MODELLING THE SURVIVAL TIME AMONG ADULT HIV/AIDS PATIENTS UNDER ANTIRETROVIRAL THERAPY IN MOI TEACHING AND REFERRAL HOSPITAL, KENYA

BY

MENGICH KIBICHII ROBERT

A THESIS SUBMITTED TO THE SCHOOL OF SCIENCE AND AEROSPACE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER OF SCIENCE IN BIOSTATISTICS

MOI UNIVERSITY

2021

DECLARATION

Declaration by Candidate

This thesis is my original work and has never been presented for a degree in any other university. No part of this thesis may be reproduced without the prior permission of the author and/or Moi University.

Signature: …… Date……..

Mengich Kibichii Robert

MSC/BS/01/16

Declaration by the Supervisors

This thesis has been submitted for examination with our approval as university supervisors.

Signature: …… Date……..

Prof. Ann Mwangi
Department of Mathematics, Physics and Computing,
School of Science and Aerospace Studies
Moi University,
ELDORET- KENYA.

Signature: … …… Date……..

Dr. Gregory Kerich
Department of Mathematics, Physics and Computing
School of Science and Aerospace Studies
Moi University,
ELDORET-KENYA.

DEDICATION

I dedicate this thesis to my late father Mr. John Kimengich Sabul, mother Mrs. Ludia Jepkering Sabul and loving sister chief inspector Maurine Jepkemboi Kutto for her financial support during the study.

Similarly, I dedicate it to my loving wife Debra J. Ronoh for her immeasurable moral support and motivation during the write up of this work.

ACKNOWLEDGMENT

I am grateful to my supervisors, Prof. Ann Mwangi and Dr. Gregory Kerich for their valuable, friendly, unlimited and meticulous guidance and supervision during this study. My heartfelt gratitude goes to Dr. Julius Koech (Lecturer University of Eldoret, Kenya) for his dedicative guidance and expertise in analysis and interpretation of medical data.

My sincere gratitude goes to the entire staff at Moi Teaching and Referral Hospital (MTRH) led by the Chief Executive Officer (CEO) Dr. Wilson Aruasa and Institutional Research and Ethics Committee (IREC) Moi University for approval of my study (refer to letters of approval in appendix 1 and 2). I am also glad to Edwin Sang, Eunice and Sarah for their major role in the access of study data.

I appreciate all the lecturers who have imparted knowledge to me this far and more so those who taught me in postgraduate studies.

I thank my mother and all my siblings for the love and financial support this far. I am greatly indebted to my postgraduate colleagues for their moral support.

In conclusion, I wish to express my whole hearted appreciation to almighty God for granting me good health, strength, and courage to carry on with this study regardless of triumphs and tribulations along my way.

ABSTRACT

Survival modelling is technique which exploits repeated measures of continuous covariates to predict explanatory variables' effects on the response factor. The survival modelling helps design interventions in the health sector, which has seen one of its applications in the management of Human Immune Virus/ Acquired Immune Deficiency Syndrome (HIV/AIDS). However, despite improvement in Anti-Retroviral Therapy (ART) interventions over the years, the observed disease effects (morbidity, progression and mortality) remain high and vary across geographical borders. The general objective of the study was to model the survival time among adult HIV/AIDS patients' under ART in Moi Teaching and Referral Hospital (MTRH) Kenya. Specifically, the study aimed at determining the mean and median survival times among this cohort of patients, fit Cox proportional hazard regression model to adult HIV/AIDS patients data and determine predictors of their survival. A retrospective study design was adopted where a target population of 10,038 patients who were on ART and were enrolled between January 2005 and January 2007 were investigated for a ten years follow-up period. Survival and hazard functions were used to determine the mean and median survival times. Kaplan Meier estimator was used to measure the overall survival trend and compare the survival time by gender. The Cox proportional hazard regression model (CPHRM) was fitted to the data using log partial likelihood function. The log rank test and 95% confidence Interval (C.I) were used to analyze the significance of the hazard ratios of each variable. The sample was constituted of 2,985(29.7%) male and 7,053(70.3%) females. From the data, censored patients consisted of 9,833(98%), while 205(2%) died. The results showed that the median and mean survival times after ART initiation were 10.00 and 9.06 years, respectively. HIV severity with Unadjusted Hazard Ratio (UHR) [UHR =0.729, p=0.032], level of education [lower UHR=0.952, p= 0.019], and ART perfect adherence [UHR=0.668, p=0.004] positively influenced patient survival time. Patient's gender [male UHR= 1.633, p< 0.001] showed negative effect on patient survival time. None of the patient's covariates were jointly significant predictor of survival time in the multivariate Cox model. However, adjusting for other factors in the model, HIV severity with Adjusted Hazard Ratio (AHR) [AHR1.18, p=0.735], underweight measured by Body Mass Index (BMI) <18.5kg/m^2 in reference to 18.5-<25kg/m^2 [AHR=1.65, p=0.847] patients' male gender [AHR=1.884, p=0.19] and ART perfect adherence at disease latter stage [AHR=1.393,p=0.498] increases the risk of mortality by 18%, 65%,88.4% and 39.3% respectively. In conclusion, ART perfect adherence enhance longer survival time in MTRH, The CPHRM fitted well to study data hence described the data optimally. HIV severity, gender, level of education and ART adherence were independent significant predictors of survival time whereas age and BMI were not. The study recommends the initiation of ART when CD4 count is at least 350mm^3, male patients should do compulsory regular screening to avoid late diagnosis and delayed presentation for ART medication and MTRH should do sensitization on the importance of perfect adherence at early stages of the disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AHR	Adjusted Hazard Ratio
AIDS	Acquired Immune Deficiency Syndrome
AMPATH	Academic Model Providing Access to Healthcare
ART	Anti-Retroviral Therapy
BMI	Body Mass Index
CD4	Cluster of Differentiation 4
CEO	Chief Executive Officer
C.I	Confidence Interval
CPHRM	Cox Proportional Hazard Regression Model
HAART	Highly Active Anti-Retroviral Therapy
HIV	Human immune Virus
IREC	Institutional Research and Ethics committee of Moi University.
LTFU	Loss to Follow Up
MIS	Management information system
MTRH	Moi Teaching and Referral Hospital
NACCK	National Aids Control Council Kenya.
NASCOP	National AIDS & STI control Programme.
PHM	Proportional Hazard Model
PLWHA	People Living With HIV/AIDS.
SAS	Statistical Analysis system
UHR	Unadjusted Hazard Ratio.
UNAIDS	United Nations for International Development
W H O	World Health Organization

DEFINITION OF TERMS

Baseline covariate: are variables whose values does not change over time for example gender.

Baseline: This is the time in which the patients are enrolled in ART for the first time and the required variable are captured i.e., the start of the follow up of individual patient where patient's covariates are measured from enrolment before he or she is initiated to Antiretroviral drugs.

Censored observation: An observation is said to be censored when the information about their survival time is unknown, i.e., occurs when individuals does not experience an event of interest but have a follow-up time that is less than the maximum possible.

Censoring: is a form of missing data, we do not know exact survival time.

Hazard function: Measures the probability that a patient fails at the time (***t***) conditioned on the fact that he/she survives to that time, denoted by $\boldsymbol{\lambda}(t)$.

Highly active antiretroviral therapy: the use of multiple antiretroviral drugs to act in a different viral target hence inhibiting viral replication hence reducing mortality and morbidity among HIV infected patients.

Median survival time: is that point in time from the time of inclusion when cumulative survival drops below 50%.

Missing data: Data that are intended to be observed, but are undetected for some reason.

Repeated measures: Are measurements belonging to the same patient but performed at different times. These repeated measures are normally done to continuous explanatory variables.

Right censored: occurs when a patient's follow-up time ends before the results of interest is detected.

Survival analysis: The study of time to event, where the outcome variable is the time until the occurrence of an event of interest which can either be death or censored.

Survival function: Is a measure of a chance that an individual patient survives beyond some specified period of time, denoted **S (t).**

Survival time: Is a measure of time until an event of interest occurs, for this study was the time between initiation of Antiretroviral Therapy (ART) and experiencing the event of interest which was either death or censored among HIV/AIDS patients on ART.

Survival: It is the absence of experience of an event of interest.

The loss to follow up: A patient who has not made contact with a clinician for three or more months since his/her last scheduled appointment and his or her where about is not known.

CHAPTER ONE

INTRODUCTION

1.1 Overview

This chapter begins by giving a brief world, regional and Kenyan HIV/AIDS history and its adverse effects and reviews survival models. Finally, it provides a statement of the problem, Justification, purpose of the study, research objectives, and the significance of the study, research questions, and limitations of the study**.**

1.2 HIV/AIDS History

1.2.1 The World and Regional History on HIV/AIDS.

The joint United Nations for International Development (UNAIDS,2019) program report on Human Immune Virus/ Acquired Immune Deficiency Syndrome (HIV/AIDS) indicates this pandemic as one of the most destructive disease in the history of humankind that has claimed the lives of infected patients.

