ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF STATISTICS

SURVIVAL AND RISK FACTORS OF HIV/TB CO-INFECTED PATIENTS UNDER ANTIRETROVIRAL THERAPY IN AMBO HOSPITAL, ETHIOPIA

HAILU REFERA

A THESIS SUBMITTED TO

THE DEPARTMENT OF STATISTICS

PRESENTED TO IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS (APPLIED STATISTICS)

ADDIS ABABA UNIVERSITY
ADDIS ABABA, ETHIOPIA
JUNE, 2012

This is to certify that the thesis prepared by Hailu Refera, entitled: *Survival and risk factors of HIV/TB co-infected patients under antiretroviral therapy in Ambo hospital, Ethiopia* and submitted in partial fulfillment of the requirements for the degree of master of science in Statistics (Applied Statistics) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Examiner _____Signature_____Date_____

Examiner_____Signature_____Date_____

Advisor_____Signature_____Date_____

_____

Chair of Department or Graduate Program Coordinator

Abstract

Survival and risk factors of HIV/TB co-infected patients under antiretroviral therapy in Ambo hospital, Ethiopia.

Hailu Refera
Addis Ababa University, 2012

HIV infection is the greatest risk factor for acquiring TB infection and developing TB. TB enhances HIV replication by accelerating the natural evolution of HIV infection and it is the most leading cause of sickness and death of PLWHIV. The objective of this study is to assess the survival and risk factors of HIV/TB co-infected patients in Ambo hospital, west Shoa zone in Oromia regional state, Ethiopia. This is a retrospective study based on cases of HIV-infected TB patients of age 15 years and above who have started anti-TB treatment between September 1, 2006 and August 31, 2011 and followed until February 29, 2012. To analyze the data descriptive statistics, univariate and multivariate analyses were used. The descriptive analysis indicates that out of the total 501 individuals, 79(15.8%) are death cases, of which 31 (39.2%) and 49 (62%) deaths occurred within six and nine months of TB treatment start. The Kaplan-Meier method was used to estimate the survival time and Cox's regression model was employed to identify the covariates that have a statistical significant effect on the survival of HIV/TB co-infected patients. The estimation of the model parameters was done by partial maximum likelihood procedures. The multivariate analysis of Cox regression model gives that initial weight, TB site, WHO clinical stage, functional status and CD4 count are significant risk factors of survival of HIV/TB co-infected patients. Furthermore the patients with lower initial weight, lower CD4 count, WHO stage III and IV, being ambulatory and bedridden are associated with higher risk of death.

Acknowledgement

Firstly, my deepest gratitude goes to Professor Eshetu Wencheko, my advisor and instructor, for his constructive comments, encouragement, tireless follow-ups that make this work a reality.

Next, I want to thank Ato Ketema Kebebew, who have provided me help and moral support in one way or another to accomplish this study.

Finally, I would like to thank the Ambo hospital administration office for allowing me to obtain data.

ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral therapy
ARV	Antiretroviral
CI	Confidence Interval
DOTS	Directly Observed Treatment, Short-Course
EPT	Extra-pulmonary TB
HAART	Highly Active Antiretroviral Therapy
HIV	Human Immunodeficiency Virus
HR	Hazard Ratio
MOH	Ministry of Health
LR	Likelihood Ratio
MLE	Maximum Likelihood Estimate/Estimator
OIs	Opportunistic Infections
PLWHA	People Living with HIV/AIDS
PT	Pulmonary TB
TB	Tuberculosis
UNAIDS	United Nations Program of HIV/AIDS
WHO	World Health Organization

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER ONE

1. INTRODUCTION

1.1 Background

The human immunodeficiency virus (HIV), causing the acquired immunodeficiency syndrome (AIDS), has created an enormous challenge to mankind since it became known or described in 1981. HIV attacks a type of immune system cell called the T-helper cell. This cell carries on its surface a protein called CD4, through which HIV uses to attach to the cell before gaining entry. The T-helper cell plays an important role in the immune system by helping to co-ordinate all the other cells to fight illnesses. HIV causes many T-helper cells to be damaged or destroyed. As a result, a major reduction in the number of T-helper cells can have a serious effect on the immune system to fight illnesses. The CD4 test measures the number of T-helper cells in the blood. The more these cells we have per cubic millimeter in the blood the stronger our immune system to fight the illnesses. A low CD4 count does not mean that one will certainly become ill, but it makes it more likely. As the immune system becomes increasingly damaged by HIV it becomes susceptible to opportunistic infections. Tuberculosis is the most common opportunistic infection complicating HIV infection especially in developing countries, and may occur at any stage in the course of immunodeficiency. TB/HIV co-infection significantly changes the natural history of both diseases and it is recognized as a major setback of both HIV and tuberculosis infection control program (Godfrey-Faussett *et al*. 2002).

1.2 Impact of HIV on Tuberculosis

HIV infection is the greatest risk factor for acquiring *Mycobacterium tuberculosis* infection and developing TB. Thus, risk of TB is dramatically increased in HIV infected patients as a result of a higher probability of either primary progression or reactivation of latent infection, as well as increasing chance of tuberculosis infection once exposed to tubercle bacilli (Jon, 2004). People with HIV infection are increasingly infected with TB because HIV weakens their immune system. HIV infection impairs cell-mediated immunity, largely through depletion of $CD4^+$ lymphocytes. The impaired immunity leads to increased number of cases of primary TB and reactivation of TB in HIV-infected people (Hoffmann, 2007). In addition, the depletion of CD4+ cells result in interruption of TB treatment and increases the risk for the development of drug resistance strains of TB which is difficult and expensive to be treated in HIV-infected patients(Davies *et al.*,1999).

Moreover, HIV-infected people who have recovered from TB have an accelerated course of HIV disease and shortened survival compared with HIV-infected people without a history of TB (Whalen C *et al.*, 1995 and Whalen CC *et al*., 2000)**.** Patients with HIV and latent tuberculosis infection have an extremely high rate of development of active tuberculosis, perhaps as much as 10% per year, as opposed to a 10% lifetime risk in otherwise healthy persons with latent tuberculosis infection (Schluger *et al.*, 1998). According to WHO report of 2009, people living with HIV (PLWH) are estimated to have a 20 times higher risk on developing tuberculosis (TB) disease compared to people living without human immunodeficiency virus (HIV) infection in countries with an HIV

prevalence of at least 1% (WHO, 2009). In addition to increasing individual susceptibility to TB following *M. tuberculosis* infection, a high burden of HIV-associated TB cases also expands *M. tuberculosis* transmission rates at the community level, threatening the health and survival of HIV-negative individuals as well (Odhiambo *et al.*,1999; Sonnenberg *et al.*, 2004). Thus, in communities with high HIV rates, TB rates have been rising dramatically even where effective TB control strategies like Directly Observed Treatment short-course (DOTS) is available. This underscored the need to improve TB control wherever HIV-co-infection is common. The other effect is discrimination of tuberculosis patients among the social, because of HIV. The link between tuberculosis and HIV/AIDS may make people equating tuberculosis with HIV/AIDS. This may lead to increasing stigma and discrimination and delay tuberculosis patients in seeking care and treatment.

1.3 Impact of Tuberculosis on HIV

Tuberculosis enhances HIV replication which leads to increased viral load by accelerating the natural evolution of HIV infection (Rosas-Taraco *et al.*, 2006). This results in more rapid progression of HIV disease. M. tuberculosis is a potent activator of the immune system, which in turn leads to profound changes in the cytokine network, providing favorable environment for viral replication (Toossi *et al.*, 1999). On the other hand, it increases the risk of progression from HIV to AIDS (Badri, 2001 and Sharma *et al.*, 2005). TB increases occurrence of other OIs. In addition, though it is preventable, treatable and curable, TB is the leading cause of sickness and death of PLWHA. The management of TB and HIV co-infected is challenging because of pill burden, increase adverse effect, drug to drug interaction and etc.

1.4 Complexity and difficulties among HIV and Tuberculosis co-infection

TB/HIV co-infection significantly changes the natural history of both diseases. This gives rise to different problems. One problem associated with patients co-infected with HIV and TB is overlapping of signs and symptoms between HIV/AIDS and tuberculosis. Clinical features of HIV/AIDS and tuberculosis are difficult to separate, both diseases present with wasting and persistent fever. In cases of tuberculosis patients co-infected with HIV/AIDS the physician tends to attribute the signs and symptoms of tuberculosis to HIV/AIDS, hence under-diagnosing tuberculosis in HIV patients and vice versa. The none-specific signs and symptoms of HIV/AIDS and its co-infection with tuberculosis make the clinical diagnosis difficult in most cases. The fact that HIV/AIDS also makes the patient susceptible to other opportunistic infections with symptoms similar to tuberculosis are among the difficulties encountered in diagnosing tuberculosis in HIV/AIDS patients.

The other impact is that two different diseases, two different modalities of diagnosis and treatment, do exist in one patient. Both diseases involve the combination of more than one drug. Treatment is for life in HIV/AIDS and for a minimum of six month in tuberculosis patients, resulting in high pill burden, many side effects and interaction of drugs. The consequence may be poor adherence to treatment and loss to follow up.

1.5 Global and Sub-Saharan African burden of HIV/TB co-infection

The global impact of the converging dual epidemics of TB and HIV remains as one of the major public health problem or challenges of our time. According to the UNAIDS annual report of 2009, one third of people living with HIV are co-infected with TB. Globally

about 11% of new adult cases of tuberculosis are HIV/AIDS co-infected where in sub-Saharan Africa about 31% of new tuberculosis cases are HIV/AIDS co-infected (Corbett *et al*., 2002&2003). In particular, WHO reported 9.27 million of the estimated new cases of TB occurred in 2007 of which 31% are in Sub-Saharan Africa and about 1.37 million (14.8%) of these being among people living with human immunodeficiency virus (PLWHIV) (WHO, 2009). In general, the majority of these cases, that is 79% HIV positive TB incident cases, occurred in Africa in 2007. In the same year, there were 456,000 TB-related deaths among HIV-positive patients accounting for 23% of the global HIV/AIDS mortality. Southeast Asia is the second most affected region with 11% of global new TB cases in 2007(WHO, 2009 and FMOHE, 2008).

Moreover, TB is the commonest infection and number one cause of death in HIV infected patients in developing countries. Other studies from sub-Saharan Africa showed that about 31% of adult TB cases were attributable to HIV infection (Corbett *et al*., 2003). Thus, worldwide TB is one of the most common opportunistic infections and leading cause of morbidity and mortality among HIV/AIDS patients accounting for about 30% of all death of HIV/AIDS patients (Raviglione, 1992, WHO, 2004 and Sharma, 2005). In another WHO report, TB killed about 0.5 million HIV infected people globally (WHO, 2009). Deaths occurring in the first few months after TB diagnosis are more likely TB related, whereas deaths occurring later are more likely to be attributable to other HIV-related illnesses (Ackah *et al*.,1995; Nunn *et al*.,1992; Elliott et al.,1995). Different studies and reports indicate that the majority of HIV/TB co-infection and deaths in the world occurs in sub-Saharan Africa countries. The proportion of HIV-infected TB patients who die during TB treatment is high especially in sub-Saharan Africa countries

(estimated to be between 6% and 39%) (Mukadi *et al.*, 2001 and WHO, 2005). The latest WHO report on global TB control indicates that there were an estimated 8.8 million incident cases and 1.4 million deaths from TB in 2010, where 1.2 million (12–14%) of the new TB cases were amongst HIV-infected people in 2010, and TB caused the death of an estimated 0.35 million HIV-infected people. The proportion of HIV-TB-co-infected persons is highest in Africa, with the African Region accounting for overall 82% of TB cases among people living with HIV worldwide in 2010 (WHO, 2011).

1.6 National burden of HIV/TB co-infection

HIV infection is now the most common predictor of TB incidence, and the other way round, TB is a common opportunistic infection in sub-Saharan Africa. Thus, these countries continue taking the leading position in HIV/TB morbidity and mortality rate, where the TB epidemic is primarily driven by HIV infection. Ethiopia is one among these countries most heavily affected by HIV and TB co-infection. The world health organization ranked Ethiopia as 7^{th} among the 22 high burden countries with TB in estimated annual incidence of 379 cases and prevalence of 643 cases per 100,000 populations (WHO, 2008) and the prevalence of HIV among TB patients is up to 41% (Demissie *et al.*, 2000a; Yassin *et al.*, 2004). The high burden of TB in Ethiopia might in part be attributed to the rapid increase of HIV infection, because available data indicate that HIV/ AIDS accounted for an estimated 32% or 141,000 total TB cases in 2005 (MOH, 2007). On the other hand, the WHO global report 2008 estimates that in Ethiopia 40% of TB patients tested for HIV are HIV-positive. In Ethiopia routine data from 44 sites in the year 2005/6 showed 41% of TB patients were HIV positive. In addition, another routine data collected in 2006/07 estimates that 31% of TB patients are HIV

positive (FMOH, 2008). TB was the cause of 76 thousands deaths in Ethiopia, out of which 30% were among HIV positive patients (WHO, 2009).

However, WHO recommends different collaborative activities for HIV/TB co-infections where one is initiation of antiretroviral therapy. Antiretroviral therapy (ART) is an essential treatment for HIV infection in order to reduce the risks of death and HIV-related morbidities, or in improvement of quality of PLWHIV. In addition, several observational studies, mostly done in resource-rich settings, have documented that HIV-infected TB patients receiving ART have higher rates of survival during and after TB treatment compared with patients who do not receive ART. Currently, the widespread use of highly active antiretroviral therapy(HAART) in developing countries has been shown to reduce the risk of treated HIV-infected persons developing TB by 70%–90%, compared with untreated individuals (Badri *et al.*, 2002; Girardi *et al.*, 2000; Jones *et al.*, 2000; Santoro-Lopes *et al.*, 2002). Although, the availability of ART has transformed HIV infection and has made a manageable disease and reduce TB incidence among individuals receiving HAART, TB risk remains elevated among those receiving treatment. Thus, the success recorded can easily be destroyed or violated by the high burden of TB co-infection in the HIV infected individual. Even after the initiation of ART, the incidence of TB among HIV infected patients remains unacceptably high (Badri *et al.*, 2002 and Fauci *et al.*, 2005). Studies conducted in some developing countries have shown that among individuals who initiated HAART, 4.8−17.6 cases per 100 person-years reported having TB, either because they are developing active TB or sub clinical TB may become apparent as the immune reconstitution inflammatory syndrome (Bonnet *et al.*, 2006).

1.7 Statement of the problem

Generally it is recommended that HIV infected patients start to take ART in order to reduce AIDS related mortality and morbidity, or to improve their quality of life. But, in most cases TB co-infection violates and disturbs this issue. Moreover, death during TB treatment and shortened life are recognized in HIV/TB co-infected patients. Hence, this study deals with survival/death and risk factors of HIV/TB co-infected patients especially during TB treatment.

Survival analysis is one of the appropriate methods to demonstrate life time and to identify risk factors. The length of survival of HIV/AIDS co-infected patients depends on time elapsed from the date of TB infection is confirmed until death or some observations with incomplete records (censored). Survival analysis is an appropriate analysis method for this study to assess survival/death and its risk factors. Hence, the research problems include: assessing and measuring survival/death status; identifying the risk factors of survival/death of HIV/TB patients; looking at implications of observing significant risk factors; fitting appropriate survival model by including the significant risk factors and evaluating the impacts of influential or risk factors on survival/death.

1.8 Objectives

The general objective of the study is to estimate the survival status of HIV/AIDS co-infected patients and to identify factors that influence their survival.

The specific objectives are

- ✓ Assessing the survival of patients

- ✓ To study the influence of predictors on survival
- ✓ To fit an appropriate model

1.9 Significance of the study

The outcome of this study would provide information about the risk factors or the most influential covariate that have significant impact on survival of HIV patients during TB treatment and to identify death risk extent of patients under these significant factors at different time during and after TB treatment.

1.10 Limitation of the study

- ✓ The study is conducted based on secondary data which might have incomplete and biased information.
- ✓ The study was restricted to adults, and results might not be applicable to infants and children.
- ✓ The study presumed that all deaths are caused by HIV/TB co-infection.
- ✓ Parts of information on individuals are missed because of censored observations.
- ✓ The data on TB site do not identify those patients infected with both pulmonary TB and Extra-pulmonary TB
- ✓ The study is based on baseline values of the variables of interest.

