ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAMME

FACULTY OF SCIENCE

STATISTICS DEPARTMENT

SURVIVAL ANALYSIS OF TIME TO TREATMENT RESUMPTION FOR CHRONIC HIV-1 PATIENTS INTERRUPTING HIGHLY ACTIVE ANTIRETROVIRAL THERAPY (HAART)

By

Berhanu Anagaw

A thesis submitted to the school of Graduate Studies of Addis Ababa University in partial fulfillment of the requirements for the Degree of Masters of Science in Biostatistics

June, 2009

ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAMME

FACULTY OF SCIENCE

STATISTICS DEPARTMENT

SURVIVAL ANALYSIS OF TIME TO TREATMENT RESUMPTION FOR CHRONIC HIV-1 PATIENTS INTERRUPTING HIGHLY ACTIVE ANTIRETROVIRAL THERAPY (HAART)

By

Berhanu Anagaw

Approved by the board of examiners:

Sileshi Fanta _____

Chairman, Department graduate committee Signature

Emmanuel G/Yohannes (Ph.D.) _____

Examiner Signature

Mekonnen Taddesse (Asst. Prof.) _____

Examiner Signature

Tel. 251-1-239465 **P.O.Box 1176 Addis Ababa, Ethiopia**

Dedication

To the memory of my brother:

Sewalem Anagaw

God may his gentle soul rest in perfect peace.

Acknowledgment

First of all, I would like to give thanks to Almighty God for his savior, without whom none of this world have been possible and who has given me the opportunity to go through this way and reach this success.

My heartfelt gratitude goes to Dr. Fentaw Abegaz, my advisor and instructor, for his immense and invaluable contribution in terms of constant advice and guidance in the best way of handling the research work, excellent scientific guidance and tireless efforts to make this work a reality.

My sincere appreciation and thanks also go to Zewditu Memorial Hospital, ART members and clinic staff, particularly to Dr. Aster Shewamare. I also thank my instructors and all the rest staffs of Department of Statistics for their unreserved knowledge sharing and cooperation. My appreciation also goes to my colleagues for their moral supports and making the atmosphere conducive during the programme and their endurance with one another at all time.

It is my deepest and warm gratitude to my family that has been a source of pride and encouragement throughout my work. I am thankful to my parents (Anagaw Wubie and Fentaye Alamneh) for their encouragement and prayer and to all of my brothers especially to Belay Anagaw, Zelalem Anagaw and his wife Tsigie Zeleke who were always there to encourage and provided me with continuing motivation and encouragement, and whose critical and constructive advice helped immensely to complete this work.

Finally, I acknowledge those that I used their materials for the completion of this study in one way or the other.

Contents

List of Tables **Page**

List of Figures **page**

List of Acronyms

ACTG - AIDS Clinical Trial Group

ADRs - Adverse Drug Reactions

AIDS - Acquired Immuno Deficiency Syndrome

ANC - Antenatal Care

ART - Antiretroviral Therapy

ARV - Antiretroviral

AZT - Azidothymidine

BSS - Behavioral Surveillance Survey

CDC - Centre for Disease Control and Prevention

DHS - Demographic and Health Survey

FDA - Food and Drug Administration

FDRE - Federal Democratic Republic of Ethiopia

FHAPCO - Federal HIV/AIDS Prevention and Control Office

FI- Fusion Inhibitor

HAART - Highly Active Antiretroviral Therapy

HIV - Human Immunodeficiency Virus

HR - Hazard Ratio

IDU - Injection Drug User

IQR - Inter Quartile Range

ITECH - International Training and Education Centre on HIV

KM - Kaplan-Meier

MOH - Ministry of Health

MSH-RPM+ - Management Science for Health-Rational Pharmaceutical Management Plus

NAT - Netherlands Australia Thailand Research Collaboration and the Thai Red Cross AIDS Research Centre

NNRTI - Non-nucleoside Reverse Transcriptase

NRTI - Nucleoside Reverse Transcriptase

OARAC - Office of AIDS Research Advisory Council

PH-Proportional Hazards

PI(s) - Protease Inhibitors

PLWHA - People Living With HIV/AIDS

PYO - Person Years of Observation

RNA - Ribonucleic Acid

RR - Relative Risk

STI - Structured Treatment Interruption

TB - Tuberculosis

TI - Treatment Interruption

TR - Treatment Resumption

UKCHIC - United Kingdom Collaborative HIV cohort

UNAIDS - Joint United Nations Programme on AIDS

VL - Viral Load

WHO - World Health Organization

ZMH - Zewditu Memorial Hospital

Abstract

Highly active antiretroviral therapy (HAART) has significantly reduced mortality caused by human immuno-deficiency virus (HIV) by enhancing the physiological and immunological ability to counterattack against the virus and increases the life expectancy. Once started, the antiretroviral treatment should be continued lifelong and adherence to this treatment should be nearly perfect to enable long-term efficacy. A continuous and life-long treatment with HAART may lead to a broad spectrum of significant toxicities. As a result many patients interrupt their HAART without the knowledge and advice of the clinicians and this cause the patient immune to degrade and even cause to death. Therefore, this study is an attempt to examine the predictive factors of interruption duration in patients who interrupt their ARV drugs treatment.

The study is a retrospective cohort study of HIV/AIDS patients under HAART but discontinued their treatment for at least one month, which comprises a total of 723 patients from Zewditu Memorial Hospital. Of these patients about 67% resumed their treatment. In the analysis of data the Kaplan-Meier survival estimator and Cox proportional hazards regression model are used. Based on the data analysis, it is found that main effects: education, baseline age, baseline weight, CD4 count at the start of HAART and prior to interruption, duration on HAART and interaction effects: marital status with disease duration and marital status with sex are important factors that are related to time to resumption of HAART.

In general, HIV/AIDS patients who have prolonged treatment interruption are characterized by having lower level of education, younger age, high weight, higher CD4 count at the start of HAART, lower CD4 count prior to treatment interruption, longer duration on treatment follow-up and currently not married and male patients with longer disease duration.

CHAPTER ONE
INTRODUCTION

1.1 Background

In countries most heavily affected by HIV life expectancy has reduced by more than 20 years, economic growth slowed and household poverty deepened. There are two types of HIV: namely HIV-1 and HIV-2, both types are transmitted mainly by sexual contact, through blood and from mother to child and they appear to cause clinically indistinguishable AIDS. However, it seems that HIV-2 is less easily transmitted and the period between initial infection and illness is longer in the case of HIV-2. World wide, the predominant virus is HIV-1, and generally when people refer to HIV without specifying the type of virus they refer to HIV-1. The relatively uncommon HIV-2 type is concentrated in West Africa and is rarely found elsewhere[1].

Almost 27 years have now elapsed since the initial description the human immunodeficiency virus type-1 (HIV-1) in 1983 (Barre-sonousssi *et al*, 1983, Broder and Gallo, 1984) and HIV-2 in (Clavel *et al*, 1986). Of these two viruses HIV-1 is the major cause of AIDS in the world today. During these years HIV infection has changed from a fatal condition to a manageable chronic illness and the development of antiretroviral therapy (ART) has been one of the dramatic advances in the history of medicine. The aims of the ART programme are to reduce HIV related morbidity, mortality, and mitigate the impact of the AIDS epidemic. However, for the vast majority of people living with HIV/AIDS, ART take decades to access in resource-poor countries where HIV continues to devastate families, communities and societies, especially the poor and the socially marginalized.

The rate of the spread of the HIV/AIDS and the damages accompanying it have reached a level which shock economists, health workers, politicians, etc that it has become a world wide issue in general and developing countries in particular. The disease being one without any cure is still accountable for economic, social and health crisis especially in developing countries. Its high prevalence and/or distribution among the youth made things even more complicated. HIV remains a global health problem of unprecedented dimensions. HIV has already caused and estimated 2.5 million deaths worldwide and has generated profound demographic changes in the most heavily affected countries.

[1] (http://www.avert.org/)

The most recent international epidemiological data contain some good news. In some countries in Asia, Latin America and Sub-Saharan Africa, the annual number of new HIV infections is falling. The estimated rate of AIDS death has also declined, in part as a result of success in expanding access to antiretroviral drugs in resource limited settings such as few laboratory facilities to test viral load and drug resistance and those that exist are faced with a shortage of trained staff and costly reagents. Yet these favorable trends are not uniformly evident, either within or between regions, underscoring the need for more comprehensive progress in implementing the effective policies and programmes.

Ethiopia is one of the countries hardest hit by HIV/AIDS epidemic. This is especially significant as the country is with a population estimated at over 77 million, the second most populous nation in sub-Saharan Africa resulting in high magnitude HIV-infected population. About 85% of the population lives in rural areas, and approximately one-fifth are aged 15-24 years. HIV/AIDS was first recognized in the country in the mid-1980's, at about the same time as in other countries in the region. Efforts to collect epidemiological data began shortly thereafter, and there are many studies from the late 1980s and 1990s reporting prevalence data and risk factors in a number of high-risk groups. In recent years, the main source of information about HIV in Ethiopia has been antenatal care (ANC) clinic based sentinel surveillance, with surveys being conducted and published at two year intervals, most recently in 2005. In addition, useful epidemiological information at national level and for specific communities can be extracted from the Demographic and Health Surveys (DHS) and Behavioral Surveillance Surveys (BSS), the most recent versions of which were also completed in 2005. The 2005 ANC surveillance based HIV prevalence among the general adult population was estimated at 3.5% that is 10.5% for urban and 1.9% for rural areas. Results of the 2005 Ethiopian Demographic and Health survey, on the other hand, indicated a national level adult prevalence rate of 1.4%, the prevalence in women was 1.9% and that of men was 0.9% in 2005 (HAPCO and GAMET, 2008).

An estimated 22 million adults and children were living with HIV in Sub-Saharan Africa at the end of 2007. In the year 2007, an estimated 1.5 million Africans died from AIDS. The epidemic has left behind some 11.6 million orphaned African children. Specifically in Ethiopia the estimated number of adults and children living with HIV/AIDS were about 980,000 and 67,000 deaths due to AIDS and 650,000 million Ethiopian children are orphaned due to this epidemic at the end of 2007 (UNAIDS, 2008).

There is no any field of medicine that has been through such dramatic developments as that of antiretroviral therapy. Few other areas have been subject to such fast and short-lived trends. Those who have experienced the rapid developments of the last few years have been through many ups and downs. Following the hope of the early years, from 1987-1990, and the modest successes with mono-therapy (Volberding *et al.*, 1990, Fischl, *et al.,* 1990), the results of the Concorde Study (Concorde, 1994) plunged both patients and clinicians into a depression that was to last for several years. Many patients, who were infected up until the mid-80s, began to die. Medical care was established, as well as more and more support groups and ambulatory nursing services. One became accustomed to AIDS and its resulting death toll.

There was, however, definite progress in the field of opportunistic infections. Cotrimoxazole, Pentamidine, Gancyclovir, Foscarnet, and Fluconazole saved many patients lives, at least in the short-term. Then, in September 1995, the results of the European-Australian DELTA Study (Delta, 1996) and the American ACTG 175 Study (Hammer *et al*, 1996) attracted attention. It became apparent that two nucleoside analogs were more effective than monotherapy (Hoffmann *et al*, 2007). As a result of these treatments mortality among the patients declined from 29.4 per 100 person-years in 1995 to 8.8 per 100 person-years in the second quarter of 1997 in USA (Palella *et al*, 1998). Both studies demonstrated that it was potentially of great importance of starting treatment immediately with two nucleoside analogs, as opposed to using the drugs successively.

Although some severely ill patients with AIDS managed to recover during these months, for many the combinations that were now - at the beginning of 1996 - widely used, came too late. Then at the mid of the year (June, 1996), the World AIDS Conference in Vancouver reported on the new “AIDS cocktails” and the strangely unscientific (and rather ridiculous) expression “highly active antiretroviral therapy” (HAART) began to spread irreversibly. Clinicians were only too happy to become infected by this enthusiasm and the introduction of HAART has transformed HIV infection in to a chronic manageable disease. Although, it still seemed no-medication ten years ago, it is now realistic to expect to control HIV for the longer term. However, the success of the antiretroviral treatment (ART) is highly dependent on the adherence to complex antiretroviral drug (ARV) regimens. Unfortunately up to 25% of the patients discontinue their HAART regimen mainly because of toxic effects.

The goal of ARV drug therapy is to improve survival; to reduce HIV associated morbidity and mortality, to increase the quality of life, to restore immune function and to achieve maximal and sustained suppression of viral replication (OARAC, 2008). In 2003 WHO/UNAIDS "3 by 5" initiative had set a target of putting 3 million people on ART by the end of 2005 in low income countries, but only 1.3 million people have started ART which is 55% of the target. By 2010 WHO has planed to put 9.8 million people on ART with the goal of providing universal access to HIV care and ART (UNAIDS, 2006).

In Ethiopia the Federal Ministry of Health has estimated 250,000 – 300,000 persons to be eligible for treatment and it had planned to put 74,000 people on ARV treatment by the end of 2006. But only about 55,000 people are getting ART at the end of 2006 (Ministry of Health, 2006). Recently, the Ministry of Health plans to make free ARV treatment available to 320,000 persons by 2008. A phase by phase approach will be used to rollout the ART program throughout the nation. HIV regional prevalence, population density and regional equity were taken as criteria to determine the number of ART sites per region (FHAPCO, 2007).

However, many patients who already harbor drug-resistant HIV require interruption of HAART due to poor compliance, poor quality of life, toxicity or development of resistance, to have a break from taking treatment every day or to re-stimulate the immune system by allowing viral loads to rise. Despite the predicted benefits of HAART interruptions such as reducing cost, time on antiretroviral therapy, and to reduce toxicities associated with therapy, there are a number of risks involved in stopping treatment, some of which can be serious. In addition to rises in viral load and falls in CD4 cell count, recent studies have shown that some treatment interruption strategies put patients at a high risk of disease progression or death.[2]

1.2 Statement of the problem

A medical crisis as socially complex and far-reaching as the HIV/AIDS pandemic stirs much controversy about what strategy will best forestall disaster. Epidemiologically, there are two approaches to managing widespread disease: prevention and treatment. Many public health programs have emphasized prevention measures over treatment provision in developing

[2] http://www.aidsmap.com/cms1032040 and 1039438

countries primarily because they are less expensive. While prevention is absolutely vital in the fight against HIV/AIDS, especially where resources are limited, this uneven focus may have promoted the appallingly inadequate medical care available to the same populations.

Fortunately, access to critical treatment is also expanding throughout the world. The efforts of governments, non-profit organizations and international funding groups have aided countless people by providing the life-prolonging therapies like HAART. However, we must also realize that the initiation of HIV/AIDS treatment is exactly that, the beginning. Because HIV/AIDS is a chronic condition, treatment must be maintained for the rest of the patient's life. This sustained adherence depends on a multitude of factors, many of which fall outside of the patient's immediate control. From economic pressures to cultural misconceptions, patients are confronted by considerable obstacles that could jeopardize their antiretroviral treatment.

A few years ago treatment interruptions was not considered as treatment options and patients who interrupt therapy generally did so of their own will, which was described as non-adherence. Drug-holidays were occasionally recommended though for example to allow resolution of drug toxicity, but it was assumed that once started, antiretroviral treatment was a life long commitment. Today antiretroviral treatment may be interrupted for the following reasons, such as: request for treatment interruption by the patient, carelessness or non-adherence, adverse effects or toxicity, change of treatment guidelines.

The major goal of HIV therapy is to maintain long-term of health of the patient, while avoiding drug toxicity and preserving viable future treatment options. Many of the studies till now has been done on the assessment of the prevalence and adherence and its determinants of patients to ART, the factors that influence the survival/death status and has been done of a person living with the virus and is under the follow up of ART. However, little has been done on patients stopping their ART and patients interrupt and resume back to ART which mainly put patients at a high risk of disease progression or death. Therefore, this study is motivated to investigate the determinants of treatment interruption duration of patients with chronic HIV-1 infection who stopped their highly active antiretroviral therapy for reasons other than treatment failure and which factors are predictive of early or late resumption to treatment.

1.3 Objectives of the Study

- **General objective**
  - To examine determinant factors of treatment resumption (TR) within a cohort of HIV-1 infected patients having stopped their treatment and resumed back to their Highly Active Antiretroviral treatments (HAART) for reasons other than treatment failure. Using survival analysis based on non-parametric and semi-parametric methods.

- **Specific objectives**:
  - To describe how the risk factors for patients who stopped their Highly Active Antiretroviral treatments affect resumption to their treatment.

  - To characterize the magnitude of risk related to socio-demographic and immunological factors on time to treatment resumption after an unstructured treatment interruption in adolescents and old age peoples with chronic HIV-1 infection.

  - To estimate time to resumption of Highly Active Antiretroviral Therapy (HAART).

1.4 Application of the Results

- ✓ The findings from this study are expected to give information to government and non-governmental organizations working in the areas of giving care, support and treatment for HIV/AIDS patients to make a policy and plan strategy on the area.

- ✓ The result helps both government and non-government organizations and different donors to understand and look ways to avoid or reduce unstructured HAART treatment interruption.

- ✓ It also gives information for patients, who are in the HAART follow-up and have good clinical/immunological records, which factors have influence on the time to treatment resumption.

- ✓ It is also believed that results of the research will be used as a basis for further study in the area.

1.5 Limitation of the study

- The study used a cross sectional study data from a single hospital. The result may not be representative of the national picture, since the study employed in one hospital with relatively well organized and started ART clinical service early, before the others start.
- There was no enough mechanism in place to trace patients lost to follow up (dropped) and dead. Therefore, the reason for some of those lost to follow up and death was not known which could bias the finding.
- The diagnosis and severity were taken as recorded on the clinical registers or records. The clinical records were very often incomplete lacking the important socio-demographic and clinical variables. It was difficult to find consistent information in the follow up report from the clinical records of patients who started developed ART. There is a tendency to record ART side effects which are severe enough to change the drug or discontinue treatment.
- Another limitation of our study is that we were unable to corroborate patients interrupting their HAART viral load measurements. Most rarely recorded viral load results we found from the patients' record were done at different private institutions. Besides not being able to assess the validity of the measurements.
- Lack of literature on our country related to the subject under study.

