 20% of the patients with LLV episodes. Comparing the VF column with the three study groups (pLLV, blips, control) in Table 8, VF is not significantly associated with blips or pLLV study groups. Kaplan-Meier survival curve was used for further analysis to compare the rate at which LLV categories (pLLV and blips) reached the VF stage. Patients with pLLV were observed to have a higher probability of reaching VF compared to blips. The Cox proportional hazard ratio (HR) of 1.082 represents the chances of patients with pLLV is 1 times higher than the patients with blips, which, along with the trend seen in survival curve is not significant. The observed trend cannot be supported as the sample size of the two LLV study groups with VF stage is too small and is not significant to draw conclusion of any appreciable differences among the groups (Figure 6).

Table 8: Association of LLV with VF. The number of individuals with pLLV, blips or control study groups having VF is shown, along with the percentage of each group belonging to their respective outcomes. The significance is indicated as P value less than 0.05. pLLV: persistent low level viremia; ART: antiretroviral therapy.

	ART Success	Virologic Failure	Total
pLLV	15 (6.6%)	2 (3.1%)	17 (5.8%)
Blips	36 (15.7%)	4 (6.3%)	40 (13.7%)
Control	178 (77.7%)	58 (90.6%)	236 (80.5%)
Total	229 (78.2%)	64 (21.8%)	293 (100%)
$\chi^2 = 5.33$, p-value = 0.0696			

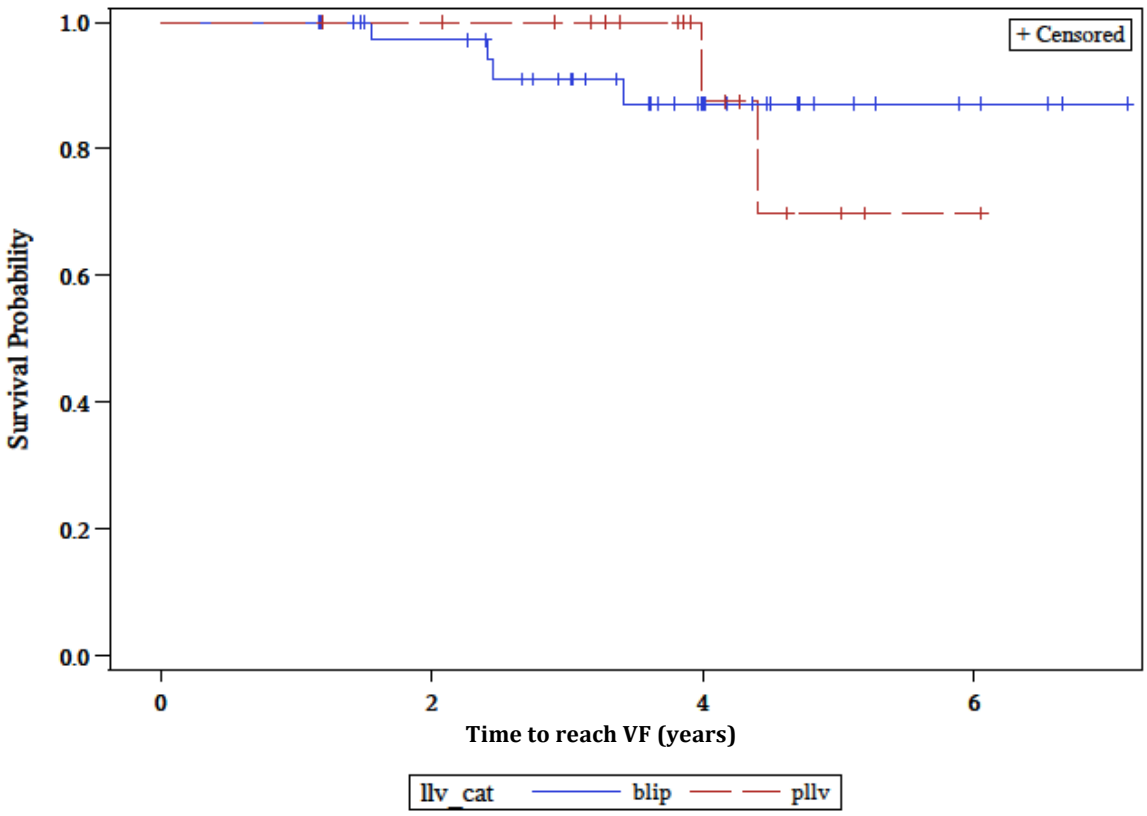


Figure 6: Kaplan-Meier survival analysis comparing individuals with blips and pLLV. The red-dashed line represents patients with pLLV while the blue line represents patients with blips. Individuals with no VF episode during their treatment were censored. pLLV: persistent low level viremia; VF: virological failure.; llv_cat: LLV category [HR = 1.082 (0.214 - 5.476), p = 0.9243]

To gain an insight of the factors that potentially play a role in viral failure, the study population of 293 patients was analyzed for correlation with various confounding variables in HIV-infected individuals with ART treatment. VF and LLV were compared using a correlation analysis of patients with good (optimal) and poor (suboptimal and poor) adherence. 76/125 (61%) patients reported more than 95% of ART pills intake (good adherence) while 49/125 (39%) reported missed doses (poor and suboptimal adherence) for the duration of their treatment. Adherence data was missing for 168 patients. Good (optimal) ART adherence significantly reduces the risk of reaching VF stage in HIV-infected children ($p<0.0001$). On the contrary, the impact of adherence was not statistically significant for LLV (pLLV plus blips) (Table 9).

Table 9: Impact of adherence on VF and LLV. The number of individuals with good (optimal) or poor (suboptimal and poor) adherence correlating with (A) virologic failure and (B) low-level viremia, with the percentage of each group belonging to their respective outcomes. The significance is indicated as P value less than 0.05. ART: antiretroviral therapy; LLV: low-level viremia.

(A)

	ART Success	Virologic Failure	Total
Good	65 (72%)	11 (31%)	76 (61%)
Poor	25 (28%)	24 (69%)	49 (39%)
Total	90 (72%)	35 (28%)	125 (100%)
$\chi^2 = 17.59$, $p < 0.0001$			

(B)

	Control	LLV	Total
Good	56 (58%)	20 (71%)	76 (61%)
Poor	41 (42%)	8 (29%)	49 (39%)
Total	97 (78%)	28 (22%)	125 (100%)
$\chi^2 = 1.71$, $p = 0.19$			

Patients with mild and severe WHO clinical stages at baseline did not have a significant association with VF or LLV (Table 10). At least 55% of patients belonging to the severe WHO staging system had VF and 42% had LLV. WHO clinical stage information was unavailable for 126 patients.

Table 10: Impact of WHO clinical stage on VF & LLV. The number of individuals from mild or severe WHO clinical stage suffering from (A) virologic failure and (B) low-level viremia, with the percentage of each group belonging to their respective outcomes. ART: antiretroviral therapy; LLV: low-level viremia.

(A)

	ART Success	Virologic Failure	Total
Severe (stage 3, 4)	54 (43%)	23 (55%)	77 (46%)
Mild (stage 1, 2)	71 (57%)	19 (45%)	90 (54%)
Total	125 (75%)	42 (25%)	167 (100%)
$\chi^2 = 1.69$, $p = 0.19$			

(B)

	Control	LLV	Total
Severe (stage 3, 4)	63 (47%)	14 (42%)	77 (46%)
Mild (stage 1, 2)	71 (53%)	19 (58%)	90 (54%)
Total	134 (80%)	33 (20%)	167 (100%)
$\chi^2 = 0.22$, $p = 0.63$			

For CD4 T-cell counts, association studies were performed between VF or LLV and CD4 counts <500 cells/mm^3, or 500-1200 cells/mm^3, or >1200 cells/mm^3. A significant increase in the risk of VF was found in patients with CD4 count of less than 500 cells/mm^3 ($p<0.05$) when compared to CD4 500-1200 cells/mm^3 or >1200 cells/mm^3. However, the rate of LLV was much lower than the control group suggesting no significant association between CD4 T-cell counts and LLV (Table 11).

Table 11: Impact of CD4 T-cell counts on VF & LLV. The association of HIV-positive (<500 cells/mm^3) and HIV-negative children (>500 cells/mm^3) was analyzed for an increased risk of (A) virologic failure and (B) low-level viremia. ART: antiretroviral therapy; LLV: low-level viremia.

(A)

	ART success	Virologic Failure	Total
<500 cells/mm^3	13 (6%)	17 (27%)	30 (10%)
500-1200 cells/mm^3	120 (52%)	27 (43%)	147 (50%)
>1200 cells/mm^3	97 (42%)	19 (30%)	116 (40%)
Total	230 (78%)	63 (22%)	293 (100%)
$\chi^2 = 24.63$, $p < 0.00001$			

(B)

	Control	LLV	Total
<500 cells/mm^3	27 (11%)	6 (11%)	33 (11%)
500-1200 cells/mm^3	134 (57%)	31 (55%)	165 (56%)
>1200 cells/mm^3	76 (32%)	19 (34%)	95 (32%)
Total	237 (81%)	56 (19%)	293 (100%)
$\chi^2 = 0.08$, $p = 0.96$			

ART usage was analyzed as individual drugs, as a first or second line regimen, and as NNRTI-drug based regimen to investigate their relationship with VF and LLV (Table 12). The drugs AZT, 3TC and ABC (NRTI class), NVP and EFV (NNRTI class) and Kaletra (PI class) were commonly used drugs. There were no reports specifying the use of drugs ddI, d4T and TDF by HIV-infected children in this cohort as part of their therapy. Table 12 indicates the total number of patients reported taking the drug, total number of patients with viral failure or LLV being on the drug or regimen along with chi-square statistics and FDR-adjusted analysis. The p-value obtained from chi-square analysis indicates a significant association of AZT and ABC with VF. The survival curve and cox proportional hazard ratio for AZT shows the drug is associated with viral suppression, reducing the risk of VF (Figure 7) while the curve for drug ABC demonstrated two times higher risk of VF with the use of this drug (Figure 8). Although the chi-square analysis and survival curve analysis indicates AZT and ABC are positively or negatively associated with VF, the adjusted values by FDR proved no significant association between any individual drugs and VF or LLV. Patients categorized by ART regimen and NNRTI-based regimen did not show a significant association with VF or LLV (Table 12).

Table 12: Impact of ART treatment on VF and LLV. The number of total patients reportedly taking the drug or regimen is described along with the number of patients reaching virologic failure and low-level viremia in the study population. The p-value from chi-square statistic and adjusted p-value using false discovery rate is shown below.

Drugs	Total	VF	P value	Adjusted p-value	LLV	P value	Adjusted p-value
Individual Drugs							
AZT	229	55	0.02	0.12	39	0.24	0.72
3TC	245	49	0.60	0.65	54	0.64	0.88
ABC	40	15	0.04	0.12	13	0.05	0.30
NVP	128	26	0.65	0.65	26	0.46	0.88
EFV	79	9	0.48	0.65	16	0.75	0.88
Kaletra	29	5	0.42	0.65	6	0.88	0.88
ddI, d4T, TDF	n/a						
ART Regimen			0.78			0.31	
NRTIs only	50	12			8		
NRTIs + NNRTIs	214	36			35		
NRTIs + PI	29	5			6		
NNRTI Prescribed							
EFV-based	79	16	0.53		15	0.88	
NVP-based	128	26	0.71		27	0.53	

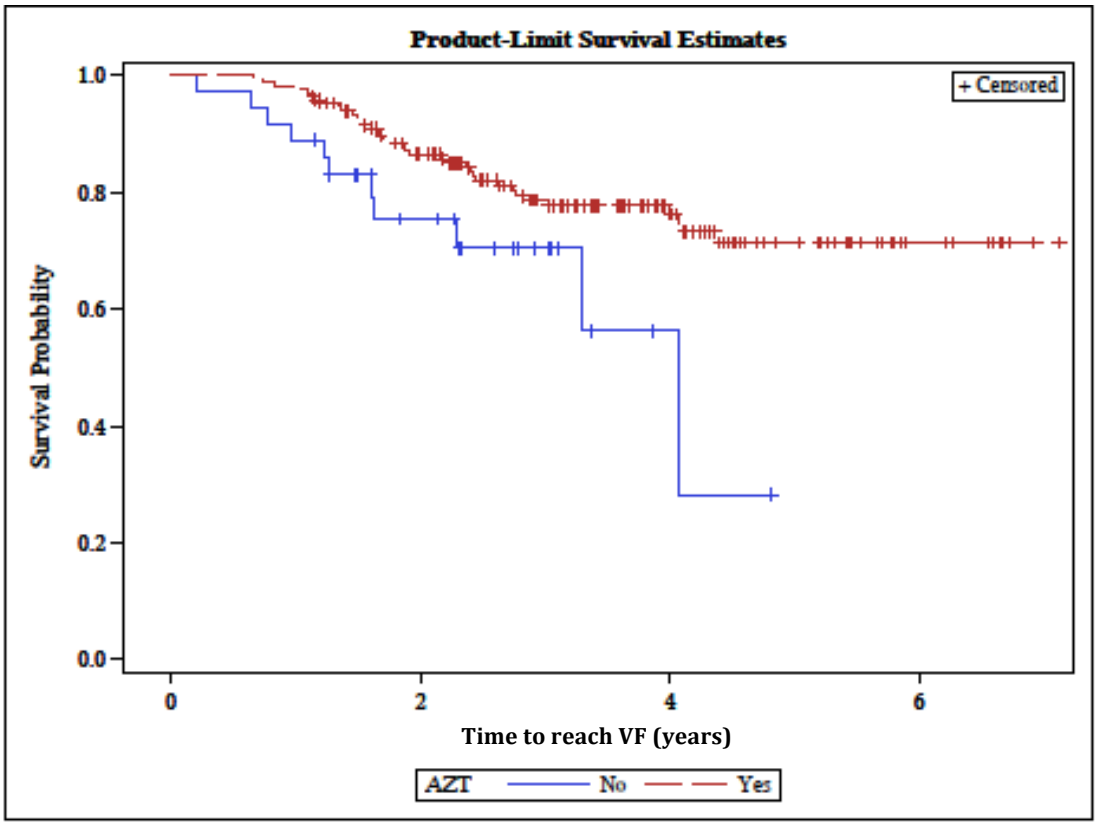


Figure 7: Kaplan-Meier survival analysis comparing individuals with and without AZT. The red-dashed line represents patients using the drug while the blue line represents patients not prescribed for Azidothymidine (AZT). Individuals with no virologic failure (VF) during their treatment were censored. [HR: 0.45 (0.23-0.89)].