CHAPTER TWO

2. REVIEW OF LITERATURE

The literature selected and discussed in this section are those that are more related and relevant to this study. Studies or research related to the survival of HIV/TB co-infection especially in Ethiopia are scarce. The literature review given below has three parts. The first part is a general overview of TB in HIV and TB co-infection. The second part concerns the effect of ART on HIV/TB co-infection and the third is about risk factors of survival/death of HIV/TB co-infected patients.

2.1 General literature on effect of TB in HIV/TB co-infection

Abera *et al* (2009) conducted an ecological study on the association between HIV and TB in Oromia regional state, Ethiopia in 2006/7. The study covered a total of 40779 cases of TB including 12818 smear positive pulmonary TB cases and 29,590 positive for HIV infection. The objective of the study was to assess the association between infection with HIV and tuberculosis. The result of the study suggests that the prevalence of HIV infection was significantly associated with the incidence of TB in Oromia region ($r=0.69$, $p<0.01$). The ecological association between different types of tuberculosis and prevalence of HIV across zones and towns in Oromia was estimated using the Spearman's correlation. The study has also shown that similar associations were also seen between prevalence of HIV infection and the incidence of smear positive tuberculosis, smear negative tuberculosis and extra-pulmonary tuberculosis.

Deribew *et al* (2009) conducted a study from February to April, 2009 in selected hospitals (Adama, Nekemte and Jimma) in Oromia regional state, Ethiopia. The study consisted of 467 HIV patients and 124 TB/HIV co-infected patients. The objective of the study was to assess the quality of life of HIV/TB co-infected patients. Using a logistic regression model the study showed that moral depression or sadness, whether or not having source of income and family support were strongly associated with the quality of life of the patients. The results obtained suggest that TB/HIV co-infected patients had a lower quality of life in all domains as compared to HIV infected patients without active TB. In co-infected patients, individuals who had depression were 8.8 times more likely to have poor physical health as compared to individuals who had no depression, with odds ratio[(OR) = 8.8(95%CI: 3.2, 23)]. Self-stigma or being discriminated was associated with a poor quality of life in the psychological domain.

Another cross sectional study was done by Deribew *et al.* (2010) in three Oromia regional state hospitals (Adama, Nekemte and Jimma), Ethiopia from February to April, 2009. The study consisted of 155 TB/HIV co-infected and 465 non-co-infected HIV patients. The aim of the study was to investigate the relationship between TB/HIV co-infection and common mental disorders (CMD). The study compared the occurrence of CMD in TB/HIV co-infected and non-co-infected HIV patients. The results of the study obtained by using logistic regression showed that TB/HIV co-infected patients had significantly (p value- 0.001) greater risk of CMD (63.7%) than the non-co-infected patients (46.7%) [OR = 1.7, (95%CI: 1.0, 2.9)].

Straetemans *et al* (2010) conducted a meta-analysis of cohort studies by selecting relevant articles. Their purpose was to assess the effect of TB on mortality in people living with HIV. Pooled overall analysis of fifteen studies estimating the effect of tuberculosis on mortality in PLWHIV has shown a Hazard ratio (HR) of 1.8 (95% CI: 1.4-2.3). However, out of fifteen pooled studies, sub-analysis of eight studies in which the cohort was not exposed to HAART showed a hazards ratio (HR) of 2.6 (95% CI: 1.8–3.6) that indicated impact of TB on HIV in co-infection.

In Thailand, a team of physicians conducted a prospective observational study and described the cause of death of patients LWHIV who have tuberculosis, after reviewing verbal autopsies or interviews of family members about events preceding death, laboratory data, and medical records. The result of the study indicated that out of 849 HIV-infected TB enrolled, 142 (17%) died, of which the cause of death was TB for 38 patients (27%) (Kevin *et al*., 2009).

Kamenju and Aboud (2011) conducted a study employing logistic regression analysis to assess the clinical presentation and prevalence of TB-HIV co-infection among patients admitted at Muhimbili National Hospital between August 2008 and July 2009 in Dar es Salaam, Tanzania. Of the 300 TB patients tested for HIV, 175 (58.3%) were HIV-infected and 97 (55.4%) of these were already on antiretroviral therapy (ART) at the time of admission. Overall, 104 (26.9%) of the TB patients admitted died. About two thirds of patients who died had PTB. There was significantly higher proportions of deaths among

HIV-infected TB patients (29.1% *versus* 15.2%) than in the HIV uninfected TB patients (*P*=0.005).

Badri *et al* (2001) conducted a prospective patient cohort study on adult (≥18 years of age) HIV-infected patients attending two HIV clinics affiliated to the University of Cape Town (New Somerset and Groote Schuur Hospitals), South Africa. The study cohort consisted of 609 patients with 158 (25.9%) cases of TB patients whose initial clinic visit was between 1992 and 1996. The objective of the study was to assess the impact of tuberculosis on HIV-1disease progression to AIDS and then to death in areas with high tuberculosis prevalence. Cox proportional hazards regression models were fitted to determine the unadjusted risk of death associated with tuberculosis and to adjust for potential confounding variables. The results obtained indicated that the more rapid progression to AIDS in tuberculosis cases remained significant after controlling for potentially confounding variables. Other factors associated with progression of AIDS were CD4+ T-lymphocyte count ≤200 cells/mm^3. Of the 609 patients, 103 (17%) died over the 5-year period: $\frac{50}{158}$ (31.7%) TB cases and $\frac{53}{451}$(11.7%) in the comparison group (non-TB cases). The Kaplan-Meier survival probability of tuberculosis cases was consistently less than that of the comparison group (69% vs. 92% at 1 year, $P < 0.0001$; 49% vs. 82% at 2 years, $P < 0.0001$; and 28% vs. 64% at 3 years, $P< 0.0001$). The multivariate Cox proportional hazards regression revealed that even after controlling for differences in baseline characteristics and known predictors of mortality in HIV infection, tuberculosis conferred an increased independent risk of death in HIV-infected patients. The association of tuberculosis with increased mortality persisted after removing the

transient reduction in the CD4+ T-lymphocyte count that may have been induced by tuberculosis (HR =1.90, 95%CI 1.16–3.1; $P< 0.01$).

2.2 Literature on effect of ART on HIV/TB co-infection

The convergence of HIV and TB pandemics continue to be collectively the most leading cause of morbidity and mortality worldwide. Highly active antiretroviral therapy (HAART) has been a critical component in combating the scourge of these two conditions as both a preemptive and therapeutic modality.

Weerawat *et al* (2006) conducted a retrospective cohort study among HIV infected patients with TB between January 2000 and December 2004 in Thailand. The objective of the study was to establish the impact of antiretroviral therapy (ART) on survival of patients co-infected with HIV and TB. Patients were categorized into ART^+ group (received ART) and ART^- group (did not receive ART) and were followed until April 2005. A total of 1003 patients were identified; 411 in ART+ group and 592 in ART^- group. The results indicated that based on the log-rank test the survival rates at 1, 2, and 3 years after TB diagnosis were 96.1%, 94.0%, and 87.7% for ART+ group and 44.4%, 19.2%, and 9.3% for ART^- group ($P< 0.001$). From the study it was concluded that antiretroviral therapy substantially reduces mortality rate among HIV/TB-co-infected patients.

In another study conducted in Thailand, the impact of ART on survival of HIV-infected TB patients was measured using propensity score analysis that adjusted for factors

associated with receiving ART. The aim of the study was to document the impact of ART on HIV-infected TB patients in public health programs in resource-limited settings. The study consisted of 626 TB patients who started and 643 who did not start ART. The result obtained showed that out of 626 HIV-infected TB patients who started ART during TB treatment, 68 (11%) died compared with 295 (46%) patients who did not receive ART (relative risk 0.24, 95% CI: 0.19 to 0.30); in patients with very low CD4 count (<10), 12 patients out of 56 (21%) who received ART died compared with 35 out of 43 (81%) who did not receive ART (relative risk 0.26, 95% CI: 0.16 to 0.44). After controlling for propensity to receive ART, the hazard ratio of death among patients treated with ART was 0.17 (95% CI: 0.12 to 0.24) (Natpatou *et al*., 2008).

Jaime *et al* (2010) also conducted a retrospective cohort study of HIV patients who developed TB from January 1995 to December 2000 in a Peruvian HIV reference centre, Peru. The patients were tracked for 24 months after TB diagnosis. One hundred patients were included in the study. The aim of the study was to estimate the survival of HIV/TB co-infected patients and the effect of ART on mortality. Survival was estimated using Kaplan-Meier method and the effect of ART on survival was evaluated using Cox proportional hazards models. Out of the study population 53 patients who received some form of ART: 31 mono/dual ART and 22 highly active antiretroviral therapy (HAART). The result of the study showed that 77% (36/47) of patients died in the no ART group, 65% (20/31) in the mono/dual ART group, and 27% (6/22) in the HAART group (p=0.0001). Multivariable analysis revealed a reduction in the risk of death of 70% [HR: 0.3, 95% CI 0.1-0.6] among mono/dual ART group patients, and 90% [HR: 0.1, 95% CI

0.06-0.4] among patients who received HAART as compared to those in the no ART control group. The study suggested that patients with TB and HIV who did not receive ART have shorter survival; survival significantly improved with the addition of ART. The risk of death is reduced due to mono/dual ART and HAART.

2.3 Literature on risk factors of survival/death of HIV/TB co-infected patients

Dembele, *et al* (2010) conducted a retrospective cohort study on four human immunodeficiency virus (HIV) treatment centers in Ouagadougou, Burkina Faso. TB incidence was measured at different intervals after HAART initiation. Cox regression models were used to identify factors associated with TB incidence. They analyzed a cohort of 2383 subjects and identified WHO clinical Stage III or IV, CD4cell count < 50 cells/mm^3 and weight < 18.5 were associated with increased risk of TB on univariable analysis. In the Cox regression, weight < 18.5 and CD4+ T-cell count < 50 cells/mm^3 at HAART initiation were independently associated with a two-fold higher risk of TB.

Mohammed *et al.* (2011) conducted a case control study in Jimma and Mettu Karl hospitals where the two hospitals serve as referral and treatment centers for HIV and TB in south-west Ethiopia from January to March, 2009. The study population consisted of 162 cases and 647 controls. Cases were adult people living with HIV/AIDS who developed active pulmonary tuberculosis and controls were people living with HIV/AIDS without active tuberculosis. The objective of the study is to identify the risk factors of active pulmonary TB among PLWHIV using multiple logistic regression model. The final multivariate model was obtained by a forward and backward variables selection procedure. Then, the result reveals that, after adjustment for potential confounders, an

initial weight less than 18.5 kg (OR=4.1; 95% CI: 2.3, 7.4), a CD4 lymphocyte count less than 200 cells/mm^3 (OR=9.8‘95% CI: 5.5, 17.5), a WHO clinical stage IV (OR=4.3; 95% CI: 2.6, 6.8) and not taking antiretroviral treatment (OR=3.1; 95%CI: 1.9,4.9), were independently associated with the development of active tuberculosis in people living with HIV/AIDS.

Tarekegn (2011) conducted retrospective study in which a total of 632 patients (316 in ART and pre- ART cohort) were followed for a median of 32.9 months in Pre-HAART and 35.4 (IQR=23.6-36.5) months in HAART. The objective of the study was to identify factors that increase the risk of TB in PLWHIV. He used Cox proportional hazard analysis. The result of the study indicated that WHO stage III or IV (HR=1.999, 95%CI=1.025-3.896, P=0.042), being bedridden (HR=4.689, 95%CI=1.715-12.819, P=0.003), and having hemoglobin level less than 10mg/dl (HR=2.497, 95%CI=1.098-5.679, P=0.036) were factors associated with increased risk of TB in PLWHIV.

In another study conducted in Nigeria, the case files of HIV/AIDS patients from January to December 2006 attending Aminu Kano Teaching Hospital, were reviewed. A total of 1320 HIV/AIDS patients had complete records and were reviewed, among which 138 (10.5%) were co-infected with TB; Pulmonary TB was diagnosed in 103 (74.6%) patients. Fifty (36.2%) co-infected patients had some type of extra-pulmonary TB (EPTB); 15 had both pulmonary TB and EPTB. The objective of the study was to identify the predictors of Tuberculosis co-infection among HIV positive Patients. Chi-square test was used for testing the significance of association between categorical variables. All

variables that were significantly associated with TB/HIV co-infection were included in a multiple logistic regression analysis in order to determine their independent effects. The result of the study reveals that marital status (OR, 2.1; 95% CI, 1.28–3.59; P = 0.04), WHO clinical stage at presentation (4.81; 1.42–8.34; P = 0.001), and baseline CD4 count (2.71; 1.51–6.21; P = 0.02) remained significant predictors after adjustment for confounding (Zubairu and Musa, 2009).

Enrico *et al*. (2001) conducted a study to assess changing clinical presentation and survival in HIV-associated Tuberculosis after HAART. They reviewed clinical charts of HIV-infected patients with pulmonary TB in two referral clinical centers in Rome, Italy. The 67 patients diagnosed in 1995 to 1996 were compared with 51 patients diagnosed in 1997 to 1998. To identify factors associated with survival, they used a Cox model including ART as a time-dependent covariate. The result of their study indicates that, survival at 1 year was 80% for patients diagnosed in 1997 to 1998 versus 65% for those diagnosed in 1995 to 1996 (*p* by log-rank = .02). Moreover, the study reveals that Age, CD4+ cell count <25/mm3 and AIDS defining illnesses before TB diagnosis were all associated with a higher risk of death.

Bhagyabati *et al* (2005) conducted cross sectional study in Manipur, India to assess the vulnerability of the HIV infected persons to tuberculosis. From the study, site of TB and low CD4 count (less than 100/cmm) reported as risk factors.

Mashimbye (2010) dealt with TB treatment outcomes in adult TB patients attending the Rixile HIV clinic in Tintswalo hospital in Bushbuckridge, South Africa. The objective of

the study was to identify factors associated with TB mortality in adult TB patients receiving TB treatment in Rixile HIV clinic from March 2003 to March 2008. Chi-square test was carried out to assess the association between TB mortality and its risk factors. Univariate and multiple logistic regression models were constructed to assess the predictors of TB mortality. The univariate analysis showed associations of TB mortality with age, sex, BMI, education and ARV treatment. However, only age, sex, and ARV treatment were found to predict mortality related to TB in multiple logistic analysis.

Catala *et al*. (2011) conducted a retrospective cohort study that included all HIV-infected TB patients reported in Barcelona between 1996 and 2006. Survival analysis was done based on the Kaplan-Meier estimator and Cox proportional hazards model. The objectives of the study were to estimate survival and to identify predictive factors and causes of death in a cohort of HIV infected TB patients in the era of HAART. In the study causes of death were classified using the International Classification of Diseases (ICD)-9 and ICD-10, and classified as AIDS related or non-AIDS-related (that is, death because of other burden than AIDS). The results have shown that out of the 792 patients included, 341 (43.1%) died during the study period. Survival was worse among patients aged >30 years (HR 1.6, 95%CI 1.1-2.1), inner-city residents (HR 1.3, 95%CI 1.1-1.7), injecting drug users (HR 1.4, 95%CI 1.1-1.8), those with a non-cavitary radiological pattern (HR 1.5, 95%CI 1.0-2.2), those with <200 CD4/μl (HR 1.8, 95%CI 1.2-2.7) and those diagnosed with AIDS prior to their TB episode (HR 1.85, 95%CI 1.4-2.2). From the study it was reported that 64.8% of the deaths were non-AIDS-related, and that poor

survival was observed despite the availability of HAART; non-AIDS-related mortality was also noted as high.

A retrospective study conducted in Thailand using the Cox proportional hazard model showed that ART was associated with lower mortality rate; gastrointestinal TB and multidrug resistant TB were associated with higher mortality rate ($P< 0.05$). Among patients in ART+ group, patients who delayed starting ART for a minimum of 6 months after TB diagnosis had a higher mortality rate than those who began ART within 6 months after TB diagnosis ($P= 0.018$, HR = 2.651, 95% CI = 1.152-6.102) (Weerawat *et al.,* 2006).