Since the first detection of HIV/AIDS in the world to date, it is estimated that 78 million people have contracted the disease and out of these more than 35 million patients have died due to AIDS and its related illness by 2016 (UNAIDS, 2019). The World Health Organization (WHO) report on HIV indicates that among the estimated 36.7 million patients living with HIV, 1.8 million were children. However, the findings showed that globally, HIV/AIDS prevalence was 0.8% among adults with 30% of this population lacking knowledge on their disease status.

The most affected population were people living in third world and developing countries, with a population close to 25.5 million living in sub-Saharan Africa and amongst these 19.4 million were living in East and south Africa, (UNAIDS, 2019).The demographic characteristics indicates that 59% of new infections affected

young people aged between 15 and 24 years. By 2016, more than half of People Living With HIV/AIDS (PLWHA) i.e. (53%) had access to life-saving treatment equal to 19.5 million up from 17 million in June 2015 and 7.5 million in 2010 (UNAIDS, 2019).

1.2.2 HIV and AIDS in Kenya

In Kenya the first HIV/AIDS case was detected in 1984. Since then, it has become one of the most significant causes of death, putting vast demands on the healthcare system and economy at large.

HIV/AIDS prevalence in Kenya has a declining trend from peak of 10.5% in 1996 to 5.9% in 2015 and 5.6% in 2016. The observation is mainly due to improved HIV/AIDS therapy and care National AIDS and STI Control Program (NASCOP, 2016).

In Kenya HIV is a generalized epidemic in the sense that it affects each segment of the population structure from toddlers to old adults. By 2015, AIDS had orphaned 660,000 children. The (NASCOP, 2016) report, estimated that, 1.6 million people were living with AIDS in Kenya with 36,000 fatalities reported from AIDS defining illnesses in 2016.

An estimated 65% of all new HIV infections occurs in nine counties predominantly in the west and coast of Kenya (NASCOP, 2016).The Anti-Retroviral Therapy (ART) scale-up services in Kenya has approximately one million HIV/AIDS patients out of 1.6 million infected people (NASCOP,2016).

The ever rising global and local statistics on prevalence and intensity of destruction of human lives through this pandemic as witnessed above, necessitate mathematical and

statistical modelling so as to establish the underlying factors and conditions that need to be tamed to minimize the effects.

1.3 Survival Models

Binary outcomes are common in medical research, where "success "may indicate that the patient is alive after treatment while "failure" shows that a patient died. Survival analysis is interested in the statistical study of such occurrence (time until event) in a group of individuals. The follow-up period of these individual patients occurs within a definite period with attention on the time in which the outcome occurs, known as failure time, survival time, or event time. The event time can be measured in hours, days, weeks, months, quarters, semiannually, annually, or years. Examples include death occurrence or re-occurrence of the disease, marriage, and divorce.

Survival models are essential in research because of their unique attributes. The factors of interest comprise of survival time, defined by either time to death or censoring, and the presence of explanatory variables. The existence of partial data regarding the survival time of some individuals complicates the analysis of survival data. This partial data was majorly as result of the right censoring. Therefore, this study tried to find out the proportions of HIV/AIDS patients who survived past the study period, the hazard rate among the survivors who surpassed the study time and effects of specific variable on the probability of survival.

In order to estimate model parameters, survival models correctly incorporate information from both uncensored and censored data. Modeling strategies for survival data falls into three categories comprising of nonparametric, parametric, and semi-parametric analysis (Mecha, Kubo, Nganga, Muiruri, Njagi, Mutisya, Odionyi, Ilovi, Wambui, Githu, Ngethe,Obimbo & Ngumi, 2016). The extent of parametric

assumption reliance influences the classifications of these models. When the true distribution is not known or difficult to approximate, non-parametric models are used which does not require assumptions about survival. These include Life tables and Kaplan Meier estimation.

When the interest is in the association between time to event and covariates, parametric and semi-parametric models are applicable. For parametric models, all the covariates in the model require a specification and full characterization of the hazard function and the Accelerated Failure Time (AFC) model is an example. The semi-parametric model, which contains both parametric and non-parametric components, is more flexible and make less assumption about the distribution of survival data. In addition to these attributes, the model still allows for estimation of relative hazard between covariates to evaluate the effects of explanatory variables with or without specification of the baseline distribution. An example is the Cox Proportional Hazard model (Cox, 1972).

The outcome of interest is composed of two important parts; the time until an event occurred and the event status. These guides in the estimation of survival probabilities and hazard ratios which are essential in describing the distribution of survival time and effects of specific variable on the survival of the patients. The survival function gives the probability of not experiencing the event and the hazard function measures potential risk of death given that an individual patient survives beyond the study period.

In regards to censoring, this research considered right censoring defined as the censoring, which occurs when an individual patient survives beyond the study time. The underlying principle behind Cox proportional hazard regression model is that

two or more individual patients have constant proportional hazard ratios which does not depend on time (Piulachs, Lozada-Benavente, Alemany & Guillén, 2017).

The covariates are classified either as fixed or time-dependent. The hazard ratio depends only on baseline covariates such as age and sex. However, when it is necessary to examine if continuous covariates contributes to the risk of death the extended Cox model is useful. (Andrinopoulou, Rizopoulos, Takkenberg & Lesaffre, 2014).

In this study, the focus was on the Cox Proportion Hazard model (Cox, 1972). The model considers the effects of numerous explanatory variables at a given time and investigates the association of survival distribution with the variables considered. Hazard function is the dependent variable at given time (t).

In order to approximate Parameter values of the Cox Proportional Hazard Regression Model (CPHRM), the partial likelihood weights are maximized. In this model the log hazard of an individual patient is a linear function of their static covariates and population -level baseline hazard function that varies over time.

Thus, the adopted CPHRM was given by: -

$$\boldsymbol{\lambda(t/x) = \beta_o(t) \exp\left(\sum_{i=1}^{n} \beta_i(x_i)\right)} \qquad (1.1)$$

Where $\beta_o(t)$ is the baseline hazard function and $\boldsymbol{\beta_i}$ is the regression coefficient of the corresponding $\boldsymbol{xi}$.

The mathematical sign that precedes the regression coefficients'$\boldsymbol{\beta_i}$, is significant in drawing inference in the sense that a positive sign for $\boldsymbol{\beta_i}$ means the risk of occurrence

of an event was higher. Conversely, a negative sign means that the risk of occurrence of an event (for this case death) was lower.

The magnitude of the regression coefficient is useful in making inference and drawing conclusions in that a value of the hazard ratio equating to one (1.00) has no effect on the risk of failure whereas less than one (1.00) decreases the hazard and greater than one (1.00) increases the failure rate.

The assumptions considered in CPHRM applied in this study include: - multiplicative effects of the covariates on the hazard ratio, the hazard ratios of two subjects' remains the same at all time and the end times (death) of the patients are independent of each other.

For successful implementation of ART care among HIV/AIDS patients, there is need for such quantitative statistical analysis in appropriate settings to be able establish determinants of survival time among AIDS patients on antiretroviral therapy in MTRH, Kenya.

1.4 Statement of the Problem

Regardless of advancement in (ART) over time, the observed disease effects persist and differs across geographical borders. Therefore, there was need to study determinants of survival time among AIDS patients on ART in Kenya to permit closer follow up and expedite targeted intervention of patients on higher risk so as to reduce mortality.

HIV/AIDS has ravaged every part of the Kenyan economy ranging from agriculture, education, health, business, and industrial sectors. Thus, to decrease mortalities associated to HIV/AIDS, Kenya has witnessed rapid expansion of free ART over the

past decade, but there inadequate information about treatment outcome and determinants of survival time among PLWHA on ART. Because most studies countrywide focused on prevention and factors that increase the chances of infection with HIV/AIDS. However, only a few studies have focused on the determinant factors of survival time, which are unique to patients enrolled in ART in Moi Teaching and Referral Hospital (MTRH) Kenya. This study addressed these issues and will thus contributes to the strengthening implementation of ART.

1.5 Justification of the Study

Despite improvement of ART services globally, there has been continuous higher level of destruction on human lives by the pandemic through deaths. This necessitates quantitative statistical modelling in appropriate setting between the variables of interest for this case survival time and explanatory variables to be able to unearth the risk factors of survival time other than strict adherence to anti-retroviral regiments in MTRH. These determining factors seem to vary across the geographical borders as postulated by earlier studies done outside our country, therefore, it was appropriate to establish these covariates in MTRH Kenya. In addition, unlike this research, few studies done within this geographical border considered shorter follow up period of not more than five years and a little sample size of less than two thousand HIV infected patients data.