CHAPTER TWO
CONCEPTS AND DEFINITIONS

2.1. Background on HIV/AIDS

HIV stands for Human Immunodeficiency Virus; reversing this wording helps illustrate its definition as it is a virus that specifically targets and devastates the immune system in humans. As with all viruses, HIV is a particle many times smaller than a cell, which contains genetic material protected by a protein envelope. Once HIV enters the body, it seeks out its target that is our immune system. The human immune system is a network of vessels and specialized junctions through which cells called leukocytes (white blood cells) monitor physiological conditions and share information about sites of damage and foreign invasion. These leukocytes come in many forms so they can successfully identify and manage a variety of physiological problems. When HIV infects humans, the cell it infects most often is CD4 cells. The virus becomes part of the cells and when the cell multiplies to fight an infection, this virus also multiplies. When someone is infected with HIV for a long time, the CD4 count goes down. This is a symptom that the immune system is weakened.

While the initial phase of the disease manifests itself with symptoms similar to the common cold, the HIV infection steadily grows in strength and severity. Eventually, it incapacitates the majority of CD4-positive T-cells (CD4) which are crucial to the function of the immune system as a whole. Proper diagnosis depends on many factors, but when an individual reaches a CD4 count lower than 200 cells/µl of blood, a significant drop from normal levels of about 1000 cells/µl, they are generally diagnosed with Acquired Immunodeficiency Syndrome (AIDS) (CDC, 1992).

The period from initial infection to AIDS take between 10-12 years on average. Clinically, the progression of HIV/AIDS is monitored by laboratory tests which report the concentration of HIV RNA called viral load (VL) and the CD4 count. However, many factors such as time of the day, fatigue, and stress which can affect the test result. Hence, it is highly recommended by experts that it is better to have the blood samples at the same time of the day and using the same laboratory.

The degradation of the immune system does not usually cause death but instead renders the body incapable of fighting off normally harmless pathogens. Opportunistic diseases result from organisms across a broad range, including: protozoa, fungi, bacteria and other viruses, which exploit the body's shattered defenses. AIDS-related illnesses affect multiple regions of the body such as the pulmonary, gastro-intestinal and neurological systems. The types of opportunistic diseases that are most prevalent in a population depend on the region; for example, tuberculosis has an extremely high incidence rate among HIV-positive individuals in many African countries whereas it is relatively rare in the U.S. Aside from opportunistic infections, the direct consequences of AIDS often include chronic diarrhea, wasting syndrome, prolonged fever, dementia and a type of cancer called Kaposi's sarcoma (Weeks and Alcamo, 2006). Consideration of these secondary infections is vital in overall treatment.

2.2 Antiretroviral Treatments

Despite the discouraging outlook of HIV/AIDS progression, hope lies in the fact that viable treatments have been developed over the last two decades to curb the lethal course of this illness. The first antiretroviral drug approved by the U.S. Food and Drug Administration (FDA) in 1987 was zidovudine commonly known as AZT (azidothymidine), originally designed to treat cancer. This drug was successful in that it slowed disease progression and instilled optimism in the HIV/AIDS-affected community. Regretfully, it was later shown that treatment with AZT did not increase overall survival rate (Institute of Medicine 2005a). This outcome spurred drug development in several different directions ultimately producing four distinct classes of drugs namely: nucleoside reverse transcriptase inhibitors (NRTI), AZT included, non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PI) and fusion inhibitors (FI).

Researchers continue to expand on current treatment strategies as well as invent completely novel approaches, yet the most significant advance in alleviating the morbidity and mortality resulting from HIV infection has not been a single drug discovery but the utilization of combination treatment. The use of three or more antiretroviral (ARV) drugs from at least two of the classes is termed highly active antiretroviral treatment (HAART or ART) and is the recommended form of treatment (Institute of Medicine, 2005b). With the advent of HAART in 1996, HIV/AIDS has begun to transform into a manageable chronic illness instead of an acutely lethal infection. The effect of this therapy was immediately apparent; the number of

AIDS related deaths decreased by 43% in the US from 1995 to 1997 (Institute of Medicine, 2005a).

2.3 Highly Active Antiretroviral Treatment (HAART)

The name given to aggressive treatment regimens used to suppress HIV viral replication and the progression of HIV disease. The usual HAART regimen combines three or more different drugs such as two nucleoside reverse transcriptase inhibitors (NRTIs) and a protease inhibitor (PI), two NRTIs and a non-nucleoside reverse transcriptase inhibitor (NNRTI) or other such combinations. These treatment regimens have proven to reduce the amount of active virus and in some cases can lower the number of active virus until it is undetectable by current blood testing techniques.

At first glance, it appears straightforward that the lower the CD4 cell count and the higher the viral load, the higher the risk of AIDS (Mellors *et al,* 1997, Lyles *et al,* 2000), and the more urgent the indication for treatment. International treatment guidelines agree that all symptomatic patients and patients with CD4 cells/µl less than 200 have to be treated. The situation is less clear for asymptomatic patients with CD4 cells/µl more than 200. Therefore, the indication for antiretroviral therapy to start or defer is based on clinical assessment, CD4 cell count and viral load. It is assumed that once a patient started, antiretroviral treatment is a life long commitment as the main goal of HAART is to prolong the life of the patient while avoiding drug toxicity. However, these days HAART may be interrupted for a number of reasons.

A treatment interruption was defined as the absence of any antiretroviral drug at least one month in a patient who was previously receiving HAART. Recently, treatment interruption has been a topic of hot discussion in terms of over possible risks (AIDS, resistance) or advantages (reduction of toxicity and costs). A distinction should be drawn between structured treatment interruptions (STIs), which are made with the knowledge of the treating physician, and unstructured drug holidays, which the patient elects to take.

The reasons for interruption of treatment can differ greatly but the strategy of structured treatment interruption (STI) in HIV infected patients has been discussed for following main reasons: at the patient's request, to stimulate anti-HIV immune response, to raise compliance

by increasing time off therapy, improving quality of life and diminishing toxicity, and in the case of occurrence of multi-resistant virus strains.

However, many of the treatment interruptions occur without the clinician's knowledge. For this reason alone, treatment interruptions are an important constituent of antiretroviral therapies, whether, as a clinician, one approves of them or not. To oppose them means to disregard the realities of treatment. Whatever the reason for interruption, viral load rebounds or CD4 cells decline will be observed with each treatment interruption cycle for many patients. These are considered as undesirable as they may result in high risk of disease progression or death. To this many clinical trials have been carried out in an attempt to come up with treatment interruption protocols that will work fore most individuals.

CHAPTER THREE
LITERATURE REVIEW

3.1 Background

Though, the implementation of ART started since 1990; it was universally recognized that access to effective HIV/AIDS treatment and care till recently was highly inequitable. In 2005, Ethiopia launched free ART; over 71,000 patients were initiated on ART by the end of November 2006 and some 241 hospitals and health centres are now providing HIV care and treatment services in all regions of the country (FHAPCO, 2007; FDRE MOH, 2002 and 2006).

The new generations of HAART offer the potential for long-term suppression of HIV replication; however, the challenge now is to encourage and enable patients to take these medicines correctly, in order to achieve their maximum effect (Morse *et al*, 1991). Adherence to this antiretroviral therapy, however, is critically important for the success of the therapy. Some researchers suggest that near-perfect adherence, i.e., higher than 95%, is necessary to achieve suppression of HIV replication (Paterson *et al*, 2000).

The consequences of not taking the drugs properly may be severe: resistance can develop rapidly, and the potential benefits of the treatment can be entirely lost. There is also the danger of cross resistance-resistance to one medication can result in the decreased effectiveness of many others, restricting future options. The main focus of this chapter is to present reviews of past studies related to HIV patients interrupting and resuming back to HAART and the different statistical methodology employed in these studies.

3.2 General literature about Statistical models used in HAART studies

In recent investigation conducted in Zewditu Memorial Hospital by Assegid and Ahmed (2007) to assess the magnitude of adverse drug reactions (ADRs) and associated factors among patients on ART; out of the total of 554 clinical records, 526(94.9%) were reviewed, at least one of the drugs was changed in 120(22.8%) of the patients, there were 73(13.7%) lost to follow up. Multiple logistic regression model is used in order to assess the independent contribution of some potential risk factors. They found that the main reasons for changing ARV drugs were toxicity/side effects 97(81%) followed by illness in 14(12 %), and

pregnancy in 4(3%). The frequency of ADRs was 124(24%). The most frequently diagnosed ADRs were Anemia 42(33.9%) followed by Peripheral Neuropathy 35 (28.2%) and Elevated ALT 31(25%).

Similarly, in the prospective study of 97 urban Ugandan patients with advanced purchasing ART, Oyugi, *et al.* (2007) using logistic regression to assess predictors of adherence and survival with viral suppression found that high rate of adherence of 90% or greater had a potential association with survival with viral load suppression. Treatment interruptions of more than 48h, largely as a result of financial or structural barriers of securing treatment, explained 90% of all missed doses and predicted drug resistance. In this study, mortality was about 10% during the first 24 weeks of treatment and most deaths occurred shortly after treatment initiation which is similar to Laurent *et al* (2005).

Factors associated with viral load rebound were identified using poison regression by Bansi *et al.* (2008). The study investigated whether previous treatment interruptions were associated with a raised risk of viral rebound in individuals who have attained virology suppression on the UK collaborative HIV cohort (UK CHIC) study and indicated that the rates of viral rebound were up to 64% higher (rate ratio 1.64; CI, 1.43, 1.88) in those who had previously interrupted therapy compared with those who had not. Patients who had interrupted at detectable viral loads had up to a 74% (1.74; CI, 1.42, 2.14) higher chance of rebounding compared with those who had not interrupted with a detectable viral load. They found no evidence to suggest interrupting treatment at an undetectable viral load was associated with viral rebound.

In addition some studies like Olsen, *et al.* (2007) used poison regression to compare incidence rates of deaths due to AIDS in patients with and without treatment interruption and to describe the relationship between a treatment interruption and disease progression. They observed 1243 interruptions and 403 deaths due to AIDS. The incidence rate of deaths due to AIDS was higher in patients with lower CD4 cell counts or higher viral loads, regardless of treatment interruption and the incidence rates was significantly higher in the treatment interruption (TI) group than in non- treatment interruption group. The factors associated with disease progression in the TI of ≥ 3 months were patients with a current CD4 count of 200 cells/µl or lower had a 2.95 fold higher incidence of new AIDS/death (P = 0.0003) while patients with a CD4 count above 350 cells/µl had almost an 80% lower incidence when

compared with patients with a CD4 count between 201 and 350 cells/µl and among patients experiencing one or more TIs, a current higher viral load was associated with a 36% increased incidence of clinical disease progression (P=0.0007).

Degu and Bernt (2005) investigated disease progression among untreated HIV-infected patients in South Ethiopia. Similarly, Degu *et al*. in (2006) attempted to find out the effect of the HAART on mortality and tuberculosis incidence rates. Both studies employed Kaplan-Meier method to determine the event-free survival, log-rank test to compare survival experience and Cox-proportional hazards regression model in order to identify predictors of death in the study and they reached the mortality rate was 15.4 per 100 person years of observation (PYO) in the HAART group and the tuberculosis (TB) incidence rates was 3.7 per 100 PYO in addition HAART resulted in a 65% decline in mortality and TB incidence rate was lower in the HAART group. A study carried out by Giuntini *et al.* (2005) examined the effect of prolonged treatment interruption in chronic HIV infection on the rate of change in the CD4 cell count during treatment interruption and found that the main predictive factor of treatment interruption seems to be the CD4 cell nadir (the lowest CD4 at the start of interruption), patients with a CD4 count greater than 350 were much less likely to resume therapy than those with a lower CD4 cell nadir greater than 350 who were under HAART in the previous time by using the statistical methods like Kaplan-Meier and linear regression.

Similarly, Cox proportional hazards model were by Fernandez-Guerrero *et al* (2005) used to determine which factors were independently related to a longer time reinitiation of HAART, the categorical variables considered were sex, exposure group, and presence of chronic hepatitis B and the continuous variables were age, duration of ART, duration of interruption, CD4 HAART. They found that the factors associated with a higher probability of reinitiating therapy were long duration of undetectable viral load during treatment (hazard ratio 1.065; P<.003) and low CD4+ cell count at the time of therapy interruption (HR, 0.998; P<.048). Similarly, Cox regression analysis were used by Wood *et al* (2006) to evaluate the impact of baseline plasma HIV RNA on the survival among patients with CD4 cell counts ≥ 200 cells/µl. In the study, analysis were restricted to HIV infected men and women who were first prescribed triple drug ART therapy through recording linkage carried out with the British Columbia division vital statistics and the primary end point in this analysis was time to death. A baseline HIV RNA ≥100,000 copies/ml was statistically associated with elevated mortality among non-adherent patients.

A study carried out by Toulson *et al* (2005) in a systematic chart or database review to identify patients with nadir CD4 cell counts ≥ 200 cells/mm^3 and without acquired immunodeficiency syndrome defining illness who underwent treatment interruption. The data collected include duration and reason for treatment duration, demographic characteristics, CD4 cell count and plasma viral load. In this study Cox proportional hazards model were used to assess characteristics associated with HAART reinitiation after treatment interruption, based on 208 of the 4461 patients underwent TI, the characteristics nadir CD4 cell count ≤ 250 cells/mm^3 (risk ratio [RR], 2.79; P< .001) was positively and independently associated with faster time to HAART reinitiation. Pogany *et al* (2007) used a similar methodology to evaluate the safety and efficacy of discontinuing HAART in patients with a CD4 cell nadir greater than 350 cells/mm^3. The study enrolled 71 patients of these patients 46(64%) interrupted and 25(36%) continued HAART with median CD4 cell counts at the start of HAART were 469 and 510 respectively. And only 5(11%) of the interrupted patients reinitiate therapy. As a result the viral load stabilizes its pre-HAART count, but the CD4 count still exceeds the pre-HAART count (563cells/mm^3).

3.3 Literature review on the specific variables used in the study

3.3.1 Socio-demographic characteristics

Demographic and social variables or characteristics such as sex, age, weight, marital status and level of education were included in many studies. These variables mainly used to compare by grouping the target population in to different categories based on descriptive statistical methods but in some studies, these were included in model based analysis and were found to have significant effect on the outcome variable.

Age at infection has been widely considered in studies concerned with identifying variables associated with HIV disease progression. Some studies like (D'Arminio *et al*, 2000) found that younger age and recent initiation of combination ART (HAART) were associated with an increased risk of treatment interruption.

A sex difference in terms of clinical response to HAART has been described but this was not consistently reported among different authors. In a cohort study involving 62 French hospitals Fardet *et al. (*2006) observed no significant difference in the timing of elective HAART initiation and clinical progression rate according to gender. On the other hand, female gender and injection drug user (IDU) were associated with a higher risk of TI. The gender difference

may be explained by women having a higher risk of toxicities as a result of lower weight (Moore *et al*, 2002).

The gender ratio of male and female patients in the HIV- NAT cohort is unlike that described in cohorts in Europe and the USA which tend to consist predominantly of male patients. After treatment initiation, male and female patients in the cohort had similar levels of risk for reaching the endpoints studied (Srasuebkul *et al*, 2007). The results of this study were similar to result from other studies which showed that men and women similarly responded to HAART (Nicastri et al, 2005).

In a study conducted on Danish HIV infected patients, the baseline body weight and plasma viral load were identified as predictive factors for discontinuation of the initial PI among patients in the randomized study on 318 patients. Patients with a lower body weight and lower plasma viral load on starting their initial PI had a higher risk of stopping the PI therapy. The risk of treatment discontinuation increased in patients by 16% per 5 kg lower body weight at baseline, (Kirk *et al*, 2001).

3.3.2 Disease duration, duration of follow up and CD4 change

In Ethiopians, low CD4 counts before the seroconversion do not lead to faster HIV disease progression by Yared *et al* (2003). They studied the immunological profile i.e. low baseline CD4 count would result in shorter survival time and used data from 149 HIV infected factory workers with mean age at intake 34 years and 69% were males. At the first HIV visit, the median CD4 count was 327 cells per microlitre. During active follow-up 29 individuals died. A continuous-time Markov model was designed and the estimated median survival time from HIV seroconversion to death was 9.1 to 13.0 years. Even if low CD4 counts, the survival time in HIV infected individuals in Ethiopians is similar to that of populations in the industrialized countries before the advent of ART. According to the study the generally low CD4 counts among Ethiopians do not imply shorter survival time, because of a slower CD4 decline.

In several studies we have seen that an unstructured treatment interruption is a major issue among children and adolescents with perinatality acquired HIV infection. Treatment interruption was associated with HIV disease progression including a rapid HIV-1 RNA increase during the first several months and a precipitous CD4 decline. Thus close monitoring is required when a TI undertaken. Importantly, nadir CD4 cell counts were reported to predict

a CD4 cell decline. The magnitude and predictors of HAART interruption in the study of 1551 subjects starting HAART, 299 (19.3%) interrupted treatment. Median (interquartile range (IQR)) duration of the TI was 189 (101-382) days. Women were more likely to have a TI than men in the same exposure group (35.8% vs 24.2% among drug users, 22.1% vs 13.3% among heterosexuals; $P < 0.05$). Higher baseline viremia and poor immunologic response to HAART were associated with higher probabilities of TI. Median individual CD4 cell loss during TI was 94 (IQR 1-220) cells/µl. Older age at HAART (>40 yr), lower pre-HAART nadir (<200 cells/µl), and lower CD4 count at start of TI (<350 cells/µl) were significantly associated with greater relative CD4 loss during duration of treatment interruption by employing Cox proportional hazards model (Thiébaut *et al*, 2005; Pellegrin *et al*, 2005; Touloumi *et al*, 2006).

Moreover, Cox proportional hazards model was applied in a similar fashion to assess characteristics associated with time to HAART reinitiation after TI (Toulson *et al*, 2005; Boschi *et al*, 2004; Mussini *et al*, 2005). A study by Toulson *et al,* (2005) indicated that the time of the reinitiation of HAART, the median CD4 cell count was 260 cells/mm^3. Among the 197 patients in the study, there were 6 deaths, none of which was attributable to the TI. A total of 81% had plasma viral loads < 50 copies/ml by 15 months of follow-up after reinitiation of HAART. In the multivariate analysis, a nadir CD4 cell count ≤ 250 cells/mm^3 (risk ratio, 2.79; P <.001) and is positively and independently associated with faster time to HAART reinitiation in adults who had HIV-1 infection and underwent treatment interruption.