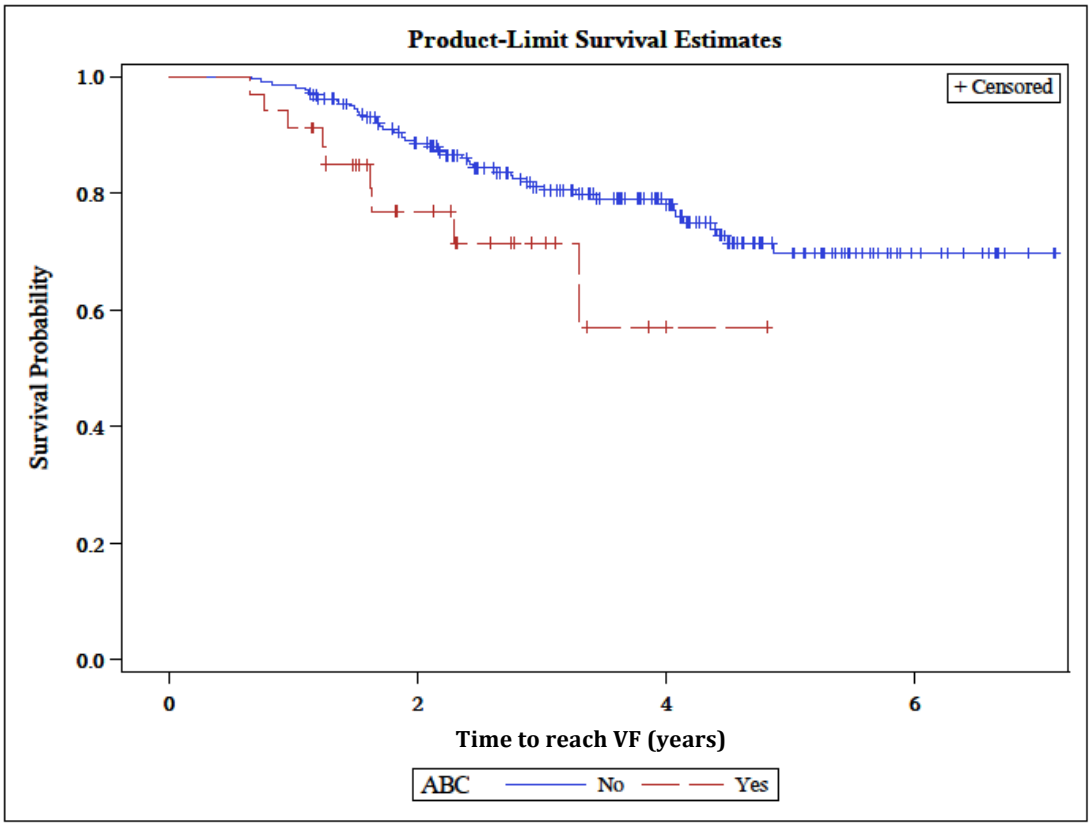


Figure 8: Kaplan-Meier survival analysis comparing individuals with and without ABC. The red-dashed line represents patients using the drug while the blue line represents patients not prescribed for Abacavir (ABC). Individuals with no virologic failure (VF) were censored. [HR: 2.148 (1.03-4.46)].

HIV-infected children co-infected with pneumonia, tuberculosis, malaria or malnutrition was investigated for an association with VF and LLV. The total number of patients with co-infections in the study population along with number of patients in each group reaching VF or LLV is stated in Table 13. Overall, 42% children had co-infection with pneumonia, 61% with tuberculosis, 20% with malaria and 23% with malnutrition. Among all the co-infections, the presence of HIV/pneumonia co-infection was found to be significantly associated with an increased risk of VF (Table 14). Kaplan-Meier curves were also constructed for each co-infection to investigate viral failure as a time-dependent variable (Figures 9-12).

Table 13: Impact of co-morbidities on VF and LLV. The number of HIV-infected individuals harboring another infection is illustrated, with the percentage of each group belonging to their respective outcomes of VF or LLV. For each co-infection, the p-value from chi-square statistic and adjusted p-value using false discovery rate is shown below. VF: virologic failure; LLV: low-level viremia.

	Total	VF	P-value	Adjusted value	LLV	P-value	Adjusted value
Pneumonia	106	47	1.9e-05	7.6e-05	18	0.17	0.36
Tuberculosis	67	17	0.98	0.98	9	0.18	0.36
Malaria	43	10	0.92	0.98	10	0.73	0.73
Malnutrition	49	10	0.65	0.98	9	0.50	0.67

Table 14: Impact of HIV/pneumonia co-infection on VF. Patients with the reported presence and absence of pneumonia as a co-infection were analyzed for their viral outcome. ART: antiretroviral therapy.

	ART success	Virologic Failure	Total
Pneumonia Absent	117 (66%)	28 (37%)	145 (58%)
Pneumonia Present	59 (34%)	47 (63%)	106 (42%)
Total	176 (70%)	75 (30%)	251 (100%)
$\chi^2 = 18.30$, p = 0.000019			

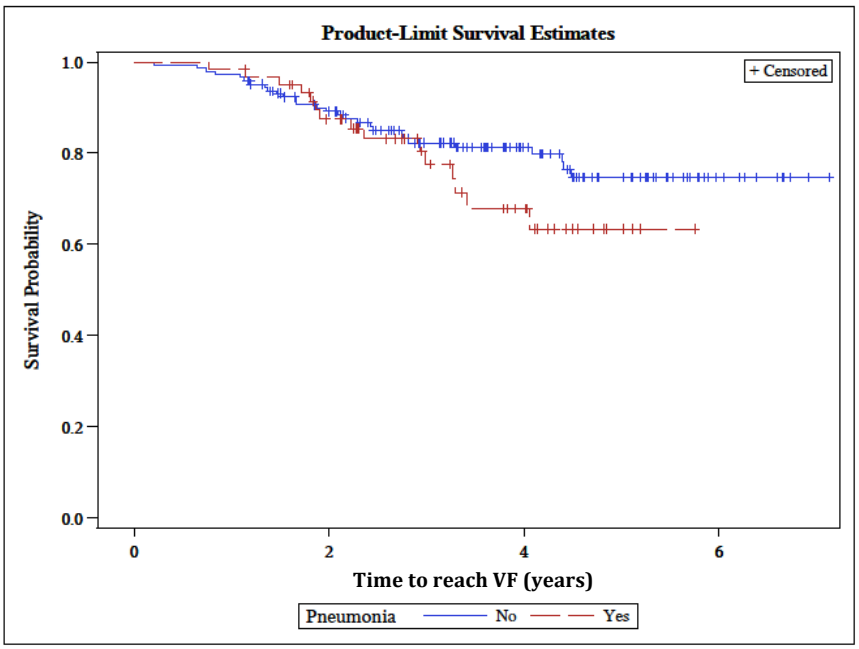


Figure 9: Kaplan Meier survival curve for the impact of pneumonia. The red-dashed line represents patients with the presence of co-infection while the blue line represents patients without pneumonia. Individuals with no virological failure (VF) during the course of their treatment were censored.

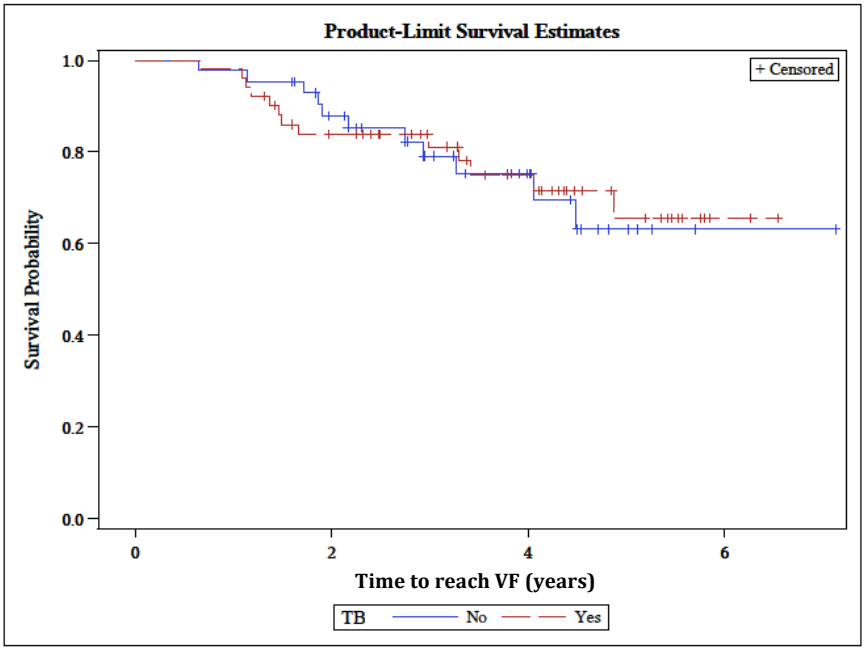


Figure 10: Kaplan Meier survival curve for the impact of tuberculosis. The red-dashed line represents patients with the presence of co-infection while the blue line represents patients without tuberculosis. Individuals with no virologic failure (VF) during the course of their treatment were censored.

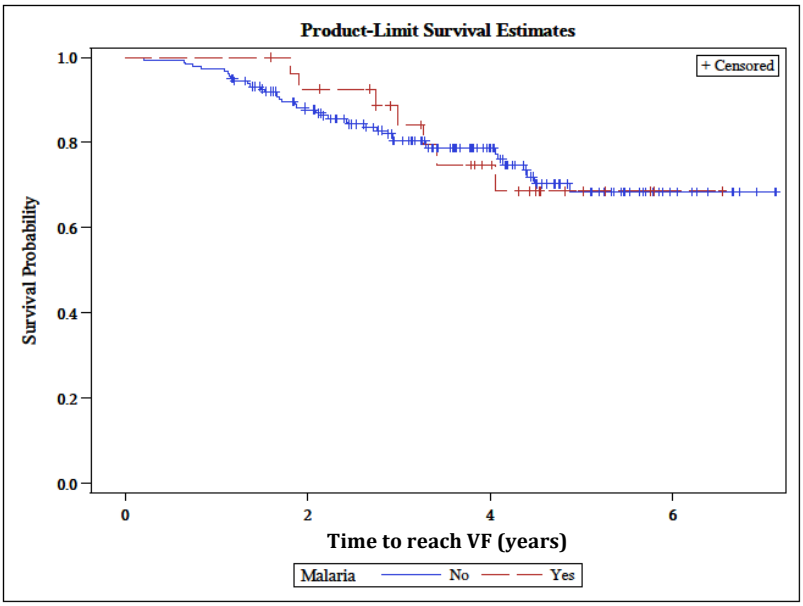


Figure 11: Kaplan Meier survival curve for the impact of malaria. The red-dashed line represents patients with the presence of co-infection while the blue line represents patients without malaria. Individuals with no virologic failure (VF) were censored.

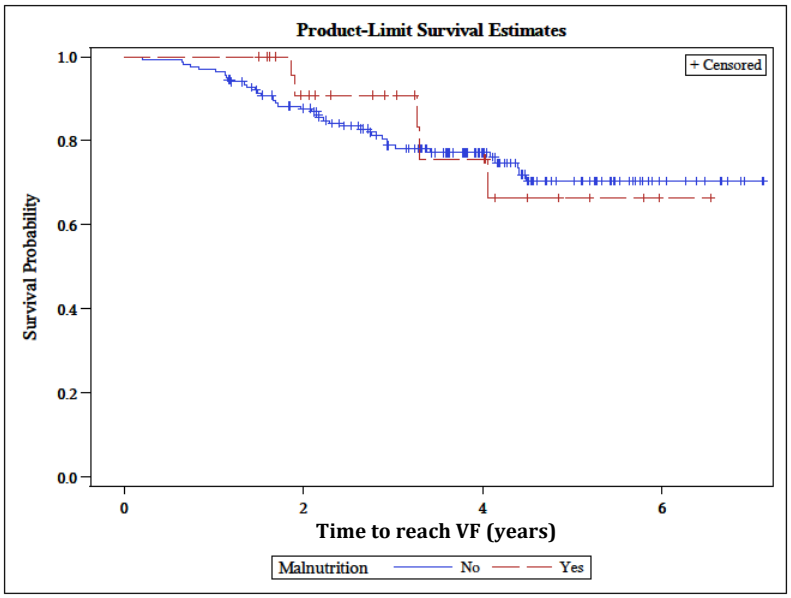


Figure 12: Kaplan Meier survival curve for the impact of malnutrition. The red-dashed line represents patients with the presence of co-infection while the blue line represents patients without malnutrition. Individuals with no virologic failure (VF) during the course of their treatment were censored.