1.6 Purpose of the Study

This study applied survival models, to identify the prognostic factors of survival time among adult HIV infected patients on ART in MTRH, Kenya. The aim was to advice on the need to control this covariates so as to optimize care through targeted and timely intervention. This would result to reduction of disease morbidity, progression

and mortality within therapy thus improves quality and length of life among these group.

1.7 Research Objectives

1.7.1 General Objective

To model the survival time among adult HIV/AIDS patients' under ART in Moi Teaching and Referral Hospital (MTRH) Kenya

1.7.2 Specific Objectives

The specific objectives were to:

1. Determine the mean and median survival times of HIV/AIDS patients on ART in MTRH.
2. Fit Cox proportional hazard regression model to adult HIV/AIDS patients' data.
3. Investigate survival time predictors for adult HIV/AIDS patients on ART in MTRH.

1.8 Significance of the Study

The impact of the (ART) program on survival time requires consistent evaluation. Successful implementation of such programs needs scientific evidence, well-studied research, and routine hospital data. Therefore, the findings from this research helps augment existing strategies in strengthening ART scale-up services.

Since this study was carried out in MTRH which is the second largest hospital in Kenya, the findings provided sufficient statistical information, specifically significant predictors of survival time to PLWHA, clinicians, ART facilitators, researchers and other interested bodies hence will serve as input for other hospitals.

The outcome of this study would further create a general awareness of the risk factors influencing HIV/AIDS patient's life span after being enrolled on ART. Hence, this would form a basis for health education for the infected patients and ART facilitators as well as an input for ART monitoring and evaluation assessment tool.

1.9 Delimitation of the Study

The study considered only PLWHA enrolled on antiretroviral therapy in (MTRH) between January 2005 and January 2007 and followed up for ten years. All Considered patients were adults of age eighteen years and above during initiation of ART regardless of the treatment category.

This study excluded patients referred from other health institutions after the initiation of free ART due incomplete data records or missing data.

1.10 Limitations to the Study

The limitations included use of data from only one hospital consequently; the findings should be interpreted carefully if anyone intents to draw inference at the national level. In addition, there was insufficient published literatures in Kenya related to the study. Finally, loss to follow up resulted to missing the data for some patients who dropped from the study.

CHAPTER TWO

LITERATURE REVIEW

2.1 General Overview of ART

The anti-retroviral medication inhibits the reproduction of HIV in the body of an infected patient. The main effect of anti-retroviral therapy is to suppress viral replication allowing the immune system to recover and protect the infected patient from developing AIDS and potential death. In terms of mortality prevention and enhanced quality of life, the clinical advantage of ART for AIDS patients indicates geographical differences with higher mortality cases in developing countries.(Nischang, Sutmuller, Gers-Huber, Audigé, Li, Rochat & Amssoms, 2012).

2.2 Survival Time

The average survival time for HIV-positive patients was 65.22 months after initiation of antiretroviral therapy (Abebe et al., 2014). Comparable studies done in Ethiopia obtained almost identical mean results of 63.7 months (Tsehaineh, 2010) and 77 months (Kebebew & Wencheko, 2012).However, according to Tsegaye and Worku,(2011), their study established a mean survival time of 43 months. The discrepancy in results (43 vs. 77) was attributed to the fact that the latter study considered relatively short study period.

Guerreiro et al.,(2002) investigated 974 HIV positive adult patients in Zambia. The results indicated that the median survival time for patients with baseline CD4count $\geq$500cell/mm^3 was 5.62 years, with CD4 count between 200 cells and 499 cells/mm^3 was 5.46 years and CD4 count <200cell/mm^3 was 3.89 years.

The observed variations in the mean and median survival times among HIV positive patients on ART across geographical divide as postulated by earlier studies

necessitate this research in order to establish on average how long can a patient survives while on ART and what time in point when cumulative survival drops below 50% in MTRH ART center.

2.3 Determinant Factors of Survival in HIV/AIDS Patients.

The notion that all demographic, socio-economic, health and risky behavioral factors may have significant relationships with survival of HIV/AIDS patients was supported by many researchers (Alomepe, 2011).

Rai, Mahapatra, Sircar, Raj, Venkatesh, Shaukat and Rewari (2013) examined the survival probability of AIDS patients using socio-demographic factors. Age was significant predictor of HIV/AID patient survival time according to their study. Similarly, Alemu, and Sebastián, (2010) in Ethiopia and Choi, Choi, Han, Kim, Kee, Kim, & Kim (2018) in Korea pointed out that gender and age of the patients were significant predictors for development of AIDS. According to this Korean study, relative to older group, younger patients had the benefit of living longer. However, the fatality rate among the male patients was two folds compared to female counterparts.

Among people who were HIV infected, insecure housing had an association to higher morbidity, mortality, low hospitalization, and low adherence to drugs (Stewart, Chan, Carusone, To, Schaefer-McDaniel, Halman, & Grimes, 2012).

In the U.S.A, patients' education status was the primary predictor of survival. According to Lee, M., & Rotheram,Borus, M. J. (2001) in Ceara state in Brazil 0.6% of HIV/AIDS patients joined university, while 1.7% joined the university in capital Fortaleza. According to these studies, students on the first year of their university education exhibited lower risk of mortality compared to those in other university

academic years. These findings indicated a positive relationship between education and longer survival among HIV/AIDS.

Opportunistic infection are common causes of fatalities in AIDS patients. The role of antiretroviral treatments is to lower the progression of HIV to AIDS among the infected patients on ART care. However, substantial HIV/AIDS morbidities and mortalities are still witnessed despite their use of antiretroviral drugs (Ciaranello, Myer, Kelly, Christensen, Daskilewicz, Doherty & Wools-Kaloustian, 2015). Pulmonary tuberculosis is one of the most common AIDS opportunistic infection affecting 30.9% of HIV diagnosed patients characterizing it as AIDS defining illness. Whereas disseminating tuberculosis related to pulmonary tuberculosis has an incidence rate of 6.4% among AIDS patients. Delayed HIV detection, treatment and concomitant existence of tuberculosis influences survival beyond the second year after infection (Zekan et al., 2008).

Croxford, Kitching, Desai, Kall, Edelstein, Skingsley and Delpech (2017), established that late diagnosis was the main predictor of death. The study attributed most deaths to AIDS-defining illness and liver diseases after conducting a study of 220 outpatients in public health in England.

Burkey, Weiser, Fehmie, Alamo-Talisuna, Sunday, Nannyunja and Chang (2014) did a study in urban Uganda which revealed that the level of education, dependents ratios, status of employment and standard wealth index indicated susceptibility to death in HIV positive patients, despite uniform access to ART.

Palombi, Mancinelli, Liotta, Narciso, & Marazzi (1997) conducted a study of 168 HIV/AIDS patients on ART in Rome. Results showed that patience functional status had an impact on longevity of HIV/AIDS patient life. Their results further implied

that bedridden and ambulatory patient's functional status would lead to decreased chances of survival.

Andrade, Shinotsuka, Ho, Carvalho, Bozza and Japiassú (2017), applied CPHRM on 421 HIV/AIDS patients' data. Results indicated that pre-admission health status, functional status and weight loss were highly correlated to the survival of PLWHA on antiretroviral drugs.

Just at the onset of AIDS, it has been established that moderate weight loss was highly correlated to increased risk of death before diagnosis of HIV and start of antiretroviral treatment. Hendricks, Dong, Tang, Ding, Spiegelman, Woods & Wanke (2003) did prospective cohort study among 678 patients living on ART in Boston between April, 1995 and August, 1997 on "nutrition for health living". The results established that weight loss was highly correlated to decreased chances of survival.

Chaisson *et al.*, (1995) used CPHRM to determine significant predictors of survival among a cohort of 1372 HIV patients treated in one urban Centre. The results showed that lower CD4 cell count and old age were correlated to decreased chances of survival. Whereas gender and stage of the disease had association.

Guerreiro *et al.*, (2002) assessed 486 adult patients' for 12 years using both parametric and non-parametric survival methods in order to ascertain determinants of survival among AIDS patients on ART. The results indicated no major difference between the male and female patient's median survival time. In addition, higher CD4 cell count and adherence to antiretroviral drug regiments had significant effects on the decreased risk of death across different age group divides of AIDS patients.

Survival analysis in South Africa (Mzileni, Longo-Mbenza & Chephe, 2008) and Petros, (2016) results pointed out that the estimated survivorship function across all levels of ARV drug adherence in AIDS patients was significantly different. *Jerene et al.,* (2006) at Abarmich Hospital established that WHO stage (IV), body mass index and loss of weight were significant predictors of survival in AIDS patients by using Cox regression.