Maruza et al (2012) conducted a retrospective cohort study between June 2007 and December 2009 with HIV-infected patients who had started anti-tuberculosis treatment in the state of Pernambuco, Brazil. Survival data were analyzed using the Kaplan-Meier estimator, the log-rank test and Cox model. The aim of the study was to estimate the probability of survival and to identify risk factors for death in cohort of PLWHIV who started TB treatment. Out of a cohort of 2310 HIV-positive individuals, the data on 333 patients who commenced treatment for TB were analyzed. The results obtained show that the mortality rate was 5.25 per 10,000 person-years (95% CI: 4.15-6.63), that out of 10,000 patients at risk about 5.25 expected to have died. The probability to survive 30 months was 74%. Risk factors for death in the study population were females, age above 30 years, having anemia, not using HAART during treatment for TB and a CD4+ lymphocyte count below 200 cells/mm^3

CHAPTER THREE

3. DATA AND METHODOLOGY

3.1 Data

The data used in this study are obtained from Ambo hospital in Ambo town in Oromia regional state of Ethiopia. This hospital is used as referral centre for different health centers in woredas in western Shoa zone.

The hospital has a separate ART clinic and units for ART follow up for PLWHIV where the units in the clinic have nurses, a pharmacy, data clerks, rooms, equipments and etc. The hospital started giving free ART service in 2006. Data for this study were extracted from the available standard national medical registers, which have been adopted by the Ministry of Health (MOH). The registers include the Pre ART register (register of patients at their first visit), the ART register (registration after ART initiation), and the follow-up patient form. The hospital also has a separate clinic and units for TB treatment and follow-ups. In this case also data regarding TB treatment and follow-ups were recorded on standard register given by MOH. For this study, both ART and TB registers have been looked through in order to record all persons who have been infected with HIV and started ART and have been under an anti-TB treatment. The resulting data set comprises all cases of HIV-infected TB patients older than 15 years who have started anti-TB treatment between September 1st 2006 and August 31st 2011 and followed for the outcomes of either the event (death) or censored (dropped out, lost to follow up, transferred out to other centers or on follow up at the end of study time). The end of the

follow-up time is February 29th 2012. The total number of cases included in the study is 501.

3.2 Variables of the study

3.2.1 The Response Variable

The response or outcome variable in this study is the survival time measured (in months) from the date of the anti-TB treatment's start until the date of the patient's death or censor.

3.2.2 Predictor Variables

The predictor variables in survival data analysis are called covariates. These are explanatory variables which are assumed to influence the survival of HIV/TB co-infected patients and are given below.

- Age in years
- Sex (Male, Female)
- Marital status (Never Married, Married, Others)
- Baseline Weight (kg)
- Level of education (No education, Primary, Secondary or above)
- Substance (Smoking, Alcohol)(no, yes)
- Baseline CD4 cell count (Cells/mm^3)
- Sites(Type) of TB (Pulmonary TB, Extra-pulmonary TB)
- WHO clinical stage (Stage I/II, Stage III or Stage IV)
- Functional Status (Working, Ambulatory or Bedridden)
- Regimen type (D4T-based, AZT-based)

Note: All predictor variables were taken at baseline values.

3.3 Methodology

3.3.1 Survival Analysis

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. By time, we mean years, months, weeks, or days from the beginning of follow-up of an individual until an event occurs; alternatively, time can refer to the age of an individual when an event occurs. By event, we mean death, disease incidence, relapse from remission, recovery (e.g., return to work) or any designated experience of interest that may happen to an individual. The problem of analyzing time-to-event data arises in several applied fields such as medicine, biology, public health, epidemiology, engineering, economics, sociology, demography and etc. The terms lifetime analysis, duration analysis, event-history analysis, failure-time analysis, reliability analysis, and transition analysis refer essentially to the same group of techniques although the emphases in certain modeling aspects could differ across disciplines (Aalen *et al*, 2008).

The use of survival analysis, as opposed to the use of other statistical method, is most important when some subjects are lost to follow up or when the period of observation is finite certain patients may not experience the event of interest over the study period. In this latter case one cannot have complete information for such individuals. These incomplete observations are referred to as being censored. Most survival analyses consider a key analytical problem of censoring. In essence, censoring occurs when we

have some information about individual survival time, but we do not know the survival time exactly.

In reality such event can occur due to the following reasons:

1. a person does not experience the event before the study ends
2. a person is lost to follow-up during the study period and
3. a person withdraws from the study for unknown/known reasons

There are three categories of censoring.

i) **Right censoring**: Survival time is said to be right censored when it is recorded from its beginning to a defined time before its end time. This type of censoring is commonly recognized survival analysis and also considered in this study.

ii) **Left censoring**: Survival time is said to be left censored if an individual develops an event of interest prior to the beginning of the study; this is not common in survival studies.

iii) **Interval censoring**: Survival time is said to be interval censored when it is only known that the event of interest occurs within an interval of time but the exact time of its occurrence is not known.

The prospect of censoring complicates research design and complicates statistical analysis. Thus, censoring creates some unusual problem in the analysis of data because such data cannot be handled properly by standard statistical methods. Researchers used different techniques to respond to the complication due to censoring unsatisfactorily. New developments in statistical theory accompanied by new development in statistical computing have changed how researchers can study such data. Survival analysis (lifetime

analysis) was developed by biostatisticians modeling human life times in the medical and biological sciences.

Once this method was developed in modeling human life time where the target event is death, it has been serving as a powerful methodology that appropriately uses data from all observations. It does not matter whether the data is uncensored or censored. Data collection can be prospective or retrospective, experimental or observational. Time can be measured continuously or discretely and explanatory variables can be continuous or categorical.

The response variable in survival analysis is survival time and is no longer limited to only time to death. It is a non-negative random variable used loosely for the time period from a starting time point to the occurrence of a any event. In this study context, survival time is the length of lifetime of HIV co-infected patients from the date of TB treatment to the death of the patient that, and is examined and modeled using survival analysis.

3.3.2 Descriptive methods for survival data

In any applied setting, a statistical analysis should begin with a thoughtful and thorough description of the data. In particular, an initial step in the analysis of a set of survival data is to present numerical or graphical summaries of the survival times in a particular group. Routine applications of standard measures of central tendency and variability will not yield estimates of the desired parameters when the data include censored observations. In summarizing survival data, the two common functions of applied are the survivor function and the hazard function (Hosmer and Lemeshow, 1999).

3.3.2.1 Survivor function S(*t*)

The survivor function is defined to be the probability that the survival time of a randomly selected subject is greater than or equal to some specified time. Thus, it gives the probability that an individual surviving beyond a specified time. Moreover, the distribution of survival time is characterized by three functions: (a) the survivorship function, (b) the probability density function, and (c) the hazard function.

Let T be a random variable associated with the survival times, *t* be the specified value of the random variable T and *f* (*t*) be the underlying probability density function of the survival time T.

The cumulative distribution function *F* (*t*), which represents the probability that a subject selected at random will have a survival time less than some stated value *t*, is given

$$F(t) = P(T<t) = \int_0^t f(u)du, \quad t \geq 0 \tag{3.1}$$

Using the above the survivor function, *S*(*t*), can be given as

$$S(t) = P(T \geq t) = 1 - F(t), \; t \geq 0 \tag{3.2}$$

From equations (3.1) and (3.2) the relationship between *f* (*t*) and *S* (*t*) can be derived as

$$f(t) = \frac{d}{dt}F(t) = \frac{d}{dt}(1 - S(t)) = \frac{-d}{dt}S(t), \; t \geq 0 \tag{3.3}$$

Theoretically, as *t* ranges from 0 to infinity, the survivor function can be graphed as a smooth curve.

Survivor functions have the characteristics that:

1. they are non-increasing
2. at time $t = 0$, $S(t) = S(0) = 1$; that is, at the start of the study, since no one has experienced the event yet, the probability of surviving past time 0 is one and

3. at time $t \to \infty$, $S(t) = S(\infty) \to 0$; that is, theoretically, if the study period increased without limit, eventually nobody would survive, so the survivor curve must eventually converge to zero.

3.3.2.2 Hazard function h (t)

The hazard function *h* (*t*) gives the instantaneous potential for failing at time *t*, given that the individual has survived up to time *t*. It is also known as the conditional failure rate in reliability, the force of mortality in demography, the intensity function in stochastic process, the age specific failure rate in epidemiology, the inverse of the Mill's ratio in economics or simply the hazard rate. In contrast to the survivor function, which focuses on failing, the hazard function focuses on not failing, that is, on the event occurring. Thus, in some sense, the hazard function can be considered as giving the opposite side of the information given by the survivor function (David *at al*, 2005).

The hazard function $h(t) \geq 0$, is given as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P\{\text{an individual fails in the time interval}(t, t+\Delta t) \text{ given survives until time } t\}}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P\{t \leq T \leq t+\Delta t \mid T \geq t\}}{\Delta t} \qquad (3.4)$$

By applying the theory of conditional probability and the relationship in equation (3.3), the hazard function can be expressed in terms of the underlying probability density function and the survivor function becomes

$$h(t) = \frac{f(t)}{S(t)} = \frac{-d}{dt} \ln S(t).$$

The corresponding cumulative hazard function H(t) is defined as

$$H(t) = \int_0^t h(u)du = -\ln S(t)$$

Then, $S(t) = \exp(-H(t))$ and $f(t) = h(t)S(t)$. (3.5)

3.3.2.3 Estimation of survivorship function

In practice, when using actual data, we usually obtain estimated survivor function and obtain curves that are step functions, rather than smooth curves.

Kaplan-Meier estimator

The Kaplan-Meier (KM) estimator of the survivorship function [Kaplan and Meier (1958)], also called the product limit estimator, is mostly used to estimate the survivor and hazard functions. It 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.

Suppose $t_1, t_2, \ldots, t_n$ be the survival times of n independent observations and $t_{(1)} \leq t_{(2)} \leq \ldots \leq t_{(m)}$, $m \leq n$ be the m distinct ordered death times.

The Kaplan-Meier estimator of the survivorship function (or survival probability) at time t, $S(t) = P(T \geq t)$ is defined as:

$$\hat{S}(t) = \prod_{t_{(i)} \leq t} \frac{n_i - d_i}{n_i} = \prod_{t_{(i)} \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{3.6}$$

with the convention that $\hat{S}(t) = 1$ for $t < t_{(1)}$. In equation (3.6), n_i is the number of individuals who are at risk of dying at time t_i, and d_i is the number of individuals who failed (died) at time t_i.

The variance of the KM survival estimator which is also known as the Greenwood's formula is

$$\mathrm{Var}\left(\hat{S}(t)\right) = \left(\hat{S}(t)\right)^2 \sum \frac{d_i}{n_i(n_i - d_i)} \tag{3.7}$$

3.3.2.4 Comparison of Survivorship Functions

After obtaining statistics which provide a description of the overall survival experience, the survival and hazard functions, one is expected to proceed with a comparison of the survivorship experience of subgroups in the data. These groups might be defined by the values of a covariate which are thought to be related to survival times. When comparing groups of subjects, it is always preferable to begin with a graphical display of the data in each group. In studies of survival time, we should graph the Kaplan-Meier estimator of the survival function for each of the groups. That is, plotting the corresponding estimates of the two survivor functions on the same axes of Kaplan-Meier estimator. In general, if the plot shows the pattern of one survivorship function lying above another, this means the group defined by the upper curve lived longer, or had a more favorable survival experience than the group defined by the lower curve. But, the statistical question is whether the observed difference seen on the plot is significant. This needs to be answered using appropriate statistical test (Hosmer and Lemeshow, 1999).

The general form of test statistic that deal with this issue is given as

$$Q = \frac{\left[\sum_{i=1}^{m} w_i(d_{1i} - \hat{e}_{1i})\right]^2}{\sum_{i=1}^{m} w_i^2 \hat{v}_{1i}} \qquad \hat{e}_{1i} = \frac{n_{1i} d_i}{n_i} \quad \text{and} \quad \hat{v}_{1i} = \frac{n_{0i} n_{1i} d_i (n_i - d_i)}{n_i^2 (ni - 1)}. \tag{3.8}$$

In equation (3.8):

m is the number of rank-ordered failure (death) times.

n_{0i} is the number of individuals at risk at observed survival time $t_{(i)}$ in group 0

n_{1i} is the number of individuals at risk at observed survival time $t_{(i)}$ in group 1

d_{0i} is the number of observed deaths in group 0

d_{1i} is the number of observed deaths in group 1

n_i is the total number of individuals or risk prior to time $t_{(i)}$

d_i is the total number of deaths at time $t_{(i)}$

w_i is the weight for censor adjustment at failure time $t_{(i)}$.

The contribution to the test statistic depends on which of the various tests is used, but each may be expressed in the form of a ratio of weighted sums over the observed survival times. Under the null hypothesis that the two survivorship functions are the same, and assuming that the censoring experience is independent of group, and that the total number of observed events and the sum of the expected number of events is large, Q follows a chi-square distribution with one degree of freedom. We can also use the test based on Q above to compare more than two groups.

The log rank test which is a special case of Q is used in this study.

Log rank test

The log rank test, sometimes called the Cox-Mantel test, is the most well known and widely used test statistic. This test statistic is based on weights equal to one, i.e. $w_i = 1$.

Therefore, the log rank test statistic becomes

$$Q_{GWt} = \frac{\left[\sum_{i=1}^{m}\left(d_{1i} - \hat{e}_{1i}\right)\right]^2}{\sum_{i=1}^{m}\hat{v}_{1i}} \qquad (3.9)$$

3.3.3 Regression Models for Survival Data

In most medical studies which give rise to survival data, supplementary information referred to as covariates or independent variables needs to be collected on each individual, so that the relationship between survival experience of individuals and various explanatory variables have to be investigated.

In the analysis of survival data, interest centers on the risk of hazard of failure at any time after the time origin of the study. As a consequence, the hazard function is modeled directly in survival analysis. There are two broad reasons to model survival data. One objective of the modeling process is to determine which combinations of potential explanatory variables affect the form of the hazard function. Another reason for modeling the hazard function is to obtain an estimate of the hazard function itself for an individual from a set of explanatory variables (Klein and Moeschberger, 1997).

A variety of models and methods have been developed for doing this sort of survival analysis using either parametric or semi-parametric approaches. One of the most popular types of regression models used in survival analysis is the Cox proportional hazard model (Cox, 1972).

3.3.3.1 The Cox Proportional Hazards Regression Model

The Cox Proportional Hazard (PH) Model is a multiple regression method and is used to evaluate the effect of multiple covariates on the survival. Cox (1972) proposed a semi-parametric model for the hazard function that allows the addition of covariates, while keeping the baseline hazards unspecified and can take only positive values and it is defined as

$$h(t, X, \beta) = h_0(t)e^{\beta' X}, \tag{3.10}$$

where $h(t, X, \beta)$ hazard function at time t with covariates $X = (X_1, X_2, ..., X_p)'$,

$h_0(t)$ is the arbitrary baseline hazard function that characterizes how the hazard function changes as a function of survival time.

$\beta = (\beta_1, \beta_2, ..., \beta_p)'$ is a column vector of *p* regression parameters associated with explanatory variables.

$e^{\beta' X}$ characterizes how the hazard function changes as a function of subject covariates.

t is the failure time.

Each individual has its own hazard function of survival time. Then, the above model becomes

$$h(t, x_i, \beta) = h_0(t)\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip}), i = 1,2,...,n, \tag{3.11}$$

where: n is total number of observations in the study.

$x_i = (x_{i1}, x_{i2}, ..., x_{ip})'$ is a column vector of measured covariates for the i^{th} individual (patient) which are assumed to affect the survival probability.

The Cox proportional hazard model is popular because it allows a flexible choice of covariates: time varying, time-independent, continuous and discrete. Two other issues that makes it popular are that it does not make any assumption about the underlying survival distribution and also does not require estimation of the baseline hazard rate, $h_0(t)$ to estimate the regression parameters.

3.3.3.2 Assumption of Cox proportional hazard model

1. The baseline hazard function $h_0(t)$ depends on t, but not on covariates $x_1, x_2, ..., x_p$.
2. The hazard ratio, e^{β}, depends on the covariates $X = (X_1, X_2, ..., X_p)'$ not on time.
3. The covariates X_i are time-independent.

Because of assumption (2) it is called a proportional hazards model. To show this issue mathematically, consider two distinct values of a continuous covariate *X*, say, x_{i1} and x_{i2}.

$$h(t, X, \beta) = h_0(t) e^{\beta' X}$$

Then, the hazard ratio becomes

$$\frac{h(t, x_1, \beta)}{h(t, x_2, \beta)} = \frac{h_0(t) e^{\beta_i x_{i1}}}{h_0(t) e^{\beta_i x_{i2}}} = e^{\beta_i (x_{i1} - x_{i2})} \tag{3.12}$$

which is clearly independent of time.