Lastly, the impact of gender was measured for its association with VF and LLV. There were 52% females and 48% males infected with HIV-1 undergoing ART. The statistical analysis resulted in a non-significant association among the variables (Table 15).

Table 15: Impact of gender on VF & LLV. Females and males were categorized according to their viral outcome (A) virologic failure and (B) low-level viremia. ART: antiretroviral therapy; LLV: low-level viremia.

(A)

	ART success	Virologic Failure	Total
Female	118 (52%)	34 (53%)	152 (52%)
Male	111 (48%)	30 (47%)	141 (48%)
Total	229 (78%)	64 (22%)	293 (100%)
$\chi^2 = 0.05$, $p = 0.82$			

(B)

	Control	LLV	Total
Female	126 (53%)	26 (46%)	152 (52%)
Male	110 (47%)	31 (54%)	141 (48%)
Total	236 (81%)	57 (19%)	293 (100%)
$\chi^2 = 1.11$, $p = 0.29$			

Chapter 4: Discussion

4.1 Amplification of low-level viremia

4.1.1 Developing the protocol

HIV-1 infected individuals on antiretroviral treatment or who naturally control the infection have very low viral loads in plasma, typically at undetectable levels. This viral load can often fluctuate due to viral replication at lower levels, viral reservoirs or drug resistance [116]. A protocol was modified from the existing NLHG protocol and further optimized to amplify low-level viremia samples to detect DRMs in pediatric clinical samples with LLV in Nairobi, Kenya. The protocol was developed to ensure its efficiency in amplifying all examined HIV subtypes at varied VL levels. The sensitivity of the PCR-based protocol described herein is highly dependent on the input volume of plasma as well as the plasma viral load. A challenge common to PCR amplification of any HIV gene is the high viral sequence diversity and limited template number when patients' VL is low due to either the early stage of infection or due to LLV. The optimization of RNA extraction method described here was restricted to the methods and reagents available in resource-limited settings and the availability of the plasma samples, which was 400 - 600 μl on average. Hence, it was necessary to limit the input plasma volume to 400 μl in order to amplify as many samples as possible. HIV-1 subtypes A1, B, C and D were selected as they are prevalent in Kenyan population. Instead of having a different set of primers for each subtype, only one set of primers covering the conserved region of HIV genome was used to amplify different subtypes, as the subtype information of the clinical samples was unknown.

The protocol was initially tested on EQAPOL samples diluted to low viral loads in order to represent the clinical samples with LLV. The one-step RT-PCR method is ideal for high-throughput detection applications with known targets as it is quicker to set up, less expensive and involves less handling of samples thus reducing pipetting errors and/or contamination. For amplification of LLV, one-step RT-PCR approach was efficient and sensitive for all the subtypes tested except for subtype D at lower levels. This was similar to the results obtained in NLHG (NML, Winnipeg) with the standard automated extraction protocol used for DRM surveillance. Thus, the two-step RT-PCR approach was chosen to evaluate its performance, specifically on subtype D. This methodology should enhance the efficiency of PCR products as random hexamer primers used for RT reaction are in a separate tube, permitting all the variants in an RNA sample to convert into cDNA that is more stable and allows archiving of samples [139]. However, this approach did not work well for any of the subtypes tested with various viral loads. It required two separate kits meaning more steps to perform the same task versus the one-step protocol. It requires elaborate setup, more reagents, increased human intervention and increased machine time. Similar to two-amplicon nested PCR, the labor-intensive approach allows higher chances of technical and/or sample handling discrepancies and opportunities for loss of sample. Perhaps, amplification with *Taq* DNA polymerase will provide better amplification efficiency compared to the SuperScript RT enzyme for quantitative PCR. The lower incubation temperature for the RT step required with SuperScript kit may allow for increased RNA secondary structure reducing the primer binding efficiency or hindering activity of the enzyme. This can be overcome by using *Taq* polymerase, which is complexed with proprietary antibody to block

polymerase activity at ambient temperatures [140]. Additionally, the RNA and/or cDNA template volumes for the two-step RT-PCR may have been too concentrated for the amplification reaction to be effective. Analysis of one-step and two-step real-time RT-PCR using SuperScript versus several other RT enzymes in previous literature found that SuperScript gave the highest overall yield and supports the use of the one-step method for increased sensitivity as compared to the two-step RT-PCR [139,141]. Another study testing the SuperScript two-step RT-PCR method obtained a low reaction efficiency and sensitivity, which supports the findings presented herein [142].

Nested PCR is used to increase the yield and specificity of amplification of the target DNA. For the two-amplicon N-PCR approach, different primers were used along with preparing two separate master mixes for each gene of interest. The primers bind to many sites in the total isolated DNA and all sites initiate single stranded DNA. A PCR product is generated only when both primers are able to initiate priming of the targeted site. Perhaps the primers misintegrated into the PCR product or were not able to bind accurately which prohibited sensitive amplification of PR and RT genes. Overall, one-step RT-PCR and one-amplicon nested PCR allowed easier processing of samples and helped minimize carryover contamination between DNA synthesis and amplification. These approaches provide greater sensitivity for the subtypes tested at low viral loads keeping the resource- limited environment in mind.

4.1.2 EQAPOL amplification validation

By targeting the highly conserved regions in the *pol* gene and incorporating degenerative bases, the primers applied to both RT-PCR and nested-PCR amplifications

accommodated all major HIV-1 subtypes in Africa. Both PCR kits were selected for the use of high fidelity polymerases and were from reputable commercial resources with excellent performance consistency. The EQAPOL samples were tested in triplicates to validate the consistency and sensitivity of the PCR protocol so that it could be applied to clinical samples. As indicated in Table 4, all subtypes were successfully extracted and consistently amplified even when the VL was as low as 50 copies/ml. However, the consistency of success rate decreased when VL ≤ 200 copies/ml, fluctuating especially for subtype D. The dose-dependent PCR success rate implies that the initial HIV copy number at the RT-PCR stage determines the final PCR outcome. Assuming a conservative HIV extraction efficiency of 80%, the theoretical HIV templates introduced into RT-PCR, as little as ~6 copies per reaction with this protocol when the VL was at 100 copies/ml. These extremely low numbers of HIV templates in the PCR reaction do not allow reliable amplification. Notably, this input copy number was solely dependent on the initial input plasma volume and the elution volume at the HIV RNA extraction step that could certainly vary due to technical or enzymatic differences during each experiment. Furthermore, PCR amplification "leaped" consistently for subtype D, meaning it amplified at 200 copies/ml but not at 500 copies/ml (a dilution lower) or at 100 copies/ml (a dilution higher). The PCR efficiency may be explained by the possibility of having intramolecular errors specific to subtype D. Primers have the ability to form secondary structures such as loops or hairpins due to complementary sequences within the primer, exhibiting severely reduced amplification efficiency [143]. It is plausible to have variations in efficiency due to the use of different thermocyclers for amplification, or technical discrepancies while diluting the templates.

Increasing the initial plasma volume when constructing serial dilution series, for RNA extraction, or reducing the elution volume during the HIV RNA extraction stage when low VL is expected can further optimize this protocol. More sensitive extraction methods such as automated extraction, silica based systems or ultracentrifugation can be adopted to significantly enhance the efficiency if resources are available. Alternatively, different primers can be used for each subtype to enhance subtype-specific sensitivity. Hence, combining all strategies should help in obtaining good quality PCR amplicons for any sample with VL $\geq$ 50 copies/μl.

4.1.3 Clinical samples from Nairobi

The amplification success rate achieved from plasma samples of HIV-1 infected children was satisfactory considering the limited resources available. Further troubleshooting attempts did not give promising results, thereby failing to amplify additional LLV or baseline samples. Except nine, all 49 plasma samples available from children had their VL > 200 copies/ml which should have amplified as this VL range had been successfully validated with the optimized protocol during standardization of the method. Several reasons could be contributory to failure in achieving approximately 80% of amplification, all of which notably were outside the control of our main research group. The plasma samples were collected almost a year prior to library preparation and were transferred to different labs, facing many freeze-thaw cycles that would have contributed to RNA degradation. The plasma samples were stored in EDTA anticoagulant that can change the concentration of the constituents to be measured and/or interfere with certain analytical methods [144]. Moreover, the clinical samples collected at the time of

blood test was expected to have plasma volume of 500 µl on average, but 13 samples had plasma volume as low as 50 µl available to which PBS had to be added to permit extraction. Low starting volume would have severely reduced the viral copy number present in the plasma volume hindering efficient extraction and amplification. Further, the enzymes and reagents were shipped from Canada to Nairobi, which might have reduced the life span of sensitive RT-PCR enzyme during long exposures to the above ambient temperatures at airport customs and at the labs before proper storage. There is also a high probability of mechanical issues with the thermocycler, as it lacked evidence of servicing from the instrument log. The constant airflow in the room from outside environment could also have contaminated the samples and/or reagents used for amplification.

These results prove that the protocol is applicable for amplifying HIV samples with unknown VL and subtypes. However, extra consideration must be given to the viral copy number in the PCR reaction when amplicons from LLV samples are subject to sequencing analysis for detecting low abundance DRMs. The protocol can be further optimized based on the sample volume and resources available, but the conditions in resource-limited settings may make it impossible to achieve consistent amplification of low copy number plasma samples.

4.2 Sequencing analysis of DRMs in LLV population

Drug resistance has become an obstacle in maximizing the clinical benefit of ART and consequently, HIV genotyping is recommended prior to ART initiation during viral rebounds and on the failure of an ART regimen. Successful PCR amplification of PR and

RT genes is essential regardless of which method is used for sequencing. The objective was to identify the mutations and their prevalence during LLV in HIV-1 infected children while receiving ART. Results from this study show that not only it is possible to detect but also to characterize virus that continues to be produced in low levels. The samples with LLV were amplified to determine whether LLV present in patients on effective ART regimen are correlated with the development of drug resistance. Out of the 31 samples that were amplified using the optimized protocol, seven had no product by the time the sequence library was prepared. The likelihood of the loss of PCR products during clean up exists. There is a high possibility of mishandling during shipping and receiving of the samples. Instead of screw-cap tubes, the samples were shipped on a 96 well plate with a parafilm on top, which is not ideal as per protocol as it is not air tight and may not retain seal integrity during sample handling or shipping, leading to sample leakage and/or contamination. Additionally, the ice packs had melted by the time the lab received samples for sequencing which can affect the efficiency of reagents used for amplification. Evaporation of solution could have the DNA adhere to the tube that may not later dissolve when resolubilized for library preparation or could fail the quality control due to low concentration or volume.

Many samples failed the quality control for library preparation via Nextera kit prior to MiSeq sequencing. Samples mainly had low concentration or less quantity of DNA specimen for further processing. When the nucleic acid was tested with Nanodrop at KMRI lab (Nairobi, Kenya) prior to shipping, the average concentration was 300 ng/µl but the sequencing laboratory measured many samples reporting low concentrations (<5 ng/µl) using Qubit. The Nanodrop measures the amount of nucleic acid and was initially

used simply because of no other resources available. Perhaps the initial working concentration was very low, which Qubit captured more accurately as it only measures double-stranded nucleic acid in the solution. The use of different instruments for quantification could cause discrepancy in the sample concentration detected. Further, it is expected to use fresh reagents and kits to perform sequencing on a previously certified machine by the sequencing laboratory but these conditions were not verified by the contract company performing the sequencing. Lack of attentiveness to such details could lead to many errors and unreliable results.

The success rate of LLV amplification assays vary between studies but the comparison between different procedures is undermined by the lack of experimental specifications of the protocol. Also, the actual number of input copies for the PCR amplification can differ substantially. A low input of viral RNA in the amplification reaction may lead to the selection of individual and possibly rare mutations [119]. This could explain minor mutations arising from the children samples detected in abundance. The most frequently found mutation in both patients was L89M in the PR gene, which is a minor mutation and usually associated with Tipranavir or Fosamprenavir, although this treatment was not provided to the children in this cohort. Perhaps it is a transmitted DRM from the treatment-experienced parents to the newly infected children, presenting a major obstacle in successful ART [145]. This DRM could also be a dominant mutation among the circulating strains in Africa. Previous literature on drug resistance profiling in non-B subtype population had commonly observed T12S (71.6 %), I15V (71.6 %), L19I (67.7 %), M36I (69.6 %), S37N (79.4 %), R41K (92.2 %), L63P (63.7 %), H69K (100 %), and I93L (91.1 %) protease polymorphisms in the adult population. Among these, L89M was

reported with 74.5% frequency and was possibly associated with resistance to Tipranavir [145]. This is consistent with another study detecting 76.1% prevalence of L89M minor protease mutation, potentially associated with resistance to Fosamprenavir, Darunavir or Lopinavir [146].