Jerene *et al.*, (2006) analyzed predictors of mortality among 162 AIDS patients on Highly Active Antiretroviral Therapy (HAART) using CPHRM and product limit function. The outcome indicated that most deaths were realized on the first month of initiation of this type of therapy. Furthermore, this research established that WHO stage IV and loss of weight were significant predictors of death among this cohort.

Ibrahim, (2007) conducted a retrospective study among 259 AIDS subjects in Adama hospital (ART clinic). The study results indicated condom and alcohol use, CD4 cell count, weight and WHO stage IV were significant predictors of survival among this cohort of patients. A comparative study carried out in Tikur Ambesa hospital (Addis Ababa) revealed that function status, age, level of adherence to ART, CD4 cell count, gender and weight were statistically significantly related to survival of HIV/AIDS patients. Getahun, (2010), in Beherdar Feleg Hiwot hospital applied product limit function on patients' data, in order to estimate survival probabilities and CPHRM to determine factors of survival in adults' HIV/AIDS patients. Results revealed 13.6% deaths rate in five years follow up time.

Petros (2016) conducted a retrospective study on 756 HIV/AIDS patients in Dilla and Hawassa University Referral Hospitals (Ethiopia) and Cox regression output

showed that survival time was significantly related to age, condom use, marital status, functional status, alcohol, WHO stage, CD4, and ARV adherence.

Sungkanuparph *et al.,* (2007) did a retrospective cohort study among patients co-infected with HIV and tuberculosis in order to determine the effects of ART on the survival of this cohort. The survival analysis methods employed included log-rank test and CPHRM. Results demonstrated that ART considerably decreases death rate in this cohort of patients. The commencement of ART within the early six months of TB detection among AIDS patients was highly related to improved survival.

Sika, Wools-Kaloustian, Mwangi, Kimaiyo, Diero, Ayuo and Musick (2010) did comparative case study between November, 2001 and December 2005 where they considered demographic, laboratory and clinical data of 527 dead and 1054 living patients who had received ART in MTRH Kenya. This study results pointed out that patients with CD4 cell count $<100mm^3$ were at higher risk of death compared to those with CD4counts $>100mm^3$. Also, the researchers established that HIV/AIDS patients attending rural clinics had a three-fold risk of death compared to their counterparts attending tertiary referral hospitals clinics. ARV drug adherence, gender, WHO staging, and hemoglobin level were correlated to the survival of HIV/AIDS patients on ART in MTRH. Whereas age, marital status, education level, employment status, and weight had no influence on survival time. The adjusted model in this study indicated that patient's gender, TB infection status, WHO stages, alcohol use, CD4 cell count, employment status and weight were significant predictors of survival time in this cohorts of patients at a 5% significance level.

The above discussed literatures emphasize on the use of survival analysis methods in HIV/AIDS studies, in particular the use of Kaplan Meier estimator and Cox

proportion hazard regression models for identification of clinical and socio-demographic variables believed to affect survival time of patients with HIV/AIDS. However, A few studies and little literatures have been postulated on determinants of survival among adult HIV/AIDS patients on ART in Kenya. Therefore, this study will contribute immensely to the development of such literatures.

2.4 Conceptual Framework

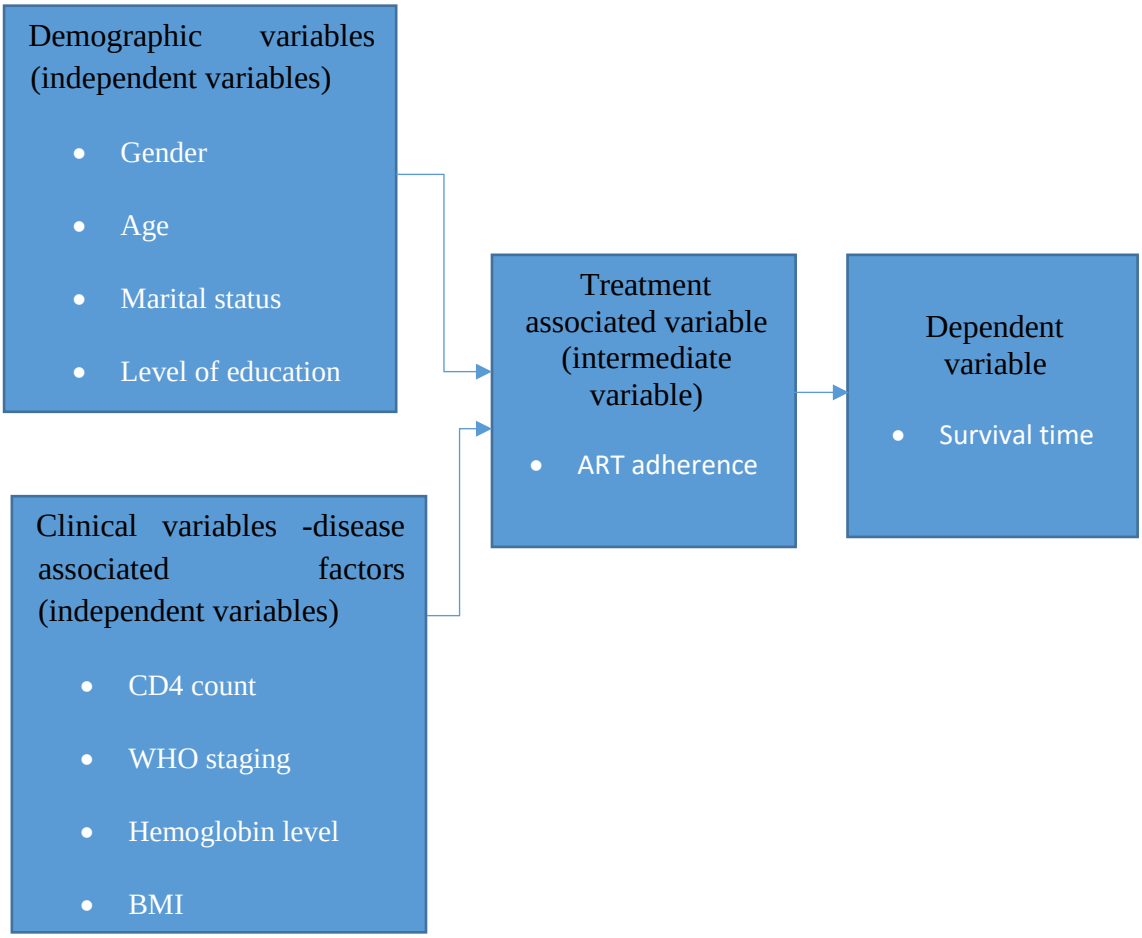

Figure 2.1: Conceptual Framework

Source: The current study

From figure 2.1 above, the explanatory/ independent variables for survival time among HIV/AIDS positive patients on ART was categorized into social & demographic variables which include: - gender, age, marital status, employment status and level of education and baseline clinical variables such as weight, CD4

count, WHO stage, hemoglobin level, and opportunistic infections. The social demographic and clinical factors interact with treatment-associated/ intermediate variable (ART adherence) to influence survival time which is the dependent variable (Bamulangeyo, 2018).

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This section explains the study design adopted, data collection instruments, survival analysis models use, ethical clearance and limitations of this study.

3.2 Study Design, Area and Population.

A longitudinal study design was adopted. The target population were HIV/AIDS patients who were 18 years and above during the start of ART in MTRH (Eldoret, Kenya). These patients were enrolled in January 2005 to January, 2007 and followed up for ten years. MTRH serves Rift Valley, Western and Nyanza regions of Kenya with Academic Model Providing Access to Healthcare (AMPATH) unit designated ART program.

Records based data from baseline and follow up visits of 10038 adult HIV/AIDS patients on ART were reviewed. The study sample consisted of 2985males and 7053 females.

In 2001, MTRH began to offer antiretroviral therapy service to PLWHA who were able to meet whole cost of treatment. However, later by the year 2005, the hospital began registering and offering free ART services for those who made the standard diagnostic requirements to begin care in line with national treatment guidelines.

3.3 Data Collection Instruments

The required data set was extracted from an existing research database at MTRH. The data obtained was de-identified before sharing to the researcher to enhance the study subject's privacy and confidentiality. The research data were retrospectively collected

from MTRH antiretroviral database management information system, in AMPATH within MTRH, Eldoret Kenya.

Therefore, patients attending AMPATH programme had a unique ART number and this was a requirement for each patient who voluntary wished to start ART according to the national treatment guidelines (NASCOP, 2016).