This reveals that the ratio of the hazard functions for two individuals with different covariate values does not vary with time.

3.3.3.3 Estimation of Parameters in proportional hazard model

The regression coefficients in the proportional hazards Cox model, which are the unknown parameters in the model, can be estimated using the method of maximum likelihood.

In Cox proportional hazards model we can estimate the vector of parameters β without having any assumptions about the baseline hazard $h_0(t)$.

Consider *n* independent individuals, the data that we need for the Cox proportional hazard model is represented by triplet

$(t_i, \delta_i, x_i), i = 1,2,..., n,$ where

t_i is the survival time for i^{th} individual

δ_i an indicator of censoring for the i^{th} individual given by 0 for censored and 1 for event/death

x_i a vector of covariates for individual i $(x_{i1}, x_{i2},..., x_{ip})$

Then, the full maximum likelihood is defined as

$$L(\beta) = \prod_{i=1}^{n} h(t_i, x_i, \beta)^{\delta_i} S(t_i, x_i, \beta) \tag{3.13}$$

where $h(t_i, x_i, \beta) = h_0(t_i) e^{\beta' x_i}$ is the hazard function for individual i

$S(t_i, x_i, \beta) = (S_0(t_i))\exp(\beta' x_i)$ is the survival function for individual i.

Then, the full maximum likelihood becomes

$$L(\beta) = \prod_{i=1}^{n} (h_0(t_i)\exp(\beta' x_i))^{\delta_i} (S_0(t_i)) exp(\beta' x_i) \tag{3.14}$$

Full maximum likelihood requires that we maximize (3.14) with respect to the unknown parameter of interest, β, and unspecified baseline hazard and survival functions. This indicates that unless we explicitly specify the baseline hazard, $h_0(t)$ we cannot obtain the maximum likelihood estimators for the full likelihood. But, Cox (1972) proposed using an expression he called “partial likelihood function” that depends only on the parameter of interest.

3.3.3.4 Partial likelihood

Suppose *m* failures occur out of *n* subjects with $m \leq n$. Suppose that $t_{(1)} \leq t_{(2)} \leq ... \leq t_{(m)}$ are the *m* distinct ordered failure times observed, and let R_i be the set of subjects at risk prior to time $t_{(i)}$. Assume that there is only a single failure at time t_i, i.e. no ties. Then

P (individual *i* has experienced an event at time $t_{(i)}$ | one event at time $t_{(i)}$)

$$= \frac{h(t, x_i)}{\sum_{j \varepsilon R_{t(i)}} h(t, x_j)} = \frac{h_0(t)\exp(\beta' x_i)}{\sum_{j \varepsilon R_{t(i)}} h_0(t)\exp(\beta' x_j)} = \frac{\exp(\beta' x_i)}{\sum_{j \varepsilon R_{t(i)}} \exp(\beta' x_j)} \tag{3.15}$$

When that there are no tied times assumed, the partial likelihood is defined over all failure time $t_{(i)}$ that $i = 1,2,..,m$ and given as

$$L_p(\beta) = \prod_{i=1}^{m} \frac{\exp(\beta' x_i)}{\sum_{j \varepsilon R_{t(i)}} \exp(\beta' x_j)} \tag{3.16}$$

where the product is over *m* distinct ordered failure times and $x_{(i)}$ denotes the value of the covariate for the subject with ordered survival time $t_{(i)}$. The log partial likelihood function is

$$l_p(\beta) = \sum_{i=1}^{m} \left[\beta' x_i - \ln\left(\sum_{j \varepsilon R_{(ti)}} \exp(\beta' x_j) \right) \right] \tag{3.17}$$

We obtain the maximum partial likelihood estimator by differentiating the right hand side of (3.17) with respect to the component of β, setting the derivative equal to zero and solving for the unknown parameters. However, this partial likelihood function methods are based on the assumption that there are no tied values among the observed survival times. But, in most real situations tied survival times are more likely to occur. In addition

to the possibility of more than one death at a time, there might also be more than one censored observations at a time of death.

A number of approaches to handle tied data have been suggested and, of these, three are used by software packages: an exact expression that is derived in Kalbfleisch and Prentice (1980) and approximations due to Breslow (1974) and Efron (1977). However, approximations derived by Breslow (1974) and Efron (1977) are designed to provide expressions that are more easily computed than the exact partial likelihood, yet that still account for the fact that ties are among the observed values of survival time. In many applied settings there is little or no practical difference between the estimators obtained from the two approximations. Because of this, and since the Breslow approximation is more commonly available and popular it is used mostly (Hosmer and Lemeshow, 1999).

The Breslow approximation

This approximation is proposed by Breslow and Peto to modify the partial likelihood and has the form

$$L_B(\beta) = \prod_{i=1}^{m} \frac{\exp(\beta' s_i)}{\left[\sum_{l \in R_{t(i)}} \exp(\beta' x_l)\right]^{d_i}} \tag{3.18}$$

where d_i the number of deaths occurred at time t_i

s_i the sum of covariates over d_i subjects at time t_i

Then, the partial log of (3.18) is given as

$$l_B(\beta) = \sum_{i=1}^{m} \left[\beta' s_i - d_i \ln \sum_{l \in R_{t(i)}} \exp(\beta' x_l) \right] \tag{3.19}$$

Breslow maximum partial likelihood estimator, adjusted for tied observation is obtained, by differentiating equation (3.19) with respect to the components of β and setting the derivative equal to zero and solving for the unknown parameters.

3.3.4 Model development

In any applied setting, performing a proportional hazard regression analysis of 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 Hosmer and Lemeshow, 1999).

Selection of covariates

The methods available to select a subset of covariates to include in a proportional hazards regression model are essentially the same as those used in any other regression model. There are three methods of selection of influential covariates. These are purposeful selection, stepwise selection (forward selection and backward elimination) and best subset selection. Survival analysis using Cox regression method begins with a thorough univariable analysis of the association between survival time and all important covariates (Hosmer and Lemeshow, 1999).

Recommendable procedure in selecting variables in the study

Hosmer and Lemeshow (1999) and Collett (2003) recommended the following procedure in variable selection.

1. Include all variables that are significant in the univariable analysis at the 25 percent level and also any other variables which are presumed to be clinically important to fit the initial multivariable model.
2. The variables that appear to be important from step 1 are then fitted together in a multivariable model. In the presence of certain variables others may cease to be important. Consequently, backward elimination is used to omit non-significant variables from the model. Once a variable has been dropped, the effect of omitting each of the remaining variables in turn should be examined.
3. Variables, that were not important on their own, and so were not under consideration in step 2, may become important in the presence of others. These variables are therefore added to the model from step 2, with forward selection method. This process may result in terms in the model determined at step 2 ceasing to be significant.
4. A final check is made to ensure that neither significant variable is eliminated from the model nor non-significant variable is included in the model. At this stage the interactions between any of the main effects currently in the model can be considered for inclusion if the inclusion significantly modifies the model.

3.3.5 Model diagnostics for Cox PH model

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. The methods for assessment of a fitted proportional hazards model are essentially the same as for other regression models (Hosmer and Lemeshow, 1999).

Residuals are used to investigate the lack of fit of a model and useful for examining different aspects of the model. The following residuals have been proposed for use by different authors in connection with the Cox regression model.

Martingale residuals: It is a slight modification of the Cox-Snell residuals and is defined as $\hat{M}_i = \delta_i - r_i$ where, δ_i is the censoring indicator and r_i is Cox-Snell residual for i^{th} individual given as $r_i = \hat{H}_i(t) = -\ln \hat{S}_i(t)$, where $\hat{H}_i(t)$ and $\hat{S}_i(t)$ are the estimated values of the cumulative hazard and survivor functions of the i^{th} individual at time *t*.

The martingale residuals have the property that $\sum_{i=1}^{n} \hat{M}_i = 0$. They are uncorrelated and with expected value zero in large samples. In this respect they have properties similar to those possessed by residuals encountered in linear regression analysis (Collett, 2003).

Schoenfeld residuals: This overcomes the problem that the above three residuals depend heavily on observed survival time and cumulative hazard function. They are computed for each individual and covariate. It follows that, the Schoenfeld residual for the i^{th} individual and k^{th} covariate is defined as:

$$\hat{s}_{ik} = \delta_i \left[X_{ik} - \frac{\sum_{j \in R(t_i)} X_{jk} \exp(\hat{\beta}' Xj)}{\sum_{j \in R(t_i)} \exp(\hat{\beta}' X_j)} \right] \qquad (3.21)$$

where, X_j is a vector of p fixed covariates for the j^{th} individual, X_{jk} is the value of k^{th} covariate on the j^{th} individual.

Because of that, Schoenfeld residuals are defined only for the uncensored observations in which case

$\hat{s}_{ik} = X_{ik} - \frac{\sum_{j \in R(t_i)} X_{jk} \exp(\hat{\beta}' X_j)}{\sum_{j \in R(t_i)} \exp(\hat{\beta}' X_j)}$ and for each covariate it must sum to zero. In addition, they are uncorrelated and with expected value zero (Schoenfeld, 1982).

Most of the model diagnostics in survival data are based on the residuals stated above.

The following are model diagnostics that are required to assess the model adequacy in this study.

3.3.5.1 Testing for the Nonlinearity of covariates

After identifying a particular set of explanatory variables on which the hazard function depends, it is desirable to check that the correct functional form has been adopted for the continuous covariates. The Martingale residuals can be plotted against covariates to detect nonlinearity. Nonlinearity is not an issue for categorical variables, so we only examine plots of martingale residuals against continuous covariate. LOESS smoothed curve can be superimposed on the scatter plots to give interpretation. If the functional form observed in using the plots has some pattern, which is non linear, the covariate can be so transformed and the martingale residuals again should be plotted against the transformed covariate. A horizontal straight line which is drawn as a reference through zero would then confirm that the appropriate transformation has been used to the covariate. In addition, if the resulting smooth plot is a straight line compared to the reference line, then it shows linearity (Collett, 2003).

3.3.5.2 Examining influential observations

Another important aspect of model evaluation is through diagnostic statistics in order to identify which subjects have an unusual configuration of covariates or observations that

have influence on the estimates of the parameters or on the fit of the model. In other words a fitted model is particularly sensitive to one or more observations in the data set. Such observations can be termed as influential observations. Conclusions from survival analyses are often framed in terms of estimates of the relative hazard, which depends on the estimated values of the coefficients in the Cox regression model. Thus, it is desirable to examine the influence of each observation on these estimates. The interest is about observations that influence estimate of hazard functions and the complete estimate of the model and identifications of these observations. This could be done by fitting the model to all n observations in the data set, and then fitting the same model to the sets of $n-1$ observations obtained by omitting each of the n observations in turn. The interest is to determine if the result would change when a particular observation is removed from the analysis (Collett, 2003).

Suppose that $l_p(\beta)$ is log partial likelihood and $\hat{\beta}_j$ is the corresponding j^{th} parameter estimate of the model containing all the *n* observations and $l_p(\beta_{-i})$ be the log partial likelihood and $\hat{\beta}_{j(-i)}$ is the j^{th} parameter estimate of the model containing only the $n-1$ observations after deleting the i^{th} observation, respectively. Then, the statistic $\Delta_i\hat{\beta}_j = \hat{\beta}_j - \hat{\beta}_{j(-i)}$, which is known as DFBETA, can be used as a measure of how the j^{th} parameter estimate would change if the i^{th} observation was deleted from the data set. On the other hand, the statistic, $LD_{i=}\, 2l_p(\beta) - 2l_p(\beta_{-i})$, which is called the likelihood displacement statistic, can be used as a measure of how the maximized partial log likelihood changes if the i^{th} observation was deleted from the data set. Observations that influence a particular parameter estimate have a large absolute value of DFBETA than

other observations in the data set. Observations that do influence the overall fit of the model are those which have large values of likelihood displacement statistics than the other observations in the data set (Collett, 2003).

3.3.5.3 Checking Cox Proportional Hazard Assumption

In order to use the Cox model, it has to be checked that the assumption of whether the effects of covariates on hazard ratio remain constant over time. This is a vital assumption of proportional hazards model and must be assessed for each covariate. Several procedures of graphical techniques and tests are proposed to investigate the proportionality assumptions in fitting the Cox model (Cox, 1972). The Schoenfeld residuals are employed to assess this assumption.

The Schoenfeld residuals graphical technique can be used to assess Cox model proportionality assumption. The technique is based on individual contributions to the log-partial likelihood and measures the difference between the covariate for the i^{th} individual and a weighted average of the covariate over the risk set at each event. To check the proportionality assumption for each covariate, we plot the scaled Schoenfeld residuals against log of survival time. If the proportional hazards assumption is satisfied, the distribution of residuals over time is random, that is, does not show a particular trend, and the smoothed plot called Locally Weighted polynomial regression (Lowess) line summarizing the residuals should be a straight line and close to the horizontal reference line. Otherwise, a plot of scaled Schoenfeld residuals for a given covariate may reveal a violation of the proportional hazards assumption (Schoenfeld, 1982).

Formal tests need to detect any time dependency in particular covariates, after allowing for the effects of explanatory variables that are known. Testing the dependency of

covariates on time is equivalent to testing for a non-zero slope in a generalized linear regression of the scaled Schoenfeld residuals on functions of time. A non-zero slope is an indication of a violation of the proportional hazard assumption. The Grambsch-Therneau test of non-proportionality uses partial residuals for the test of proportional hazards assumption. In order to use this test for the i^{th} covariate, Grambsch and Therneau (1994) propose a time-varying coefficient as

$$\beta_i(t) = \beta_i + \gamma_i g_i(t) \tag{3.22}$$

where $\beta_i(t)$ is time varying coefficient, β_i is constant, and $g_i(t)$ is some specified function of time, usually $g_i(t) = \ln(t)$;

Then, the Cox proportional hazard model for time varying coefficient

with $g_i(t) = \ln(t)$ is defined as

$h(t, x_i, \beta_i(t)) = h_0(t)\exp(\beta_i(t)x)$, by substituting for $\beta_i(t)$ and $g_i(t)$ becomes

$$= h_0(t)\exp(\beta_i + \gamma_i \ln t)x$$

$$= h_0(t)\exp(\beta_i x + \gamma_i(\ln t)x). \tag{3.23}$$

Equation (3.23) is the proportional hazards model with the interaction term, $x\ln(t)$ and main effect x_i. To test the significance of the interaction term $x\ln(t)$, we perform the test: $H_0 : \gamma = 0$ versus $H_1 : \gamma \neq 0$ and we use likelihood-based tests like Wald test. If γ = 0 is not rejected, $\beta_i's$ are not time varying coefficients and hence the proportional hazards assumption is satisfied. If γ = 0 is rejected then the proportional hazards assumption is not satisfied, that leads to the need of other methods that cope with time-dependency (Schoenfeld, 1982).

CHAPTER FOUR

4. STATISTICAL DATA ANALYSIS AND DISCUSSION

4.1 Introduction

In this section we present results of data analysis, discussion and interpretation. The first part presents summary statistics of factors considered in this study. The second part compares the survival time in different groups. Descriptive survival analysis is investigated in comparing the survival time in different groups. The third part is about fitting the model. Then, the adequacy of the model is investigated. Finally, the results are discussed and interpreted. The statistical packages SPSS, SAS and STATA are employed to analyze the data.

4.2 Summary Statistics

The table below shows summary of demographic and health factors.

Table 4.1: Demographic and health factors of categorical covariates by HIV/TB co-infection under ART in Ambo hospital.

Demographic and Health factors		Status of censoring or event			
		Total	Death	Censored	Event/Death Percentage
Sex	Female	248	41	207	16.5
	Male	253	38	215	15.0
Marital status	Never Married	97	20	77	20.6
	Married	291	39	252	13.4
	Others	113	20	93	17.7
Level of Education	No Education	144	21	123	14.6
	Primary	216	35	181	16.2
	Secondary or above	141	23	118	16.3
Substance use (Tobacco, Alcohol)	NO	306	46	260	15.0
	YES	195	33	162	16.9
Type(site) of TB	Extra-Pulmonary	181	17	164	9.4
	Pulmonary	320	62	258	19.4
Functional Status	Ambulatory	120	27	93	22.5

	Bedridden	44	19	25	43.2
	Working	337	33	304	9.8
WHO clinical stage	Stage I or II	99	5	94	5.1
	Stage III	307	49	258	16.0
	Stage IV	95	25	70	26.3
Regimen type	AZT- based	215	30	185	14.0
	D4T- based	286	49	237	17.1

Table 4.2 shows the summary statistics of continuous variables considered in this study.