A study conducted by Saitho *et al.* (2008) at four academic centres in the United States between January 2000 and September 2004 using the Wilcoxon-Rank sum test for group comparison and binomial regression analysis methods were used to compare the proportion of patients who started treatment interruption or not and to know the clinical outcomes after an unstructured treatment interruption among 405 patients with perinatal HIV-1 infection. Based on the analysis 72 (17.8%) experienced a treatment interruption during the observation period. Medication fatigue was the most common reason for a treatment interruption. In the finding, there is a persistent negative correlation between CD4 count gains from nadir to the time of treatment interruption and changes in CD4 count during treatment interruption. And 48 (67%) patients resumed antiretroviral medications. According to the result there was a continuous CD4 cell percentage decrease and plasma HIV-1 RNA increases during the observation period. Overall, 7 (10%) patients were admitted to hospital; 2 (3%) patients experienced an

AIDS-defining illness. Patients who gained the greatest CD4 cell percentage while receiving antiretroviral therapy experienced the greatest CD4 cell percentage decline during the first year of a TI. A similar finding was reported for adults who had long-term HIV infection and discontinued antiretroviral therapy (Mata *et al*, 2005, Nacher *et al,* 2006).

In a Swiss HIV cohort study to assess the impact of occasional short interruptions of HAART four different endpoints were considered. These were death, U.S. Centers for Disease Control and Prevention (CDC) stages B and C, and CD4 cell count increase >= 50×10^6 cells/µl). The study used Cox-time dependent proportional hazard regression model and Wilcoxon Mann-Whitney test for group comparison; and found that a high baseline viraemia (hazard ratio (HR), 1.2; P >0.01), high baseline CD4 cell count (HR, 1.07 per 100×10^6 cells/µl; P > 0.01), injecting drug use (HR, 2.0; P > 0.01), and low education (HR, 1.7; P > 0.01) were associated with a higher probability of having treatment interruptions. In patients who have initiated HAART indicates that treatment interruptions of 1-3 months do not seem to have an overall deleterious effect, particularly if the CD4 cell count is high and the viraemia is low. This result by Taffe *et al*, (2002) was somehow reassuring as drug intolerances, treatment failures, and the wish for 'drug holidays' are likely to increase in the future, with longer duration of treatment or in patients attaining long period of follow up.

Lamid *et al* (2007) examined the outcome of patients with chronic HIV-1 infection whom their treatments were stopped for reasons other than treatment failure through their rebound of plasma viral load. They employed the Kaplan-Meier and the Cox proportional hazards regression model and found that decrease in CD4 cell count and clinical events has characterized and pointed out that percentage change in CD4 count, duration of interruption and disease duration are among the important covariates that have a significant impact on a considerable decrease in CD4 counts and the experience of high rise viral load rebound which is detrimental to their living.

A previous study results from EuroSIDA study showed that both age and lower CD4 count were associated with a slower increase in CD4 count at 12 months of HAART (Florence *et al*, 2003). In this study seven hundred and eighty patients were included. A low CD4 count response was observed in 225 patients (29%). The risk factors for this condition were older age, lower CD4 count at baseline, higher increase from the nadir to baseline CD4 count and lower pVL at baseline as analyzed by multiple and logistic regressions.

The study of Touloumi *et al. (* 2008) based on the rates and determinants of virologic and immunologic response to HAART resumption after treatment interruption using Kaplan-Meier survival curves and Cox proportional hazards models in 281 persons with HIV-1 treated reported that median treatment duration of 18.4 months before treatment interruption, of these patients 259 resumed HAART, and subjects with low CD4 count decrease during treatment interruption, aged older than 40 years and those with the same HAART as their pre-TI regimen had lower CD4 count increase during the first three months of treatment resumption. The majority (86%) of individuals reinitiate therapy achieved HIV RNA > 500 copies/µl. Another study conducted on 185 patients, a survival analysis using Cox proportional hazards regression model was employed and showed that TI duration was shorter if CD4 nadir was below 250 cells per mm3 before TI (relative hazard, 2.10), older patients age greater than 40 were 1.7 times more likely to resume treatment earlier than those under 40 years of age at TI (Fillaux *et al*, 2006).

In Ethiopia to determine the prevalence and factors associated with defaulting from antiretroviral treatment (ART) in Jimma University specialized hospital was studied by Kebede *et al.* (2008). During the study period of the register review out of 1270 patients registered as starting ART, 706 (55.6%) were female, 55 (4.3%) were younger than 15 years and 355 (28%) missed one or more clinical appointments, these comprised 75(5.9%) dead; 173(13.6%) defaulted; 101(8.0%) transferred out and 6(0.5%) restarted their treatment, of the total defaulters 108(62.4%) were traced where as 65(37.6%) of the defaulters were not back to their treatment. The reason for unsuccessful tracing was because of their register (file) showed an incorrect address, no telephone access and due to other reasons. Using logistic-regression model, they found that the factors that were significantly associated with ART defaulting were: taking hard drugs (cocaine, cannabis and IV drugs); drinking alcohol most of the time; being bedridden; living outside Jimma town and having an HIV negative partner (or of unknown HIV status).

To estimate the prevalence and to identify factors associated with HAART discontinuation and modification in Ugandans by Kiguba *et al. (*2007) based on 686 HIV patients using logistic regression found that duration on HAART, being unmarried and more than three or more months on therapy were the variables that determine the discontinuation and modification of HAART.

In this chapter we have reviewed the various factors that are associated with Highly Active Antiretroviral Treatment (HAART) interruption and resumption and also the statistical methods employed to address it. In Ethiopia, few studies have been done on the outcome of HIV/AIDS patients who interrupted ART. In this regard, this study is an attempt to investigate factors related to ART resumption.

CHAPTER 4
DATA AND METHODOLOGY

4.1 The data

The data for this study were obtained from Zewditu Memorial Hospital (ZMH) ART Department, Addis Ababa, Ethiopia. Zewditu Memorial Hospital ART Clinic is one of the earliest ART sites in the country which was started as HIV Clinic in 1991, before the introduction of HAART in the country. On July 2003, the Hospital started providing ART treatment for peoples living with HIV/AIDS (PLWHA) who are financially strong and have the capability to pay for ART according to the national treatment guidelines. And subsequently it also provides for those patients who can not afford the price and produce a free paper for assistance from their place of residence "Kebele". Later on, by the year 2005 the hospital started delivering the service of follow-up and giving the medication to any patient who is eligible to start HAART.

ZMH ART clinic is staffed with five physicians, one pharmacist, two pharmacy technicians, five counselor nurses and six laboratory technicians working full time. All are trained on ART treatment and adherence counseling which is mandatory to work as ART Team. The Hospital is under the Addis Ababa Administration Health Bureau and often gets technical and financial assistance from CDC Ethiopia, ITECH, John Hopkins University, WHO and Management Science for Health-Rational Pharmaceutical Management Plus (MSH-RPM+). In addition to adherence counseling and ART for AIDS patients from Addis Ababa, it also provides screening, follow up and referral services for TB patients.

Thus, ZMH was taken as the research area since it has sufficient number of patients on ART as it started treatment follow up earlier than the other similar public health institutions. In addition the hospital is well staffed and organized, and has relatively good recording of clinical events using patient flow up chart provided by the Ministry of Health.

This study reviews patient's intake forms and follow up cards of HIV patients taking HAART in ZMH ART clinic. The patients' chart include the patient intake forms and follow up cards, which are prepared by Federal Ministry of Health (FMOH) to be uniformly used by clinicians to simply identify and document clinical and laboratory variables. Thus, this study uses

secondary data obtained from patients follow up records. Variables are collected by using the patient's identification number or the laboratory code without any direct contact with the patients. According the International and National treatment guideline, since it requires special consideration, all patient records will be kept in a confidential way during and after the study.

The data were gathered from a cohort study of patients with chronic HIV-1 patients with age higher than eighteen years old and being placed on highly active antiretroviral treatment (HAART) for at least three months. In the study patients with undetectable viral load and CD4 count at treatments interruption (drug holiday) and have their treatments interrupted for at least one month are considered.

The total number of patients on HAART incorporates up to 8360 individuals with the data extracted from the HIV cohort database. This cohort contains the demographic, laboratory and clinical information of all HIV patients in follow-up in between the year 1998 to 2007, including a detailed antiretroviral therapy history. Among these patients around 723 patients interrupted their treatment or they stopped their treatment at least once. Of those who interrupted their HAART 487 resumed back to their treatment follow up and 236 of the interrupters did not resume back to treatment.

4.2 The variables in the study

In this study, several variables which are expected to be associated with the duration of interruption of HAART among patients living with HIV-1/AIDS and which may cause early or late resumption and even the indirect consequence of interruption like death or dropout from follow-up will be considered.

4.2.1 Dependent variable (outcome variable)

The outcome variable used in this analysis is the length of drug holidays attained by the patient that is time to resumption to treatment. It is defined as the difference between the date of treatment resumption (TR) and the date the patient started treatment interruption (TI). It is measured in days.

4.2.2 Independent (predictor) variables or covariates are:

Several predictors are considered in this study to investigate the important determinant factors for the resumption of treatment in patients interrupting their HAART. Some of these variables are categorical and others are continuous:

- Age in this study is considered as the age of the patient at the start of HAART and patients with age >= 18 years are included.
- Gender of the patient (Male, Female).
- Weight of the patient at the start of HAART.
- Marital status of the patient at the start of treatment and it is categorized as currently married and currently not married (single, separated, divorced and widowed).
- Level of education of the patients is categorized as the level in which it is attained by the patients at the time of start of therapy and is classified as no education (illiterate), primary (educated up to 8^{th} grade), secondary and above (patients who attend up to 12 and above grade 12)
- Religion of the patient classified as Orthodox, Muslim and others (Protestant, Catholic and Others).
- Disease duration in this study is obtained by taking the difference between the date in which the patient started treatment interruption and the date he/she is tested positive, as it is recorded in the follow-up register format and is measured in days.
- Duration of follow up is defined as difference between the date the patient interrupted treatment and the first date the patient enrolled for HAART and is measured in days.
- The CD4 count of the patient at the start of Highly Active Antiretroviral Treatment (CD4HAART).
- Prior CD4 count the last CD4 cell count registered for the patient prior to treatment interruption (CD4Labvalue).

4.3 The methodology

Survival analysis is a statistical method designed to study the amount of time an experimental unit survives, or the study of time between entry into observation and a subsequent event. The statistical approach to be used in this study is the analysis of time-to-event data which are related with individual time elapse in certain situation or state. The methodological development of survival analysis or time-to-event data analysis is the product of long process, which has undergone a great impulse in the last 60 years. Researchers from the biological,

medical and epidemiological (public health), demography, economics, engineering as well as from industry have stimulated this process with their scientific problems.

The term "survival data" has been used in a broad sense for data involving time to the occurrence of a certain event. This event may be death, the appearance of a tumor, the development of some disease, recurrence of a disease, conception, cessation of smoking, and so forth. In the past, application of the statistical methods for survival data analysis have been extended beyond biomedical and reliability research to other fields, such as the social sciences and business. For example, we could look at the duration of a first marriage (sociology), the length of subscription to a newspaper or a magazine (marketing), and so on. The study of survival data was previously focused on predicting the probability of survival or mean lifetime, and comparing the survival distributions of experimental animals or of human patients under different conditions. In recent years, the identification of risk and/or prognostic factors related to survival and the development of disease have become important applications of survival analysis.

As the uses of survival analysis grew, parametric models gave way to nonparametric and semi-parametric approaches for their appeal in dealing with the ever-growing field of clinical trials in medical research. Survival analysis was well suited for such work because medical intervention follow-up studies could start without all experimental units enrolled at start of observation time and could end before all experimental units had experienced an event. This is extremely important because even in the best-developed studies, there will be subjects who choose to quit participating, who move too far away to follow, or who will die from some unrelated event. The researcher was no longer forced to withdraw the experimental unit and all associating data from the study; instead techniques called censoring enabled researchers to analyze incomplete data due to delayed entry or withdrawal from the study. This was important in allowing each experimental unit to contribute all of the information possible to the model for the amount of time the researcher was able to observe the unit.

Survival analysis consists of a set of specialized statistical techniques used to study response time data. In analyzing such data, the main objects are to determine the length of time interval spent in a state and the transition probabilities from the current state to the entered state. The interest of this statistical tool is mainly focused on two distinguishing features of time to event data. Primarily, duration times are non-negative values usually exhibiting highly skewed

distribution and therefore assumption of normality may be violated. Secondly, censoring may occur or the true duration is not always observed or known, that is, some subjects potentially being unobserved for the full time to failure.

The main characteristic of these data is the issue of censoring which occurs when the periods of time of event occurrence for some individuals can not be completely observed. The process of censoring and truncation make these data unsuitable to analyze with traditional regression method and hence, the appropriate techniques and analyses, usually called Survival Analysis. Details on various estimation methods developed in survival data analysis taken censoring and truncation in to account can be obtained in Hosmer and Lemeshow (1998) and Marubini and Valsecchi (1995) among others.

However, it is censoring which makes standard statistical techniques unsuitable for analyzing survival data. As mentioned before, censoring is said to occur when the end-point of interest has not been observed by the end of data collection. It occurs, for example, when some patients survive to the end of the trial investigating time to death; when a certain type of cancer does not occur again after surgical removal; when a patient has died from an unrelated cause to the one being investigated; and when a patient is lost to follow-up.

4.3.1 Censoring

The time period confinement for survival data gives rise to considerations specific to survival analysis, censoring and truncation. A censored observation is one whose value is incomplete due to random factors for each subject. Censoring can appear in various forms and the most common forms are explained below:

(i) **Right Censored**: The most common form of incomplete data is right censoring. An observation is said to be right censoring if it is recorded from its beginning until a well defined time before its end time. For instance, if HIV-1 patient is followed until he has a viral load high than 1000 copies/μl and is followed without experiencing this scenario until the end the observation period, and then this patient is known to be right censored. In other words, an observation is said to be right censored if it begins at time t = 0 and terminate before the outcome of interest is observed.

(ii) **Left Censored**: An observation is said to be left censored if all that is known is that the individual developed the event of interest prior to the beginning of the study. An observation is said to be left censored if the event of interest has already occurred when observation begins. This situation is less common in survival studies and is often not a focus.

(iii) **Interval Censored**: An observation is categorized into interval censored if it is only known that the event of interest occurs within an interval of time without the knowledge of when exactly it occurs. Interval censoring occurs in clinical trial where patients have periodic follow-ups and in industrial experiments where equipment items are inspected periodically, etc.

4.3.2 Descriptive methods for survival data

In any applied setting, a statistical analysis should begin with a thoughtful and thorough univariate description of the data. And this description includes life table and Kaplan-Meier survival function estimation which are used for the estimation of the distribution of survival time from all observations available.

The Survival Function

The cumulative distribution function (cdf) is very useful in describing the continuous probability distribution of a random variable, such as time, in a survival analysis. The cdf of a random variable T, denoted $F_T(t)$, is defined by $F_T(t) = P_T(T \leq t)$.

The survival function is defined as the probability of a subject at risk surviving beyond time t. Let $T \geq 0$ have a pdf f(t) and cdf F(t). Then the survival function takes on the following form:

$$S(t) = P\{T > t\} = 1 - F(t)$$

That is, the survival function gives the probability of surviving or being event-free beyond time t. Because S(t) is a probability, it is positive and ranges from 0 to 1. It is defined as S(0) = 1 and as t approaches ∞, S(t) approaches 0.

Median Survival Time: Median survival time *m* is defined as the quantity *m* satisfying $S(m) = 0.5$. Sometimes denoted by $t_{0.5}$. If S(t) is not strictly decreasing, *m* is the smallest one such that $S(m) \leq 0.5$ or $t_{med=0.5} = S^{-1}(0.5)$.

The Hazard Function

The hazard function is also known as failure rate, force of infection, force of mortality, conditional failure rate, intensity function, or simply hazard rate and it is defined as the probability that an individual fails at time t, conditional on the fact that he or she has survived to that time. It therefore, represents the instantaneous failure rate for an individual surviving to time t. For $h(t) \geq 0$, the hazard function h(t) is given by the following:

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{p\{an\ individual\ fails\ in\ the\ time\ \text{int}erval\ (t, t+\Delta t) \mid alive\ at\ t\}}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{p\{t \leq T \leq t+\Delta t\}}{\Delta t} \\ &= P\{t < T < (t+\Delta t) \mid T > t\} \\ &= f(t) \Big/ (1 - F(t)) \\ &= f(t) \Big/ S(t) \end{aligned}$$

The hazard function describes the concept of the risk of an outcome (e.g., death, failure, hospitalization) in an interval after time t, conditional on the subject having survived to time t. The hazard function seems to be more intuitive to use in survival analysis than the pdf because it attempts to quantify the instantaneous risk that an event will take place at time t given that the subject survived to time t.

4.3.3 Non-parametric methods in survival analysis

Kaplan-Meier survival function

The Kaplan-Meier (KM) estimator, or product limit estimator, is the estimator used by most software packages. The KM estimator incorporates information from all of the observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. When there is no censoring, the estimator is simply the sample proportion of observations with event times greater than t. The technique becomes a little more complicated but still manageable when censored times are included.

The KM estimator is a nonparametric estimator of the survivor function S(t).

$$\hat{S}(t) = \prod_{t_{(j)} \leq t} (1 - \frac{d_j}{n_j}) \quad \text{---------- (4.1)}$$

Where d_j is the number of individuals who experience the event or failure at time $t_{(j)}$, and n_j is the number of individuals who have not yet experienced the event at that time and are

therefore still at risk for experiencing it. The Kaplan-Meier estimator (4.1) is a step function with jumps at the observed event times. The size of the jump at a certain event time $t_{(j)} = t_i$ depends on the number of events observed at t_i, as well as on the pattern of the censored event times before t_i.

4.3.4 Comparison of survival curves

In clinical research one is concerned not only with estimating the survival function but, more often, with the comparison of the life experience of two or more groups of subjects differing for a given characteristic or randomly allocated to different treatments. After providing a description of the overall survival experience in the study, we usually turn our attention to a comparison of the survivorship experience in key subjects in the data. The simplest way of comparing the survival times obtained from two or more groups is to plot the Kaplan-Meier curves for these groups on the same graph. However, this graph does not allow us to say, with any confidence, whether or not there is a real difference between the groups. The observed difference may be a true difference, but equally, it could also be due merely to chance variation. Assessing whether or not there is a real difference between groups can only be done, with any degree of confidence, by utilizing statistical tests.

Since survival data are typically right skewed, we would likely use rank-based non-parametric tests followed by estimates and confidence intervals of the medians or other quantiles with in groups. Modifications of these procedures are required when censored observations are present in the data. When we compare groups of subjects, it is always good to begin with a graphical display of the data in each group.