3.4 Study Variables

The variables which were considered in this research were classified into dependent and independent variables

3.4.1 Dependent variable

The dependent variable was survival time of HIV/ AIDS patients'

3.4.2 Independent Variables

The predictor variables and their levels were recorded as follows: -

Table 3.1: Independent variables considered and the levels

Variable	Descriptions	Coding	Data type
Gender	The gender of the respondents	0-Female 1-Male	Binary
Age group (Years)	This describes the age group of the respondents	0-18-29 1-30-39 2 ≥40	Ordinal
Marital status	The marital status of the patients	0-Never married 1-Married 2-Divorced/Separated 3-Widowed	Nominal
Education Status	The education of the patients under study	0-Primary 1-Secondary - Tertiary	Ordinal
Employment status	The employment status of the respondents	0-None 1-Paid employee 2-Peasant 3-Other	Ordinal
Body Mass Index (BMI)	The ratio of body weight to the square of patient's height of the respondents in kg/m^2	0-< 18.5-Under weight 1->18.5 - < 25-Normal weight 2-25 - < 30 Over weight 3->30- Obesity	Ordinal
WHO stage	The WHO staging level of the respondents	0-Stage I 1-Stage II 1-Stage III 2-Stage IV	Ordinal
Drug Adherence	Drug adherence status of the patients	0-perfect 1-non- perfect	Binary
House holds	The number of people in the household	0= 1-2 1= 3-4 2> 4	Ordinal
CD4 count (HIV severity)	The Cell Differential four count of the patients	0-<350 1>=350	Binary
Opportunistic infections	Patient having opportunistic infections	0- yes 1- no	Binary
Functional status	The body functional status of the patient	0-Bedridden 1-Ambulatory working	Nominal
Drug abuse	The use of alcohol and other drugs abuse	0-yes 1-no	Binary

3.5 Data Analysis Techniques

Data cleaning was done using Statistical Analysis System (SAS) version 13 applying iteration SAS codes in (Appendix 3) with the aim of removing outliers hence remaining with the desired patients' characteristics before being analyzed. Data analysis was carried out at two levels using R- software version 4.0.2. First, using iteration R codes in (Appendix 4), a descriptive baseline summary of the HIV/AIDS patient's demographic and clinical characteristics was made using frequency distributions and summary statistics where appropriate. Secondly, an analysis of disparities in survival times of demographic and clinical characteristics was done using the log rank for both unadjusted and adjusted Cox models. The purpose of the analysis was to establish predictor variables of survival in HIV/AIDS patients on antiretroviral treatment.

3.6 Survival Function and Hazard Function

3.6.1 To Determine Mean, Median Survival Times.

Let **T** be a non-negative valued random variable for failure time (survival time, lifetime). The meaning of **T** is the time from the start of therapy to the occurrence of an event i.e. death or censoring.

De Castro, Cancho and Rodrigues, (2010) defined survival function as a measure of the likelihood that an individual patient survives longer than some defined time (***t***) denoted by ***S* (*t*)**. The survival function ***S* (*t*)**, measures the chances of **T** surpassing the defined time ***t***. Suppose that the random variable **T** has a cumulative distribution function ***F (t)*** with underlying probability density function ***f (t)*** then according to De Castro et al., (2010) survival function ***S*** (***t***) is therefore given by:

$$S(t) = \Pr(T \geq t) = 1 - \Pr(T < t) \qquad (3.1)$$

The cumulative distribution function ***F (t)*** is given as *F (t)* =Pr (T≤ *t*), while the survival function is typically derived using the following equations: -

$S(t) = \Pr(T>t) = \int_{t=10}^{\infty} f(u)du$

$S(t) = 1 - \int_{t=10}^{\infty} f(u)du$

$S(t) = 1 - F(t)$ (3.2)

Properties of survival function *S (t)* are:

- $S(t)=1$ if $t \leq 0$, $S(\infty) = \lim_{t\to\infty} S(t) = 0$
- *S(t)* is non-increasing in *t*
- *S(t)* is left continuous

The hazard function denoted by $\lambda(t)$ measures the instantaneous failure rate, provided that a patient has survived up to time *t.* The hazard function is also referred to as conditional failure rate or simply hazard rate and is defined as the likelihood of a patient failing at time *t*, subject to the fact that he or she has survived to that time.

From De Castro et al., (2010), for λ (t) ≥0, the hazard function λ (t) is given by the following:

$$\lambda(t) = \Pr(T = t | T \geq t = \frac{\Pr(T=t)}{\Pr(T \geq t)} = \frac{f(t)}{s(t)} \quad (3.3)$$

Let Λ (·) be the cumulative hazard function.

$$\Lambda(t) = \int_{t=0}^{t=10} \lambda(u)du \quad (3.4)$$

The survival function in terms of the hazard function is given by

$$\Lambda(t) = e^{-\int_{t=0}^{t=10} \lambda(u)du} = e^{-\Lambda(t)} \quad (3.5)$$

The probability density function (pdf) in terms of the hazard function can also be derived

$$f(t) = \lambda(t) \exp \int_0^t \lambda(u)du \quad (3.6)$$

The average survival time is given as:

$$E(T) = \mu = \int_0^\infty tf(t)dt = \int_0^\infty S(t)dt \quad (3.7)$$

While the median survival time is estimated by the following equation (3.8)

$$S(t_{median}) = 0.5,\ t_{median} = S^{-1}(0.5) \quad (3.8)$$

3.7 Non-Parametric Estimator of Survival Function

3.7.1 Kaplan-Meier Estimator

The Kaplan-Meier estimator is a non-parametric or distribution-free survival model because there is no particular assumption to be made about the underlying distribution of the survival times. This research applied both random right censoring and independent censoring.

The Kaplan-Meier (KM) estimator, also referred to as the product-limit estimator, integrates data from all available observations, both uncensored and censored, by considering any point in time as a series of steps identified by the survival and censored times observed. On the other hand, the estimator is essentially the sample proportion of observations with event times greater than *t.* where there is no censoring.

Suppose we have a sample of n independent observations with their survival times denoted by t_1, t_2,…,t_n and indicator of censoring by δ_1, δ_2, …, δ_n where δ_i = 1, event/death and where δ_i= 0 for censored observations. The survival data are denoted by (t_i, δ_i), i= 1, 2… n. The first step in obtaining the Kaplan-Meier estimator is to order the survival times as $t_{(1)} < t_{(2)} \ldots <$ t (n) respectively. Subsequently, suppose that among the *n* observations, there are $m \leq n$ failures that occurred at distinct *m* times. Thus, the parameters used to estimate the product limit are given as follows,

$d_{(j)}$ = failure or death numbers at time t (j)

N (j) = number of individuals at risk at t (j)

The product limit estimator of survival function at time t can thus be derived as: -

$$\hat{S}_{KM}(t) = (1- \frac{d_{(1)}}{N_{(1)}}) (1- \frac{d_{(2)}}{N_{(2)}}) \ldots (1- \frac{d_{(j-1)}}{N_{(j-1)}})$$

$$= \prod_{t(j)<t}^{k} (1 - \frac{d_{(j)}}{N(j)}) \text{ for } t_{(k)} \leq t < t_{(k+1)}, \text{ k =1, 2, 3, … , m} \qquad (3.9)$$

According to Kaplan-Meier (1958).

If t = 0, S (0) = 1, this means that at time 0 all the subjects are alive.

In order to estimate the variance of the survival function using Kaplan-Meier estimator the greenwood formula applied, given by:

$$\hat{\text{var}}(\text{S (t)} = \left[\hat{s(t)} \right]^2 \sum_{t_{(i)} \leq t} \frac{di}{ni(ni - di)} \qquad (3.10)$$

The (1-α) *95% confidence interval for the survival function (S_{KM} (t)) at time t is:

$$\hat{S}_{KM}(t) \pm z_{\alpha/2} \text{S.e}\,(\hat{S}_{km}(t)) \quad (3.11)$$

The survival function only changes at failure time. This implies that decreasing phase function takes a leap at the time of failure and the conditional probability of survival is all the time 1.00. As a result, the chance of living remains unchanged from the previous survival probability.

3.8 The Cox Proportional Hazard Model

3.8.1 To determine the survival time predictors for adult HIV/AIDS patients on ART

The CPHRM is the basic regression model applied on survival data. This model was used to investigate the determinant factors of survival time in HIV/AIDS patients on ART in MTRH. This model was introduced by Cox (1972). The CPHRM is multiple regression approach to analyze the connection between survival times and one or more predictor variables or covariates. These predictor variables may dependent or independent of time.

The Cox proportional hazard regression model is function of a set of explanatory variables ($\underline{x}$) and baseline hazard function. This model characterizes how the hazard function changes as a function of survival time and r {(X', β), β} as a function of subject covariates.

The data structure was in the following form: **{(**t_1, δ_1, x_1**) …, (**t_n, δ_n, x_n**)}**

Where: t_i is the observed survival time for ith individual.

δ_i is an indicator of censoring.

x_i explanatory variables.