From the experimental results, multiple drug resistant mutations from two or more drug classes were noticed for each patient as expected [147]. Despite the fact that 50% of the patients (1/2) with pLLV developed NRTI inhibitor-based resistance to Abacavir with D67E mutation, the sample size is too small to statistically conclude that DRMs are prevalent in patients with pLLV. Both patients had DRMs in RT and PR regions and did not experience virologic failure during their therapy, implying DRMs do not increase the risk of reaching VF, but this cannot be validated due to an insufficient sample size.

Low input number can increase the potential for PCR and sequencing errors. However, exclusion of PCR or sequencing errors was not possible due to the low quantity of sample availability. In addition, the low copy number of HIV templates in the PCR reaction does not allow for reliable identification and quantification of DRMs at low levels. For instance, if the patient VL is 50 copies/ml then the initial HIV template introduced into the RT-PCR step is ~3 copies theoretically and all progeny amplicons are derived from them. The input copy number is critical in detecting low abundance variants with NGS. It should not be expected to reliably detect any DR variants present at 33% or lower frequencies with this limited template number. Besides the extremely low sample number, this limitation may explain the lack of positive correlation between low abundance DRM and VF in LLV samples.

Although the patient's regimen at the time of specimen collection included 2 NRTI and 1 NNRTI, and 2 NRTI and 1 PI drug classes, only NRTI-based drug resistant mutation were identified. Lower resistance to PIs and NNRTIs is expected when resistance is analyzed from DNA samples [147]. Most of the detected mutations were not attributable to the ART regimen prescribed at the time of sample collection. DRMs found in DNA can reveal archived or transmitted mutations, therefore these mutations were probably selected under regimens taken before the manifestation of pLLV and were picked up during blood sampling [147]. Thus, the impact of these mutations on the efficacy of ART remains unknown. A relationship between DRMs on LLV is also not apparent. Low-level virus production in patients undergoing ART can occur over several years without direct relation to the resistance mutations to the regimen. Along with previous studies conducted in the adult population [67,91,112], these results agree with a previous study profiling DRMs in children [90]. Perhaps the most plausible explanation is intermittent release from the latent reservoir with limited replication at a level that is insufficient to generate DRMs, or that early transient viremia may be associated with a reversion to wild-type viruses [116].

To detect DRMs, only the patients with LLV were chosen but not the patients with sustained undetectable viral load from an effective ART. It is important to emphasize that the limited sampling from a small number of patients does not allow excluding ongoing viral replication in patients who have maintained suppression of viral load with ART. It is possible that the virus with DRMs will eventually increase when the level of viremia in these patients with undetectable VL reaches detectable range [69,71]. Unavailability of the longitudinal samples for amplification prevented the investigation of

the presence of baseline mutations and if the new mutations have developed over time. It is possible that significant mutations may have occurred in those patients for whom amplification could not be accomplished. It is also possible that there were resistance mutations present in a minority of the viral species that were unable to be detected due to low viral load of the plasma sample, preventing reliable amplification from low copy number. Thus, the study may have underestimated the amount of resistance. Indeed, this is a problem with all forms of resistance testing. Furthermore, patients were not excluded based on the therapies they were receiving, perhaps a controlled trial will be more definitive in proving that resistance testing at LLV is a useful strategy to prevent disease progression. For future studies, a controlled investigation with a large sample size along with longitudinal samples should be chosen to better understand the impact of DRMs on LLV in children population. Due to an unknown long-term ART evolution, DRM testing prior to prescribing ART regimen and developing approaches to eliminate the source of residual viremia is still vital to virus eradication. The need to understand the point at which minority resistant viral population may be clinically relevant still exists. More studies are needed to determine a cutoff value for the detection of minority variants, below which the risk of failure declines. This threshold may be specific for each mutation and ART regimen.

4.3 Statistical analysis for the risk factors of LLV and VF

In a cohort of 293 patients selected on the basis of stringent viral load criteria during the inclusion period, approximately 20% patients were found to have low-level viremia during an average of three years follow-up. The frequency of transient or

persistent viral rebound was lower than expected but in accordance with recent literature showing 10-40% of patients who had undetectable plasma HIV-RNA levels during ART reaching either persistent LLV or blips [65,67]. The rate of virologic failure was found to be 22% in the HIV-infected children cohort. Interestingly, the VF rate reported herein is lower than 26% previously reported for children [11,114]. In contrast, a previous study on adult populations (above 18 years old) presented only 6.2% VF among treated patients [110] while other studies reported 12.9% to 16% viral failure cases [109,116]. The actual rates varied across these studies due to differences in the study designs, population and assays used.

To investigate the risk factors associating with LLV and VF, a database was created for all the HIV-infected children receiving ART in the Leo Toto program. There were 293 patients filtered via inclusion/exclusion criteria's applied for this study and were further categorized according to their definitions for LLV status or virologic failure outcome. The relationship between LLV and VF was measured using standard parametric and survival analysis. Statistical analysis was also performed on a number of epidemiological and clinical factors such as adherence, WHO clinical stage, CD4 T-cell counts, individual ART drugs and regimen, co-infections, gender and therapy stage to investigate an association with VF as well as low level viremia.

pLLV and blips are not significantly associated with VF.

Association studies between patients who experienced LLV and VF were performed on HIV-infected children aged 0-17 years old. Patients in the control group were more highly associated with VF whereas more than half the patients in LLV study

groups sustained undetectable viremia (Table 8). The Kaplan-Meier survival curve implies that patients with pLLV represent a higher probability of VF as compared to the patients with transient blips for a similar duration (Figure 6). However, the trend is statistically insignificant and it is hard to conclude a correlation due to small sample size of patients with pLLV or blips associating with VF. Only two pLLV patients progressed to VF while four patients with blips lead to VF. Perhaps the lack of drug resistance in clinical samples from the cohort or other confounding factors further explains the reduced rate of LLV and elevated viral suppression in this HIV-infected children cohort.

Reports from previous studies suggest that blips occur frequently with sustained viral suppression and does not reflect viral replication nor predict VF [65,148]. This is in agreement with numerous studies that did not report an increased risk of VF in patients with LLV [60,67,71,72,110,118,128,148,149]. In contrast, there have been studies on adult population that found a significant increase in risk of VF in patients with pLLV and/or blips, of which many used definitions for LLV with a different upper VL cutoff ranging from 200 – 500 copies/ml [66,112,117,125] and different viral load cutoff or a single viral rebound to define virologic failure [61,65,109,117]. The results of currently available literature suggest defining the viral load cutoff of LLV and VF is important, explaining heterogeneous findings in the previous literature examining similar correlations.

This reduced virological failure or a higher viral suppression rate in LTP children cohort can occur due to several reasons. Similar to this study, published literature that failed to show a similar association either had a relatively small sample size of patients with low-level viremia or had a rare development of viral failure overall [150,151]. As

the pre-treatment CD4 T-cell count was in the healthy range of 1017 cell/mm^3 in this cohort, it is likely that low viral replication levels do not necessarily imply a substantial increase in plasma viral load, instead, reflects intermittent output from infected cells with a few cells becoming infected [137]. Moreover, as the subtype information of the participants was unavailable, virological efficacy of ART can be influenced by the viral diversity of different subtypes or CRF strains present in the children population [11]. As most of the patients were on the first-line regimen, this low level of VF supports the efficacy of the first-line treatment. Another possible reason is that patients had previous knowledge and concerns about the disease progression due to family history so they made regular visits to LTP center in order to benefit with free drugs and proper counseling provided by the program. It is possible that the risk of VF following LLV has been decreasing in recent years because of the use of newer ART or active switching of therapy before virological failure in some patients, but further studies will be required to investigate this [117].

Optimal adherence is associated with reduced rate of virological failure.

In this study, patients with optimal adherence were significantly associated with decreased VF rate during ART (Table 9). These patients had a higher chance of maintaining undetectable viral load below 50 copies/ml as compared to individuals with suboptimal or poor adherence as they strictly adhere to more than 95% of their therapy. A well-established relationship between virologic failure and lapses in adherence has been shown previously [109]. On the other hand, residual low-level viremia was not related to the adherence level of combination therapy ($p>0.05$). The subjects with a higher chance

of experiencing LLV had more adherence to the therapy. Adherence studies have addressed viral rebound in general but not LLV specifically. Some studies have linked LLV to adherence [67,152], but this has not been consistently reported [71,119]. Differences in these results may represent precise monitoring of adherence. Due to the large size of this cohort with limited resources, adherence was monitored via random pill counts from patients. This method is only moderately reliable as patients in this cohort have sold ART drugs in need of money or simply avoided taking the prescribed medicine due to stigma or inconvenience. Perhaps using an electronic medication event monitoring system to measure adherence is more reliable than manually recording when patients had administered a dose of their regimen [67]. Moreover, longitudinal measurements were lacking in patient reports from which data was extracted, providing only one data entry on adherence without detailed information on the stage of therapy. Although longitudinal adherence could not be monitored, the results support that adherence associated issues play a central part in LLV and VF development. Evaluating adherence during a shorter and more clinically relevant time period around LLV events would be more relevant for further investigating its association with LLV.

The WHO clinical stages are not associated with the rate of VF or LLV.

In this study, the prevalence of virological failure among WHO defined severe clinical stages (stage 3 and 4) was found to be 55% and is higher than the 45% VF rate in patients with mild clinical stages (Table 10). This finding was not found to be statistically significant even though the trends were as expected. A report on a children cohort stated that majority had advanced HIV disease at initiation of treatment as indicated by their

WHO stage that later lead to mortality [115]. This supports the hypothesis that severe clinical staging enhances the risk of disease progression possibly to viral failure. Similarly, LLV was not significantly associated with WHO clinical stages. Hypothetically, a more advanced clinical stage prior to treatment should increase the risk for viral rebound to LLV and/or VF during ART because of a greater diversity of HIV quasispecies, more extensive viral reservoir, and exhaustion of immune effector responses [60,151]. Previous studies have presented an association of severe CDC clinical stage with LLV [151] but studies specifically investigating the impact of WHO clinical stage on LLV are lacking. As the WHO staging is solely dependent on typical HIV-defining symptoms of the patients to estimate the start of ART in resource-limited areas, these symptoms can be misapprehended or under-represented by healthcare workers or nurses if proper training is not achieved. The individuals attending clinics are supposed to be assessed for WHO clinical stage prior to the treatment initiation but this data may have differed up to a few months before or after ART initiation. A study in Sub-Saharan Africa has reported that WHO staging system misses a high proportion of individuals who are ART eligible by CD4 count, providing inaccurate data [153]. These findings highlight the complexities of the WHO staging system for which in-depth medical knowledge, experience, and ability to interpret diagnostics is required. Thus, an alternative simplified and cost-effective CD4 testing approaches need to be developed.

Low CD4-T cell counts increase the chances of reaching virological failure.

CD4 T-cell counts measured from HIV-infected children on ART treatment for an average duration of three years were associated with the occurrence of VF (Table 11).

From the 30 patients categorized in <500 cells/mm^3, the increase in the risk of VF frequency (17/30) was significantly associated when compared to the group with ART success. This finding is anticipated because depletion of CD4 immune cells promotes a significantly elevated viral load in severely infected patients [60,109]. This association pattern was not apparent between CD4 cell counts and LLV. The frequency of patients had increased in the control group for all three CD4 count sub-categories, indicating the detection of LLV was not significantly linked to CD4 T-cell counts. While studies have signaled that the CD4 T-cell trajectory affects viremia [126-129], there does not appear to be a clinically meaningful impact of LLV on the trajectory according to other studies [62,118,151]. It is predictable to observe the majority of children showing a higher CD4 cell count and representing a healthier immune response after receiving treatment, compared to the total of 33 children with a poor immune cell count of less than 500 cell/mm^3. Children with severe infection for a longer duration may have caused a permanent change to the immune system and lymphoid tissues that could contribute to poor long-term status [154]. Moreover, children may exhibit slower recovery rate due to immunologic burden and possibly poor adherence. The certainty of CD4 count as a predictor of LLV is limited by the overall low rates of children facing low-level viremia.

Individual ART drugs and regimen are not significantly associated with VF or LLV.

Children with HIV-infected parents are initially enrolled in the LTP clinics and ART treatment is initiated once they develop symptoms or display reduced CD4 counts. The drug d4T (Stavudine) is an NRTI that has been discontinued due to the high frequency of serious toxicity. The NRTI drugs ddI (Didanosine) and TDF (Tenofovir)

that are generally prescribed for first-line regimen was not used in this cohort. Rather, AZT (n=229) and 3TC (n=245) were commonly seen NRTI drugs, followed by ABC (n=40). The NNRTI drug NVP was more commonly prescribed (n=128) in comparison with EFV (n=79). There was only one option available for a PI-based regimen that included the drug Kaletra (n=29). Independent analyses indicated a significant association of drugs AZT and ABC with VF. Patients receiving the drug AZT had a higher rate of viral suppression and reduced probability of reaching the stage of virological failure (Figure 7). The hazard ratio of 0.45 and small confidence interval indicates the chance of these patients reaching VF is only 0.45 times more than the patients who are not on AZT, suggesting the impact is not significant. On the other hand, patients receiving the drug ABC had two times higher probability of reaching viral failure. Perhaps the previous findings of the occurrence of DRMs from drug ABC, confirmed by the amplification of LLV in children samples, plays a role in increasing the risk of VF. However, the contradictory results of independent drug analyses could arise from the false discovery rate. Taking into account the drastically different median CD4 counts among patients initiating ART observed in this study, it would be expected that FDR models controlling these differences would lead to shifts in estimates observed from unadjusted analyses. It is notable that drug analyses of adjusted models largely yielded consistent findings to those observed in primary analyses. In summary, the six drugs prescribed to HIV-infected children were not associated with either VF or LLV (Table 12).