Table 4.2: Summary statistics of continuous variables included in the study

Patient Status	Continuous Variables	Mean	Standard deviation	Min.	Max.	Median	Q_1	Q_3
Censored	Time	32.53	20.03	1	78.66	31.89	15.37	46.51
	Age	33.05	9.08	17	78	31	27	38
	Weight	49.44	8.28	15	78	49	44	55
	CD4	178.54	125.13	6	699	159	80	244.75
Event/Death	Time	12.60	14.26	0.9	67.83	7.03	4.86	14.56
	Age	31.44	7.57	19	57	30	26	35
	Weight	40.28	11.6	16	64	42	30.5	49
	CD4	84.47	72.65	2	514	69	40	106
Overall	Time	29.38	20.55	0.9	78.66	27.7	9.56	44.6
	Age	32.79	8.87	17	78	31	27	38
	Weight	47.99	9.48	15	78	48	42.5	54
	CD4	163.71	123.22	2	699	137	65	226

4.3 Descriptive survival analysis

The patients were followed up for a median period of 27.7 months. The minimum follow-up time was 0.9 month and the maximum was 78.66 months. 79 subjects (15.8%) died during the follow up time, of which 5 (6.3%), 31 (39.2%) and 49 (62%) deaths occurred within two months, six months and nine months of TB treatment initiation, respectively. Next, we investigate whether the observed differences in data summary among different

factors are statistically significant or not with the help of Kaplan-Meier survival estimates and the log-rank test. The following graph of the estimate of overall Kaplan-Meier survivor function reveals that most of the deaths occurred in the earlier months of TB treatment initiation.

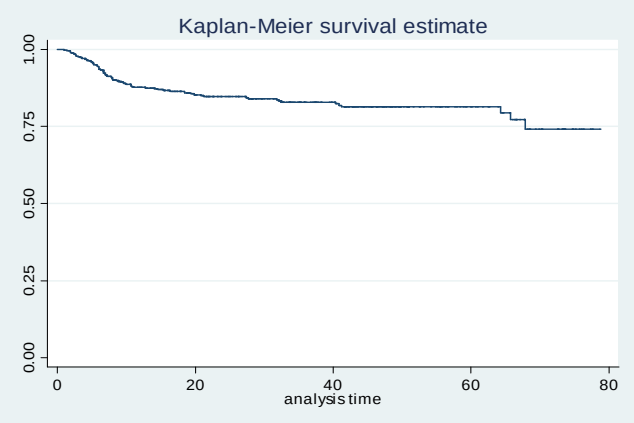


Figure 4.1: The plot of the overall estimate of Kaplan-Meier survivor function of HIV/TB co-infected patients under ART in Ambo hospital.

The Kaplan-Meier estimator survival curve gives the estimate of survivor function among different strata or groups of a covariates to make comparisons. Separate graphs of the estimates of the Kaplan-Meier survivor functions are constructed for different categorical covariates. In general, the pattern that one survivorship function lying above another means the group defined by the upper curve has a better survival than the group defined by the lower curve. From the graph for WHO clinical stage, functional status and TB type (Figures 4.2 a, b, c), there are clear differences among the various groups. However, the differences is not clear among sex, marital status, level of education, substance use and regimen type. For detail see Appendix D (Fig a-h).

a) b)

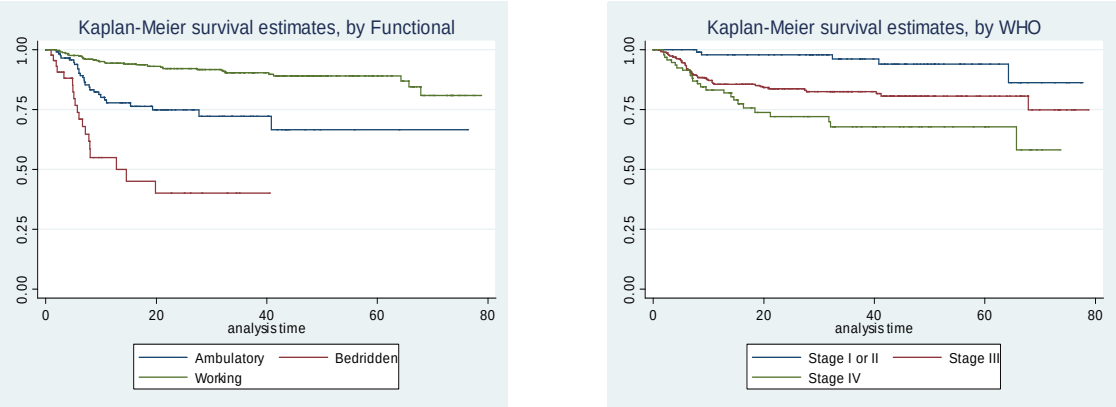


Fig. 4.2: The plot of the estimate of Kaplan-Meier survivor function of HIV/TB co-infected patients under ART in Ambo hospital a) functional status and b) WHO clinical stage.

Figure 4.2a shows that patients with working functional status have better survival than both ambulatory and bedridden; the ambulatory survive longer than the bedridden. Poor survival among bedridden patients is recognized at the beginning. Unlike bedridden patients, ambulatory patients show improvement after TB treatment is completed. Figure 4.2b shows that patients with WHO clinical stage of IV have the worst survival experience; where stages I and II have better survival compared to III.

c)

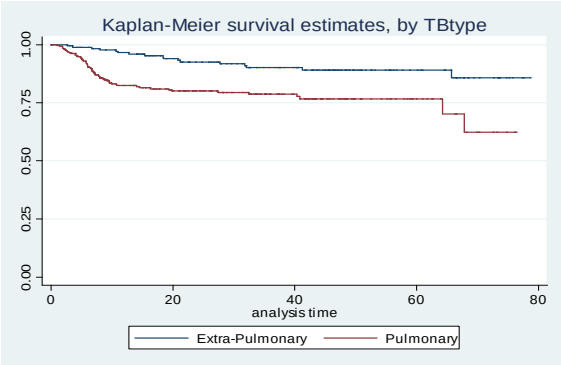


Fig. 4.2: The plot of the estimate of Kaplan-Meier survivor function of HIV/TB co-infected patients under ART in Ambo hospital c) TB site

Patients with Extra-pulmonary TB have better survival than those with Pulmonary TB.

To explore whether there are significant differences among different groups of categorical predictors we employ the log-rank statistical test. The null hypothesis being tested is that there is no difference between the survival curves.

Table 4.3: Results of the Log-rank test for the categorical variables of HIV/TB co-infected patients under ART in Ambo hospital.

Test of Equality over Strata			
Covariates	Chi-Square	DF	Pr >Chi-Square
Sex	1.1306	1	0.2876
Marital	4.3399	2	0.1142
Education	0.3914	2	0.8223
Substance	0.4633	1	0.4961
TB type	14.7560	1	0.0001
Functional	86.2667	2	<.0001
WHO	19.0185	2	<.0001
Regimen	1.5249	1	0.2169

The log-rank test indicates that statistically there is a significant difference of survival experience among groups of TB type (site), Functional status and WHO clinical stage. On the other hand, there are statistically no significant difference in survival/death experiencing among groups of the remaining categorical covariates sex, marital status, education, substance, and regimen.

4.4 Results of the Cox proportional hazards model

In statistical modeling, when the number of variables is relatively large, it can be computationally expensive to fit all possible models. Thus, one of the options is fitting a

multivariable model containing the variables that are significant at a modest level of significance in a univariable analysis. This is done in order to identify statistically significantly factors that influence the survival/death of HIV/TB co-infected patients using a multivariate Cox proportional hazard model.

The model will be constructed by first identifying factors which are significant at p-value 0.25 in univariate Cox proportional hazard analysis. Wald chi-square test is used to test the significance. Accordingly, Table 4.4 is the summary of univariate analysis is used to select potential predictors for further analysis.

Table 4.4: Results of the univariable proportional hazards Cox regression model of HIV/TB co-infected patients under ART in Ambo hospital.

Analysis of Maximum Likelihood Estimates									
Covariates	DF	$\hat{\beta}$	$s.e(\hat{\beta})$	Chi-Square	-2LL	Pr> ChiSq	$\widehat{HR}$	Lik. Rat. Sig	Score sig
Sex	1	0.239	0.22549	1.124	924.77	0.2899	1.270	0.2888	0.2878
Age	1	-0.023	0.01391	2.643	923.06	0.1040	0.978	0.0921	0.1038
Weight	1	-0.121	0.01294	88.019	839.21	<.0001	0.886	<.0001	<.0001
Marital	2			4.249	921.81	0.1194		0.1299	0.1143
Marital Married	1	-0.558	0.27530	4.103		0.0428	0.573	0.8200	0.8223
Marital Others	1	-0.269	0.31652	0.726		0.3942	0.764	0.4986	0.4962
Education	2			0.390	925.49	0.8227		0.8200	0.8223
Education Primary	1	0.130	0.27617	0.223		0.6368	1.139		
Education Secondary or above	1	0.184	0.30315	0.367		0.5448	1.202		
Substance	1	-0.155	0.22817	0.462	925.45	0.4966	0.856	0.4986	0.4962
TB Site	1	1.019	0.27621	13.598	909.89	0.0002	2.769	<.0001	<.0001
Functional	2			63.410	869.86	<.0001		<.0001	<.0001
Functional Ambulatory	1	1.257	0.26657	22.239		<.0001	3.515		
Functional Bedridden	1	2.358	0.30082	61.449		<.0001	10.571		

WHO	2			16.018	904.54	<.0003		<.0001	<.0001
WHO III	1	1.365	0.46975	8.439		0.0037	3.914		
WHO IV	1	1.906	0.49055	15.098		.0001	6.727		
CD4	1	-0.011	0.00174	42.722	856.99	<.0001	0.989	<.0001	<.0001
Regimen	1	-0.286	0.23238	1.514	924.35	0.2186	0.751	0.2140	0.2170
The value of -2LL for the null model is 925.893									

Consequently, the most important subset of these predictors to be included in the multivariable model will be selected based on their contribution to the maximized log partial likelihood of the model (-2*LL)*. The summary above reveals that the highest reduction in $-2LL(\hat{\beta})$ is observed for weight that reduced the value for the null/empty model, from 925.893 to 839.214, the difference is 86.679 and it is statistically significant (p-value <0.0001). The next highest change is obtained for CD4 of 68.99 followed by functional status 56.037.

Proceeding in this manner covariates will be eliminated in order of the magnitude in which they increase the $-2LL(\hat{\beta})$. Thus, using the Wald chi-square test, the predictors that are found to be significant, and will be considered for the next multivariable analysis at p-value 0.25 are Age, Baseline weight, marital status, TB type, Functional status, WHO clinical stage, baseline CD4 and regimen type. Furthermore it looks that weight, TB site, functional status, WHO clinical stage and CD4 count have strong associations with survival/death of PLWHIV co-infected with TB at p-value less than 0.05. Since keeping the covariates sex, education and substance in the model does not bring significant changes in the value of $-2LL(\hat{\beta})$, these variables become the first to be removed from the multivariable model. We proceed fitting the initial multiple Cox

proportional model by including the eight covariates that are significant at univariate analysis. Then, another multivariable Cox proportional regression will be fitted by eliminating those covariates that are not significant at 0.05 significance level. Accordingly, from the total of eight covariates, the prescribed ART regimen (p-value < 0.7618) is eliminated from the model. Following the same procedure age and marital status are not significant at p-values of 0.3934 and 0.1232, respectively. Hence, at this point we have a multivariate model which includes the five covariates weight, TB type, Functional status, WHO clinical stage and CD4. These covariates are significant at 5% significance level [see Appendix A Tables: 1-4].

The next step is assessing the importance of the variables which were not found to be significant in the univariable analysis as predictors or useful confounder of survival experience of patients and their effects. In addition, the effect of those variables not significant in multivariable analysis is also investigated. These variables are added one at a time in the model containing the five variables significant at 5% significance level. Then, the significance of their improvement on $-2LL(\hat{\beta})$ is dealt with. The result of the analysis reveals that, none of those variables were found to be significant and therefore they cannot be retained in the model. Thus, variables that were neither significant at univariable analysis nor at multivariable analysis are not confounders of the main factors in the preliminary model [see Appendix B, Tables: 1-6].

The final step in model development strategy is consideration of interaction terms that may be useful in the improvement of the model. All possible interactions of the variables of the preliminary multivariate model are formed to see if the interaction effects can

increase or decrease the survival time of TB patients. The Wald test is used to assess the significance of reasonable and possible interactions. The null hypothesis to be tested is that the model with only main effect fits the model equally well as the model having the main effects and their interactions as predictors. The decision for rejection of null hypothesis is reached if $-2LL_2-(-2LL_1) > \chi^2(\alpha=0.05) = 3.84$. Thus, the interaction of each variable is assessed. Accordingly, none of the variables significantly interact with the other variables [see Appendix B, Table: 7]. Therefore, the final model will be the one which contains only the main effects given in Table 4.5

Table 4.5: Estimated values of the coefficients, hazard ratios, 95% CI for the hazard ratio and P-values of the explanatory variables on fitting the proportional hazards model to the data from HIV/TB coinfection patients under ART in Ambo hospital.

Analysis of Maximum Likelihood Estimates								
Covariates	DF	$\hat{\beta}$	$s.e(\hat{\beta})$	Chi-Square	Pr> ChiSq	Estimated Hazard Ratio	95% CI for Hazard Ratio	
Weight	1	-0.08927	0.01318	45.8943	<.0001	0.915	0.891	0.939
TB type PT	1	0.84590	0.28778	8.6400	0.0033	2.330	1.326	4.096
Ref. group(extra-pulmonary)								
Functional	2			9.6562	0.0080			
Ambulatory	1	0.72137	0.27913	6.6787	0.0098	2.057	1.190	3.555
Bedridden	1	0.88486	0.32360	7.4772	0.0062	2.423	1.285	4.568
Ref. group(working)								
WHO	2			9.1172	0.0105			
WHO III	1	1.07090	0.47767	5.0263	0.0250	2.918	1.144	7.442
WHO IV	1	1.48525	0.50314	8.7141	0.0032	4.416	1.647	11.839
Ref. group(WHOI/II)								
CD4	1	-0.00838	0.00175	22.9328	<.0001	0.992	0.988	0.995

4.5 Model Diagnostics

The goal of statistical model development is to obtain which best describes the “middle” of the data. Having identified the final preliminary 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, it is desirable to determine whether a fitted Cox regression model adequately describes the data. In this study, three kinds of diagnostics that are, violation of the assumption of proportional hazards, effect of influential observations and nonlinearity in the relationship between the log hazard and the covariates are assessed.

4.5.1 Assessment of the proportional hazards assumption

One of the key assumptions of the Cox proportional hazard model is proportionality of hazards. The proportional hazards assumption, which asserts that the hazard ratios are constant overtime, is vital to the interpretation and use of a fitted proportional hazards model. That means, the risk of failure must be the same no matter how long subjects have been followed. In order to test this assumption, the extended Cox model is employed and a graphical display is used to substantiate the same. Thus, in this study, using a test based on the interaction of the covariates with the log of time and also using the plot of the scaled Schoenfeld residuals the assumption are used to see if the assumption of proportionality is violated or not.

Therefore, one of the statistical tests for proportional hazards assumption is to generate time varying covariates by creating interactions of the predictors and a function of survival times, usually covariate times log of time, and including these in the model. If any of the time dependent covariates are significant then those predictors do not show a

proportional effect over the study period. That is the proportional hazard assumption fails to hold. The result of the test is given in Appendix C (Table 1). The table shows that the Wald chi-square value and corresponding p-values for each covariate. Since the p-value of the Wald test is greater than 0.05 for all covariates, there no evidence against the proportionality of hazard assumption.

In addition, the assumption of proportionality was also assessed graphically by plotting the scaled Schoenfeld residuals of each covariate against log time. All interactions of covariates with the logarithm of survival times are modeled together with the main effects; and Wald statistic is used to test the significance of the interaction terms at 5% level of significance. The result of the test indicates that none of the coefficients of interaction terms are significant at 5% level (i.e. higher p-value) (Appendix D: Table1).