Let $\underline{X'}$= ($x_1, x_2, ...x_p$) and $\underline{\beta}$ = ($\beta_1, \beta_2, \ldots, \beta_p$)

For r ($\underline{X'}$,$\underline{\beta}$) = exp ($\underline{X'}$,$\underline{\beta}$) then the hazard ratio is:

$$\lambda \text{ (t, } \underline{X'}, \underline{\beta}) = \lambda_0(t)\exp(\underline{X'},\underline{\beta}) \quad (3.12)$$

The model in equation (3.11) is known as the Cox Proportional Hazard Model.

Where: $\underline{X'}$ are P x 1 vector of predictor variables.

$\underline{\beta}$ Is a vector of 1 x P vectors of parameters.

Where X_i's is the vector of values of the explanatory variables for the i[th] individual at time t and $\underline{\beta}$ is the vector of unknown regression parameters that are assumed to be the same for all individuals in the study, which measures the influence of the covariate on the survival experience. The Cox model formula has the property that if X'**s** is entirely equal to zero, the formula reduces to the baseline hazard function. This property of the Cox model is the reason why $\lambda_0\ (t)$it is called the baseline function.

From De Castro et al, (2010), it can be noted that the cumulative hazard function at time t for a subject is given by:

$$\Lambda(t,x,\beta) = \int_0^t \lambda(u,x,\beta)du = r\ (X',\beta)\int_0^t \lambda_0\ (u)du = r\ (\ X', \beta, \Lambda_0(\text{t}) \quad (3.13)$$

The survival function for the semi parametric hazard model is

$$\text{S (t,}X', \beta) = \left[S_0(T)\exp(\text{X}', \beta)\right] \quad (3.14)$$

And in the equation, $S_0(T)$ is the baseline survival function.

The CPHRM is widely and commonly used on survival data and can handle covariate of interest that are both time-dependent and independent. R statistical software was used to analyze patients' data.

3.8.2 Fitting the Cox Proportional Hazard Regression Model

To obtain the maximum likelihood (ML) estimates of the CPHRM parameters the likelihood function (L) is maximized. The likelihood function is a mathematical expression which describes the joint probability of obtaining the data actually observed on the subjects in the study as a function of the unknown parameters ($\underline{\beta}$'s) in the model being considered. The L is sometimes written notational as L ($\underline{\beta}$) where $\underline{\beta}$ denotes the collection of unknown parameters.

Suppose the survival data is represented by (t_i, δ_i, X_i) for i=1, 2… n where t_i the length of time a subject is observed (survival time), δ_i an indicator of censoring for the ith individual and $X_i{}'$ a vector of covariates for the i^{th} individual. The likelihood for right-censored data includes both the survival and hazard functions given by equation (3.15)

$$L(\underline{\beta}) = \prod_{i=1}^{m} [\lambda, X_i, \beta)^{\delta i}] \; S(t, X_i, \beta) \qquad (3.15)$$

The proposed partial likelihood function avoids specification of the baseline hazard function, treating it as a parameter of nuisance and excluding it from the equation of estimation due to the its assumption that there are no tied values among the observed survival times. Suppose we have m distinct failure times and let $X_{(i)}$ is the vector of covariates at ordered failure time t (i). We define the Partial Likelihood by the following expression:

$$L_P(\beta) = \prod_i^m \left[\frac{e^{x_i,\beta}}{\sum_{j \in Rt_{(i)}} e^{X_i\beta}} \right]^{d_i} \quad (3.16)$$

Where d_i is the number of deaths, we assume there are no deaths so excluded for d_i=0. And, R (t_i) the set of subjects at risk at the time just before t_i (t_i-0). And the summation in the denominator is over all subjects in the risk set at time t_i denoted by R (t_i)

The log partial likelihood function is given as:

$$L_P(\underline{\beta}) = \sum_{i=1}^{m} \{X_i\beta - \ln[\sum_{j \in R(t_i)} e^{X_j,\beta}]\}$$

We obtain the maximum partial likelihood estimates by differentiating the above equation with respect to β, setting the derivative equal to zero and solving for the unknown parameter.

Interpretation of the fitted CPHRM was based on the hazard function i.e., $e^{\hat{\beta}}$ is the maximum partial likelihood estimator of β.

The (1-**α**) *100% confidence interval (C.I) for the parameter estimates is given by equation (3.16) below.

$$\hat{\beta} \pm Z\alpha_{/2} \text{ S.e } (\hat{\beta}) \quad (3.17)$$

The hazard ratios confidence intervals are obtained by exponentiation of the parameter estimates.

$$\exp \widehat{\beta} \pm Z\alpha_{/2} \text{ S.e } (\hat{\beta})$$

3.9 Ethical Clearance, Data Safety, and Confidentiality

The ethical clearances were obtained from MTRH and Institutional Research Ethics Committee (Moi University) refer to letters of approval in appendix 1 and 2. However, because the study was conducted using secondary data individual pediatric patients' consent was not required. The retrieved research database used only patients' identification numbers that were uniquely linked to the patient. However, the patient’s identification numbers were changed to serial numbers by the assigned supervisor responsible for ART data in Ampath before sent to the researcher encrypted laptop for security and confidentiality.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents the results of the study for each specific objective and its discussions.

4.2 Determining the Baseline Characteristics, Mean and Median Survival Times of HIV patients on ART at MTRH

Table 4.1: Summary of the distribution of HIV patients per variable.

CHARACTERISTIC TABLE			
Variable Label	**Total, n (%)**	**Male, n (%)**	**Female, n (%)**
ARV Perfect Adherence	6792(67.7)	2077(69.6)	4715(66.9)
Death	205(2)	90(3)	115(1.6)
WHO Stage at enrollment			
Stage 1	3778(37.6)	887(29.7)	2891(41)
Stage 2	1825(18.2)	527(17.6)	1298(18.4)
Stage 3	2869(28.6)	1044(35)	1825(25.9)
Stage 4	338(3.4)	153(5.1)	185(2.6)
Age at Enrollment, (in years)	37.1±8.9	40.3±8.9	35.8±8.6
Age Category at Enrollment (Yrs.)			
<30	2263(22.5)	310(10.4)	1953(27.7)
30 - <40	4283(42.7)	1258(42.1)	3025(42.9)
40 - <50	2583(25.7)	997(33.4)	1586(22.5)
>50	907(9)	421(14.1)	486(6.9)
Age at ART Therapy, (in years)	38.7±9	41.4±8.9	37.6±8.8
Age Category at ART Therapy Start (Yrs)			
<30	1675(16.7)	231(7.7)	1444 (20.5)
30 - <40	4278(42.6)	1205(40.4)	3073 (43.6)
40 -< 55	3558(35.5)	1312(43.9)	2246(31.9)
>55	525(5.2)	238 (8)	287 (4.1)
HIV Severity	4629(46.1)	1646 (55.1)	2983(42.3)
School Years Completed	8.7±3.5	9.5±3.6	8.4±3.4
Level of Education			
Primary	4761(47.4)	1242(41.6)	3519(49.9)
Secondary	3252(32.4)	1092(36.6)	2160(30.6)
Tertiary	771(7.7)	379(12.7)	392(5.6)
BMI kg/m^2	21.8±3.6	20.9±2.8	22.2±3.8

BMI Category, kg/m^2			
<18.5	159(1.6)	58(1.9)	101(1.4)
18.5 - <25	533(5.3)	191(6.4)	342(4.9)
25 - <30	138(1.4)	28(0.9)	110(1.6)
>30	16(0.2)	0(0)	16(0.2)
CD4 count, cells/mm^3	243.7±211.3	191.7±171	266.1±222.8
CD4 Category, cells/mm^3			
350≤	6051(60.3)	2035(68.2)	4016(57)
>350	1951(19.4)	373(12.5)	1578(22.4)
Person Years	9.1±2	9.3±1.7	8.9±2.1

From (Table 4.1), a total of 10,038 out of 10,195 HIV/AIDS patients on ART were considered in the study representing 98.5% of the study population. The 157(1.5%) patient's data left out in this study had extreme characteristics (outliers). These patients (10038) adhered to ART services in MTRH during the study period, with male and female patients being 2985(29.7%) and 7053(70.3%) respectively. However, 6792 (67.7%) patients perfectly adhered to the ARV regimens and out of these data records, 2077 (69.6%) males and 4715 (66.9%) females were detected respectively.

The number of patients who died during the study period were 205 accounting for 2% of the target population. The deaths comprised 90 males and 115 females representing 3% and 1.6% respectively.

The patients' distribution based on the WHO stages during enrolment for those who perfectly adhered to the ART were as follows; 3778(37.6%), 1825(18.2%), 2869(28.6%) and 338 (3.4%) for stages 1, 2, 3, and 4 respectively. 887 (29.7%) males and 2891(41%) females were enrolled on WHO stage 1, while 527 (17.6%) males and 1298 (18.4%) females were enrolled in stage 2. 1044 (35%) males and 1825 (25.98%) females were enrolled in WHO stage 3 and finally 153(5.1%) males and 185(2.6%) females were enrolled in the WHO stage 4.