Among the various non-parametric tests one can find in the statistical literature, the Mantel-Haenzel (1959) test, currently called the “log-rank” test will be used. Nowadays the Kaplan-Meier method for estimating survival curves and the log-Rank test for comparing two estimated survival curves are the most frequently used statistical tools in medical reports on survival data.

Log-rank test

The log rank test, developed by Mantel and Haenszel, is a non-parametric test for comparing two or more independent survival curves. Since it is a non-parametric test, no assumptions about the distributional form of the data need to be made. This test is however most powerful when used for non-overlapping survival curves. This test can be generalized to accommodate other tests that are equally used sometime in practice such as Generalized Wilcoxon test, Tarone-Ware test, and Peto-Peto Prentice test. Each of these tests uses different weight to adjust for censoring that is often encountered in survival data.

The log rank test statistic for comparing two groups is given by:

$$Q = \frac{\left[\sum_{i=1}^{m}(d_{1i} - \hat{e}_{1i})\right]^2}{\sum_{i=1}^{m}\hat{v}_{1i}}$$

Where:

m is the number of rank ordered event times.

d_{1i} is the observed number of events (the number of patients resumed to treatment) in group1 at event time t_i.

$\hat{e}_{1i} = \frac{n_{1i}d_i}{n_i}$ is the expected no of events corresponding to d_{1i}.

n_{1i} is the number of individuals at risk in group1 just prior to event time t_i.

$\hat{v}_{1i} = \frac{n_{1i}n_{0i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$ is the variance of the number of events d_{1i} at time t_i

n_i is the number of individuals at risk in both groups1 and 2 just prior to event time t_i

The log rank test statistic Q has an approximation of chi-square distribution with one degree of freedom for large samples. The null hypothesis of equality of survival functions will be rejected for large values of Q. The log rank test can be extended for comparing three or more groups of survival experience. See details in Hosmer and Lemeshow (1998).

4.3.5 Regression Models for Survival Data

In most medical studies which give rise to survival data, supplementary information is collected on each individual so that the relationship between the survival experience of individuals and various explanatory variables may be investigated. A variety of models and

methods have been developed for doing this sort of survival analysis using either parametric or semi-parametric approaches. Semi-parametric models are models that parametrically specify the functional relationship between the lifetime of an individual and his characteristics (demographic, socio-economic, etc.) but leave the actual distribution of lifetimes arbitrary. The most popular of the semi-parametric models is the Proportional hazards model. It has the property that the ratio of the hazards depends on the values of their explanatory variables, say $X_1, X_2, ...,$ but does not depend on time t.

The semi-parametric Cox-proportional Hazards Model

We can specify the density function of a parametric distribution or we can specify the hazard function. The advantage of the latter approach is that we directly address the aging process, but as shown previously, it does not easily lend to itself to the use of scatter plots to motivate regression models. The latter approach may also be preferred in a setting where the end products of the statistical analysis are estimated parameters that compare the survival experience of the selected subgroups. By specifying a model through the hazard function, we may address specific questions such as how survival is related to the subject's characteristics or the covariates.

Cox's (1972) paper took a different approach to standard parametric survival analysis and extended the methods of the non-parametric Kaplan-Meier estimates to regression type arguments for life-table analyses. Cox advanced to prediction of survival time in individual subjects by only utilizing variables covarying with survival and ignoring the baseline hazard of individuals. He did this by making no assumptions about the baseline hazard of individuals and only assumed that the hazard functions of different individuals remained proportional and constant over time. When there are several explanatory variables, and in particular when some of these are continuous, it is much more useful to use a regression method such as Cox rather than a KM approach.

Cox introduced the semi-parametric proportional hazards model to cater for covariate effects for single event failures. This model is valid under the assumption of proportional hazards. Cox (1972) observed that if proportional hazards assumption holds (or is assumed to hold), then it is possible to estimate the effect parameter(s) without any consideration of the hazard function.

There are several reasons in which Cox's proportional hazards modeling was chosen to explain the effect of covariates on time until event. They are discussed below and include: the relative risk, no parametric assumptions, hazard function, the use of the partial likelihood function, and the estimates of survivor function.

Relative Risk

The simple interpretation given by the Cox model as "relative risk" type ratio is very desirable in explaining the risk of event for a certain covariate. For example, when we have a two-level covariate with a value of 0 or 1, the hazard ratio becomes e^{β}. If the value of the coefficient is $\beta = \ln(3)$ then it is simply saying that the subjects labeled with 1 are three times more likely to have an event than the subjects labeled with 0. In this way we have a measure of difference between our exposure cohorts instead of simply knowing whether they were different.

No Parametric Assumptions

Another attractive feature of Cox regression is that, it does not choose the density function of a parametric distribution. This means that Cox's semi-parametric modeling allows for no assumptions to be made about the parametric distribution of the survival times, making the method considerably more robust. Instead, the researcher must only validate the assumption that the hazards are proportional over time. The proportional hazards assumption refers to the fact that the hazard functions are multiplicatively related. That is, their ratio is assumed constant over survival time. In other words, the Cox proportional hazards model assumes that changes in the hazard of any subject over time will always be proportional to changes in the hazard of any other subject and to changes in the underlying hazard over time.

The hazard function

The data in survival analysis based on the sample size n, consists of $(t_i, \delta_i, \mathbf{X}_i)$, i=1,2,…,n, where t_i is the time on the study for the i^{th} individual, δ_i is the event indicator (δ_i=1 if the event has occurred and δ_i=0 if it is censored (the lifetime may be right, left or interval censored), and $\mathbf{X}_i$ is the vector of covariates or the risk factors for the i^{th} individual that may affect for instance the outcome of patients interrupting their HAART, i.e. time to resumption of HAART.

The Cox proportional hazards model is generally given by:

$$h(t,\mathbf{X}_i,\boldsymbol{\beta}) = h_0(t)\exp(\boldsymbol{\beta}'\mathbf{X}_i) \quad \text{---------- (4.2)}$$

Where $h_0(t)$ is the baseline hazard function at time *t*, $\mathbf{X}_i' = (X_{1i}, X_{2i}, ..., X_{Ki})$ for $i = 1,2,...,n$ is a vector of measured covariates for the i^{th} individual at time *t*, and $\boldsymbol{\beta}'$ is a $1 \times k$ vector of unknown regression parameters that are assumed to be the same for all individuals in the study, which measures the influence of the covariate on the survival experience with β_i representing the increase in the log hazards as X_i increases one unit relative to the baseline hazard function. This model is referred to in the literature by a variety of terms, such as the Cox model, the Cox proportional hazards model or simply the proportional hazards model. The hazard function in equation (4.2) depends on both time and the associated covariates, but through two separate factors: the first is a function of time only which is left arbitrary, but is assumed to be the same for all the subjects, the second is a quantity which depends on the individual covariates.

From the representation in equation (4.2) one can notice a couple of features. First, if $X_i = 0$ then the hazard function for the i^{th} individual is the baseline hazard function. It's the hazard function in the absence of covariates or when all of the coefficients of the covariates are assumed to be zero. Second, if we divide both sides by $h_0(t)$, we get equation (4.3) which shows where the term proportional comes from. Since for each individual, $e^{\mathbf{X}_i'\boldsymbol{\beta}}$ is constant across time, equation (4.5) shows that at every value of t, the i^{th} individual's log hazard function is constant proportion of the baseline hazard. Very loosely speaking, this implies that each individual's hazard function is "parallel" to the $h_0(t)$.

$$\frac{h_i(t,\mathbf{X}_i)}{h_0(t,0)} = \frac{h_0(t)\exp(\boldsymbol{\beta}'\mathbf{X}_i)}{h_0(t)} = e^{\boldsymbol{\beta}'\mathbf{X}_i} \quad \text{---------- (4.3)}$$

The Cox model is often called proportional hazards model because, if we look at two independent subjects with covariate values $\mathbf{X_1}$ and $\mathbf{X_2}$, the ratio of their hazard functions at time t is:

$$\frac{h(t,\mathbf{X}_1)}{h\ (t,\mathbf{X}_2)} = \frac{h_0(t)\exp(\boldsymbol{\beta}'\mathbf{X_1})}{h_0(t)\exp(\boldsymbol{\beta}'\mathbf{X_2})} = \exp\left[\boldsymbol{\beta}'(\mathbf{X_1} - \mathbf{X_2})\right] \quad \text{---------- (4.4)}$$

Which is constant and does not vary over time, that is, the ratio does not depend on t and the hazard rates are proportional. The Cox proportional hazards model can equally be regarded as

linear model, as a linear combination of the covariates for the logarithm transformation of the hazard ratio given by:

$$\log\left\{\frac{h(t,\mathbf{X})}{h_0(t)}\right\} = \boldsymbol{\beta}'\mathbf{X} \text{ ---------- (4.5)}$$

Note that the cumulative hazard function is given by:

$$H(t) = H_0(t)\exp(\boldsymbol{\beta}'\mathbf{X}) \text{ ---------- (4.6)}$$

Consequently, from the proportional hazard function, we obtained the survivor function given by:

$$S(t,\mathbf{X},\boldsymbol{\beta}) = \left[S_0(t)\right]^{\exp(\boldsymbol{\beta}'\mathbf{X})} \text{ ---------- (4.7)}$$

Where, $S_0(t)$ is the baseline survival function.

Estimation of the regression parameters using partial likelihood

Since $h_0(t)$ is not specified parametrically, it is not possible to use an ordinary likelihood to estimate the regression coefficients $\boldsymbol{\beta}$. The arbitrary function $h_0(t)$ is a nuisance function, and the aim is to estimate $\boldsymbol{\beta}$ on the basis of the information conveyed by the observed data without having to involve $h_0(t)$. Cox (1972) argued conditionally on the set of observed failures and described the data with a function depending on $\boldsymbol{\beta}$ only. Consider a sample of n subjects and suppose a total of m failures occur, with m generally smaller than n, due to the presence of censoring. Let $t_{(1)} < t_{(2)} < \ldots < t_{(m)}$ be the m distinct ordered failure times observed and let R(t) be the set of subjects, at risk at time t, who are not failed and under observation just before t. With a slight change of notation, we indicate with j the label of the subject who fails at $t_{(j)}$ so that its vector of covariates is $\mathbf{X}_j$. In general, $\mathbf{X}_i$ the vectors of covariates for the i^{th} subject and the covariates have a constant value in time. The probability that an individual with covariates $\mathbf{X}$ fails in the small interval $(t, t+\Delta t)$, given the set at risk at t, is:

$$\frac{h(t_{(j)},\mathbf{X}_j)\Delta t}{\sum_{i \in R(t_{(j)})} h(t_{(j)},\mathbf{X}_i)\Delta t}$$

It follows that the function describing the failure pattern is the product of m terms, one for each observed failure time.

$$L(h_0(t),\beta) = \prod_{j=1}^{m} \frac{h(t_{(j)},\mathbf{X}_j)\,\Delta t}{\sum_{i \in R_j} h(t_{(j)},\mathbf{X}_i)\,\Delta t}$$

Where the hazard function is defined by (4.2) and $R_j = R(t_{(j)})$. Given expression (4.2), the baseline function $h_0(t)\,\Delta t$ cancels out and the product above simplifies to:

$$L = L(\boldsymbol{\beta}) = \prod_{j=1}^{m} \frac{\exp(\boldsymbol{\beta}'\mathbf{X}_j)}{\sum_{i \in R_j} \exp(\boldsymbol{\beta}'\mathbf{X}_i)} \quad \text{---------- (4.8)}$$

Where $L(\boldsymbol{\beta})$ in equation (4.8) depends on the unknown parameters $\boldsymbol{\beta}$ is referred to as the partial likelihood.

The partial likelihood given by equation (4.8), although it describes only part of the data, could be regarded as a likelihood function allowing the estimation of $\boldsymbol{\beta}$ with standard procedures. In general, large sample properties like normality and consistency of maximum likelihood estimators of $\boldsymbol{\beta}$ based on partial likelihood have been shown to be the same as those of any estimator from complete likelihood (Cox, 1975; Anderson and Gill, 1982).

The asymptomatic theory of maximum likelihood estimation requires that the likelihood function satisfies some "regularity conditions" which are met in most applications. The regression coefficients $\boldsymbol{\beta}$ are estimated by the values $\hat{\boldsymbol{\beta}}$ which maximize the partial likelihood $L(\hat{\boldsymbol{\beta}})$ or equivalently its logarithm $LL(\hat{\boldsymbol{\beta}})$:

$$LL(\boldsymbol{\beta}) = \sum_{j}^{m} \left\{ \boldsymbol{\beta}'\mathbf{X}_j - \ln\left[\sum_{i \in R_j} \exp(\boldsymbol{\beta}'\mathbf{X}_i) \right] \right\} = \sum_{j=1}^{m} l_j \quad \text{----------(*)}$$

Where l_j is the contribution of the log-likelihood corresponding to the failure time $t_{(j)}$. The values $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, ..., \hat{\beta}_K)$ are obtained by equating to zero the K first derivatives of log likelihood function with respect to β_k $(k = 1, ..., K)$.

The partial likelihood (4.8) has been written for t as a continuous variable. However, even when t is continuous in theory, tied observations may occur in the practice because of the measurement units of time. Provided that ties are few in the number, it is possible to deal satisfactorily with the problem by modifying the likelihood (4.8). The approach which is routinely used is the one proposed by Peto in the discussion of Cox's paper (1972) and Breslow (1974) as cited in Marubini and Valsecchi (1995). Let d_j be the number of failures observed at time $t_{(j)}$ and $\mathbf{S}_j$ be the sum of the covariate vectors of the d_j subjects who fail.

The logarithm of the partial likelihood is written as if the d_j failures occurred in any order and ignoring the fact that the risk set should successively be administered by one subject at each failure. It is

$$LL(\boldsymbol{\beta}) = \sum_{j}^{m} \left\{ \boldsymbol{\beta}' \mathbf{S}_j - d_j \ln \left[\sum_{i \in R_j} \exp(\boldsymbol{\beta}' \mathbf{X}_i) \right] \right\} = \sum_{j=1}^{m} l_j \text{ ----------(**)}$$

And $LL(\boldsymbol{\beta})$ (**) reduces to (*) when no ties are present ($d_j = 1$ and $\mathbf{S}_j = \mathbf{X}_j$).

In this study to estimate survival function we will use the non-parametric Kaplan-Meier or product limit estimation method and for comparisons of survival estimates the Log-Rank and Generalized-Wilcoxon test will be considered. Moreover, to investigate the effect of factors or covariates on the time to an event in the follow-up study, we consider the Cox's proportional hazards regression model. For the analysis, we use the application of different software packages such as ***SAS 9.2, STATA 10 and SPSS 15***.

4.4. Model Building or model development

In performing proportional hazards regression analysis for survival data requires a number of critical decisions. It is likely that we will have data on more covariates than we can reasonably expect to include in the model, so we must decide on a method to select a subset of the total number of covariates. When selecting a subset of the covariates, we must consider such issues as clinical importance and statistical significance.

Before any model could be fitted, it is a statistical tradition to investigate which variable(s) goes into the model by using conventional selection procedure. The methods available to select a subset of the covariates to include in a proportional hazards regression model are essentially the same as those used in the other regression models, like purposeful selection, stepwise (forward selection and backward elimination) and best subsets selection.

In this study, model building starts from univariate analysis as suggested by Collet (1994), Collect recommended the approach of first doing a univariate analysis to "screen" out potentially significant variables for consideration in the multivariate model in order to identify the importance of each predictor. All variables that are significant at 25% level, the modest level of significance for bivariate regression from one explanatory univariate regression model are taken into multivariable model where backward selection approach is used with

10% significant level of stay in the model. Variables that are selected at this stage are taken to stage three of the analysis where variables that are not significant in stage one are added one at a time and forward selection procedure is used with 5% significant level of entry into the model. The fourth stage involves combination of all variables that are significant at stage three in addition with their possible interactions using stepwise selection procedure with 10% significant level of entry and stay in the model and if the interaction is significant, but not the main effect of the covariate, we include both the interaction and the main effect in the final model even if the main effect is not significant. According to the Hierarchic principle, if a model contains interaction terms, the corresponding lower order terms should also be included in the model. The final variables selected at this stage are then pruned to have the final model.

4.5 Assessment of Model Adequacy

Model-based inferences depend completely on the fitted statistical model. For these inferences to be "valid" in any sense of the word, the fitted model must provide an adequate summary of the data upon which it is based. Some of the methods for the assessment of a fitted proportional hazards model can equally used for parametric regression models. There are basically four requirements for model adequacy considered in this study. They are:

(i) Methods for testing the assumption of proportional hazards. The proportional hazards assumption is vital to the interpretation and use of a fitted proportional hazards model. This is an assessment of how extent the two curves are equidistant over time. The proportional hazards model has a log-hazards functional form given by.

$$\ln\left[h(t,\mathbf{X},\boldsymbol{\beta})\right]=\ln\left[h_0(t)\right]+\mathbf{X}'\boldsymbol{\beta} \quad \text{---------- (4.9)}$$

This function has two parts, the log of the baseline hazard function $\ln\left[h_0(t)\right]$ and the linear predictor, $\mathbf{X}'\boldsymbol{\beta}$.

There are, effectively, an infinite number of ways the model in (4.9) can be changed to yield non-proportional hazard function or log hazard functions that are not equidistant. As a result, a large number of tests and procedures have been proposed. However, recent development work by Grambsch and Therneau (1994) and simulation comparison by Ng'andu (1997) have shown that one easily performed test and associated graph yield a powerful and effective method for the examining this critical assumption.

This formal test is used to detect any time dependency in particular covariates, after allowing for the effects of explanatory variables that are known. Time-dependent covariates are covariates whose values change over time generated by creating interactions of the predictors and a function of survival time (time, or log (*t*)) and this can be added, to examine the assumption of proportional hazards in Cox regression model. Kleinbaum (1996) suggests that if any of the time dependent covariates are significant then those predictors are not proportional and this indicates a violation of the proportionality assumption for that specific predictor Grambsch and Therneau (1994) consider an alternative to the model in (4.9), originally proposed by Schoenfeld (1982) that has the following form of time varying coefficient:

$$\beta_j(t) = \beta_j + \gamma_j g_j(t) \quad \text{---------- (4.10)}$$

Where $g_j(t)$ a specified function of time and γ_j is the coefficient or slope of the equation. When $g_j(t) = \ln(t)$, equation (4.10) reduces to the expression in equation (4.11).

$$\beta_j(t) X_j = \beta_j X_j + \gamma_j X_j \ln(t) \quad \text{---------- (4.11)}$$

The rationale for this model is that the effect a covariate may change over the period of follow-up. To test the Cox proportional assumption holds or not we test whether the coefficient γ_j is zero or different from zero: if it is different from zero the proportional assumption fails to hold but if it not significant then the assumption is satisfied. Another method of testing the validity of the Cox proportionality assumption is to plot the scaled schoenfeld residuals against the log of time and, if this plot shows some trend the assumption is violated, where as if the plot demonstrates randomly distributed around the reference line then the assumption is satisfied.