A stratification of the study patients by ART regimen was performed for further association analysis. Typically, two NRTI and one NNRTI are prescribed as a first-line

regimen and two NRTI with one PI is used as a second-line regimen, as recommended by WHO. Among the study population, the most commonly used was NNRTI-based regimen (n=214), followed by PIs (n=29). This trend is similar to a study done by Gisela et al. on an adult population [155]. Patients receiving PIs as a second-line regimen (NRTIs and PIs) were limited to the drug Kaletra as an accessible option. The number of patients progressing to VF (n=5) and LLV (n=6) in this category is similar to the patients with VF and LLV from the drug Kaletra, proposing that the drug itself is a cause for viral progression. A PI-based regimen has been previously associated as a factor with increased risk of LLV [108]. Surprisingly, 50 patients were observed receiving NRTI monotherapy, indicating the cohort was not strictly obeying WHO recommended guidelines. Patients receiving NNRTIs were further categorized into an EFV-based and NVP-based regimen, neither of which signified an impact on VF or LLV. Thus, a significant trend or impact was not associated with VF or LLV based on the ART classes and NNRTI drug-based regimen. In fact, the drug regimen favors viral suppression leading to a high virological efficacy of ART in the majority of the study population. This would have been a contributing factor for the lack of patients progressing to salvage therapy or mortality in this cohort, although information was missing for 27 individuals regarding their therapy stage.

These drugs were provided to the children due to high accessibility and less toxicity. Several studies have already shown that NNRTI-based regimen may result in a better suppression of plasma viral load [155]. The data suggests that this pediatric program in a resource-limited setting can be both clinically effective and sustainable in therapeutic responses, patient survival, disease retention, and regimen durability. The

response to ART is comparable to other pediatric cohort studies. One of the cohort presented 84% of virologic suppression [115] while another children cohort achieved 49% virological suppression [156]. A multi-site pediatric cohort in the United States reported a rate of 34% viral suppression among children between 0-17 years of age, similar to the presented LTP cohort [157]. The difference between the rates of viral suppression or failure in various studies may be due to a different population, viral subtypes, treatment regimen used, personal awareness and social factors. Perhaps, HIV-infected caregivers are more knowledgeable and provide a more informed treatment support and management for their children resulting in the better outcome. Thus, they are more likely to access available health services and counsel their children for proper adherence to treatment. Further research is necessary to investigate these possibilities as it may help HIV-infected individuals prevent disease progression, globally.

Co-infection with pneumonia is significantly associated with higher VF rate.

Children with HIV-infection were investigated for the presence of co-infections with pneumonia, tuberculosis, malaria or malnutrition and their impact on VF and LLV. A high prevalence of co-morbidities in this resource-limited cohort was apparent. The presence of pneumonia-HIV co-infection in 106 patients led to a 63% VF rate (Table 14). This significant increase in VF is further supported by the survival curve where the probability of patients reaching VF increased with pneumonia (Figure 9), thus proving that the presence of pneumonia in HIV-infected children was significantly associated with a higher risk of VF. Approximately 67 children in total had tuberculosis, 43 had malaria, and 49 were facing mild to severe malnutrition along with HIV-infection.

Kaplan-Meier survival analysis for these co-infections (Figures 9-11) except for tuberculosis trended a higher probability of reaching VF in the patients presenting the respective co-infection. Although it was observed from the survival curves that the risk of VF is higher with these co-infections, the p-value obtained from standard parametric test and survival test was not statistically significant. Further, the presence of pneumonia, tuberculosis, malaria or malnutrition does not have an impact on low-level viremia in HIV-infected children.

A dynamic interaction exists between viral replication and HIV-infected patients with co-infection as a result of the relationship between viral life cycle and the activation state of cells that supports viral replication. Acquiring recurrent co-infections has the potential to weaken the immune system by depleting CD4 cells, thus promoting viral replication and an accelerated viral or immune failure [158]. This supports the findings of pneumonia increasing the risk of VF in HIV-infected children. Development of tuberculosis leads to an increase in plasma HIV load, enhancing the chances of HIV genotypic diversification. Several studies have associated the occurrence of HIV/tuberculosis infection as a major contributor to VF [109,159-161]. Similarly, a study found higher plasma RNA load in HIV-infected individuals with malaria compared to HIV-infected controls with similar percentage CD4 cells [158]. The cell-mediated immune response is also compromised in malnourished children. Data from observational studies suggest that growth failure due to micronutrient status has been associated with the occurrence of secondary infections, high viral load in blood and breast milk, faster rate of progression to AIDS and shorter survival time in HIV-infected children [21]. The results presented in this study could differ due to missing information of up to 183

individuals, which could have potentially changed the findings. An optimal adherence to ART and high viral suppression noticed in the children of this cohort can also present lower VF and LLV rates, even with co-morbidities. It has been reported that a month after receiving malaria treatment, the proportion of macrophage-derived virus in plasma decreased to levels comparable to those in HIV-infected control without malaria [158]. Subjects treated for tuberculosis also documented a reduction in plasma HIV viral load [162]. However, studies have yet to establish if the long-term chemoprophylaxis for co-infections would have a survival benefit for children on ART.

Gender and therapy stage are not significant risk factors for VF or LLV.

In general, gender was not established to be a significant factor for VF and LLV (Table 15). A higher number of females (53%) were found progressing to VF than males (47%), however the difference is minor and insignificant. Among the patients attaining LLV episodes, higher frequency of males (54%) exhibited low viral load compared to females (46%) though this observation is not statistically significant for its impact on LLV. In published studies of perinatal-infected adolescents, there is usually equal gender distribution similar to the LTP cohort [45]. A French cohort demonstrated greater virological suppression rates in female adolescents [163], while studies in the United States reported improved viral suppression in males [164,165]. On the other hand, there are studies supporting no significant association with gender and occurrence of LLV rebound [110]. Contradictory findings regarding the impact of gender on virological suppression or rebound warrants further investigation.

Furthermore, small sample size (n=3) of patients in the second-line therapy stage is not statistically significant to analyze the virological significance of therapy stage concerning VF and LLV. It would be expected for patients in second-line to progress to VF and perhaps LLV as the presence of possible DRMs and immune exhaustion will allow the HIV-RNA to elevate in severely infected children. A possible reason for a low sample size is the efficacy of ART in this cohort, preventing disease progression and an advanced therapy stage. Alternatively, the number of patients receiving second line therapy must have been underrepresented or inaccurately recorded since there were at least 29 patients receiving PI-based regimen that is typically given as a second-line regimen. A strict guideline for recording patient data should be proposed to avoid errors in a patient's HIV-related medical history.

4.4 Limitations

Despite the large sample size of the study population, persistent LLV remains a relatively rare occurrence in this study. Though, it likely indicates that children maintained viral suppression via optimal adherence to their prescribed combination therapy. The small sample size of patients with LLV hindered the statistical significance of risk factors. Perhaps stratification of pLLV and blips will give a better insight of association with several variables but this requires a larger sample size. There were also a limited number of post-LLV outcomes with virological failure for analysis, making it difficult to detect a significant association between previously studied risk factors for VF. Confounding variables may also have biased the results because groups of patients might have differed in unmeasured characteristics, although most important variables associated

with virological failure were considered in the modeling strategy. Limited power to discern drug specific differences along with the observational and retrospective nature of this study made it impossible to avoid ART changes promoting heterogeneity in the regimen. Even though all the children in this cohort have transmitted HIV-infection from their mothers and are supposed to be on ART from birth, many were not started on therapy as late as 17 years of age. The gap in monitoring these patients would have underrepresented HIV related medical history, along with the mortality rate in the cohort. Moreover, patient's data was initially recorded on spreadsheets, which can be error prone leading to typing mistakes and introducing errors in the database. Missing data for some variables from the selected study population meeting inclusion/exclusion criteria also limits the power for data analysis. If data was collected with standard blood test intervals of 6 months for all the variables, more children could have been included in the study with fewer discrepancies to account for. It would have been interesting to analyze the longitudinal effect of adherence, regimen and co-morbidities categorized by different age groups and the duration of treatment received (after 1 yr., 2 yrs. or 3 yrs. of ART). Additionally, the correlation among all the significant variables presented in this study should be further analyzed.

Other limitations that affected the investigation of DRM prevalence in patients with LLV included the lack of longitudinal plasma samples affecting analysis of baseline mutations and new mutations arising at LLV and VF events. Amplification of all the clinical samples was hampered by the amount of available plasma along with long-term storage of LLV plasma samples accompanied with multiple freeze-thaw conditions in the resource-limited settings. A small quantity of plasma sample was also prohibitive to

duplicate analysis of samples to exclude PCR and sequencing errors. A controlled trial would have been more definitive in proving resistance testing at LLV as a useful strategy.

Chapter 5: Conclusion & Significance

This work was conducted to better understand the association of VF and LLV to various risk factors in HIV-1 infected children undergoing ART. To probe the impact DRMs may have on LLV, I developed a protocol that could extract and allow amplification of LLV samples in resource-limited settings. Due to limited equipment available at the KMRI lab in Nairobi, the protocol was developed with easily accessible kits to extract and amplify low viral loads on unknown HIV-1 subtypes A1, B, C and D. This protocol was found to be highly sensitive in successful amplification of HIV-1 protease and first 233 amino acids of reverse transcriptase genomic region when validated with EQAPOL samples. However, this method did display inconsistency in amplification of HIV-1 subtype D. Implementation of the protocol in Nairobi yielded 63% amplification rate on plasma samples of children with LLV despite the multiple challenges faced including lack of sufficient plasma volume, missing samples, lack of certification assuring the efficacy of equipment's, improper storage of shipped reagents, and perhaps degradation of enzymes. The amplified samples were sent to a company to perform Illumina MiSeq sequencing and then HyDRA pipeline was used to identify DRMs. The limited sample size did not support an association of DRMs with LLV or VF, instead indicated the presence of transmitted or archival mutations. This study supports analysis of DRMs using low-cost sensitive assays to evaluate LLV and clinical outcomes. Understanding the principles of HIV DR is essential in future treatment strategies, especially in cases where clinical decisions such as therapy switches are pending. The protocol proves to be reproducible and efficient for samples with low copy numbers and can be used by other researchers to study low abundant DRMs in LLV in resource-

limited countries. This protocol may also find application in routine HIV DR assays that employ PCR amplification of HIV *pol* gene for genotyping with expanded dynamic range.

Several clinical and epidemiological risk factors including adherence, WHO clinical stage, CD4 T-cell counts, individual ART drugs and regimen, co-infections, gender and therapy stage were also investigated for an association with LLV or VF. A database was created for HIV-infected children undergoing ART in the LTP cohort, and 293 children were selected after sanitizing the database and applying exclusion criteria for patient selection. The standardized format of the database structured for this study can be used by clinics in LTP cohort and possibly other cohorts providing ART to HIV-infected individuals to successfully monitor and manage their clinical outcomes in the future. Statistical significance of association studies was performed using chi-square test, Kaplan Meier survival curves or Cox proportional hazard models. This study demonstrated that low-level viremia is not significantly associated with virologic failure, whereas, adherence, CD4 T-cell counts and co-infection with pneumonia are significantly associated with virological failure in ART-treated HIV-1 infected children. There was no significant relationship among the risk factors explored for its correlation with LLV. The findings from this study contribute to the knowledge required to guide clinical management of HIV-infected children population. Many previous studies failed to account for adherence as a risk factor, which was assessed in the present study strongly implying the benefits of optimal adherence. The risk factors identified to impact VF can be used as predictors of VF in resource-limited settings where routine plasma viral load monitoring is not possible. It is likely that the clinical significance of pLLV and blips on

virologic failure and clinical outcomes in children population will continue to be a topic of interest. The importance of LLV will need to be reviewed periodically, as both the efficacy and accessibility of ART regimens improve and as HIV DR measurement becomes more common in resource-limited settings.