The result reveals the non-significance of time-dependent covariates. On the other hand, there are no covariates which show a trend/pattern with the time that indicates the hazard ratios will be constant over the study time. This shows that there is no sufficient evidence to reject the null hypothesis that the coefficients of the time varying covariates (interaction terms) are zero. This suggests that there is no enough evidence against proportionality assumption to hold.

Furthermore, plotting the scaled Schoenfeld residuals of each covariate against log time will be used to check whether the assumption of proportional hazards is violated or not. Clearly, a close look of this plot indicates that the residuals are random and LOESS curve are smooth and horizontal with zero slope. This also suggests that the plots support proportionality assumption to hold [see Appendix D, Fig: a-g].

4.5.2 Assessment of Influential Observations

The next step we follow in evaluation of regression diagnostic is to see if there are any observations that have undue influence on the estimates of the Cox regression parameters, or have an unexpected influence on the fit of the model. The DFBETA statistic for measuring the influence of the i^{th} observation is defined as the one-step approximation to the difference in the MLE of the regression parameter vector with i^{th} and the MLE of the regression parameter vector without the i^{th} observation. As a result, the first five largest changes in parameter estimates are shown in Table 2 of Appendix C.

4.5.3 Assessment of linearity of covariates in the model

In the process of model development, it is relevant to check whether the correct functional form of the continuous covariate holds in the model proposed to describe the data. A number of techniques are available, which are designed to determine whether the data support the hypothesis that the effect of the covariate is linear in the log hazard. As a result in this study graphical technique of the plots of the martingale residuals are used to assess the linearity of relation of continuous covariate in which the correct functional form is understood. Thus, the plot of the two continuous variables in the model, weight and CD4 is given in Figure 4.3(a, b).

a) b)

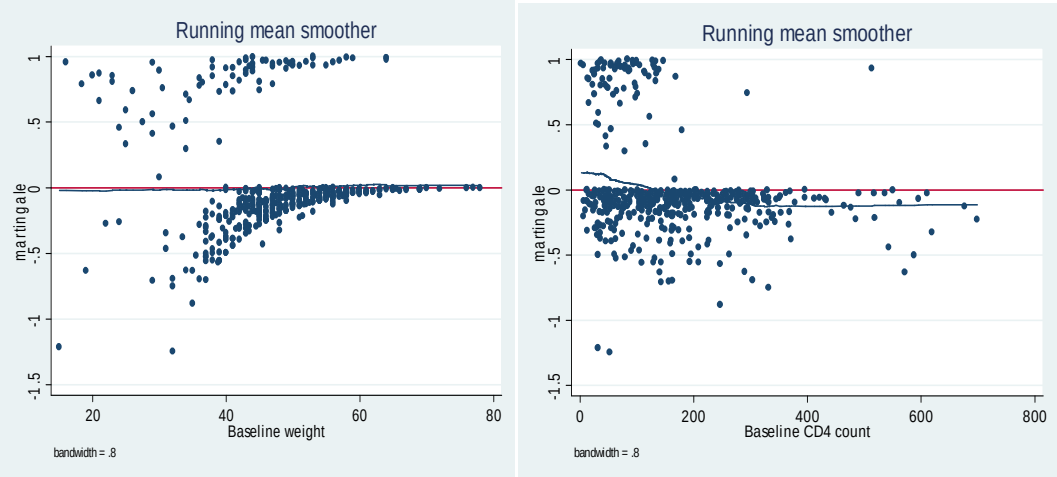


Figure 4.4: Plots of Martingale residuals for the continuous covariates (a) Baseline weight and (b) Baseline CD4 count

The figures show the plot of martingale residuals versus each covariate. For both of the covariates, baseline weight and CD4, the plots do not show systematic patterns or trend and the resulting smoothed plots (LOESS) are approximately horizontal straight lines. Therefore the plots of martingale residual confirm that initial weight and CD4 of a patient have an approximate linear relationship with the survival time.

At this point we can conclude that the model containing covariates Baseline weight, TB type, functional status, WHO clinical stage and baseline CD4 is an appropriate model to describe the data, since it has passed all tests of fitting model. In addition, results of the likelihood ratio, score and Wald tests for model goodness of fit displayed in Table 4 of Appendix A, Table 4 suggest that the model is good fit (i.e. significant at 5% level of significance). In general, the covariates of the model with respective property are given in Table 4.5 above.

4.6 Interpretation and discussion of the results

4.6.1 Interpretation of the results

The interpretation of the result from the fitted final model is based on the hazard ratios. The coefficient of the categorical covariates is interpreted as the logarithm of the ratio of the hazard of death to the baseline (reference group) hazard. That is, they are interpreted by comparing the reference group with others. Similarly, the coefficient for a continuous explanatory variable indicates the estimated change in the logarithm of the hazard ratio for a unit increase in the value of the respective covariate when the remaining covariates in the model are under control. Accordingly, the interpretation of the covariates included in the Cox proportional hazard model of HIV/TB co-infected patients in the case of Ambo hospital is as follows.

Suppose we take an increment of weight by kg to make comparisons. Then, the estimated hazard ratio for a five kg increase in initial weight will be 0.64 =exp (5*-0.08927). This means then the risk of death is 36% lower for a patient whose weight is 5kg higher than another patient controlling for other covariates.

The covariate CD4 is statistically significant influence on the survival of the patients. The associated estimated hazard ratio, $\widehat{HR}$, is 0.992 [95% CI: 0.988-0.995, p<0.0001]. For appropriateness of interpretation we take the estimated hazard ratio for a 50 cells/mm^3 increase in the baseline CD4 cell count which is 0.657=exp (-0.00838*50) showing that patients whose CD4 cell count is higher by 50cells/mm^3 die at lower hazard rate of 34.3% controlling for other variables in the model .

TB site or location has a significant impact on survival/death status of patients in which estimated hazard ratio of pulmonary TB in comparison to the extra-pulmonary TB is $\widehat{HR}$, 2.33 [95%CI: 1.326 - 4.096, p=0.0033]. The result reveals that the rate of dying among patients with pulmonary TB is 2.33 times higher than with extra-pulmonary TB patients controlling for other variables in the model.

Patients with ambulatory and bedridden functional status have $\widehat{HR}$=2.057 [95% CI: 1.190-3.555, p=0.0098] and $\widehat{HR}$ = 2.423 [95% CI: 1.285-4.568, p=0.0062] respectively. These results reveal that the risk of death in ambulatory and bedridden patients are 2 times and 2.4 times higher than that of patients with working functional status. On the other hand, patients with bedridden status in comparison to ambulatory have $\widehat{HR}$ =1.18=exp (0.88486-0.72137). This result indicates that patients with bedridden are 18% more likely to die than those with ambulatory functional status controlling for other variables in the model.

For patients in WHO stages III and IV $\widehat{HR}$ = 2.918 [95% CI: 1.144-7.442] and $\widehat{HR}$ =4.416 [95% CI: 1.647-11.839] in comparison to that of stage I and II controlling for other variables in the model. This indicates that the hazard of death is almost 3 times higher in stage III and 4.4 times higher in stage IV as compared with stage I and II. On the other hand, the estimated hazard ratio of stage IV compared to stage III is 1.51 =exp (1.48525-1.07090). This suggests, patients in stage IV are 51% more likely to die than patients in stage III.

4.6.2 Discussion of the results

The study by Mukadi *et al.*(2001) and WHO report (2005) reveal that the proportion of HIV-infected TB patients who die during TB treatment in sub-Saharan African countries is estimated to be between 6% and 39%. The current study found that initial weight, baseline CD4, TB site or location, functional status and WHO clinical stage are significant predictors of survival/death of HIV/TB co-infected patients under ART.

The impact of CD4 cell count on survival rate has been assessed by several studies indicating that the depletion of CD4 cell count is associated with high risk of death. Dembele *et al* (2010) conducted in Ouagadougou, Burkina Faso a retrospective cohort study that shows CD4+ T-cell count < 50 cells/mm^3 at HAART initiation were independently associated with a two-fold higher risk of TB. A study conducted in southwest Ethiopia by Mohammed *et al* (2011) indicates that a CD4 lymphocyte count less than 200 cells/mm^3 (OR=9.8, 95% CI: 5.5, 17.5) was associated with the development of active tuberculosis in people living with HIV/AIDS. Another study conducted in South Africa by Badri *et al* (2001) suggests that CD4+ T-lymphocyte count ≤200 cells/mm^3 was associated with progression of AIDS. A study conducted in Barcelona, Spain by Catala *et al* (2011) reveals that survival was worse among patients with <200/mm^3 CD4 cells (HR 1.8, 95%CI 1.2-2.7). The study by Maruza et al (2012) conducted in Brazil indicate CD4<200/mm^3 is a risk factor of death of HIV/TB co-infected patients. The finding of the current study also agrees with the above sited findings with the estimated hazard ratio $\widehat{HR}$, 0.992 [95% CI: 0.988-0.995, p<0.0001].

A retrospective study conducted by (Tarekegn, 2011) in Hawassa University referral hospital, Ethiopia, shows being bedridden functional status is one of risk factors of TB that shortens the survival by progressing the HIV infection among PLWHIV. The current study indicates being ambulatory and bedridden functional status are significantly affect the survival of HIV/TB co-infected patients with $\widehat{HR}$=2.057 [95% CI: 1.190-3.555, p=0.0098] and $\widehat{HR}$ = 2.423 [95% CI: 1.285-4.568, p=0.0062] respectively.

A study by Dembele *et al* (2010) suggested that WHO clinical Stages III and IV have significant impact on the survival by increasing the incidence of TB among PLWHIV. Mohammed *et al* (2011) show that WHO clinical stage has significant influence on the risk of mortality associated with TB in PLWHIV. Zubairu and Musa (2009) also indicated that WHO clinical stage is a risk factor of HIV/TB co-infection. The current study shows that the advanced WHO clinical stage III and IV are the risk factors of mortality among HIV/TB co-infected patients with estimated hazard ratio $\widehat{HR} = 2.918$ [95% CI: 1.144-7.442] and $\widehat{HR}$ =4.416 [95% CI: 1.647-11.839] respectively.

In Mohammed *et al* (2011) initial weight is reported as a risk factor of TB activation among PLWHIV. A study by Mashimbye (2010) conducted in South Africa suggested that initial weight is predicts the mortality of HIV/TB co-infection. The current study also identifies initial weight as a risk factor with estimated hazard ratio $\widehat{HR}$ = 0.915 [95% CI: 0.891-0.939].

Several studies show pulmonary TB more occur in PLWHIV than extra pulmonary TB. Zubairu and Musa (2009) show that among HIV co-infected patients pulmonary TB is more serious than extra-pulmonary as a cause of death. A study conducted in Dar es Salaam, Tanzania by Kamenju and Aboud (2011) suggested that, out of the total 300 TB patients tested for HIV, in which 175 (58.3%) were HIV-infected overall 104 of the TB patients that were admitted died. About two thirds of patients who died had Pulmonary TB in which a significantly higher proportion of deaths occurred among HIV-infected TB patients (29.1% *versus* 15.2%) than in the HIV uninfected TB patients (P=0.005). The study by Bhagyabati *et al* (2005) in Manipur, India suggested that site of TB is a risk factor of vulnerability of the HIV infected persons to TB. The current study also identified site of TB and Pulmonary TB significantly affect the survival of HIV/TB co-infected patients with estimated hazard ratio $\widehat{HR}$, 2.33 [95%CI: 1.326-4.096, p=0.0033].

CHAPTER FIVE

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In this study, the survival/ death status of HIV/TB co-infected patients who received TB treatment at Ambo hospital from September 1st 2006 to August 31st 2011has been studied. The Kaplan-Meier method is used to estimate the survival time of patients after commencement of TB treatment. The mortality rate was very high in the earlier months of TB treatment initiation and stabilized in later stages; 31 and 49 deaths occurred within six and nine months of TB treatment initiation, respectively. Using Cox proportional hazard model covariates that significantly influence the survival of PLWHIV co-infected patients are identified. Five covariates that are identified to affect the survival of the patients significantly at 0.05 level are initial weight, TB site, functional status, WHO clinical stage and baseline CD4 cell count.

5.2 Recommendations

- ✓ TB test must be done for people living with HIV in order to mitigate further deterioration of their health.
- ✓ Health workers should be cautious when a patient has lower initial weight and lower baseline CD4 cell count, and a patient is bedridden and is in WHO clinical stages III or IV.
- ✓ HIV/TB co-infected people should have awareness about the hazard of the risk factors identified in this study.

REFERENCES

1. Abera B, Kate F, Zelalem H, Andrew F. The association of TB with HIV infection in Oromia Regional National State, Ethiopia in 2006/7. Journal of Ethiopian Health Dev 2009; 23(1): 63- 67.
2. Aalen O, Borgan O, Gjessing H (2008). Survival and Event History Analysis. Springer-Verlag, New York .
3. Ackah AN, Coulibaly D, Digbeu H, Diallo K, Vetter KM, Coulibaly IM, et al. Response to treatment, mortality, and CD4 lymphocyte counts in HIV-infected persons with tuberculosis in Abidjan, Cote d'Ivoire. Lancet 1995; 345:607–10.
4. Bhagyabati *et al.* HIV and TB Co-infection. Journal, Indian Academy of Clinical Medicine 2005; 6(3): 220-3.
5. Bonnet MM, Pinoges LL, Varaine FF, Oberhauser BB, O'Brien DD, Kebede YY, Hewison CC, Zachariah RR, and Ferradini LL. Tuberculosis after HAART initiation in HIV positive patients from five countries with a high tuberculosis burden. AIDS 2006; 20:1275-9.
6. Catala L, Orcau A, Garcia de Olalla P, Millet J-P, Rodriguez-Mondragon A and Cayla J.A. Survival of a large cohort of HIV-infected tuberculosis patients in the era of highly active antiretroviral treatment. The International Journal of Tuberculosis and Lung Diseases 2011; 5(2):263-269.
7. Collett D. Modelling Survival Data for Medical Research (2003, 2nd Edition). Chapman-Hall, London.
8. Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC and Dye C: The growing burden of tuberculosis global trends and interactions with the HIV epidemic. Arch Intern Med 2003; 165:1009-1021.
9. Davies GR, Connolly C, Sturm AW, McAdam, KPWJ, Wilkinson D. Twice-weekly, directly observed treatment for HIV-infected and uninfected tuberculosis patients: cohort study in rural South Africa. AIDS 1999; 137: 811-817
10. Demissie M, Lindtjorn B, and Tegbaru B. HIV infection in tuberculosis patients in Addis Ababa. Ethiop J Health Dev 2000; 14: 277-82.
11. Deribew A, Markos T, Yohannes H, Nebiyu N, Shallo D, Ajeme W, Tefera B, Ludwig A and Robert C. Tuberculosis and HIV co-infection: its impact on quality of

life. Health and Quality of Life Outcomes 2009;**7**:105. Available at Website: http://www.hqlo.com/content/7/1/105.

12. Deribew Amara, Markos Tesfaye, Yohannes Hailmichael, Ludwig Apers, Gemeda Abebe, Luc Duchateau and Robert Colebunders. Common mental disorders in TB/HIV co-infected patients in Ethiopia BMC Infectious Diseases 2010, 10:201. Available at Website: http://www.biomedcentral.com/1471-2334/10/201.
13. EL Corbett, RW Steketee, FO ter Kuile, AS Latif, A Kamali, RJ Hayes: HIV-/AIDS and the control of other infectious diseases in Africa. Lancet 2002; 359:2177-87.
14. EL Corbett, CJ Watt, N Walker, D Maher, BG Williams, MC Raviglione, C Dye: The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch Intern Med 2003;163:1009-21.
15. Elliott AM, Halwiindi B, Hayes RJ. The impact of human immunodeficiency virus on mortality of patients treated for tuberculosis in a cohort study in Zambia. Trans R Soc Trop Med Hyg. 1995;89:78–82.
16. Enrico Girardi, Fabrizio Palmieri, Antonella Cingolani, Adriana Ammassari, Nicola Petrosillo, Laura Gillini, Daniela Zinzi, Andrea De Luca, Andrea Antinori and Giuseppe Ippolito. Changing Clinical Presentation and Survival in HIV-Associated Tuberculosis After Highly Active Antiretroviral Therapy. Journal of Acquired Immune Deficiency Syndromes 2001; 26(4), 326-31
17. Fauci AS, Lane HC, Braunwald E, Fauci As, Hauser SL, Longo DL, Jameson JL, editors. Harrison's Principles of Internal Medicine (16th Edition) McGraw-Hill, New York 2005; pp. 1076-139.
18. FMOHE. Tuberculosis, Leprosy and TB/HIV prevention and control program: Manual Fourth Edition 2008;. Addis Ababa, Ethiopia.
19. Girardi E, Antonucci G, Vanacore P, Libanore M, Errante I, Matteelli A, and Ippolito G. Impact of combination antiretroviral therapy on the risk of tuberculosis among persons with HIV infection. AIDS 2000;14:1985-91.
20. Godfrey-Fausset P, Maher D, Mukadi YD, Nunn P, Perriens J and Raviglione M. How human immunodeficiency virus voluntary testing can contribute to tuberculosis control. Bull World Health Organ 2002, 80:939-945.