(ii) Goodness-of-fit: as in all regression analyses, some measures analogous to R^2 may be of interest to measure model performance. Cox and Snell (1989) proposed model assessment using R^2 similar to the one used in linear regression which is given by:

$$R^2 = 1 - \exp\left\lfloor {}^{2}\!/_{n} (LL_0 - LL_{\hat{\beta}}) \right\rfloor \quad \text{---------- (4.12)}$$

Where LL_0 is the (partial) log likelihood for zero model or without covariates, $LL_{\hat{\beta}}$ is the log likelihood including covariates, n is the number of subjects included in the study. To check the measure of goodness of fit for the final model in addition to R^2 we use tests like: the partial likelihood ratio, Wald and Score tests.

The partial likelihood ratio (LR) test: to use this we need to fit both the unrestricted and the restricted models. We shall obtain the value of the log-partial likelihood function $LL_P(\hat{\boldsymbol{\beta}})$ in the unrestricted model and $LL_P(\boldsymbol{\beta}=\mathbf{0})$ when the model imposes the restrictions under H_0. The test statistic for H_0 is based on the difference of the log-likelihood values. Under H_0, the statistic is asymptotically distributed as χ^2 with a number P degrees of freedom.

$$Q_{LR} = 2\left[LL_P(\beta) - LL(0)\right] \sim \chi^2(P)$$

The Wald test: this requires fitting the unrestricted model, and is based on the partial likelihood estimator $\hat{\boldsymbol{\beta}}$. The test statistic is

$$Q_W = \hat{\boldsymbol{\beta}}' I_{p\times p}{}^{-1}(\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\beta}} \sim \chi^2(P)$$

The quadratic form of the above equation requires the inverse of the variance-covariance estimates corresponding to the P parameters in H_0 matrix $I_{p\times p}$ and, under H_0 is asymptotically distributed as χ^2 with P degrees of freedom.

The Score test: this test is performed after fitting the restricted model and obtaining $LL(\boldsymbol{\beta}=\mathbf{0})$. It is based on the gradient of this function at $\boldsymbol{\beta}=\mathbf{0}$ i.e, on the score vector of P components.

$$Q_S = U'_{Ho}\, I_{p\times p}{}^{-1}(0) U_{Ho} \sim \chi^2(P)$$

Where $I_{p\times p}{}^{-1}(.)$ indicates the matrix of dimension p×p, extracted from the inverse of the observed information matrix evaluated at $\boldsymbol{\beta}=\mathbf{0}$ and U_{Ho} is the score function under H_0. It has approximately a χ^2 distribution with P degrees of freedom.

(iii) Subject-specific diagnostic statistics that extends the notions of leverage and influence to the proportional hazards model and

(iv) Testing for linearity of the covariates in Cox-Proportional regression model.

Under the proportional hazards model, residuals play a central role in evaluating the model assessment and adequacy of item (iii) and (iv) above. The following residual diagnostics are considered in assessing model adequacy in this study. The first and most important one is the scaled (weighted) Schoenfeld residuals which are useful to check the proportionality of the covariates over time that is to check the validity of the proportional hazards assumption. If the model fits well then the residuals are randomly distributed without any systematic pattern around the zero line, reference line.

The second one is the score residual which is a weighted average of the distance of the value, X_{ij}, to the risk set means, X_{wjk}, with the weights taken as the change in the martingale residual, that is used in assessing subject-specific diagnostic or the influential subjects by observing how large the deviation is. The larger the deviation the more distant the residual is to the mean. The plot of the score residuals looks like a basic hourglass shape, fanning out from its narrowest point at approximately the mean of the covariate.

The third one, which is used to check the linearity of the covariates to be included in the model, is the Martingale residuals. Although Martingale residuals share many of the properties possessed by residuals encountered in other situations, such as in linear regression analysis, they are not symmetrically distributed about zero, even when the fitted model is correct. The skewness makes plots based on the residuals difficult to interpret.

4.6 Covariate adjusted survival curves

Cox's proportional hazards regression model is a useful statistical tool for the analysis of survival data from longitudinal studies. This multivariate method compares the survival experience between two or more exposure groups while allowing for simultaneous adjustment of confounding due to one or more covariates. In addition, to the summary regression statistics, further insight on the exposure response relationship can be gained by visually examining the covariates–adjusted survival curves in the respective comparison groups. Covariates–adjusted survival curves are usually computed by the average covariate method which is given above in equation 4.7. Detailed discussion is found in Hosmer and Lemeshow (1998) among others.

CHAPTER FIVE
STATISTICAL DATA ANALYSIS AND DISCUSSION

5.1 Introduction

The data comprised of 723 patients who were enrolled in ART delivery programme at Zewditu memorial hospital but interrupted their treatment for at least 3 months. The outcome response is time to treatment resumption. Among the patients 67.1% (487) resumed their treatment and the rest 32.9% (236) were dead and loss to follow-up from the cohort study. The median time to resumption of treatment is estimated to be 377 days with interquartile range (IQR) of (361 - 393) days. In ART-delivery programme in ZMH, Table 5.1 shows that 32.9% of 723 patients were dead or lost to follow-up from the cohort study. The percentage of patients that are dead and lost to follow up is less than the patients that resume to treatment.

5.1.1 Descriptive Statistics

Summary results for socio-demographic variables included in this study are discussed below. Table 5.1 displays the summary statistics for categorical variables. The minimum and maximum ages for male and female patients in this study are respectively 18 and 60 years. The median age for female patients is 31 years and that of male is 38 years. Similarly, the mean age for females is 33.80 years with standard deviation of 0.473 and that of male is 39.43 with a standard deviation of 0.557.

The average CD4 count at the start of treatment for male is 123.36, with a maximum of 296, while female's average CD4 count is 130.31 with a maximum of 315. In addition the CD4 count prior to the interruption for both sexes increases as a result of treatment with HAART and attained up to the a maximum CD4 count of 598 by female patients and 586 by male patients. When we consider the median duration of follow-up for patients with treatment resumption is 245 days with IQR (175 - 394) days and for those who did not resume is 280 days with IQR (194 - 486) days. (Table A1.1, Annex 1).

Table 5.1: Summary results of different socio-demographic variables of chronic HIV-1 patients with and without HAART resumption for the data from ZMH

Demographic variables		Patients without HAART resumption n(%)	Patients with HAART resumption n(%)	Total n(%)
Sex	Female	135(18.67)	265(36.65)	400(55.33)
	Male	101(13.97)	222(30.7)	323(44.67)
Marital status	Currently Married	87(12.03)	170(23.51)	257(35.54)
	Currently not married	149(20.61)	317(43.85)	466(64.45)
Level of education	No education	45(6.22)	105(14.53)	150(20.75)
	Primary education	78(10.79)	150(20.75)	228(31.54)
	Secondary education and above	113(15.63)	232(32.09)	345(47.72)
Religion	Orthodox	132(18.26)	220(30.43)	352(48.69)
	Muslim and others	104(14.38)	267(36.93)	371(51.31)

As can be seen in Table 5.1 female interrupters (55.33%) are slightly more than that of males (44.67%) while 66% of female patients and 69% of male patients resume back to treatment. Patients with no education in the level of education group interrupt their treatment less (about 20.75%) compared to patients who have primary (31.54%) or secondary and above (47.72%) level of education.

Patients who are currently unmarried interrupt treatment more than that of currently married, that is, the number of patients who are currently unmarried and who interrupt their treatment is about 64.45%, more than currently married patients. From the currently unmarried patients 43.85% of them resume the therapy. In addition, comparing patients with respect to their

religion, patients from Orthodox religion interrupt treatment slightly less as compared to the Muslims and others. But, patients from Orthodox religion who did not resume to therapy (about 37.5%) are greater than patients from Muslims and others group (28%). Next, we discuss whether the observed differences using descriptive measures are statistically significant or not with the help of Kaplan-Meier survival estimates and the log rank test.

Plots of Kaplan-Meier survival estimates displayed in Figure B2.2, Annex 2 can be used for visual comparison of the survival experience of the HIV-1 patients on HAART interruption and resumption based on the various categories of socio-demographic characteristics. Moreover, Log-rank test of equality across categories is used to explore whether the difference in survival experience is significant or not. The results of the test with a modest level of significance will help to screen out the categorical variables for inclusion in the multivariable model.

5.1.2 Comparison of survival experience by socio-demographic variables

In order to investigate if there is significant difference between the time to resumption to treatment by gender, Kaplan-Meier survivor estimates for the two gender groups are plotted in Figure B2.2, Annex 2. This figure shows that the curves are not much different from each other indicating that the time to treatment resumption for the male and female patients may not be different significantly. Statistical test of significance is made by using log-rank test (Table 5.2) and this shows that there is no statistical significant difference between male and female patients with respect to time to treatment resumption (Chi-square = 0.1, d.f.=1, p-value = 0.3957). The median survival time for time to treatment resumption of female patients who interrupt their therapy is 381 days and for male patients is 367 days.

Similar analysis is performed to investigate difference in the time to treatment resumption among patients with respect to marital status: currently not married and currently married. Observation from the Kaplan-Meier curves in Figure B2.2, Annex2 show that the curves overlap each other indicating that time to resumption to treatment may be the same for these groups. Log-rank test (Table 5.2) shows that there is no evidence against the null hypothesis of no difference in the time to treatment resumption with Chi-square value of 1.61 at 1 degrees of freedom (p-value = 0.2042).

Table 5.2.The Log-rank test for the comparison of survival experience on chronic HIV-1 patients at ZMH using socio-demographic variables

Categorical covariates	Chi-Square	Df.	Sig.
Sex	0.72	1	0.3957
Marital Status	1.61	1	0.2042
Level of education	2.68	2	0.2615
Religion	1.21	1	0.2707

Df.- degree of freedom

The other categorical variable included in the study is level of education of the patients. As the result depicts that there is no statistical significant difference in the time to treatment resumption among the different levels of educational status attained by the patients: no education, primary and secondary or above. In addition, religion of the patient has no any significant impact on the difference in the survival times of patients' interrupted and resumed to treatment.

5.2. Survival experience with respect to the continuous covariates

Considering the continuous covariates first we reported the median values of each continuous covariate as follows. The median duration of follow up for treatment is 256 days with IQR (180 - 409), similarly the median CD4 count of patients interrupting their therapy when enrolled in the follow up for the first time (CD4 at the start of HAART) is 119 cells/µl IQR (76 - 178) where as their CD4 count at the end of follow up and at the start of interruption (CD4 prior to interruption) is 199 cells/µl IQR (121 - 272) (Table A1.1, Annex1).

We next used Cox proportional hazards model to assess the relationship between each covariates and time to treatment resumption. In handling ties in the covariates, Breslow's method is adopted in case there may be many tied failure time. Table 5.3 shows the result of the univariate analysis along with the contribution to the likelihood function -2 log likelihood

(-2LOGL), and p-values. As can be seen from Table 5.3, covariates like religion, weight of the patient, duration of disease and duration of follow-up are statistically significant ($P < 0.05$).

Some of the covariates to be included in the multiple covariates model to investigate the joint effect of covariates on time to resumption of treatment. Using a modest level of significance of 25% to include potential covariates, from the categorical variables presented in Table 5.3 religion, marital status and education and from the continuous covariates presented in the same Table 5.3 age of the patient, weight of the patient, the CD4 count at the start of treatment, disease duration and duration of follow-up are chosen to be included in the multiple covariates model for further investigation. In addition, the other categorical and continuous covariates sex and CD4 prior to interruption are purposefully retained even though they have large P-values.

Table 5.3: Single covariate analysis of Cox proportional hazards on time to resumption of treatment for chronic HIV-1 patients who interrupt their HAART at ZMH

Parameter		DF	Parameter Estimate	Standard Error	Pr > Chi Sq	Hazard Ratio	95% CI for HR		-2 LOG
Sex	(F)	1	-0.0776	0.0913	0.3952	0.925	0.7736	1.1067	5193.92
	(M)	0	.	.	.	.	------	------	
Marital	(married)	1	-0.1214	0.0959	0.2057	0.886	0.7340	1.0689	5193.015
	(not married)	0	.	.	.	.	------	------	
Level educ	(no education)	1	0.1915	0.1185	0.1063	1.211	0.9510	1.5277	5192.017
	(Primary)	1	0.0834	0.1054	0.429	1.087	0.8841	1.3364	
	(Secondary and above)	0	.	.	.	.	------	-------	
Religion	(Orthodox)	1	-0.2405	0.1107	0.0298	0.786	0.6329	0.9767	5190.124
	(Others)	0	.	.	.	.	------	-------	
Age		1	-0.717	0.457	0.1168	0.993	0.1993	1.1957	5192.144
Weight		1	0.1643	0.0539	0.0023	1.017	1.0604	1.3099	5185.394
CD4 at the start of HAART		1	0.0959	0.0698	0.1694	1.001	0.9599	1.2619	5192.771
CD4 prior to interruption		1	0.2587	0.4549	0.5695	1.2952	0.5310	3.1592	5194.315
Disease duration		1	0.140	0.0209	<0.0001	1.1503	1.1041	1.1983	5153.325
Duration of follow-up		1	0.116	0.0202	<0.0001	1.1229	1.0794	1.1683	5166.599

5.3 Multiple covariates analysis

The problem with any single covariate analysis approach is that it ignores the possibility that a collection of covariates, each of which is weakly associated with the outcome, can become an important predictor of the outcome variable biologically or clinically when taken with the other covariates together. The results of the multiple covariates Cox proportional hazards model for all covariates is displayed in Table 5.4.

Table 5.4: The parameter estimates, standard errors and the hazard ratios of the Cox-proportional hazards model for chronic HIV-1 patients interrupting their HAART at ZMH

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% CI for HR	
Sex	(F)	1	-0.0926	0.1010	0.8409	0.3591	0.9116	0.7479	1.1110
	(M)	0	.	.	.	.	1	-------	--------
Marital	(married)	1	-0.0489	0.0991	0.2435	0.6217	0.9523	0.7841	1.1565
	(not married)	0	.	.	.	.	1	-------	-------
Leveduc	(no education)	1	0.2556	0.1232	4.2996	0.0381	1.2912	1.0141	1.6439
	(primary)	1	0.0676	0.1071	0.3989	0.5277	1.0700	0.8674	1.3198
	(Secondary and above)	0	.	.	.	.	1	-------	-------
Religion	(orthodox)	1	-0.0352	0.0942	0.1399	0.7084	0.9654	0.8026	1.1611
	(others)	0	.	.	.	.	1	-------	-------
Age		1	-0.1044	0.0509	4.2057	0.0403	0.9009	0.8153	0.9954
Weight		1	0.1418	0.0581	5.9507	0.0147	1.1523	1.0283	1.2913
CD4 at the start of HAART		1	0.260	0.0936	7.7054	0.0055	1.2969	1.0796	1.5580
CD4 prior to interruption		1	-0.234	0.0710	10.8769	0.0010	0.7914	0.6886	0.9094
Disease duration		1	0.106	0.0227	21.6354	<0.0001	1.1118	1.0634	1.1624
Duration of follow-up		1	0.130	0.0267	23.5670	<0.0001	1.1388	1.0807	1.2001

In addition, interaction terms which have feasible, meaningful interpretations and statistically significant effects on the time to treatment resumption at 5% level of significance are considered and were incorporated in to the model. The preliminary final model with the plausible interaction and main effects that are significant at 5% level are presented in Table 5.5.

Table 5.5: The parameter estimates, standard errors and the hazard ratios of the Cox-proportional hazards model including the interaction terms for chronic HIV-1 patients interrupting their HAART at ZMH

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% CI for HR	
Sex	1	0.3973	0.2677	2.2027	0.1378	1.4878	0.8803	2.5145
Level of educ (no educ)	1	0.2339	0.1149	4.1445	0.0418	1.2635	1.0087	1.5826
Marital status	1	-0.1191	0.1307	0.8313	0.3619	0.8877	0.6871	1.1468
Age	1	-0.1362	0.0529	6.6309	0.010	0.8727	0.7867	0.9680
Weight	1	0.1707	0.0583	8.5684	0.0034	1.1861	1.0581	1.3297
CD4 at start of HAART	1	0.237	0.0946	6.2869	0.0122	1.2674	1.0529	1.5257
CD4 prior to interruption	1	-0.234	0.0703	11.1159	0.0009	0.7914	0.6895	0.9083
Disease duration	1	-0.444	0.6054	0.538	0.4633	0.6415	0.1958	2.1013
Duration of follow-up	1	0.138	0.0265	27.1041	<0.0001	1.1480	1.0899	1.2091
Sex*Marital status	1	-0.0549	0.0249	4.8593	0.0275	0.9466	0.9016	0.9939
Disease dur*Marital	1	0.0722	0.0261	7.6278	0.0057	1.0748	1.0211	1.1314

5.4 Model diagnostics

After arriving at the most parsimonious preliminary final model, the next step and most important in statistical analysis is to diagnose the fit of the model. As in the case for a linear or generalized linear model, it is desirable to determine whether a fitted Cox regression model adequately describes the data. In this study, four kinds of diagnostics: for violation of the assumption of proportional hazards; for outliers and influential observations; for goodness-of-fit using R^2 and for nonlinearity in the relationship between the log hazard and the covariates are discussed below.

5.4.1 Checking the linearity of covariates in the model

In order to assess the linearity assumption on the part of the covariates, we plot martingale residuals compared by excluding the covariate to be checked for linearity against the values of the covariate. The scatter plots or the smoothed curve using lowess can be used to check the linearity assumption. As can be seen from the plots in Figure C3.1 (Annex 3), there is no definite pattern in the scatter plots and the smoothed curve is almost a horizontal line through the origin. These are indicators of approximately linearity in the covariates. Thus, there are no

signs of nonlinearity in all of the continuous covariates indicating that the assumption of linearity is more or less fulfilled by all the covariates.