The distribution of patients based on age categories during enrolment to ART and the start of the therapy were as follows; below 30 years were 2263 (22.5%) and 1675 (16.7%) respectively, between 30 and 40 were 4283(42.7%) and 4278 (42.6%) , the total number of patients enrolled in age category between 40 and 50 years were 2583(25.7%) and a total of 3558(35.5%) began ART therapy between the age of 40 and 55 years. 907 (9%)Patients with ages 50 years and above were enrolled to ART and 525(5.2%) began ART when they were 55years and above .

Furthermore, patients were classified into three levels of education attainment during enrolment it was found out that those with primary, secondary and tertiary education levels were 4761(47.4%), 3252(32.4%) and 771(7.7%) respectively.

The body mass index was categorized into patients with less than 18.5 kg/m^2 were 159(1.6%) between 18.5-25 kg/m^2 were 553(5.3%), 25-30 kg/m^2were 138(1.4%) and with more than 30kg/m^2 were 16(0.2%)

Patients were classified into two categories of cluster of differentiation 4 i.e., those with at most 350$cells/mm^3$ CD count were 6051(60.3%) and with at least 350$cell/mm^3$ were 1951(19.4%).

Table 4.2: Summary of Survival Time per gender of HIV positive patients on ART in MTRH.

	Minimum	1st Quarter	Median	Mean	3rd Quarter	Maximum
Male	0.019	10	10	9.347	10	10
Female	0.025	8.949	10	8.939	10	10
Overall	0.019	9.456	10	9.06	10	10

Table (4.2) was obtained using equations 3.7 and 3.8. The estimated mean and median survival times for patients who adhered to ART services in MTRH was 9.06 and 10 years respectively. There was no difference in the median survival time for both genders but male recorded a mean survival time of 9.347 years against 8.939 years for their counterpart.

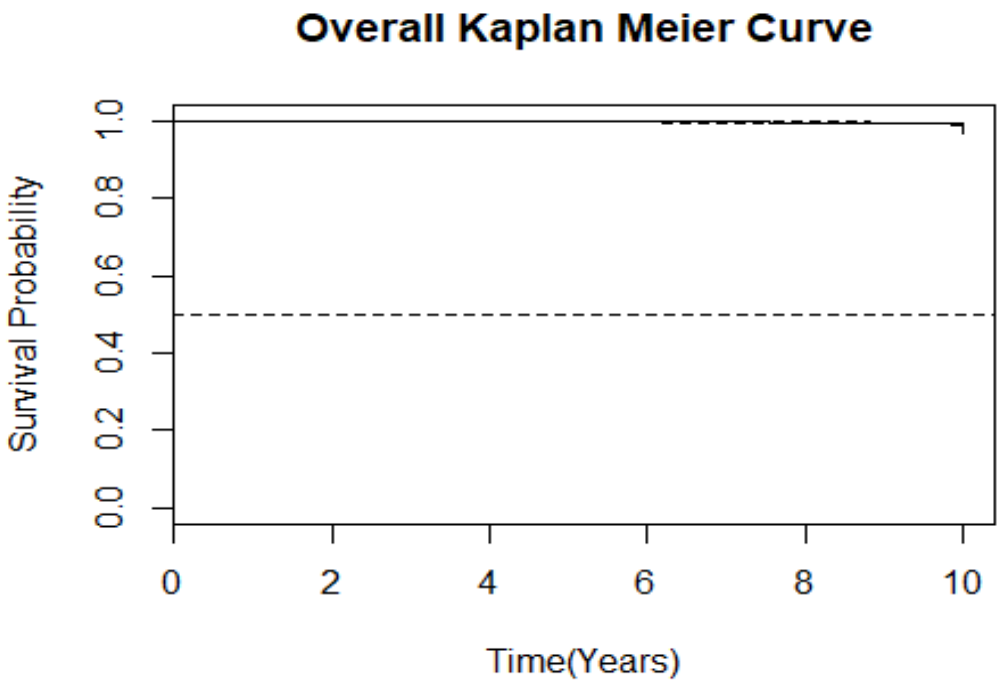

Figure 4.1: Overall Kaplan-Meier curve for HIV positive patients on ART in MTRH

The overall Kaplan Meier curve was obtained from equation (3.9). The results indicated an almost horizontal line, implying that most patients in this study were

censored as result of their adherence to the antiretroviral therapy which gave them moderately good chances of survival as shown in Figure (4.1) above. Also, it can be observed that all the patients who perfectly adhered to therapy survives for at least six years before facing the risk of death, this was evidenced by first descent of the Kaplan Meier curve at the sixth year.

4.3 Comparison of Gender Survival Probabilities

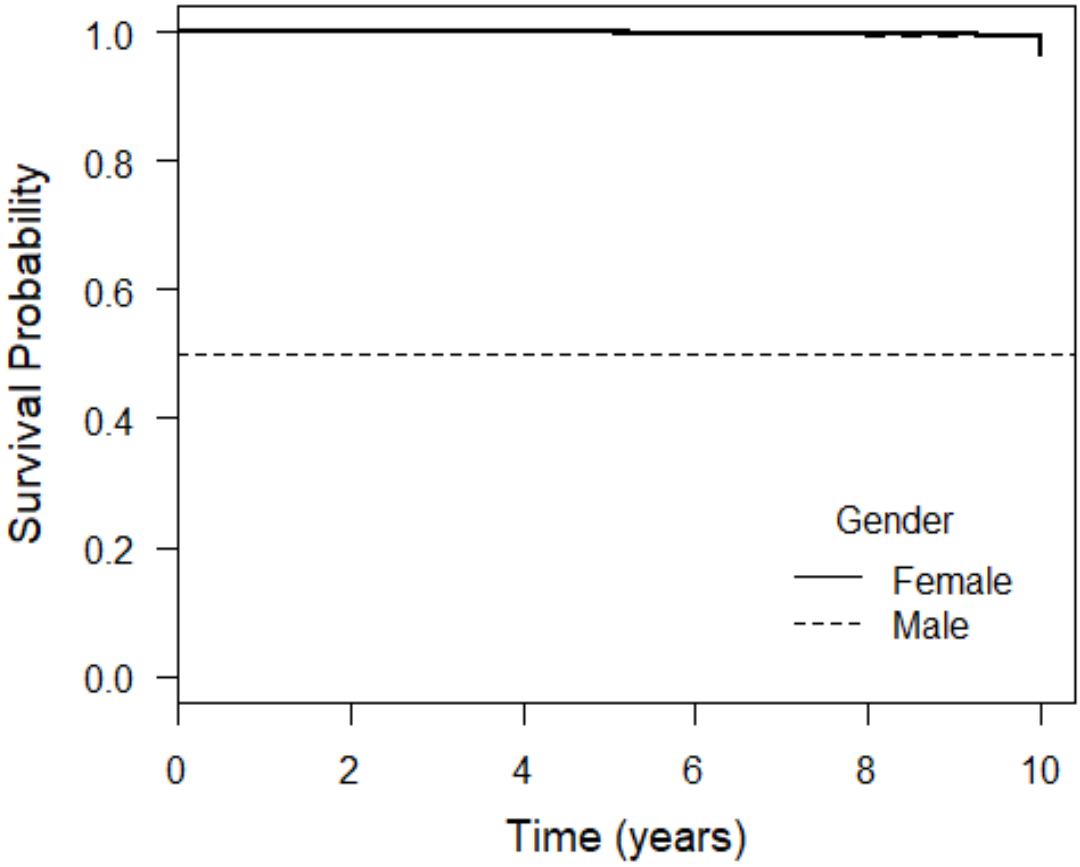

Figure 4.2: The Kaplan Meier Curve showing comparison of survival probabilities by gender

Figure 4.2 above, shows that the survival curves are significantly different. The log-rank test gives p-value = 0.000399. This means that the female patients had higher probability of survival as compared to the male since it was observed that the female curve was on the top of the male curve as shown in Figure 4.2. The male HIV/AIDS

patients experienced earlier deaths than the female patients because the male curve descended prior to the one belonging to their counterpart.

In addition, more than half of each gender population survived beyond the study period, this shown by the Kaplan curve not intersecting with the median probability line.

4.4 Fitting the Cox proportional hazard model

In order to study the relationship between survival time and predictor variables, a regression modeling approach to survival data using Cox proportional hazard regression model was used with the aim of estimating the regression coefficients, performing statistical test, construction of confidence intervals and drawing inference based on the hazard function. Model development and its adequacy assessment were done before drawing inference from the results of the fitted model. The main objective of model development was to attain a model which describes the data optimally.