Chapter 6: Literature Cited

1 . Chiu IM, Yaniv A, Dahlberg JE, et al. Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. *Nature* 1985;317:366-8

2 . Wain-Hobson S, Alizon M, Montagnier L. Relationship of AIDS to other retroviruses. *Nature* 1985;313:743

3 . Levy JA. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* 1993;57:183-289

4 . Center ANR. Guide for HIV/AIDS Clinical Care. HIV Classification: CDC and WHO Staging Systems. *AETC National Resource Center* 2014

5 . Signorelli D, Joseph RJ,2nd. Epidemiology of HIV and AIDS. A retrospective look. *Clin Podiatr Med Surg* 1998;15:179-87

6 . HIV - Origin of HIV & AIDS. Available at: http://www.avert.org/history-hiv-aids/origin. , 2016

7 . The top 10 causes of death. Available at: http://www.who.int/mediacentre/factsheets/fs310/en/. , 2016

8 . Global Health Observatory Data: HIV/AIDS. Available at: http://www.who.int/gho/hiv/en/. , 2016

9 . HIV/AIDS: Adolescents and young people - Current status + progress. Available at: http://data.unicef.org/hiv-aids/adolescents-young-people.html. , 216

10 . HIV/AIDS : When to start ART in children. Available at: http://www.who.int/hiv/pub/guidelines/arv2013/art/statartchildren/en/index1.html. , 2016

11 . Chaix ML, Rouet F, Kouakoussui KA, et al. Genotypic human immunodeficiency virus type 1 drug resistance in highly active antiretroviral therapy-treated children in Abidjan, Cote d'Ivoire. *Pediatr Infect Dis J* 2005;24:1072-6 doi:00006454-200512000-00010 [pii]

13 . Ministry of Health RoK. AIDS in Kenya: trends, interventions and impact 2005

14 . Hemelaar J, Gouws E, Ghys PD, et al. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006;20:W13-23 doi:10.1097/01.aids.0000247564.73009.bc [doi]

15 . Santos AF, Soares MA. HIV Genetic Diversity and Drug Resistance. *Viruses* 2010;2:503-31 doi:10.3390/v2020503 [doi]

16 . Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol* 2007;81:10209-19 doi:JVI.00872-07 [pii]

17 . Burns DN, Mofenson LM. Paediatric HIV-1 infection. *Lancet* 1999;354 Suppl 2:SII1-6

18 . Watts FA, Vickerman C.P. Are people using condoms: current evidence from sub-Saharan Africa and Asia. *Reproductive Health Matters* 2004;12:176-177

19 . Levy JA. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* 1993;57:183-289

20 . HIV prevention: The basic of HIV prevention. Available at: https://aidsinfo.nih.gov/education-materials/fact-sheets/20/48/the-basics-of-hiv-prevention.

21 . Fawzi W, Villamore E. Nutrition, Immunity and Infection in infants and children: Malnutritioin and HIV infection. 45th Nutrition Workshop.:43

22 . Vernazza PL, Eron JJ, Fiscus SA, et al. Sexual transmission of HIV: infectiousness and prevention. *AIDS* 1999;13:155-66

23 . UNICEF Fact Sheet: Children orphaned by AIDS in sub-Saharan Africa. Available at: http://data.unaids.org/publications/fact-sheets03/fs_orphans_africa_en.pdf.

24 . Global HIV/AIDS overview. Available at: https://www.aids.gov/federal-resources/around-the-world/global-aids-overview/.

25 . Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* 2000;342:921-9 doi:10.1056/NEJM200003303421303 [doi]

26 . Chan DJ. Factors affecting sexual transmission of HIV-1: current evidence and implications for prevention. *Curr HIV Res* 2005;3:223-41

27 . HIV/AIDS: HIV Basics. Available at: http://www.cdc.gov/hiv/basics. , 2016

28 . Coombs RW, Reichelderfer PS, Landay AL. Recent observations on HIV type-1 infection in the genital tract of men and women. *AIDS* 2003;17:455-80 doi:10.1097/01.aids.0000042970.95433.f9 [doi]

29 . Fauci A. Immunopathogenic mechanisms of human immunodeficiency virus disease: implications for therapy. *Am J Med* 1995;99:59S-60S

30 . Brenchley JM, Price DA, Douek DC. HIV disease: fallout from a mucosal catastrophe?. *Nat Immunol* 2006;7:235-9 doi:ni1316 [pii]

31 . Miller CJ, Li Q, Abel K, et al. Propagation and dissemination of infection after vaginal transmission of simian immunodeficiency virus. *J Virol* 2005;79:9217-27 doi:79/14/9217 [pii]

32 . Gartner S, Markovits P, Markovitz DM, et al. The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science* 1986;233:215-9

33 . Palmisano L, Vella S. A brief history of antiretroviral therapy of HIV infection: success and challenges. *Ann Ist Super Sanita* 2011;47:44-8 doi:10.4415/ANN_11_01_10 [doi]

34 . Health care: HIV/AIDS. Available at: http://www.hiv.va.gov/index.asp.

35 . Haase AT. Targeting early infection to prevent HIV-1 mucosal transmission. *Nature* 2010;464:217-23 doi:10.1038/nature08757 [doi]

36 . Fauci AS. Pathogenesis of HIV disease: opportunities for new prevention interventions. *Clin Infect Dis* 2007;45 Suppl 4:S206-12 doi:10.1086/522540 [doi]

37 . Vaishnav YN, Wong-Staal F. The biochemistry of AIDS. *Annu Rev Biochem* 1991;60:577-630 doi:10.1146/annurev.bi.60.070191.003045 [doi]

38 . Frankel AD, Young JA. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* 1998;67:1-25 doi:10.1146/annurev.biochem.67.1.1 [doi]

39 . Coffin JM. The virology of AIDS: 1990. *AIDS* 1990;4 Suppl 1:S1-8

40 . Gianella S, Richman DD. Minority variants of drug-resistant HIV. *J Infect Dis* 2010;202:657-66 doi:10.1086/655397 [doi]

41 . Requejo HI. Worldwide molecular epidemiology of HIV. *Rev Saude Publica* 2006;40:331-45 doi:S0034-89102006000200023 [pii]

42 . Rambaut A, Posada D, Crandall KA, et al. The causes and consequences of HIV evolution. *Nat Rev Genet* 2004;5:52-61 doi:10.1038/nrg1246 [doi]

43 . Broder S. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res* 2010;85:1-18 doi:10.1016/j.antiviral.2009.10.002 [doi]

44 . Santos AF, Soares MA. HIV Genetic Diversity and Drug Resistance. *Viruses* 2010;2:503-31 doi:10.3390/v2020503 [doi]

45 . Agwu AL, Fairlie L. Antiretroviral treatment, management challenges and outcomes in perinatally HIV-infected adolescents. *J Int AIDS Soc* 2013;16:18579 doi:10.7448/IAS.16.1.18579 [doi]

46 . Antiretroviral Therapy Cohort Collaboration. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Lancet* 2008;372:293-9 doi:10.1016/S0140-6736(08)61113-7 [doi]

47 . Bygrave H, Mtangirwa J, Ncube K, et al. Antiretroviral therapy outcomes among adolescents and youth in rural Zimbabwe. *PLoS One* 2012;7:e52856 doi:10.1371/journal.pone.0052856 [doi]

48 . Friis-Moller N, Reiss P, Sabin CA, et al. Class of antiretroviral drugs and the risk of myocardial infarction. *N Engl J Med* 2007;356:1723-35 doi:356/17/1723 [pii]

49 . Lowenthal ED, Kitkata SB, Marukutira T, et al. Perinatally acquired HIV infection in adolescents from Sub-Saharan Africa: a review of emerging challenges. *Lancet Infect Dis* 2014;14 doi:10.1016/ S1473-3099(13)70363-3

50 . Schinazi R, Larder B, Mellors J. Resistance table: mutations in retroviral genes associated with drug resistance. *Int Antiviral News* 1997;5:129-142

51 . Hogg RS, Bangsberg DR, Lima VD, et al. Emergence of drug resistance is associated with an increased risk of death among patients first starting HAART. *PLoS Med* 2006;3:e356 doi:04-PLME-RA-0322R2 [pii]

52 . Rosenberg PS, Goedert JJ, Biggar RJ. Effect of age at seroconversion on the natural AIDS incubation distribution. Multicenter Hemophilia Cohort Study and the International Registry of Seroconverters. *AIDS* 1994;8:803-10

53 . WHO. Interim WHO clinical staging of HIV/AIDS and HIV/AIDS case definitions for surveillance (African Region). *WHO* 2005

54 . Weinberg JL, Kovarik CL. The WHO Clinical Staging System for HIV/AIDS. *Virtual Mentor* 2010;12:202-6 doi:10.1001/virtualmentor.2010.12.3.cprl1-1003 [doi]

55 . USAID. Summary table of HIV treatment regimens: Pediatric and adult treatment guidelines for pepfar focus and non-focus countries. *AIDSTAR-One* 2012

56 . Bennett DE, Bertagnolio S, Sutherland D, et al. The World Health Organization's global strategy for prevention and assessment of HIV drug resistance. *Antivir Ther* 2008;13 Suppl 2:1-13

57 . Alexander CS, Montessori V, Wynhoven B, et al. Prevalence and response to antiretroviral therapy of non-B subtypes of HIV in antiretroviral-naive individuals in British Columbia. *Antivir Ther* 2002;7:31-5

58 . Bocket L, Cheret A, Deuffic-Burban S, et al. Impact of human immunodeficiency virus type 1 subtype on first-line antiretroviral therapy effectiveness. *Antivir Ther* 2005;10:247-54

59 . Bannister WP, Ruiz L, Loveday C, et al. HIV-1 subtypes and response to combination antiretroviral therapy in Europe. *Antivir Ther* 2006;11:707-15

60 . Ryscavage P, Kelly S, Li JZ, et al. Significance and clinical management of persistent low-level viremia and very-low-level viremia in HIV-1-infected patients. *Antimicrob Agents Chemother* 2014;58:3585-98 doi:10.1128/AAC.00076-14 [doi]

61 . Sungkanuparph S, Groger RK, Overton ET, et al. Persistent low-level viraemia and virological failure in HIV-1-infected patients treated with highly active antiretroviral therapy. *HIV Med* 2006;7:437-41 doi:HIV403 [pii]

62 . Taiwo B, Gallien S, Aga E, et al. Antiretroviral drug resistance in HIV-1-infected patients experiencing persistent low-level viremia during first-line therapy. *J Infect Dis* 2011;204:515-20 doi:10.1093/infdis/jir353 [doi]

63 . Pernas B, Grandal M, Pertega S, et al. Any impact of blips and low-level viraemia episodes among HIV-infected patients with sustained virological suppression on ART?. *J Antimicrob Chemother* 2016;71:1051-5 doi:10.1093/jac/dkv433 [doi]

64 . Crowell TA, Phanuphak N, Pinyakorn S, et al. Virologic failure is uncommon after treatment initiation during acute HIV infection. *AIDS* 2016;30:1943-50 doi:10.1097/QAD.0000000000001148 [doi]

65 . Grennan JT, Loutfy MR, Su D, et al. Magnitude of virologic blips is associated with a higher risk for virologic rebound in HIV-infected individuals: a recurrent events analysis. *J Infect Dis* 2012;205:1230-8 doi:10.1093/infdis/jis104 [doi]

66 . Hofstra LM, Mudrikova T, Stam AJ, et al. Residual viremia is preceding viral blips and persistent low-level viremia in treated HIV-1 patients. *PLoS One* 2014;9:e110749 doi:10.1371/journal.pone.0110749 [doi]

67 . Podsadecki TJ, Vrijens BC, Tousset EP, et al. Decreased adherence to antiretroviral therapy observed prior to transient human immunodeficiency virus type 1 viremia. *J Infect Dis* 2007;196:1773-8 doi:10.1086/523704 [doi]

68 . Greub G, Cozzi-Lepri A, Ledergerber B, et al. Intermittent and sustained low-level HIV viral rebound in patients receiving potent antiretroviral therapy. *AIDS* 2002;16:1967-9

69 . Gunthard HF, Frost SD, Leigh-Brown AJ, et al. Evolution of envelope sequences of human immunodeficiency virus type 1 in cellular reservoirs in the setting of potent antiviral therapy. *J Virol* 1999;73:9404-12

70 . Ramratnam B, Mittler JE, Zhang L, et al. The decay of the latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of residual viral replication during prolonged anti-retroviral therapy. *Nat Med* 2000;6:82-5 doi:10.1038/71577 [doi]

71 . Martinez-Picado J, DePasquale MP, Kartsonis N, et al. Antiretroviral resistance during successful therapy of HIV type 1 infection. *Proc Natl Acad Sci U S A* 2000;97:10948-53 doi:97/20/10948 [pii]

72 . Havlir DV, Marschner IC, Hirsch MS, et al. Maintenance antiretroviral therapies in HIV infected patients with undetectable plasma HIV RNA after triple-drug therapy. AIDS Clinical Trials Group Study 343 Team. *N Engl J Med* 1998;339:1261-8 doi:10.1056/NEJM199810293391801 [doi]

73 . Wit FW, van Leeuwen R, Weverling GJ, et al. Outcome and predictors of failure of highly active antiretroviral therapy: one-year follow-up of a cohort of human immunodeficiency virus type 1-infected persons. *J Infect Dis* 1999;179:790-8 doi:JID980873 [pii]

74 . Havlir DV, Bassett R, Levitan D, et al. Prevalence and predictive value of intermittent viremia with combination hiv therapy. *JAMA* 2001;286:171-9 doi:joc01773 [pii]

75 . Ware NC, Idoko J, Kaaya S, et al. Explaining adherence success in sub-Saharan Africa: an ethnographic study. *PLoS Med* 2009;6:e11 doi:10.1371/journal.pmed.1000011 [doi]

76 . Haberer J, Mellins C. Paediatric adherence to HIV antiretroviral therapy. *Current HIV/AIDS Report* 2009;6:194-200

77 . Johnson VA, Calvez V, Gunthard HF, et al. 2011 update of the drug resistance mutations in HIV-1. *Top Antivir Med* 2011;19:156-64

78 . Johnson JA, Li JF, Wei X, et al. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. *PLoS Med* 2008;5:e158 doi:10.1371/journal.pmed.0050158 [doi]

79 . Stanford University Drug Resistanct Database. Available at: http://hivdb.stanford.edu/#.