5.4.2 Checking the proportionality of covariates in the model

One of the main assumptions of the Cox proportional hazard model is proportionality of hazards. The adequacy of the preliminary final model is checked for the validity of proportional hazards assumption using test based on the interaction of the covariates with the log of time and also using the plot of the scaled schoenfeld residuals. The formal test applied to the model presented in Table 5.6 shows that the time-dependent covariates (interaction of covariates with logarithm of time) were not significant and the global fit test also shows that all the covariates were not significant which justifies the PH assumption holds at 5% level of significance. The plot of the scaled schoenfeld in Figure C3.2 (Annex 3) depicts that the residuals are random without any systematic pattern and the smoothed plot looks straight line without any departure from the horizontal line. This also implies that there is no violation of proportional hazards assumption.

Table 5.6: Statistical test for proportional hazards assumption (PH) of the covariates and their interaction with log of time (duration of interruption), for the data in chronic HIV-1 patients who interrupt their therapy at ZMH.

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Age	1	0.0250	0.0368	0.4602	0.4975	1.0253
Sex	1	-1.1679	0.7822	2.2294	0.1354	0.3110
Marital status	1	-0.6797	0.4812	1.9952	0.1578	0.5068
Level of educ1	1	-0.9630	0.8514	1.2793	0.258	0.3818
Weight	1	0.0316	0.0402	0.6158	0.4326	1.0321
CD4 at start of HAART	1	0.0391	0.0628	0.3872	0.5338	1.0399
CD4 prior to interruption	1	-0.223	0.419	0.2823	0.5952	0.8001
Disease duration	1	0.0525	0.16	0.107	0.7436	1.0539
Duration of follow-up	1	0.0308	0.0144	4.5978	0.032	1.0313
Marital Status *Sex	1	-0.1071	0.1215	0.7767	0.3782	0.8985
Disease duration *Marital	1	0.0718	0.0271	7.0459	0.0079	1.0744
Age*lnt	1	-0.0718	0.066	1.1845	0.2764	0.9307
Sex*lnt	1	-0.0199	0.0689	0.0835	0.7726	0.9803
Marital status*lnt	1	0.0982	0.0796	1.521	0.2175	1.1032
Level of educ1*lnt	1	0.2182	0.1521	2.0583	0.1514	1.2439
Weight*lnt	1	-0.253	0.72	0.1233	0.7255	0.7765

CD4 at start HAART*lnt	1	-0.0326	0.113	0.0828	0.7735	0.9679
CD4 prior to interruption*lnt	1	-0.0145	0.7867	0.0003	0.9853	0.9856
Disease duration*lnt	1	-0.1713	0.2648	0.4184	0.5177	0.8426
Duration of follow-up*lnt	1	-0.033	0.0277	1.4194	0.2335	0.9675

Label	Wald Chi-Square	DF	Pr > ChiSq
Proportionality_test	9.6535	9	0.3793

5.4.3 Checking for influential and outlier observations

Furthermore, a thorough evaluation of regression diagnostic statistic to identify, if any, subjects that have undue influence on the estimates of the Cox regression parameters, or have an unusual configuration of the covariates, or have an unexpected influence on the fit of the model is carried out using score residuals. Leverages, similar to what is obtained in linear and logistic regression, are also adapted into proportional hazards regression through the score residuals, to examine if there are subjects with undue influence on the fit. This is done through the plots shown for each covariate included in the model. This plot compares the magnitudes of the largest covariate values, it may be positive or negative, to the mean of the covariate, as we can see from the plot it appears like a fan shape. If the plot fans out from the narrow point that is approximately from the mean of the covariate then it suggests that none of the observations is terribly influential in the study. As can be observed in Figure C3.3 (Annex 3) some patients have a large spike and these patients are suspected to have undue influence on the parameter estimates. To check their influence the suspected subjects were removed one at a time and model is refitted. There were no large change in the model estimates and hence these patients are not as such influential outliers and then retained in the final model.

5.4.4 Goodness of fit

One method of checking goodness of fit of the model is to use R^2. In proportional hazards regression model as in all regression analyses there is no single, simple method of calculating and interpreting R^2, because in Cox proportional hazards model, R^2 depends on the proportion of the censored observations in the data. A perfectly adequate model may have what, at face value, seems like a terribly low R^2 due to high percent of censored data (Hosmer

and Lemeshow, 1998). Therefore, for the model fitted in this study the value of R^2 is $1-\exp\left[{}^{1}\!/\!_{723}(5102.406-5194.365)\right]=$ 0.1194. In addition, results of the Likelihood Ratio, Score and Wald tests for model goodness of fit displayed in Table 5.7, suggests that model is good fit, i.e. significant at 5% level of significance.

Table 5.7: The Likelihood Ratio, Score and Wald tests for overall measures of goodness of fit of the final model

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	92.2296	11	<.0001
Score	103.4609	11	<.0001
Wald	101.4117	11	<.0001

5.5 Interpretation and Discussion of the results

We present the results of a retrospective cohort study dealing with the interruption and resumption of Highly Active Antiretroviral Therapy (HAART) in chronic HIV-1 infected patients with age greater than 18 years.

This study tries to look the rate at which patients resume to treatment follow-up over time, the determinant predictive factors on the time to resumption in patients interrupting their HAART. The survival experience of patients in the different categories of demographic variables is estimated using the Kaplan-Meier method and compared using log rank test. The predictive factors for resumption were assessed using the Cox proportional hazard models.

In this study about 67% (487) patients resumed their HAART after at least three months of treatment interruption and about 33% (236) patients remained without resuming to their therapy which are censored observations. Even though, 487 of the patients restarted Highly Active Antiretroviral treatment (HAART), due to a lack and absence of information on structured treatment interruption in HIV-infected patients, the results we present can only be compared to those who did not resume to treatment.

The Cox's PH model fitted using complete case analysis found six main effects and two interactions that are statistically significantly associated with the hazard of having treatment resumption. The six main effects are education, baseline age, baseline weight, CD4 count at

the start of treatment, CD4 count prior to treatment interruption and duration of follow-up. The two interaction terms are between marital status and disease duration i.e., the interaction between a continuous and a dichotomous and between sex and marital status which is interaction between two dichotomous covariates. The fitted model satisfy the assumption of proportional hazards as assessed through plots of scaled schoenfeld residual and the tests based on the interaction of covariates with log of time.

The Cox regression coefficients in the final model are interpreted as follows. The first factor which affects time to treatment resumption is the level of education of the patient. The hazard rate of time to resumption for Patients who have no education is about 26% higher than for patients who have secondary or above education (HR 1.264, P=0.0418, CI 1.0087, 1.5826) after adjusting for other covariates. The confidence interval indicates that the rate could actually be as much as 58% higher than or as little as 1% higher. The baseline age at the start of treatment decreases the hazard of time to resumption by 13% (HR 0.8727, P< 0.010, CI 0.7867, 0.9680) adjusting for other effects in the model. That is for every one year increase in the baseline age of patients, the hazard rate decreases by 13% controlling the effects of all other covariates in the model. The 95% confidence interval suggests that the rate could be as much as 21 percent lower up to 3 percent lower. Weight is the other covariate which has a significant impact on the time to treatment resumption and increases the hazard of time to treatment resumption by 19% (HR 1.1861, P=0.0034, CI 1.0529, 1.5257). The confidence interval indicates that for one kg increases in weight the hazard rate could increase as much as 52% or as little as 5%.

Similarly, when we consider the immunological factors that have impact on the time to treatment resumption are the following: CD4 count at starting of HAART increases the hazards of treatment resumption by 27% (HR 1.2674, P=0.0122, CI 1.0529, 1.5257) that is, by 27% when adjusting for other effects in the model for patient with a unit increase in CD4 at starting HAART. The 95 percent confidence suggests that an increased rate of reinitiation to treatment as high as 52% and as lower as 5%. Moreover, CD4 count prior to treatment interruption decreases the hazard of time to resumption by 21%, HR = 0.7914 (P=0.0009, CI 0.6895, 0.9083). The confidence interval shows that the rate may goes down to 31% and to 9% down. Also, the duration of treatment follow-up increases the hazards of having resumption to treatment increases by a factor HR 1.1480 (P<0.0001, CI 1.0899, 1.2091), that is, it increases by 14% for a patient with a unit increase in the duration of follow-up adjusting

for other covariates. The 95 percent confidence interval indicates that the hazard rate of treatment resumption tends to a maximum of 20 percent and a minimum of 8 percent.

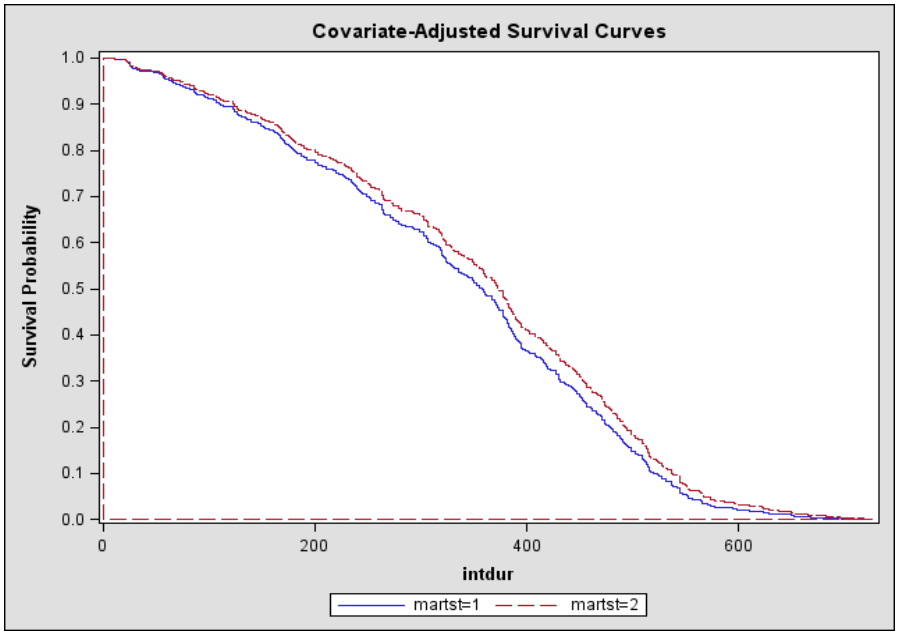
From the result we found that as the duration of treatment follow up increases, patients tend to interrupt for a longer period of time i.e., the time to treatment resumption will become longer. The reason is that because of the effect of drug toxicity, drug-drug and/or drug-food interactions, adverse ART drug reactions and in some patents drug resistant viruses are the causes for longer time to treatment resumption. Similarly, as the weight increases the time to treatment reinitiation for patients interrupting their HAART also increases. Due to improvement in the weight, health condition and suppression of the plasma viral load of HIV-1 patients on HAART follow-up, these patients feel as they are safe and went for prolong time to resumption of HAART. The other one is CD4 count prior to treatment interruption decreases the period of treatment reinitiation after suspension of HAART. This is because of the incidence of opportunistic infections such as Kaposi's sarcoma, pulmonary TB, bacterial pneumonia are some of the OIs and in addition to these OIs rebound of drug resistant viraemia are the cause for early reinitiation of HAART.

Finally, we consider the interaction of covariates which have a significant impact on the time to treatment resumption; one of these interaction terms is the interaction between duration of disease and marital status. Since marital status is categorized into two as currently married and not currently married, we present the hazard ratios for disease duration at each group of marital status. So, the interaction term reduces the rate of hazards to treatment resumption by 36% in married patients as the disease duration increases by one day. Whereas, the hazard of treatment resumption for patients who are currently not married, reduced by 31% for a one day increase in disease duration. Hence, as the duration of disease increases, currently married patients resumed treatment at slightly higher rate than currently not married patients. The other one is the interaction between the two dichotomous covariates sex and marital status. Similarly, the result suggests that the hazard rate differs between male and female patients at each group of marital status. That is, currently unmarried female patients resume to treatment at a rate that is about 49% higher than currently unmarried male patients. Whereas currently married female patients reinitiate their treatment at higher rate than male married patients (about 41% higher).

5.6 Covariate adjusted survival curves

The covariate adjusted survival curves based on marital is presented in Figure 5.1 below. The figure suggests that the survival experience of currently married patients after adjusted for the covariates is slightly better than currently unmarried patients.

Figure 5.1: Graph of the covariate adjusted survivorship functions for the marital status computed from the final model in Table 5.6



CHAPTER SIX
CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

Due to the unforgiving nature of the virus, HAART is complex and requires a high level of diligence and endurance. If the prescribed treatment is not accurately adhered to, several very bad consequences result. It is understood that duration of treatment interruption allows HIV to rebound and overwhelm the immune system. With these physiological defenses impaired, other pathogens can thrive and cause severe deterioration of the individual's health, eventually leading to death if left untreated. In addition to showing how it is a detriment to the patient, this paper discussed how the outcome of drug holiday affects the patients time to resumption of treatment. Poor adherence negates this preventative effect of HAART by allowing the re-escalation of viral concentrations. In addition, treatment interruption may result in the development of drug-resistant strains of HIV that poses an intense and unrelenting threat to public health worldwide.

In other words, resistance is directly correlated with interruption of treatment and may be the direst consequence of improper medication compliance. In this study, the median time to treatment resumption or duration of interruption is estimated to be around 377 days using Kaplan-Meier survivorship estimation. The full impact of the consequences of treatment interruption can not be predicted with any assurance, but in every scenario, the outlook is grave. For these reasons, special attention must be paid to the challenges affecting HAART discontinuation to insure the best possible outcome for the patient and avoid the deterioration of HIV/AIDS healthcare.

Since continuous follow-up to HAART is difficult even under the most ideal circumstances; unfortunately, the majority of people infected with HIV live in the most resource-constrained and underserved regions in the world. Clinical/immunological factors play a primary role in the maintenance of HAART. This study highlighted how the physiological and immunological factors influence a patient's inclination to stop and resume HAART. That is, the study identified education, baseline age, baseline weight, CD4 count at the start of treatment, CD4 prior to treatment interruption, duration of follow-up and the interactions

between marital status and sex and marital status and disease duration are important factors related to time to resumption of treatment.

The societal factors surrounding HAART resumption are tremendously diverse and complex. None of these challenges can be dissected away and dealt with separately addressing the discontinuation and resumption on HAART and this study is an attempt to jointly model and investigate demographic and immunological factors on time to treatment resumption. As a result, HIV/AIDS patients who interrupted their HAART took longer time to resume treatment are characterized by having lower level of education, younger age, higher weight, higher CD4 count at the start of HAART, lower CD4 count prior to treatment interruption, longer duration of treatment follow-up, currently not married male patients at longer disease duration.

6.2 Recommendations

In our country, the high prevalence and adherence of HIV/AIDS patients increases the need and use of HAART. The government of Ethiopia has made an effort to establish an integrated National ART Program to deliver free ARV drugs across the nation. Thus, the national and international NGO's are also providing the ARV drugs in different health institutions freely for improving the quality of life of people living with HIV/AIDS (PLWHA) and mitigating some of the impact of the epidemic. But in parallel with this, due to many reasons patients are not taking HAART properly so that they encountered different HAART complications as well. Inspite of the principle of "treatment to all" by the government there is poor quality of implementation of this treatment because of poor compliance of the patients.

In order to know the causes of patients interrupted or stopped treatment, data on the HAART interruption tracing visit should be provided in a proper and manageable form. Data could then be analyzed based on the true outcome status of patients at the HAART tracing visit at different health centres. Through this data of HAART tracing visit, the status of patients found to be alive, dead and possible reasons for unsuccessful tracing could be known. Easy assess to this information will go along way in tracing the causes of patients who stopped treatment.

According to the results of this study the main predictive factors for the duration of treatment interruption are more of the clinical variables. So, the clinicians are expected to perform the

maximum efforts to change the behavior and attitude of patients taking HAART, i.e, changing the behavior of patients towards use of ART medication should be given emphasis. The following measures should be taken on drug supply management; patient's proper use of HAART and delivery of quality services to promote the well-being of HAART service users.

- Since the ARV drugs treatment cause adverse effects on some patients because of long duration of follow-up, it would be advisable for patients to interrupt their treatment according to the principle of structured treatment interruption (STI) with continuous assessment of clinical examination (CD4 and viral load count).

- Moreover, for patients taking HAART, good clinical/immunological records like increasing weight and higher CD4 count at the start of HAART do not mean that the patient is "safe", so continuous attendance is expected from the patient to improve his/her quality of life and to reduce mortality due to stopping therapy, Since HAART is a life long treatment.

- As there was difficulty to collect data on patients who interrupt their HAART because of incomplete addresses to trace back the interrupted patients and then to know their status after interruption, it is suggested that a complete recode of addresses of patients be properly compiled. Since this will help further studies to be conducted in this area.

References:

Anderson P.K. and Gill R.D. (1982) Cox's regression model for counting process: a large sample study; Annals of statistics, 10: 1100-1120.

Assegid Tassew and Ahmed Ali (2007): Assessment of the magnitude of Adverse Drug Reactions (ADRs) and associated factors among patients on ART; the case of Zewditu Memorial Hospital.

AVERT, AVERTing HIV and AIDS (2008): Types of HIV/AIDS causing viruses http://www.avert.org/

Bansi LK, Benzie AA, Phillips AN, Portsmouth S, Hill T, Leen C, Schwenk A, Johnson M, Anderson J, Gilson R, Easterbrook P, Gazzard B, Fisher M, Orkin C, Porter K, Pillay D, Taylor GP, Walsh JC, Sabin CA (2008): Are previous treatment interruptions associated with higher viral rebound rates in patients with viral suppression? AIDS, 22:349–356

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L (1983): Isolation of a T-lymphotropic retrovirus from a patient at risk for AIDS. Science, 220: 868-71. http://amedeo.com/lit.php?id=6189183

Boschi A, Tinelli C, Ortolani P, Moscatelli G, Morigi G and Arlotti M (2004): CD4+ cell-count-guided treatment interruptions in chronic HIV-infected patients with good response to highly active antiretroviral therapy. AIDS: 18:2381–2389.

Broder S and Gallo RC (1984): A pathogenic retrovirus (HTLV-III) linked to AIDS. New England Journal of Med. 311:1292-7

Centers for Disease Control and Prevention (1992): "1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults." *MMWR Recomm Rep* 41:1-19

Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira O. (1986): Isolation of a new human retrovirus from West African patients with AIDS. Science, 233: 343

Collet, D. (1994) Modeling Survival Data in Medical Research. 1st edition, Chapman & Hall, London, UK.