21. Hoffmann C. Opportunistic Infections: In Bernd Sebastian Kamps, ed. HIV Medicine 2007, 15th Edition. Flying Publisher, Paris, p. 389-477.
22. Hosmer, D.W. and Lemeshow S. (1999). Applied Survival Analysis. John Wiley and Sons, Inc., New York.
23. Jaime A, Jorge O, David A, Adrian V and Raul S. Effect of antiretroviral therapy on survival of HIV-infected tuberculosis patients in Peru. Rev Panam Infectol 2010; 12(2):37-43.
24. Jon F. Tuberculosis. In: Jonathan C, Willium P. eds. Infectious Diseases. Vol. 1, 2nd ed. Spain: Mosby; 2004.
25. Jones JL, Hanson DL, Dworkin MS, and DeCock KM. HIV-associated tuberculosis in the era of highly active antiretroviral therapy. The Adult/Adolescent Spectrum of HIV Disease Group. Int J Tuberc Lung Dis 2000; 4:1026-31.
26. Kevin P. Cain, Thanomsak A, Channawong B, Somsak A, Wiroj M, Chawin S, Sriprapa N, Wanchai S, Theerawit T, and Jay K. Varma. Causes of Death in HIV-infected Persons Who Have Tuberculosis, Thailand. Emerging Infectious Diseases 2009; 15(2), 258-64.
27. Klein, J.P. and Moeschberger, M.L. (1997). Survival Analysis techniques for censured and truncated data. Springer-Verlag, New York.
28. Mashimbye Lawrence (2010). Tuberculosis (TB) treatment outcomes in adult TB patients attending a rural HIV clinic in South Africa (Bushbuckridge). Available at; Website: http://hdl.handle.net/10539/7982.
29. Mazura, M; Albuquerque, M. F. P. M; Braga, M. C; Barbosa, M. T. S; Byington, R; Coimbra, I; Moura, L. V; Batista, J. D. L; Diniz, G. T. N; Miranda-Filho, D. B; Lacerda, H. R; Rodrigues, L. C; Ximenes, R. A. A. Survival of HIV-infected patients after starting tuberculosis treatment: a prospective cohort study. International Journal of Tuberculosis and Lung Diseases 2012;16(5): 618-624.
30. M Badri, R Ehrlich, R Wood, T Pulerwitz, G Maartens: Association between tuberculosis and HIV disease progression in a high tuberculosis prevalence area. Int J Tuberc Lung Dis 2001, 5(3):225-32.
31. M Badri, D Wilson, R Wood. Effect of highly active antiretroviral therapy on incidence of tuberculosis in South Africa: a cohort study. Lancet 2002; 359: 2059-64.

32. MC Raviglione, JP Narain, A Kochi. HIV-associated tuberculosis in developing countries: clinical features, diagnosis, and treatment. Bull World Health Organ 1992; 70:515-26.
33. Dembele M, N Salerl A.C.C.Calvalho, T.Saouadogo, A.D.Hlen, I.Zabsonre, S.T.Koala, J.Simpore and A.Mattelli. Incidence of tuberculosis after HAART initiation in a cohort of HIV-positive patients in Burkina Faso. Int J Tuberculosis, 2010; 14(3):318–323
34. Mohammed T. *et al*. Risk factors of active tuberculosis in people living with HIV/AIDS in southwest Ethiopia: a case control study. Ethiopian J Health Sci. 2011; 21(2): 131-9
35. Mukadi YD, Maher D, Harries A. Tuberculosis case fatality rates in high HIV prevalence populations in sub-Saharan Africa. AIDS. 2001; 15:143–52.
36. Natpatou S, Kevin P, Patcharin S, Sriprapa N, Norio Y, Wanpen W, Surin S, Wanchai S, Samroui K, Sakon I and Jay K. Antiretroviral Therapy for HIV-Infected Tuberculosis Patients Saves Lives but Needs to Be Used More Frequently in Thailand. Journal of Acquired Immune Deficiency Syndromes 2008; 48(2):181–189.
37. Nunn P, Brindle R, Carpenter L. Cohort study of human immunodeficiency virus infection in patients with tuberculosis in Nairobi, Kenya. Am Rev Respir Dis. 1992; 146: 849–54.
38. Odhiambo JA, Borgdorff MW, Kiambih FM, Kibuga DK, Kwamanga DO, Ng'ang'a L, Agwanda R, Kalisvaart NA, Misljenovic O, Nagelkerke NJ, and Bosman M. Tuberculosis and the HIV epidemic: increasing annual risk of tuberculosis infection in Kenya, 1986-1996. Am J Public Health 1999; 89:1078-82.
39. Pili Kamenju and Said Aboud. Tuberculosis-HIV co-infection among patients admitted at Muhimbili National Hospital in Dar es salaam, Tanzania. Tanzania Journal of Health Research 2011; 13(1): 25-31.
40. Rosas-Taraco AG, Arce-Mendoza AY, Caballero-Olin G, and Salinas-Carmona MC. Mycobacterium tuberculosis upregulates coreceptors CCR5 and CXCR4 while HIV modulates CD14 favoring concurrent infection. AIDS Res Hum Retroviruses 2006; 22:45-51.

41. Santoro-Lopes G, de Pinho AM, Harrison LH, and Schechter M. Reduced risk of tuberculosis among Brazilian patients with advanced human immunodeficiency virus infection treated with highly active antiretroviral therapy. Clin Infect Dis 2002; 34:543-6.
42. Schluger NW, and Rom WN. The host immune response to tuberculosis. Am J Respir Crit Care Med 1998;157:679-91.
43. Schoenfeld, D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika 1982; 7: 145-153.
44. SK Sharma, A Mohan, T Kadhiravan: HIV-TB co-infection: epidemiology, diagnosis and management. Indian J Med Res 2005; 121:550-67.
45. Sonnenberg P, Glynn JR, Fielding K, Murray J, Godfrey-Faussett P, and Shearer S. HIV and pulmonary tuberculosis: the impact goes beyond those infected with HIV. AIDS 2004; 18:657-62.
46. Straetemans M, Bierrenbach AL, Nagelkerke N, Glaziou P, van der Werf MJ. The Effect of Tuberculosis on Mortality in HIV Positive People: A Meta-Analysis. PLoS ONE 2010; 5(12): 1-10: e15241. doi:10.1371/journal.pone.0015241
47. Tarekegn S, 2011. The effect of HAART on incidence of tuberculosis among HIV infected patients in Hawassa university referral hospital, South Ethiopia. Addis Ababa University unpublished MPH thesis.
48. Toossi Z, Xia L, Wu M, and Salvekar A. Transcriptional activation of HIV by Mycobacterium tuberculosis in human monocytes. Clin Exp Immunol 1999; 117:324-30.
49. Whalen C, Horsburgh CR, Hom D, Lahart C, Simberkoff M, Ellner JJ. Accelerated course of human immunodeficiency virus infection after tuberculosis. Am J Respir Crit Care Med 1995; 151:129–35.
50. Whalen CC, Nsubaga P, Okwerea A, et al. Impact of pulmonary tuberculosis on survival of HIV-infected adults: a prospective epidemiologic study in Uganda. AIDS 2000; 4:1219–28.
51. Weerawat M, Suthat C, Supeda T, Achara C and Somnuek S. Survival rate and risk factors of mortality among HIV/Tuberculosis-co-infected patients with and without

Antiretroviral Therapy. Journal of Acquired Immune Deficiency Syndromes 2006; 43(1):42-46.

52. World Health Organization. Interim policy on collaborative TB/HIV activities. Geneva 2004. WHO/HTM/TB/2004.330 [cited 2008 Jun 22]. Available from http://www.who.int/tb/publications/tbhiv_interim_policy/en/index.html.
53. World Health Organization. Global Tuberculosis Control: surveillance, planning, financing. WHO report. Geneva, Switzerland 2005. WHO/HTM/TB/2005.349.
54. World Health Organization, Global Tuberculosis Control: surveillance, planning, financing. Geneva, Switzerland. WHO/HTM/TB/2008.393.
55. World Health Organization Global Tuberculosis Control: Epidemiology, Strategy, Financing: WHO Report 2009. Geneva, Switzerland.
56. World Health Organization. Global Tuberculosis Control, 2011.
57. Yassin MA, Takele L, Gebresenbet S, Girma E, Lera M, Lendebo E, and Cuevas LE. HIV and tuberculosis co-infection in the southern region of Ethiopia: a prospective epidemiological study. Scand J Infect Dis. 2004, 36:670-3.
58. Zubairu L and Musa B. Prevalence and Predictors of Tuberculosis Co-infection among HIV-Seropositive Patients Attending the Aminu Kano Teaching Hospital, Northern Nigeria. J Epidemiol 2009; 19(2):81-87.

APPENDIX A: Results of the multivariable proportional hazards Cox regression model

Table 1: Results of the multivariable proportional hazards Cox regression model containing the variables significant at 20 - 25% level in the univariable proportional hazards Cox regression model.

Analysis of Maximum Likelihood Estimates						
Covariates	DF	$\hat{\beta}$	$s.e(\hat{\beta})$	Chi-Square	Pr > ChiSq	Hazard Ratio
Age	1	-0.01333	0.01562	0.7281	0.3935	0.987
Weight	1	-0.08701	0.01333	42.6028	<.0001	0.917
Marital	2			4.1198	0.1275	
Marital M	1	-0.36387	0.29947	1.4764	0.2243	0.695
Marital O	1	0.20655	0.35309	0.3422	0.5586	1.229
TB type PT	1	0.87829	0.28850	9.2681	0.0023	2.407
Functional	2			10.9261	0.0042	
Functional A	1	0.70819	0.28016	6.3897	0.0115	2.030
Functional B	1	1.01011	0.33265	9.2204	0.0024	2.746
WHO	2			9.7493	0.0076	
WHO III	1	1.15581	0.48181	5.7545	0.0164	3.177
WHO IV	1	1.56225	0.50802	9.4566	0.0021	4.770
CD4	1	-0.00846	0.00177	22.8955	<.0001	0.992
regimen 1	1	-0.07484	0.24691	0.0919	0.7618	0.928

Table 2: Results of the multivariable proportional hazards Cox regression model after eliminating the variable regimen from the multivariable proportional hazards Cox regression model in Table 1.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	174.6783	10	<.0001
Score	189.1490	10	<.0001
Wald	138.6041	10	<.0001

Covariates	DF	Wald Chi-Square	Pr > ChiSq
Age	1	0.7284	0.3934
Weight	1	42.7898	<.0001
Marital	2	4.6803	0.0963
TB type	1	9.2682	0.0023
Functional	2	10.8117	0.0045
WHO	2	9.8812	0.0072
CD4	1	22.8530	<.0001

Table 3: Results of the multivariable proportional hazards Cox regression model after eliminating the variable age from the multivariable proportional hazards Cox regression model in Table 2.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	174.2538	9	<.0001
Score	188.9647	9	<.0001
Wald	138.5465	9	<.0001

Covariates	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	43.7922	<.0001
Marital	2	4.1886	0.1232
TB type	1	8.9651	0.0028
Functional	2	10.8299	0.0044
WHO	2	9.6574	0.0080
CD4	1	23.6296	<.0001

Table 4: Results of the multivariable proportional hazards Cox regression model after eliminating the variable marital status from the multivariable proportional hazards Cox regression model in Table 3.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	169.6776	7	<.0001
Score	187.3912	7	<.0001
Wald	139.9440	7	<.0001

Covariates	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	45.8943	<.0001
TB type	1	8.6400	0.0033
Functional	2	9.6562	0.0080
WHO	2	9.1172	0.0105
CD4	1	22.9328	<.0001

APPENDIX B: The result of multivariable Cox hazard model those variables not significant in both the univariable analysis and in multivariable analysis fitted by included in the model containing those variable significant in multivariate analysis one at a time.

I. Result of multivariate Cox proportional hazard model containing variables in Table 4(Appendix A) including those not significant in univariable analysis one at a time.

Table 1: When sex is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	167.4708	8	<.0001
Score	178.0618	8	<.0001
Wald	131.0964	8	<.0001

Covariate	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	47.1745	<.0001
TB type	1	9.3769	0.0022
Functional	2	9.5453	0.0085
WHO	2	9.4060	0.0091
CD4	1	22.7577	<.0001
Sex	1	1.2926	0.2556

Table 2: When Education level is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	169.7904	9	<.0001
Score	177.6014	9	<.0001
Wald	127.3224	9	<.0001

Covariate	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	46.9313	<.0001
TB type	1	8.0657	0.0045
Functional	2	11.6916	0.0029
WHO	2	8.8904	0.0117
CD4	1	24.8488	<.0001
Education	2	3.5764	0.1673

Table 3: When substance use is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	166.2940	8	<.0001
Score	175.6641	8	<.0001
Wald	130.9986	8	<.0001

Effect	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	45.6855	<.0001
TB type	1	8.3361	0.0039
Functional	2	9.7319	0.0077
WHO	2	8.7934	0.0123
CD4	1	23.0695	<.0001
Substance	1	0.1097	0.7405

II. When those factors not significant in multivariate model are included to the multivariate model of covariate in Table 4(Appendix A) one at a time

Table 4: when age is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	166.4956	8	<.0001
Score	176.0626	8	<.0001
Wald	131.7699	8	<.0001

Effect	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	43.8445	<.0001
TB type	1	8.7365	0.0031
Functional	2	9.7008	0.0078
WHO	2	9.3191	0.0095
CD4	1	22.2869	<.0001
Age	1	0.3033	0.5818

Table 5: When marital status is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	170.3309	9	<.0001
Score	177.2209	9	<.0001
Wald	130.6623	9	<.0001

Effect	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	43.7922	<.0001
TB type	1	8.9651	0.0028
Functional	2	10.8299	0.0044
WHO	2	9.6574	0.0080
CD4	1	23.6296	<.0001
Marital	2	4.1886	0.1232

Table 6: When regimen is included

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	166.7704	8	<.0001
Score	175.6493	8	<.0001
Wald	130.0990	8	<.0001

Covariate	DF	Wald Chi-Square	Pr > ChiSq
Weight	1	45.8581	<.0001
TB type	1	8.5669	0.0034
Functional	2	10.1154	0.0064
WHO	2	8.9574	0.0113
CD4	1	23.1742	<.0001
Regimen	1	0.5783	0.4470

Table 7: Summary of interaction of two covariates.

Model Fit Statistics				
Interactions	$-2LL_2$ With main effects only	$-2LL_1$ With main effects and interaction	$-2LL_2-(-2LL_1)$	Decision
Weight*TB type	829.388	827.923	1.465	Do not reject
Weight*Functional	810.928	810.562	0.366	Do not reject
Weight*WHO	826.278	825.626	0.652	Do not reject
Weight*CD4	790.268	789.600	0.668	Do not reject
TB type*Functional	861.491	861.198	0.293	Do not reject
TB type*WHO	884.156	882.027	2.129	Do not reject
TB type*CD4	845.387	843.506	1.881	Do not reject
Functional*WHO	852.867	849.689	3.178	Do not reject
Functional*CD4	829.704	828.543	1.161	Do not reject
WHO*CD4	841.314	840.335	0.979	Do not reject

APPENDIX C: Results of Model diagnostics

Table 1: Result of test of proportionality assumption

Analysis of Maximum Likelihood Estimates						
Covariates	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
Weight	1	-0.056	0.043	1.733	0.188	0.945
Weight*Log time	1	0.025	0.042	0.37	0.543	1.026
TB site	1	0.823	1.328	0.378	0.539	2.277
TB type*Log time	1	-0.400	1.075	0.139	0.709	0.670
Functional	2			0.220	0.896	
Functional A	1	-0.601	1.328	0.205	0.651	0.548
Functional B	1	-0.595	1.573	0.143	0.705	0.551
Functional*Log time	2			0.375	0.829	
Functional A*Log time	1	0.246	1.433	0.30	0.8633	1.279
Functional B*Log time	1	-0.641	1.254	0.261	0.609	0.527
WHO	2			0.822	0.663	
WHO III	1	0.941	2.055	0.210	0.647	2.564

WHO IV	1	0.119	2.116	0.003	0.955	1.126
WHO*Log time	2			1.154	0.562	
WHO III*Log time	1	-0.118	1.633	0.005	0.943	0.889
WHO IV*Log time	1	-0.930	0.906	1.053	0.305	0.395
CD4	1	-0.014	0.007	4.425	0.055	0.986
CD4*Log time	1	0.008	0.005	2.587	0.108	1.008
Overall test	7			5.0695	0.6515	

Table 2: The five highest differences in the parameter estimates of the variables included in the model in Table 4, Appendix A when the data value for each patient is in turn deleted from the model.