80 . Eshleman SH, Hoover DR, Chen S, et al. Nevirapine (NVP) resistance in women with HIV-1 subtype C, compared with subtypes A and D, after the administration of single-dose NVP. *J Infect Dis* 2005;192:30-6 doi:JID34089 [pii]

81 . Flys TS, Chen S, Jones DC, et al. Quantitative analysis of HIV-1 variants with the K103N resistance mutation after single-dose nevirapine in women with HIV-1 subtypes A, C, and D. *J Acquir Immune Defic Syndr* 2006;42:610-3 doi:10.1097/01.qai.0000221686.67810.20 [doi]

82 . Loubser S, Balfe P, Sherman G, et al. Decay of K103N mutants in cellular DNA and plasma RNA after single-dose nevirapine to reduce mother-to-child HIV transmission. *AIDS* 2006;20:995-1002 doi:10.1097/01.aids.0000222071.60620.1d [doi]

83 . Izopet J, Massip P, Souyris C, et al. Shift in HIV resistance genotype after treatment interruption and short-term antiviral effect following a new salvage regimen. *AIDS* 2000;14:2247-55

84 . Lwembe R, Ochieng W, Panikulam A, et al. Anti-retroviral drug resistance-associated mutations among non-subtype B HIV-1-infected Kenyan children with treatment failure. *J Med Virol* 2007;79:865-72 doi:10.1002/jmv.20912 [doi]

85 . Roquebert B, Malet I, Wirden M, et al. Role of HIV-1 minority populations on resistance mutational pattern evolution and susceptibility to protease inhibitors. *AIDS* 2006;20:287-9 doi:10.1097/01.aids.0000202650.03279.69 [doi]

86 . Wamalwa DC, Lehman DA, Benki-Nugent S, et al. Long-term virologic response and genotypic resistance mutations in HIV-1 infected Kenyan children on combination antiretroviral therapy. *J Acquir Immune Defic Syndr* 2013;62:267-74 doi:10.1097/QAI.0b013e31827b4ac8 [doi]

87 . Sigaloff KC, Calis JC, Geelen SP, et al. HIV-1-resistance-associated mutations after failure of first-line antiretroviral treatment among children in resource-poor regions: a systematic review. *Lancet Infect Dis* 2011;11:769-79 doi:10.1016/S1473-3099(11)70141-4 [doi]

88 . Lowenthal ED, Marukutira T, Tshume O, et al. Prediction of HIV Virologic Failure Among Adolescents Using the Pediatric Symptom Checklist. *AIDS Behav* 2015;19:2044-8 doi:10.1007/s10461-015-1061-7 [doi]

89 . Kieffer TL, Finucane MM, Nettles RE, et al. Genotypic analysis of HIV-1 drug resistance at the limit of detection: virus production without evolution in treated adults with undetectable HIV loads. *J Infect Dis* 2004;189:1452-65 doi:10.1086/382488 [doi]

90 . Hermankova M, Ray,S.C., Ruff, C., Powell-Davis M, et al. HIV-1 drug resistance profiles in children and adults with viral load of . *JAMA* 2001;286:196-207

91 . Vancoillie L, Mortier V, Demecheleer E, et al. Drug resistance is rarely the cause or consequence of long-term persistent low-level viraemia in HIV-1-infected patients on ART. *Antivir Ther* 2015;20:789-94 doi:10.3851/IMP2966 [doi]

92 . Mens H, Kearney M, Wiegand A, et al. Amplifying and quantifying HIV-1 RNA in HIV infected individuals with viral loads below the limit of detection by standard clinical assays. *J Vis Exp* 2011;(55). pii: 2960. doi:10.3791/2960 doi:10.3791/2960 [doi]

93 . Hatano H, Delwart EL, Norris PJ, et al. Evidence of persistent low-level viremia in long-term HAART-suppressed, HIV-infected individuals. *AIDS* 2010;24:2535-9 doi:10.1097/QAD.0b013e32833dba03 [doi]

94 . Villahermosa ML, Thomson M, Vazquez de Parga E, et al. Improved conditions for extraction and amplification of human immunodeficiency virus type 1 RNA from plasma samples with low viral load. *J Hum Virol* 2000;3:27-34

95 . Metzner KJ, Bonhoeffer S, Fischer M, et al. Emergence of minor populations of human immunodeficiency virus type 1 carrying the M184V and L90M mutations in subjects undergoing structured treatment interruptions. *J Infect Dis* 2003;188:1433-43 doi:JID30769 [pii]

96 . Karlsson AC, Lindback S, Gaines H, et al. Characterization of the viral population during primary HIV-1 infection. *AIDS* 1998;12:839-47

97 . Charpentier C, Dwyer DE, Mammano F, et al. Role of minority populations of human immunodeficiency virus type 1 in the evolution of viral resistance to protease inhibitors. *J Virol* 2004;78:4234-47

98 . Petropoulos CJ, Parkin NT, Limoli KL, et al. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother* 2000;44:920-8

99 . Nissley DV, Halvas EK, Hoppman NL, et al. Sensitive phenotypic detection of minor drug-resistant human immunodeficiency virus type 1 reverse transcriptase variants. *J Clin Microbiol* 2005;43:5696-704 doi:43/11/5696 [pii]

100 . Paredes R, Lalama CM, Ribaudo HJ, et al. Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. *J Infect Dis* 2010;201:662-71 doi:10.1086/650543 [doi]

101 . Lalonde MS, Troyer RM, Syed AR, et al. Sensitive oligonucleotide ligation assay for low-level detection of nevirapine resistance mutations in human

immunodeficiency virus type 1 quasispecies. *J Clin Microbiol* 2007;45:2604-15 doi:JCM.00431-07 [pii]

102 . Beck IA, Mahalanabis M, Pepper G, et al. Rapid and sensitive oligonucleotide ligation assay for detection of mutations in human immunodeficiency virus type 1 associated with high-level resistance to protease inhibitors. *J Clin Microbiol* 2002;40:1413-9

103 . Stuyver L, Wyseur A, Rombout A, et al. Line probe assay for rapid detection of drug-selected mutations in the human immunodeficiency virus type 1 reverse transcriptase gene. *Antimicrob Agents Chemother* 1997;41:284-91

105 . Ji H, Masse N, Tyler S, et al. HIV drug resistance surveillance using pooled pyrosequencing. *PLoS One* 2010;5:e9263 doi:10.1371/journal.pone.0009263 [doi]

106 . Jojic V, Hertz T, Jojic N. Population sequencing using short reads: HIV as a case study. *Pac Symp Biocomput* 2008:114-25

107 . Ryscavage P, Kelly S, Li JZ, et al. Significance and clinical management of persistent low-level viremia and very-low-level viremia in HIV-1-infected patients. *Antimicrob Agents Chemother* 2014;58:3585-98 doi:10.1128/AAC.00076-14 [doi]

108 . Konstantopoulos C, Ribaudo H, Ragland K, et al. Antiretroviral regimen and suboptimal medication adherence are associated with low-level human immunodeficiency virus viremia. *Open Forum Infect Dis* 2015;2:ofu119 doi:10.1093/ofid/ofu119 [doi]

109 . Rajian M, Gill PS, Chaudhary U. Prevalence of virological failure amongst WHO-defined immunological failure HIV patients on first line of highly active antiretroviral therapy in a tertiary care hospital in Haryana, India. *Int J Res Med Sci* 2016;4:1613-1619 doi:10.18203/2320-6012.ijrms20161236

110 . Geretti AM, Smith C, Haberl A, et al. Determinants of virological failure after successful viral load suppression in first-line highly active antiretroviral therapy. *Antivir Ther* 2008;13:927-36

111 . Tang MW, Shafer RW. HIV-1 antiretroviral resistance: scientific principles and clinical applications. *Drugs* 2012;72:e1-25 doi:10.2165/11633630-000000000-00000 [doi]

112 . Vandenhende MA, Bellecave P, Recordon-Pinson P, et al. Prevalence and evolution of low frequency HIV drug resistance mutations detected by ultra deep sequencing in patients experiencing first line antiretroviral therapy failure. *PLoS One* 2014;9:e86771 doi:10.1371/journal.pone.0086771 [doi]

113 . Svarovskaia ES, Margot NA, Bae AS, et al. Low-level K65R mutation in HIV-1 reverse transcriptase of treatment-experienced patients exposed to abacavir or didanosine. *J Acquir Immune Defic Syndr* 2007;46:174-80 doi:10.1097/QAI.0b013e31814258c0 [doi]

114 . Kamya MR, Mayanja-Kizza H, Kambugu A, et al. Predictors of long-term viral failure among ugandan children and adults treated with antiretroviral therapy. *J Acquir Immune Defic Syndr* 2007;46:187-93 doi:10.1097/QAI.0b013e31814278c0 [doi]

115 . Reddi A, Leeper SC, Grobler AC, et al. Preliminary outcomes of a paediatric highly active antiretroviral therapy cohort from KwaZulu-Natal, South Africa. *BMC Pediatr* 2007;7:13 doi:1471-2431-7-13 [pii]

116 . Barennes H, Guillet S, Limsreng S, et al. Virological failure and HIV-1 drug resistance mutations among naive and antiretroviral pre-treated patients entering the ESTHER program of Calmette Hospital in Cambodia. *PLoS One* 2014;9:e105736 doi:10.1371/journal.pone.0105736 [doi]

117 . Laprise C, de Pokomandy A, Baril JG, et al. Virologic failure following persistent low-level viremia in a cohort of HIV-positive patients: results from 12 years of observation. *Clin Infect Dis* 2013;57:1489-96 doi:10.1093/cid/cit529 [doi]

118 . Mira JA, Macias J, Nogales C, et al. Transient rebounds of low-level viraemia among HIV-infected patients under HAART are not associated with virological or immunological failure. *Antivir Ther* 2002;7:251-6

119 . Nettles RE, Kieffer TL, Kwon P, et al. Intermittent HIV-1 viremia (Blips) and drug resistance in patients receiving HAART. *JAMA* 2005;293:817-29 doi:293/7/817 [pii]

120 . Peuchant O, Thiebaut R, Capdepont S, et al. Transmission of HIV-1 minority-resistant variants and response to first-line antiretroviral therapy. *AIDS* 2008;22:1417-23 doi:10.1097/QAD.0b013e3283034953 [doi]

121 . Blanco JL, Montaner JS, Marconi VC, et al. Lower prevalence of drug resistance mutations at first-line virological failure to first-line therapy with atripla vs. tenofovir + emtricitabine/lamivudine + efavirenz administered on a multiple tablet therapy. *AIDS* 2014;28:2531-9 doi:10.1097/QAD.0000000000000424 [doi]

122 . Johnson JA, Li JF, Wei X, et al. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. *PLoS Med* 2008;5:e158 doi:10.1371/journal.pmed.0050158 [doi]

123 . Castor D, Vlahov D, Hoover DR, et al. The relationship between genotypic sensitivity score and treatment outcomes in late stage HIV disease after supervised HAART. *J Med Virol* 2009;81:1323-35 doi:10.1002/jmv.21500 [doi]

124 . Gonzalez-Serna A, Min JE, Woods C, et al. Performance of HIV-1 drug resistance testing at low-level viremia and its ability to predict future virologic outcomes and viral evolution in treatment-naive individuals. *Clin Infect Dis* 2014;58:1165-73 doi:10.1093/cid/ciu019 [doi]

125 . Sorstedt E, Nilsson S, Blaxhult A, et al. Viral blips during suppressive antiretroviral treatment are associated with high baseline HIV-1 RNA levels. *BMC Infect Dis* 2016;16:305,016-1628-6 doi:10.1186/s12879-016-1628-6 [doi]

126 . Deeks SG, Barbour JD, Grant RM, et al. Duration and predictors of CD4 T-cell gains in patients who continue combination therapy despite detectable plasma viremia. *AIDS* 2002;16:201-7

127 . Karlsson AC, Younger SR, Martin JN, et al. Immunologic and virologic evolution during periods of intermittent and persistent low-level viremia. *AIDS* 2004;18:981-9 doi:00002030-200404300-00005 [pii]

128 . Sklar PA, Ward DJ, Baker RK, et al. Prevalence and clinical correlates of HIV viremia ('blips') in patients with previous suppression below the limits of quantification. *AIDS* 2002;16:2035-41

129 . Ostrowski SR, Katzenstein TL, Thim PT, et al. Low-level viremia and proviral DNA impede immune reconstitution in HIV-1-infected patients receiving highly active antiretroviral therapy. *J Infect Dis* 2005;191:348-57 doi:JID33223 [pii]

130 . Biomerieux NucliSENS easyMAG Overview. Available at: http://www.biomerieux-diagnostics.com/nuclisensr-easymag. Accessed Feb 20, 2016

131 . QIAamp Viral RNA Mini Handbook Fourth Edition. Accessed Feb 20, 2016

132 . Zhou Z, Wagar N, DeVos JR, et al. Optimization of a low cost and broadly sensitive genotyping assay for HIV-1 drug resistance surveillance and monitoring in resource-limited settings. *PLoS One* 2011;6:e28184 doi:10.1371/journal.pone.0028184 [doi]

133 . Yang C, McNulty A, Diallo K, et al. Development and application of a broadly sensitive dried-blood-spot-based genotyping assay for global surveillance of HIV-1 drug resistance. *J Clin Microbiol* 2010;48:3158-64 doi:10.1128/JCM.00564-10 [doi]

134 . Qiagen QIAxcel advanced system. Available at: https://www.qiagen.com/ca/shop/automated-solutions/dna-analysis/qiaxcel-advanced-system/#productdetails.