Concorde (1994): MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. Lancet, 343:871-81. http://amedeo.com/lit.php?id=7908356

Cox D.R. (1972): Regression models and life tables (with discussion), journal of Royal statistical society, 34: 187-220

Cox D.R. (1975): partial likelihood: Biometrica, 62, 269-276

Cox D.R. and Snell E.J. (1989) The Analysis of Binary Data. 2nd edition, Chapman and Hall, Londin, UK.

D'Arminio Monforte A, Lepri AC, Rezza G, Pezzotti P, Antinori A, Phillips AN, Angarano G, Colangeli V, De Luca A, Ippolito G, Caggese L, Soscia F, Filice G, Gritti F, Narciso P, Tirelli U and Moroni M (2000): Insights into the reasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naive patients. AIDS; 14: 499–507.

Degu Jerene and Bernt Lindtjørn (2005): Disease Progression among UntreatedHIV- Infected Patients in South Ethiopia: Implications for Patient Care: Medscape General Medicine; 7(3):

Degu Jerene, Are Næss and Bernt Lindtjørn (2006): Antiretroviral therapy at a district hospital in Ethiopia prevents death and tuberculosis in a cohort of HIV patients: AIDS Research and Therapy; 3(10)

Delta, V.S. (1996). A randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. Lancet, 348:283-91.
http://amedeo.com/lit.php?id=8709686

Ethiopia HIV/AIDS Prevention & Control Office (HAPCO) and Global HIV/AIDS Monitoring and Evaluation Team (GAMET) (2008): HIVAIDS in Ethiopia-An Epidemiological Synthesis; the World Bank Global HIV/AIDS Program http://www.worldbank/aids

Fardet L, Mary-Krause M, Heard I, Partisani M and Costagliola D. (2006): Influence of gender and HIV transmission group on initial highly active antiretroviral therapy prescription and treatment response. HIV Med; 7:520–9.

Federal HIV/AIDS Prevention and Control Office Federal Ministry of Health (FHAPCO) (2007): Guidelines for implementation of antiretroviral therapy programme in Ethiopia; Addis Ababa, Ethiopia.

Fernandez-Guerrero ML, Rivas P, Molina M, Garcia R, and De Go´ rgolas M (2005): Long-Term Follow-Up of Asymptomatic HIV-Infected Patients Who Discontinued Antiretroviral Therapy; Clinical Infectious Diseases ; 41:390–4

Fillaux J, Delpierre C, Alvarez M, Jaafar A, Marchou B, Massip P, Cuzin L (2006): Predictive factors of treatment interruption duration in a cohort of HIV-1 infected patients with CD4 count greater than 350 cells per mm^3: Médecine et maladies infectieuses; 36: 335–339

Fischl MA, Parker CB, Pettinelli C, Wulfsohn M, Hirsch MS, Collier AC, Antoniskis D, Ho M, Richman DD, Fuchs E, (1990): A randomized controlled trial of a reduced daily dose of zidovudine in patients with the acquired immunodeficiency syndrome. N Engl JMed; 3:1009-14. http://amedeo.com/lit.php?id=1977079

Florence E, Lundgren J, Dreezen C, Fisher M , Kirk O, Blaxhult A, Panos G, Katlama C, Vella S and Phillips A (2003): Factors associated with a reduced CD4 lymphocyte count response to HAART despite full viral suppression in the EuroSIDA study. HIV Med; 4:255–262.

Giuntini R, Marttinelli C, Ambu S, Bartolozzi D, Farese A, Boddi V and Leoncini F: (2005) Prolonged treatment interruption in chronic HIV infection: a new strategy? AIDS, 19:209–215.

Grambsch P.M. and Therneau T.M. (1994). Proportional hazards tests in diagnostics based on weighted residuals. Biometrica, 81: 515-526.

Hammer SM, Katzenstein DA, Hughes MD, Gundacker H, Schooley RT, Haubrich RH, Henry WK, Lederman MM, Phair JP, Niu M, Hirsch MS, Merigan TC (1996). A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. N Engl J Med, 335:1081-90. http://amedeo.com/lit.php?id=8813038

Hoffmann C, Rockstroh JK, and Kamps BS, (2007): HIV Medicine; Flying Publisher, 15th edition. http://www.hivmedicine.com/textbook/nw.htm.

Hosmer D.W. and Lemeshow S. (1998): Applied Survival Analysis: Regression Modeling of Time to Event Data; Wiley Series in Probability and Statistics. Canada.

Institute of Medicine (2005a): Public financing and delivery of HIV/AIDS care: securing the legacy of Ryan White. Washington, DC: National Academies Press.

Institute of Medicine (2005b): Scaling up treatment for the global AIDS pandemic: challenges and opportunities. Washington, DC: National Academies Press.

Interrupting HIV treatment and Treatment interruption strategies. http://www.aidsmap.com/cms1032040 and 1039438

Kebede D, Fikre H, Sibhatu B, Alemayehu A and Biruk KB (2008): Defaulters from antiretroviral treatment in Jimma University Specialized Hospital, Southwest Ethiopia: Tropical Medicine and International Health; 13(3): 328–333

Kiguba R, Byakika-Tusiime J, Karamagi C,Ssali F, Mugyenyi P and Katabira E (2007): Discontinuation and Modification of Highly Active Antiretroviral Therapy in HIV-Infected Ugandans Prevalence and Associated Factors: J Acquir Immune Defic Syndr; 45:218–223

Kirk O, Gerstoft J, Pedersen C, Nielsen H, Obel N, Katzenstein TL, Mathiesen L and Lundgren JD (2001): Low body weight and type of protease inhibitor predict discontinuation and treatment-limiting adverse drug reactions among HIV-infected patients starting a protease inhibitor regimen: consistent results from a randomized trial and an observational cohort. HIV Medicine; 2, 43-51

Kleinbaum D. G. (1996). Survival Analysis: A Self-Learning Text; Springer- Verlag, Statistics in the Health Sciences, New York.

Lamid MA, Ziv S, Florence E and Menten J (2007): Outcome of Chronic Hiv-1 Patients who interrupt their Highly Active Antiretroviral Treatments. http://hdl.handle.net/1942/3506

Laurent C, Ngom Gueye NF, Ndour CT, Gueye PM, Diouf M, Diakhate N, Touré Kane NC, Laniece I, Ndir A, Vergne L, Ndoye I, Mboup S, Sow PS, Delporte E (2005): Long-term benefits of highly active antiretroviral therapy in Senegalese HIV-1-infected adults: J Acquir Immune Defic Syndr; 38 (1):14–17.

Lyles RH, Munoz A, Yamashita TE, Bazmi H, Detels R, Rinaldo CR, Margolick JB, Phair JP, and Mellors JW (2000): Natural history of HIV type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. J Infect Dis, 181:872-880. http://amedeo.com/lit.php?id=10720507

Mantel N. and Haenszel W. (1959): Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719-748

Marubini E. and Valsecchi M. G. (1995): Analyzing Survival data from clinical trials and observational studies, statistics in practice; John Wiley and Sons, Chichester, U.K.

Mata RC, Viciana P, de Alarcon A, Lopez-Cortes LF, Gomez-Vera J, Trastoy M, Cisneros JM (2005): Discontinuation of antiretroviral therapy in patients with

chronic HIV infection: clinical, virologic, and immunologic consequences. AIDS Patient Care STDS.19:550–562

Mellors JW, Munoz AM, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, Kingsley LA, Todd JA, Saah AJ, Detels R, Phair JP and Rinaldo CR Jr.(1997). Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Ann Intern Med. 126:946-954. http://amedeo.com/lit.php?id=9182471

Ministry of Health (2002 and 2006): AIDS in Ethiopia, Sixth Report, Addis- Ababa, Ethiopia.

Moore AL, Sabin CA, Johnson MA and, Phillips AN (2002): Gender and clinical outcomes after starting highly active antiretroviral treatment: a cohort study: J Acquir Immune Defic Syndr; 29: 197–202.

Morse EV, Simon PM, Coburn M, Hyslop N, Greenspan D & Balson PM (1991). Determinants of subject compliance within an experimental anti-HIV drug protocol. *Social Science and Medicine*, 32: 1161-1167.

Mussini C, Bedini A, Borghi V, Guaraldi G and Esposito R, Barchi E, Cozzi-Lepri A, Phillips AN Ortolani P, Bratt G, Eriksson LE, Sighinolfi L, Cossarizza A, d'Arminio Monforte A, De Luca A Di Giambenedetto S and Antinori A (2005): CD4 cell-monitored treatment interruption in patients with a CD4 cell count > 500 × 106 cells/l; AIDS, 19:287–294

Nacher M, El Guedj M, Tania Vaz T, Nasser V, Randrianjohany A, Alvarez F, Sobesky M, Magnien C, And Couppiép P (2006): Risk Factors For Follow-Up Interruption Of Hiv Patients In French Guiana: Am. J. Trop. Med. Hyg., 74(5), pp. 915–917

Ng'andu N.H. (1997): an empirical comparison of statistical tests for assessing the proportional hazards assumption of Cox's model, statistics in medicine, 16: 611-626

Nicastri E, Angeletti C, Palmisano L, Sarmatic L, Antonio Chiesib A, Andrea Geracib A, Andreonic M and Vellab S (2005): Gender differences in clinical progression of HIV-1 infected individuals during long term highly active antiretroviral therapy: AIDS; 19: 577–583.

Office of AIDS Research Advisory Council (OARAC) (2008): Guidelines for the use of Antiretroviral Agents in HIV-1 Infected Adults and Adolescents. http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf.

Olsen CH, Mocroft A, Kirk O, Vella S, Blaxhult, A, Clumeck N, Fisher M, Katlama C, Phillips AN and Lundgren JD (2007): Interruption of combination antiretroviral therapy and risk of clinical disease progression to AIDS or death; British HIV Association, HIV Medicine 8, 96–104

Oyugia JH, Byakika-Tusiime J, Ragland K, Laeyendeckere O, Mugerwab R, Kityoc C, Mugyenyic P, Quinne TC and Bangsberg DR (2007): Treatment interruptions predict resistance in HIV-positive individuals purchasing fixed-dose combination antiretroviral therapy in Kampala, Uganda; AIDS, 21:965–971

Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD (1998): Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection HIV Outpatient Study Investigators; New England Journal of Medicine; 338: 853–860.

Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, Wagener MM and Singh N (2000). Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med; 133:21-30.

Pellegrin I, Thie´baut R, Blanco P, Viallard JF, Schrive M, Merel P, Cheˆne G, Fleury H, Moreau JF and Pellegrin JC (2005): Can Highly Active Antiretroviral Therapy be Interrupted in Patients With Sustained Moderate HIV RNA and >400 CD4þ Cells/*m*l? Impact on Immunovirological Parameters: Journal of Medical Virology 77:164–172

Poga´ny K, Vanvalkengoed IG, Prins JM, Nieuwkerk PT, Ende I, Kauffmann RH, Kroon FP, Verbon A, Nievaard MF, Lange J M.A, and Brinkman K (2007): Effects of Active Treatment Discontinuation in Patients With a CD4+ T-Cell Nadir Greater Than 350 Cells/mm3 48-Week Treatment Interruption in Early Starters Netherlands Study: J Acquir Immune Defic Syndr ;44:395–400

Saitoh A, Foca M, Viani RM, Heffernan-Vacca S, Vaida F, Lujan-Zilbermann J, Emmanuel PJ, Deville JG, Stephen A and Spector SA (2008): Clinical Outcomes After an Unstructured Treatment Interruption in Children and Adolescents With Perinatally Acquired HIV Infection; PEDIATRICS Vol. 121 No. 3; 513-521

Schoenfeld D. (1982): partial residuals for the proportional hazards regression model. Biometrica, 69: 239-241

Srasuebkul P, Ungsedhapand C, Ruxrungtham K, Boyd MA, Phanuphak P, Cooper DA, and Law MG (2007): Predictive factors for immunological and virological endpoints in Thai patients receiving combination antiretroviral treatment; HIV Medicine; 8, 46–54

Taffe P, Rickenbach M, Hirschel B, Opravil M, Furrer H, Janin P, Bugnon F, Ledergerber B, Wagels T, Sudre P, and the Swiss HIV Cohort Study (2002): Impact of occasional short interruptions of HAART on the progression of HIV infection, results from a cohort study; AIDS 16 (5):747 –755

Thie´baut R, Pellegrin I, Cheˆneb G, Viallarde JF, Fleury H, Moreauc JF, Pellegrine JC and Blanco P (2005): Immunological markers after long-term treatment interruption in chronically HIV-1 infected patients with CD4 cell count above 400×10^6 cells/l: AIDS, 19:53–61

Touloumi G, Pantazis N, Antoniou A, Stirnadel HA, Walker SA, and Porter K (2006): Highly Active Antiretroviral Therapy Interruption: Predictors and Virologic and immunologic Consequences: J Acquir Immune Defic Syndr; 42: 554 – 561

Touloumi G, Pantazis N, Stirnadel HA,Walker AS, Boufassa F, Vanhems P and Porter K (2008): Rates and Determinants of Virologic and Immunological Response to HAART Resumption After Treatment Interruption in HIV-1 Clinical Practice:

Toulson AR, Richard Harrigan R, Heath K, Yip B, Brumme ZL, Harris M, Hogg RS, and Montaner JSG (2005): Treatment Interruption of Highly Active Antiretroviral Therapy in Patients with Nadir CD4 Cell Counts 1200 Cells/mm^3: The Journal of Infectious Diseases; 192:1787–93

United Nations Programme on HIV/AIDS (UNAIDS) 2006 and 2008: Report on the Global HIV/AIDS Epidemic, Geneva.

Volberding PA, Lagakos SW, Koch MA, Pettinelli C, Myers MW, Booth DK, Balfour H, Reichman RC, Bartlett JA., Hirsch MS (1990). Zidovudine in asymptomatic HIV infection, a controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. N Engl J Med. 322:941-9. http://amedeo.com/lit.php?id=1969115

Weeks BS and Alcamo IC (2006): the biological basis. Sudbury, AIDS; Mass.: Jones and Bartlett Publishers.

Wood E, Hogg RS, Yip B, Moore D, Harrigan PR, and Montaner J S.G (2006): Impact of baseline viral load and adherence on survival of HIV-infected adults with baseline CD4 cell counts ≥200 cells/ml: AIDS; 20:1117–1123

Yared M, Geskus RB, Hendriks JC, Tsehaynesh M, Miedema F, Coutinho RA and Dukers NH (2003): Low CD4 counts before seroconversion did not lead to faster HIV disease progression in Ethiopia. http://gateway.nlm.nih.gov/MeetingAbstracts/ma?f=102271534.html

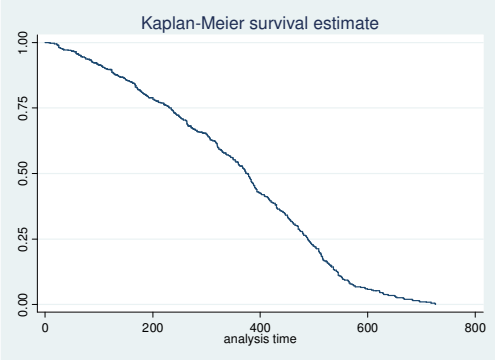
Annex 1: Descriptive measures of the continuous covariates along with the status of patients

Table A1.1: Descriptive Characteristics by Subsequent HAART Status of chronic HIV-1 patients with and without treatment resumption for the data at ZMH for the continuous covariates

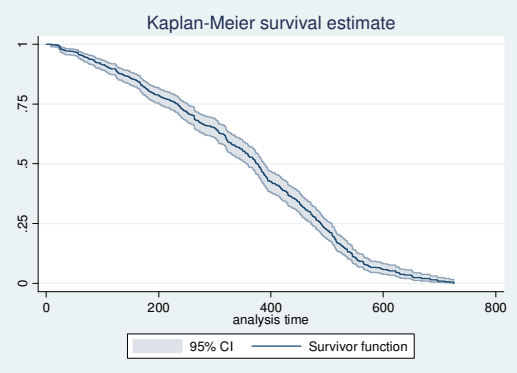
Status	Continuous covariates	Age	Weight	CD4 HAART	CD4 Labvalue	Disease duration	Duration of followup
Patients without treatment resumption	Mean	37.01	52.96	141.17	279.43	392.13	391.64
	St.Dev.	10.49	8.47	63.11	111.53	221.17	281.97
	Minimum	18	35	16	51	114	92
	Maximum	60	71	315	598	996	1552
	1st quartile	28	46	95	205.5	224.5	193.5
	2nd quartile	35	53	135	261.5	313.5	280
	3rd quartile	44	59	187	343.5	514.5	485.5
Patients with treatment resumption	Mean	35.98	51.88	120.44	175.35	408.82	300.04
	St.Dev.	9.90	8.80	64.43	93.93	201.52	177.11
	Minimum	18	34	22	20	98	89
	Maximum	60	72	290	586	1060	1214
	1st quartile	29	45	66	103	245	175
	2nd quartile	35	52	112	161	373	245
	3rd quartile	42	58	166	229	538	394
Total	Mean	36.31	52.23	127.21	209.33	403.38	329.94
	St.Dev.	10.09	8.70	64.69	111.24	208.14	220.10
	Minimum	18	34	16	20	98	89
	Maximum	60	72	315	598	1060	1552
	1st quartile	29	46	76	121	240	180
	2nd quartile	35	52	119	199	353	256
	3rd quartile	42	59	178	272	532	409

Annex 2: The Kaplan-Meier survival function estimates

Figure B2.1: Kaplan-Meier survival estimator of failure time and its 95% CI plot of chronic HIV-1 patients with treatment interruptions at ZMH

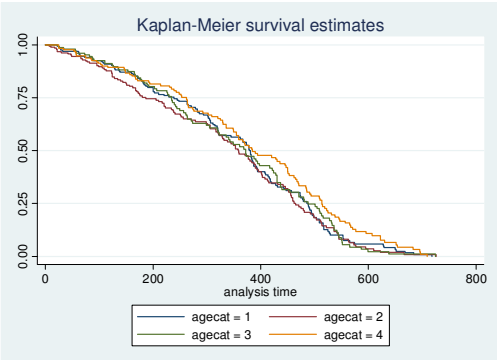


A) The KM survival function estimate for failure time of patients interrupting their HAART