Covariates	**Deleted Observation** (i)	$\Delta j(-i) = \hat{\beta}_j - \hat{\beta}_{j(-i)}$	$\left\|\Delta_{j(-i)} = \hat{\beta}_j - \hat{\beta}_{j(-i)}\right\|$
Weight	251	0.00301183	0.00301183
	72	0.00305898	0.00305898
	49	0.00318868	0.00318868
	42	0.00320662	0.00320662
	3	0.00381524	0.00381524
TB type	155	0.0337255	0.0337255
	461	0.0364544	0.0364544
	342	0.0414983	0.0414983
	284	0.0474521	0.0474521
	225	0.0549297	0.0549297
Functional Status (Ambulatory)	73	0.0389255	0.0389255
	166	0.0427295	0.0427295
	91	0.0477142	0.0477142
	251	0.0564406	0.0564406
	345	0.0606001	0.0606001
Functional Status (Bedridden)	347	0.0497107	0.0497107
	44	0.0527764	0.0527764
	42	0.0641815	0.0641815
	100	0.0656055	0.0656055
	3	0.0660085	0.0660085
WHO Clinical Stage III	483	0.0474269	0.0474269
	370	0.0520541	0.0520541
	243	0.0666383	0.0666383
	258	0.0808293	0.0808293
	358	0.0963211	0.0963211
WHO Clinical	286	0.0590756	0.0590756

Stage IV	177	0.0596506	0.0596506
	243	0.0641821	0.0641821
	258	0.0829050	0.0829050
	358	0.0976362	0.0976362
CD4	479	0.000241708	0.000241708
	91	0.000252334	0.000252334
	477	0.000359948	0.000359948
	341	0.000495403	0.000495403
	347	0.001204738	0.001204738

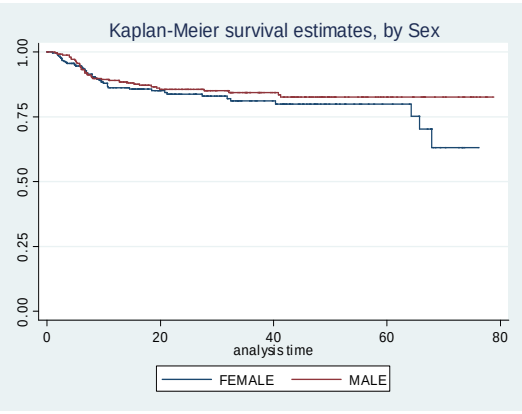
Table 3: The five highest likelihood displacement values when each observation is in turn deleted from the model in Table 4, Appendix A

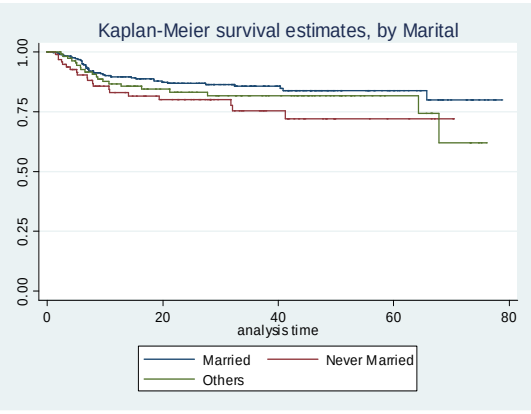
Likelihood Displacement Statistics	
$LD_i = 2l_p(\beta) - 2l_p(\beta_{-i})$	Deleted observation (i)
0.163059	53
0.205253	1
0.214839	8
0.223742	66
0.525962	45

APPENDIX D: Figures

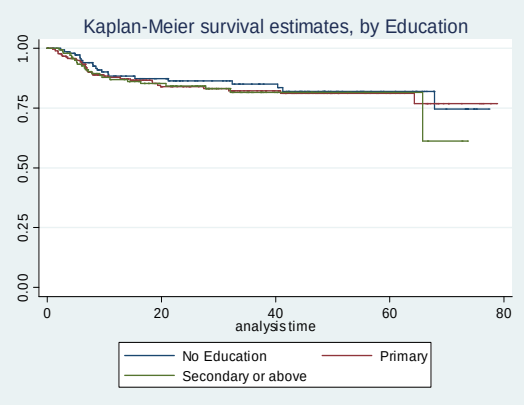
Figure 1 (a – h): Plots of Kaplan-Meier survivor functions different categories or groups

a) b)

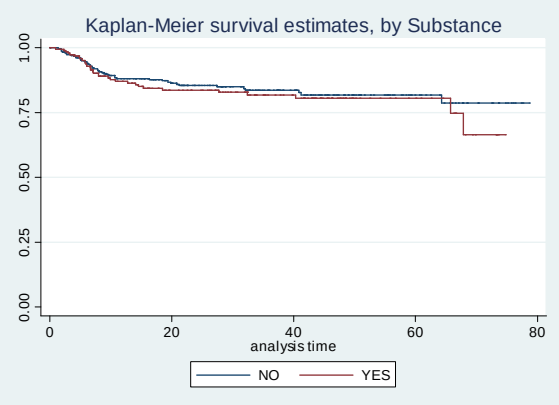




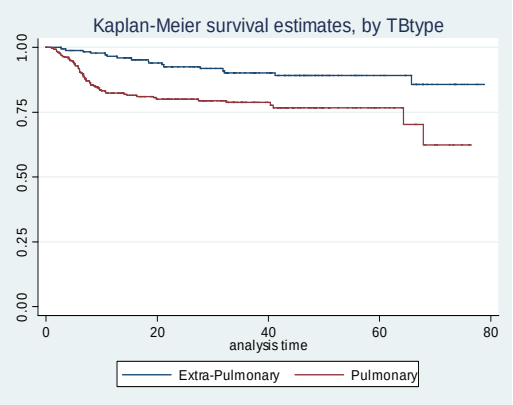
c)


Kaplan-Meier survival estimates, by Education
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
No Education
Primary
Secondary or above

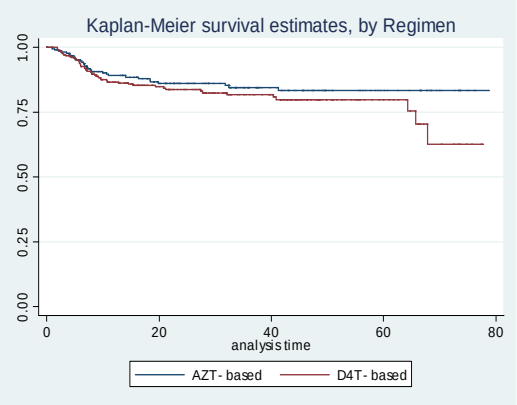
d)


Kaplan-Meier survival estimates, by Substance
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
NO
YES

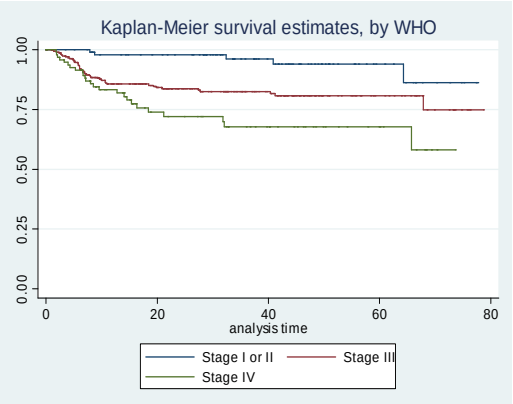
e)


Kaplan-Meier survival estimates, by TBtype
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
Extra-Pulmonary
Pulmonary

f)


Kaplan-Meier survival estimates, by Regimen
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
AZT- based
D4T- based

g)


Kaplan-Meier survival estimates, by WHO
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
Stage I or II
Stage III
Stage IV

h)

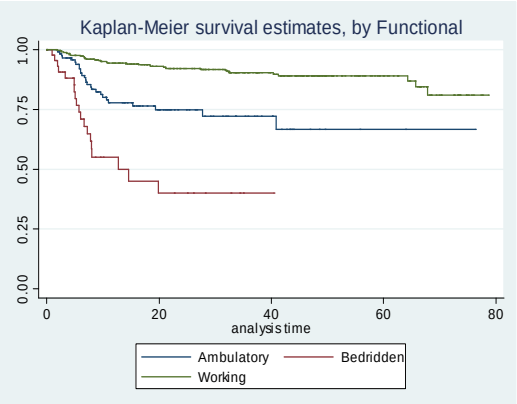
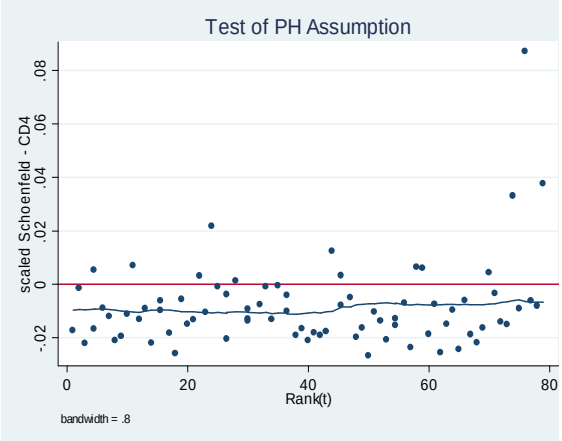
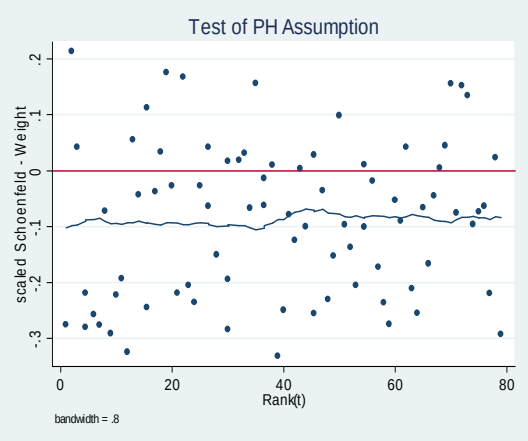

Kaplan-Meier survival estimates, by Functional
0.00
0.25
0.50
0.75
1.00
0
20
40
60
80
analysis time
Ambulatory
Bedridden
Working

Figure 2: (a – g): Graphs of the scaled Schoenfeld residuals and their LOESS smooth curves obtained from the model in Table 4(Appendix A) for the covariates weight, TB site, functional status, WHO clinical stage and CD4 cell count. The line that passes through zero is the reference line.

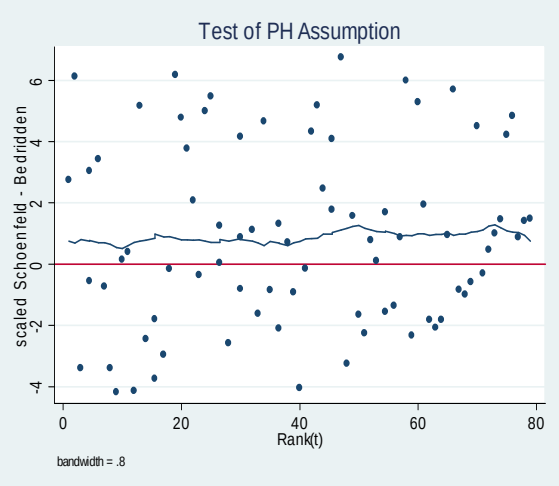
a)



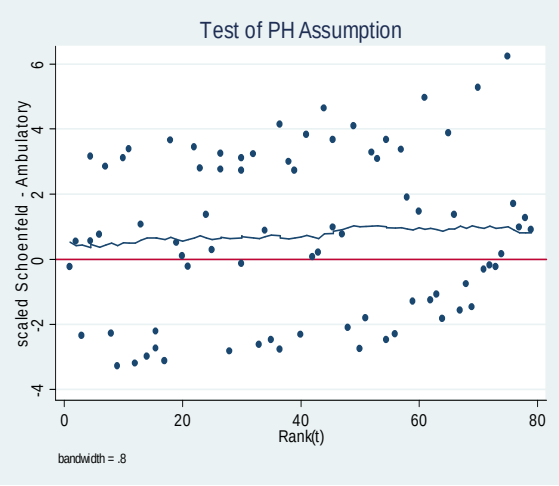
b)



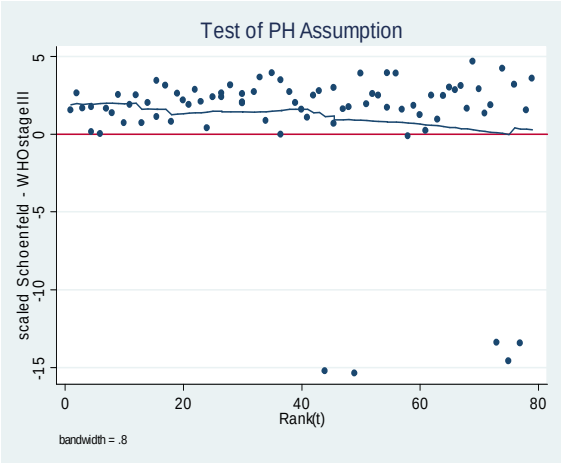
c)



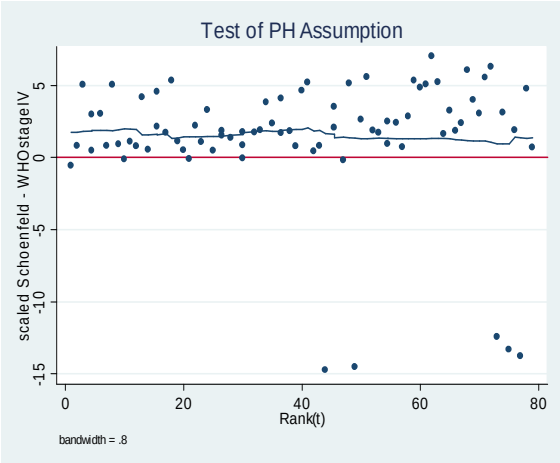
d)



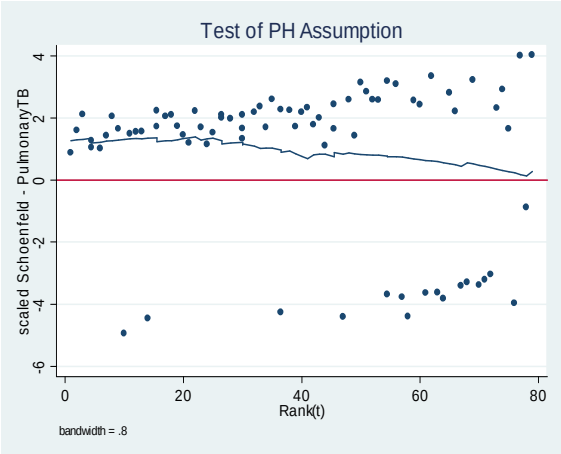
e)


Test of PH Assumption
scaled Schoenfeld - WHOStageIII
5
0
-5
-10
-15
0
20
40
60
80
Rank(t)
bandwidth = .8

f)


Test of PH Assumption
scaled Schoenfeld - WHOStageIV
5
0
-5
-10
-15
0
20
40
60
80
Rank(t)
bandwidth = .8

g)


Test of PH Assumption
scaled Schoenfeld - PulmonaryTB
4
2
0
-2
-4
-6
0
20
40
60
80
Rank(t)
bandwidth = .8