135 . Nextera XT DNA: Sample Preparation Guide. *illumina* 2012

136 . Ponstingl H. Wellcome Trust Sanger Institute: SMALT. *Genome Research Ltd* 2010-2015

137 . Katzenstein TL, Ullum H, Roge BT, et al. Virological and immunological profiles among patients with undetectable viral load followed prospectively for 24 months. *HIV Med* 2003;4:53-61 doi:0119 [pii]

138 . Young J, Rickenbach M, Calmy A, et al. Transient detectable viremia and the risk of viral rebound in patients from the Swiss HIV Cohort Study. *BMC Infect Dis* 2015;15:382,015-1120-8 doi:10.1186/s12879-015-1120-8 [doi]

139 . Wacker MJ, Godard MP. Analysis of one-step and two-step real-time RT-PCR using SuperScript III. *J Biomol Tech* 2005;16:266-71 doi:16/3/266 [pii]

140 . . SuperScript III Platinum Two-Step qRT-PCR kit. *invitrogen* 2010

141 . Stahlberg A, Kubista M, Pfaffl M. Comparison of reverse transcriptases in gene expression analysis. *Clin Chem* 2004;50:1678-80 doi:10.1373/clinchem.2004.035469 [doi]

142 . Peters IR, Helps CR, Hall EJ, et al. Real-time RT-PCR: considerations for efficient and sensitive assay design. *J Immunol Methods* 2004;286:203-17 doi:10.1016/j.jim.2004.01.003 [doi]

143 . Singh VK, Govindarajan R, Nail S, et al. The effect of hairpin structure on PCR amplification efficiency. *Molecular Biology Today* 2000;1:67-69

144 . Mohri M, Shakeri H, Zadeh SL. Effects of common anticoagulatns (heparin, citrate and EDTA) on routine plasma biochemistry of cattle. *Clin Pathol* 2007;16:204-209 doi:10.1007/s00580-006-0664-9

145 . Azam M, Malik A, Rizvi M, et al. Emergence of drug resistance-associated mutations in HIV-1 subtype C protease gene in north India. *Virus Genes* 2013;47:422-8 doi:10.1007/s11262-013-0961-8 [doi]

146 . Vairo F, Nicastri E, Liuzzi G, et al. HIV-1 drug resistance in recently HIV-infected pregnant mother's naive to antiretroviral therapy in Dodoma urban, Tanzania. *BMC Infect Dis* 2013;13:439,2334-13-439 doi:10.1186/1471-2334-13-439 [doi]

147 . de Mulder M, Yebra G, Martin L, et al. Drug resistance prevalence and HIV-1 variant characterization in the naive and pretreated HIV-1-infected paediatric population in Madrid, Spain. *J Antimicrob Chemother* 2011;66:2362-71 doi:10.1093/jac/dkr305 [doi]

148 . Raboud JM, Rae S, Woods R, et al. Consecutive rebounds in plasma viral load are associated with virological failure at 52 weeks among HIV-infected patients. *AIDS* 2002;16:1627-32

149 . Ibrahim K, Recordon-Pinson P, Malvy D, et al. Intermittent HIV-1 viremia (Blips) and virological failure in a cohort of people living with HIV from Sao Paulo, Brazil. *AIDS Patient Care and STDs* 2012;26:512-515

150 . Gianotti N, Galli L, Racca S, et al. Residual viraemia does not influence 1 year virological rebound in HIV-infected patients with HIV RNA persistently below 50 copies/mL. *J Antimicrob Chemother* 2011;67:213-217

151 . Charpentier C, Landman R, Laouenan C, et al. Persistent low-level HIV-1 RNA between 20 and 50 copies/mL in antiretroviral-treated patients: associated factors and virological outcome. *J Antimicrob Chemother* 2012;67:2231-5 doi:10.1093/jac/dks191 [doi]

152 . Masquelier B, Pereira E, Peytavin G, et al. Intermittent viremia during first-line, protease inhibitors-containing therapy: significance and relationship with drug resistance. *J Clin Virol* 2005;33:75-8 doi:S1386-6532(04)00346-4 [pii]

153 . Munthali C, Taegtmeyer M, Garner PG, et al. Diagnostic accuracy of the WHO clinical staging system for defining eligibility for ART in sub-Saharan Africa: a systematic review and meta-analysis. *J Int AIDS Soc* 2014;17:18932 doi:10.7448/IAS.17.1.18932 [doi]

154 . Lewis J, Walker AS, Castro H, et al. Age and CD4 count at initiation of antiretroviral therapy in HIV-infected children: effects on long-term T-cell reconstitution. *J Infect Dis* 2012;205:548-56 doi:10.1093/infdis/jir787 [doi]

155 . Leierer G, Grabmeier-Pfistershammer K, Steuer A, et al. Factors Associated with Low-Level Viraemia and Virological Failure: Results from the Austrian HIV Cohort Study. *PLoS One* 2015;10:e0142923 doi:10.1371/journal.pone.0142923 [doi]

156 . Fassinou P, Elenga N, Rouet F, et al. Highly active antiretroviral therapies among HIV-1-infected children in Abidjan, Cote d'Ivoire. *AIDS* 2004;18:1905-13 doi:00002030-200409240-00006 [pii]

157 . Rutstein RM, Gebo KA, Flynn PM, et al. Immunologic function and virologic suppression among children with perinatally acquired HIV Infection on highly active antiretroviral therapy. *Med Care* 2005;43:III15-22 doi:00005650-200509001-00003 [pii]

158 . Lawn SD. AIDS in Africa: the impact of coinfections on the pathogenesis of HIV-1 infection. *J Infect* 2004;48:1-12 doi:S0163445303001683 [pii]

159 . von Reyn CF, Kimambo S, Mtei L, et al. Disseminated tuberculosis in human immunodeficiency virus infection: ineffective immunity, polyclonal disease and high mortality. *Int J Tuberc Lung Dis* 2011;15:1087-92 doi:10.5588/ijtld.10.0517 [doi]

160 . Bello EJ, Correia AF, Marins JR, et al. Predictors of Virologic Failure in HIV/AIDS Patients Treated with Highly Active Antiretroviral Therapy in Brasilia, Brazil During 2002-2008. *Drug Target Insights* 2011;5:33-41 doi:10.4137/DTI.S7527 [doi]

161 . Ahoua L, Guenther G, Pinoges L, et al. Risk factors for virological failure and subtherapeutic antiretroviral drug concentrations in HIV-positive adults treated in rural northwestern Uganda. *BMC infectious diseases* 2009;9

162 . Sulkowski M, Chaisson R, Karp C, et al. The effect of acute infectious illnesses on plasma human immunodeficiency virus type I load and the expression of serologic markers of immune activation among HIV-infected adults. *J Infect Dis* 1998;178:1642-1648

163 . Dollfus C, Le Chenadec J, Faye A, et al. Long-term outcomes in adolescents perinatally infected with HIV-1 and followed up since birth in the French perinatal cohort (EPF/ANRS CO10). *Clin Infect Dis* 2010;51:214-24 doi:10.1086/653674 [doi]

164 . Van Dyke RB, Patel K, Siberry GK, et al. Antiretroviral treatment of US children with perinatally acquired HIV infection: temporal changes in therapy between 1991 and 2009 and predictors of immunologic and virologic outcomes. *J Acquir Immune Defic Syndr* 2011;57:165-73 doi:10.1097/QAI.0b013e318215c7b1 [doi]

165 . Williams PL, Storm D, Montepiedra G, et al. Predictors of adherence to antiretroviral medications in children and adolescents with HIV infection. *Pediatrics* 2006;118:e1745-57 doi:peds.2006-0493 [pii]

Chapter 7: Appendices

Appendix 1: Selection and correlation of data variables

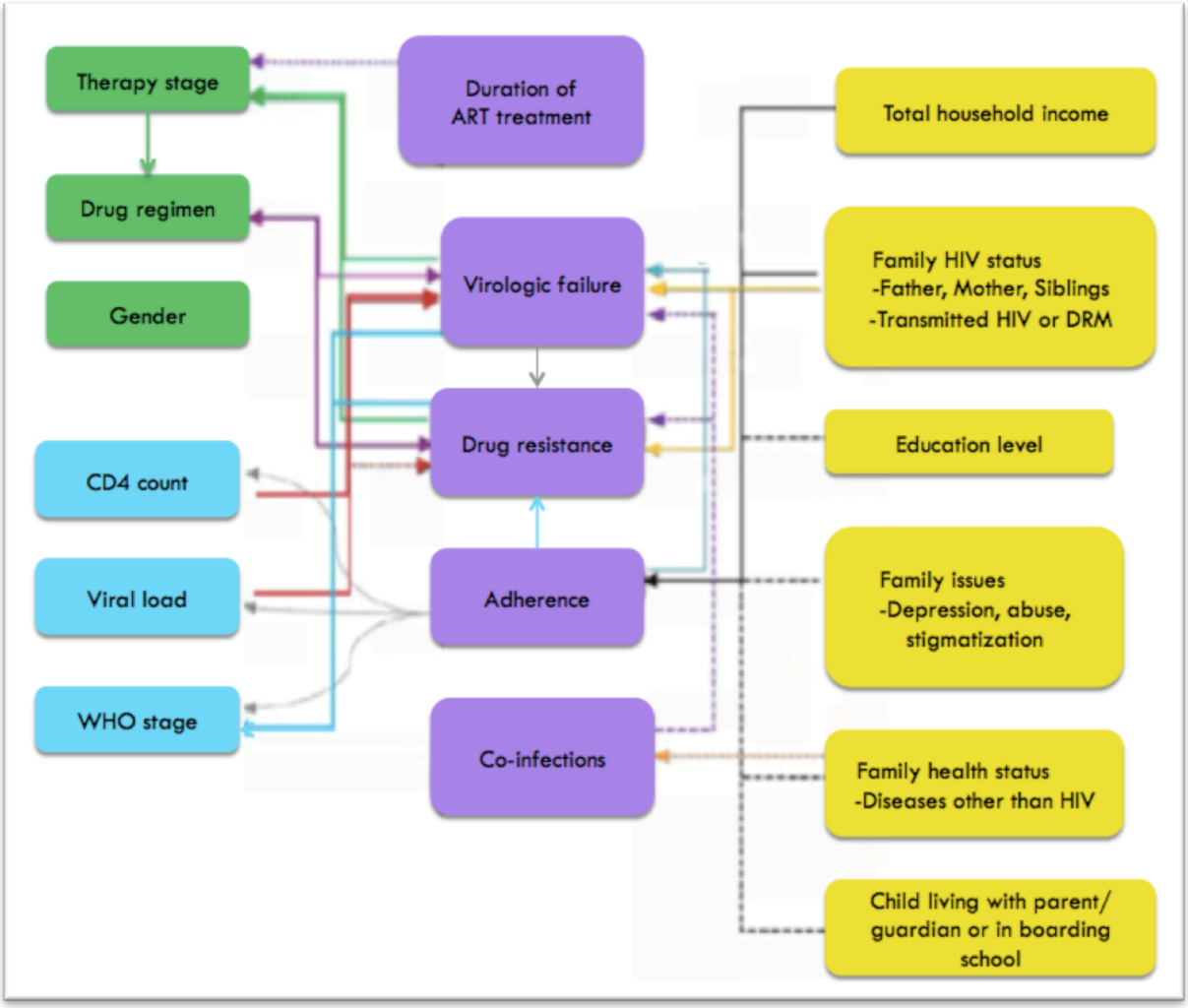


Figure A1: Data variables flow diagram. This was used to get an understanding of the relationship between key variables from raw database and to further standardize the variables for "clean database" to achieve the objectives of this study. An association is shown between variables that have a direct and indirect impact on virologic failure. Green = Independent variable, Blue = Dependent variable, Purple = Derived variables, Yellow = Social data, straight line= major effect, dashed line = minor effect.

Appendix 2: HyDRA parameters

Table A1: HyDRA pipeline parameters used for DRM identification.

Consensus percent	20
Mutation database	Stanford HIV DR database and SDRM 2009
Target coverage	1000
Length cutoff	100
Score cutoff	30
Error rate	0.0021
Minimum variant quality	30
Minimum read depth	100
Minimum allele count	5
Minimum AA frequency	0.01

Appendix 3: Variables selected for statistical analysis

Table A2: List of existing and derived variables for statistical analysis.

	Existing Variables	**Derived Variables**
1	Client admission number	Individual order in dataset
2	Clinic	Unique id
3	Gender	Age
4	Date of birth	Duration of ART
5	ART start date	Number of blood tests per patient
6	ART stop date	Viral load category
7	ART restart/switch date	Low level viremia category
8	Date of admission	Adherence category
9	Blood test date	WHO stage category
10	CD4 T-cell count	Persistent low level viremia interval
11	Viral load	Time to start ART
12	Adherence	Blood test interval
13	WHO stage	Year of birth
14	Therapy stage	
15	AZT	
16	D4T	
17	3TC	
18	NVP	
19	EFV	
20	DDL	
21	ABC	
22	Kaletra/LPV	
23	TDF	
24	Pneumonia	
25	TB	
26	Malaria	
27	Malnutrition	

Appendix 4: Manuscripts from the project

Gupta, Shivani; Taylor, Tracy; Aileen, Patterson; Brookes, James; Liang, Binhua; Bullar, Jared; Sandstrom, Paul; Domselaar, Gary V.; Ji, Hezhao. A robust PCR protocol for HIV drug resistance testing on low-level viremia samples. [In preparation]