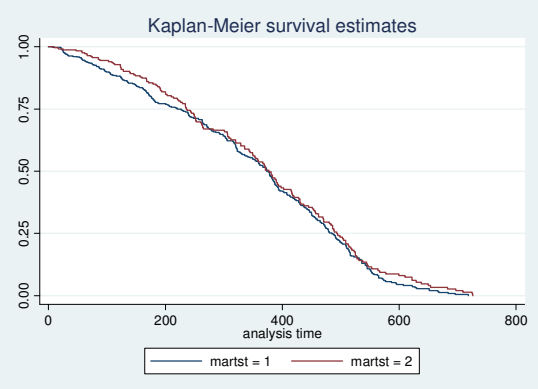


B) The KM survival function estimate and 95% CI for failure time of patients interrupting their HAART

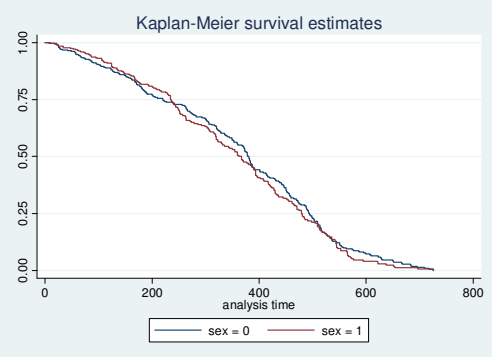
Figure B2.2: Kaplan Meier survivor estimates for categorical variables (i.e. sex, marital status, level of education, religion, age, weight, CD4HAART and CD4Labvalue) for chronic HIV-1 patients who interrupt their actively antiretroviral treatment at ZMH



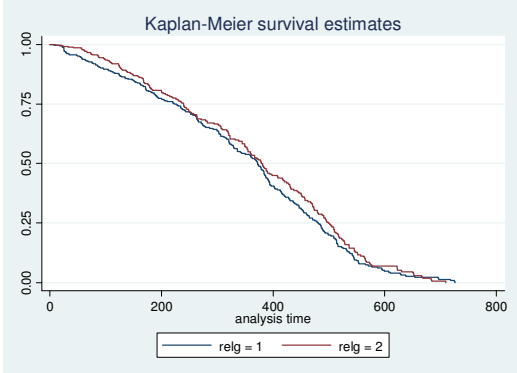
A) The graph of Kaplan-Meier survival estimate to compare the categories of age of the patients



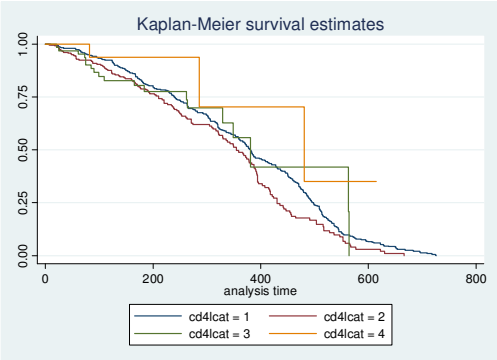
B) The graph of Kaplan-Meier survival estimate of the categories of the Marital Status of the patients



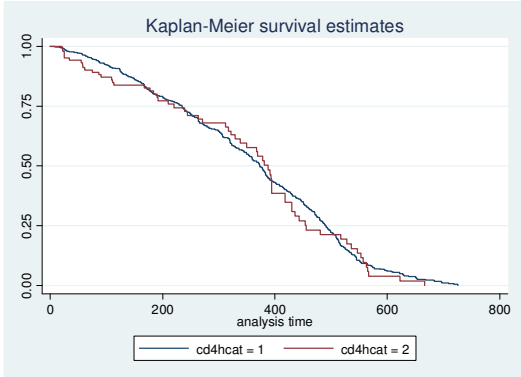
C) The graph of Kaplan-Meier survival estimate of the two sexes of the patients



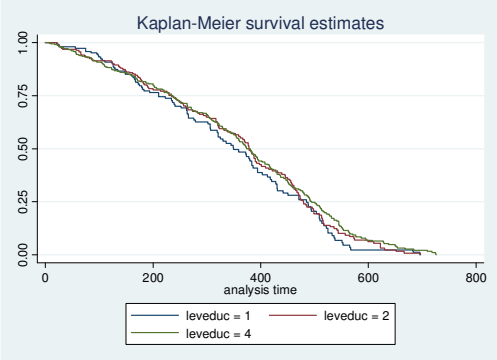
F) The graph of Kaplan-Meier survival estimate of the categories of the religion of the patients



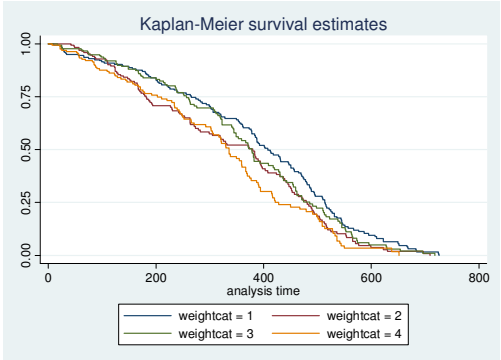
D) The graph of Kaplan-Meier survival estimate of the patients CD4labvalue count group



G) The graph of Kaplan-Meier survival estimate of the patients for different group of CD4haart count of patients



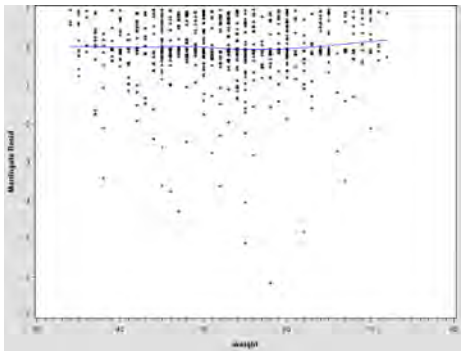
E) The graph of Kaplan-Meier survival estimate of the categories of the level of education of the patients



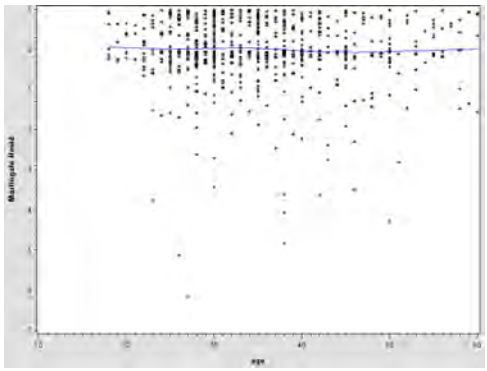
H) The graph of Kaplan-Meier survival estimate to compare the categories of weight of the patients.

Annex 3: Residuals used for model assessment

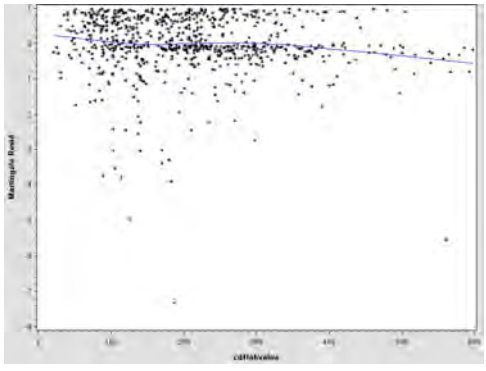
Figure C3.1: Plot of Martingale residuals against continuous variables to check linearity for chronic HIV-1 patients who interrupt their actively antiretroviral treatment at ZMH in Cox proportional hazards model



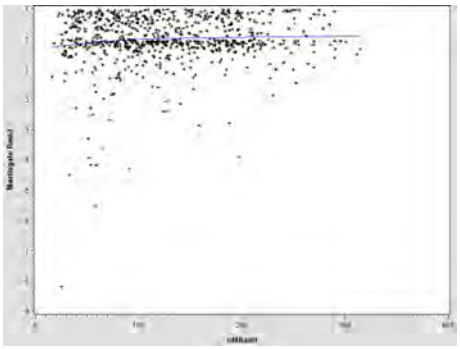
A) The plot of the martingale residuals against the excluded covariate weight



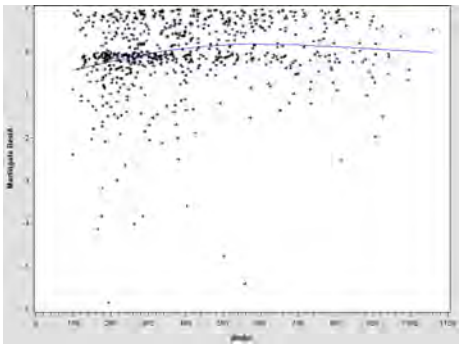
B) The plot of the martingale residuals against the excluded covariate age



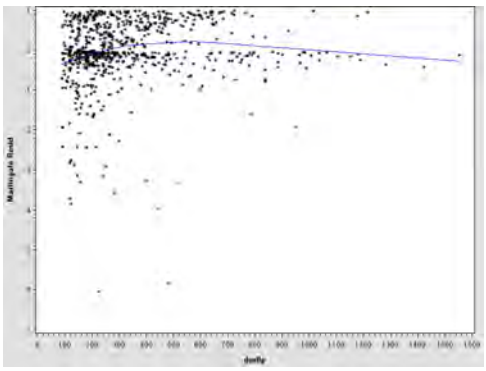
C) The plot of the martingale residuals against the excluded covariate CD4Labvalue



D) The plot of the martingale residuals against the excluded covariate C D4HAART



E) The plot of martingale residuals against the excluded covariate disease duration

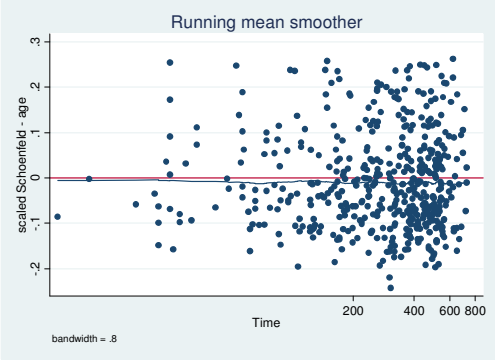


F) The plot of martingale residuals against the excluded covariate duration of follow-up

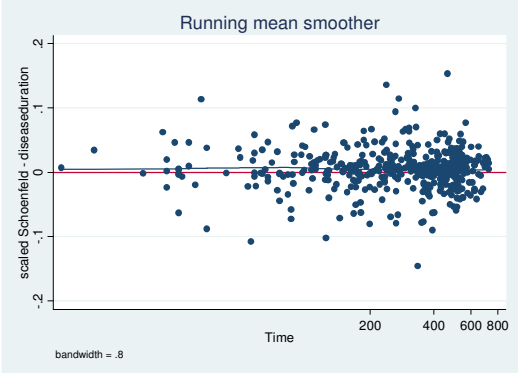
Figure C3.2: Plots of scaled Schoenfeld residuals against transformed time for each covariate in Cox Proportional Hazards Model fit for chronic HIV-1 patients who interrupt their active antiretroviral treatment at ZMH

A) The plot of Scaled Schoenfeld residual for weight to check the validity of the PH assumption

D) The plot of Scaled Schoenfeld residual for CD4HAART to check the validity of the PH assumption



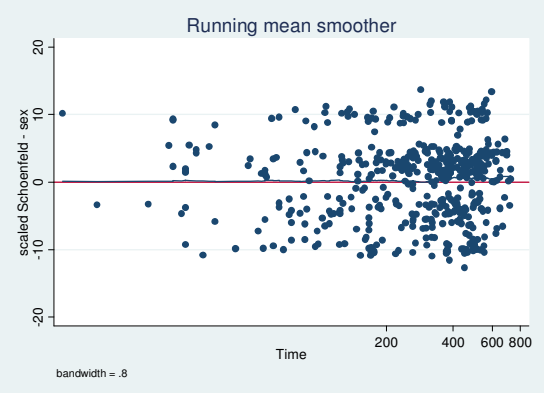
B) The plot of Scaled Schoenfeld residual for Age to check the validity of the PH assumption



E) The plot of Scaled Schoenfeld residual for disease duration to check the validity of the PH assumption

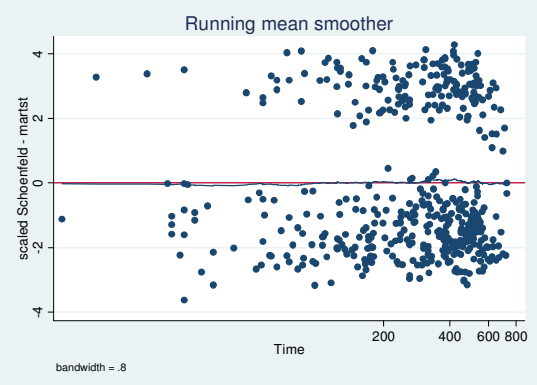
C) The plot of Scaled Schoenfeld residual for CD4labvalue to check the validity of the PH assumption

F) The plot of Scaled Schoenfeld residual for duration of follow-up to check the validity of the PH assumption



G) The plot of Scaled Schoenfeld residual for sex to check the validity of the PH assumption

J) The plot of Scaled Schoenfeld residual for sex and marital status to check the validity of the PH assumption

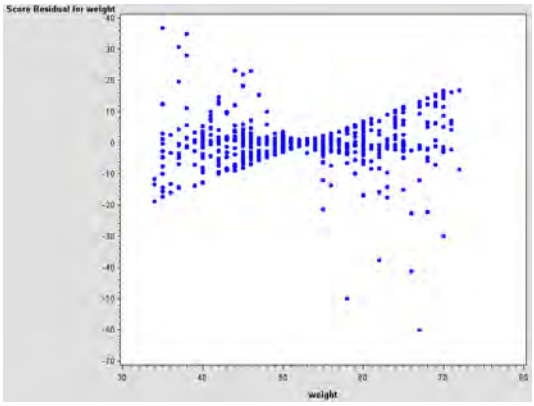


H) The plot of Scaled Schoenfeld residual for marital status to check the validity of the PH assumption

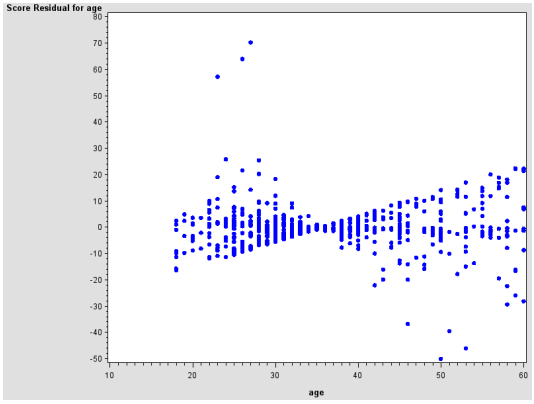
K) The plot of Scaled Schoenfeld residual for disease duration and marital status to check the validity of the PH assumption

I) The plot of Scaled Schoenfeld residual for no education to check the validity of the PH assumption

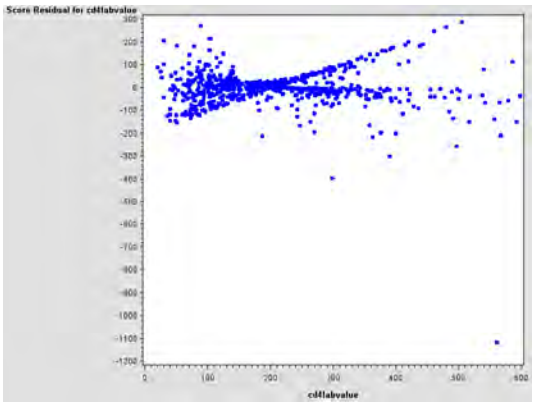
Figure C3.3: Plots of score residuals for each covariate in Cox Proportional Hazards Model fit for chronic HIV-1 patients who interrupt their active antiretroviral treatment at ZMH



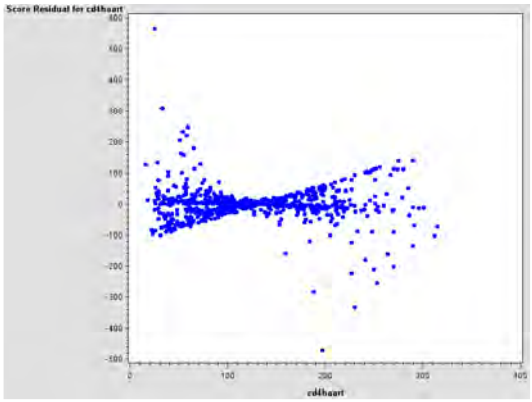
A) The Score residuals for weight to detect the existence of influential observations



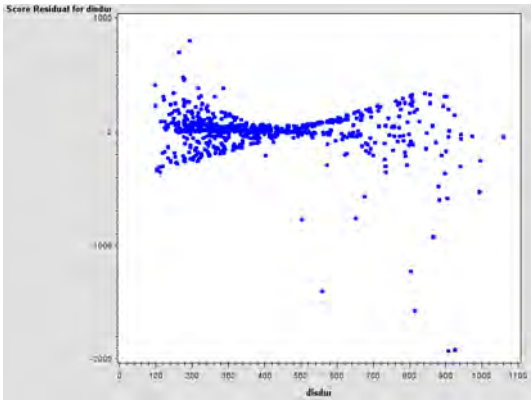
B) The Score residuals for age to detect the existence of influential observations



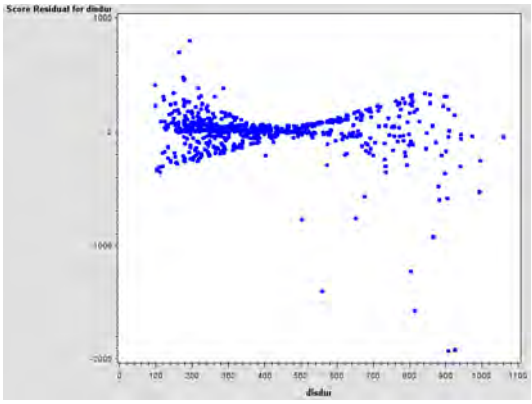
C) The Score residuals for CD4Labvalue to detect the existence of influential observations



D) The score residuals for CD4HAART to check for the existence of influential observations



E) The score residuals for Disease duration to check for the existence of influential observations



F) The score residuals for Duration of follow-up to check for the existence of influential observations

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all source materials used for the thesis have been duly acknowledged.

Name: Berhanu Anagaw Wubie

Signature: ----------

Place: Faculty of Science, Addis Ababa University

Date: ----------

This thesis has been submitted for examination with my approval as a University advisor.

Name: Fentaw Abegaz (Ph.D.)

Signature: ----------

Date: ----------