4.4.1 Model development

The first procedure in Model development process was to select variable which were important for the study. The selection process was achieved by considering clinical importance of the variables, statistical significance and adjustment for confounding.

Once the variables were selected the univariate Cox proportional hazard regression model was fitted to the selected covariate data using log likelihood and partial log likelihood functions from equations (3.15) and (3.16) respectively. This resulted to unadjusted Cox proportional model in Table (4.5). The confidence intervals for the univariate model were constructed using equation (3.17).

Purposeful selection of the covariates was the most useful method for selecting variables to the initial multivariable model. This began with a model that contained all variables which were significant in univariate at 20-25% level in relation to survival time among HIV/AIDS patients enrolled at Ampath center in MTRH.

The six variables found statistically significant in univariate analysis were the candidates for inclusion in the initial multivariate Cox model. These variables included; HIV severity, age, BMI, level of education, gender and ARV adherence. HIV severity variable was created by merging the CD4 count and WHO staging to avoid multi colinearity of the variables. However, marital status, number of households, employment status, presence of opportunistic infections, patient functional status, drug abuse and condom use were not significant at 20-25% level of significance and hence were excluded from initial multivariate Cox model.

Continuous covariates i.e., age, CD4 count and BMI were categorized in basis of quartile and median values. Their significance was studied in two schemes, in categorical and continuous form as in Table (4.5). These continuous covariates were significant in both coding schemes at 5% significance level hence selected for inclusion in the initial adjusted Cox multivariate model in continuous form.

By method of likelihood ratio tests, HIV severity, age, BMI, level of education, gender and ARV regimens adherence were found statistically significant variables.

After fitting the initial multivariate model using equations (3.15), (3.16) and (3.17). The second procedure in modelling process was to check assumption of linearity for the continuous covariates in the model i.e., age, BMI and CD4 count using smoothed plots of Martingale residuals. The plots demonstrated linearity i.e., the plots were

found to be random, showing no systematic patterns and approximated a straight line for each continuous covariate. As a result, they were linear in the model.

The final step in model development process was the consideration of interaction term for inclusion in the model for purpose of improving inference and obtaining a more realistic model. This step began with creation of a list of plausible interactions formed between the six variables of the initial Cox model that were found to be statistically significant at 5% level in prediction of the hazard rates.

Each significant interaction was added back to the initial Cox multivariate model and fitted again using forward likelihood function.

Therefore, the Cox model in Table (4.5) remains as the "final model" but interpretation based on this model should not be made until its fit and adherence to the model assumption are checked.

4.4.2 Model assessment

In order to evaluate how well the fitted Cox regression model describes the data set, model adequacy was assessed on the preliminary final model which fulfilled the model development stages. The prerequisites for model assessment included testing the assumption of the proportional hazards, checking the presence of leverage and measuring the overall goodness of fit of the model.

The basic assumption of the proportional hazard model was that the hazard ratios were constant over time. This implies that the risk of failure was the same no matter how long the subjects were followed up. The proportional hazard assumption was tested through creation of an interaction between variables and logarithm of time (time dependent covariates) and checking their significance in predicting hazard

ratios. The results in Table (4.5) demonstrates that all time dependent covariates were not significant, this assures that there was 'proportional hazard.' Furthermore, to test for the assumption above for each covariate, the function *cox.zph( )* in the r codes was used to correlate the corresponding sets of scaled Schoenfeld residuals with time and performed global test for the model as whole as shown by Table 4.3 below

Table 4.3: Test for proportional hazard assumption for each covariate and global test

Variable	Chi-square (chisq.)	Degrees of freedom (d.f)	p-value
HIV Severity	0.124	1	0.725
Age category	2.023	2	0.364
BMI category	2.689	2	0.261
Level of education (lower)	2.286	1	0.131
Gender (male)	3.426	1	0.064
ARV perfect adherence	0.624	1	0.429
Global test	10.187	8	0.252

From the output of Table 4.3, the test for independence between scaled Schoenfeld residuals and time was not statistically significant. In addition, global test is not statistically significant. This confirms that the final model met proportional hazard assumption.

Graphically checking of proportional hazard was made by plotting scaled Schoenfeld residuals of the continuous variables with the corresponding survival time through lowess plots. The graph showed that each plot was random, smooth and had approximately zero slopes. The plots had parallel shape pattern, implying that there was no time dependent covariate, hence no violation of proportional hazard assumption.

The next important aspect of model evaluation was thorough examination of regression diagnostic statistics to identify which subjects had unusual configuration

of covariates i.e., to identify which subjects exerts undue influence on the estimates of the parameters and had undue influence on the fit of the model. An outlier is an extreme observation and by examining Scaled Score residuals and normal probability plots are helpful in identifying outliers.

The effects of outliers on the proportional hazard regression were easily checked by dropping these points and refitting the regression equation. In this study outliers were examined by Scaled Score residuals.

The final step in assessment of the model adequacy was the measure the overall goodness of fit. All measures depend on the proportion of values that are censored. A perfect adequate model has low R^2 due to presence of censored data.

4.5 Investigating Predictors of Survival Time among HIV patients on ART at MTRH

This section shows how to investigate the final multivariate Cox proportional hazard model that fulfilled model development and assessment procedures. In this study, the Cox model had HIV severity, age, Body Mass Index (BMI), level of education and ARV perfect adherence as its explanatory or predictor variables and time to the death as the dependent variable also known as survival time. The Cox regression parameters **($\underline{\beta}$)** were obtained by taking natural logarithm (ln) of the respective proportional hazard values of the variables. However, in the cases where variables had categories, the parameter was treated first as stratum and in order to obtain its value, the hazard values were multiplied for all categories within the variable and then takes natural logarithm of the product to obtain the parameter.

Table 4.4: Variables and parameter estimates for both unadjusted and adjusted Cox proportional model

Covariate of interest	Symbol	Unadjusted parameter(β)	Adjusted parameter (β)
HIV severity (yes)	X_1	-0.32	+ 0.17
Age	X_2	-0.09	-1.22
BMI	X_3	-0.231	+0.35
Level of education (lower)	X_4	-0.05	-0.07
Gender (sex) male	X_5	+ 0.49	+ 0.63
ARV perfect adherence (yes)	X_6	-0.40	+0.33

4.5.1 Unadjusted cox proportional hazard model.

From Table 4.4 above, parameters' estimates, the unadjusted Cox proportion hazard regression models can thus be written as: -

$\ln\{\lambda_t\} = -0.32X_1$ $\quad \beta_1 = -0.32$

$\ln\{\lambda_t\} = -0.09X_2$ $\quad \beta_2 = -0.09$

$\ln\{\lambda_t\} = -0.23X_3$ $\quad \beta_3 = -0.23$

$\ln\{\lambda_t\} = -0.05X_4$ $\quad \beta_4 = -0.05$

$\ln\{\lambda_t\} = +0.49X_5$ $\quad \beta_5 = +0.49$

$\ln\{\lambda_t\} = -0.40X_6$ $\quad \beta_6 = -0.40$

Only gender (male, X_5) of the patients' indicated worse prognosis i.e., it was associated with increased risk of death or decreased survival because its parameter had positive coefficient whereas HIV severity (X_1), age (X_2), BMI (X_3), lower level of education (X_4) and perfect adherence to antiretroviral drugs (X_6) variables had protective effect (i.e., associated with improved survival time), since their coefficients were negative.

4.5.2 Adjusted Cox proportional hazard model

$\ln\{\lambda_t\} = 0.17X_1 - 1.22X_2 + 0.35X_3 - 0.07X_4 + 0.63X_5 + 0.33X_6$

In the adjusted model above, HIV severity (X_1), BMI(X_3), gender (X_5), and ARV perfect adherence (X_6) had positive parameter coefficients which indicated that these predictor variables are associated with worse prognosis or decreased survival among HIV/AIDS patients on ART in MTRH. While age(X_2) and lower-level education (X_4) had negative parameter coefficient indicating that it had protecting effect or associated with improved survival or decreased risk of the death among HIV/ AIDS patients.

The change in parameter coefficients estimates sign (from negative to positive) for HIV severity, BMI and ARV perfect adherence was due to the multicollinearity between the above variables and male gender variable.

Table 4.5: Unadjusted/Adjusted Hazard Ratios of factors associated with longer survival time among HIV patients on ART in MTRH

Covariates of interest	Unadjusted HR (95% C.I)	P-value	Adjusted HR (95% C.I)	P-value
HIV Severity: Yes	0.729(0.547,0.972)	0.032	1.18(0.451,3.085)	0.735
Age Cat-<30	Ref		Ref	
Age Cat-30-40	0.824(0.565,1.201)	0.314	0.459(0.148,1.424)	0.178
Age Cat-40+	1.114(0.77,1.609)	0.567	0.644(0.207,2.003)	0.447
BMI Cat-18.5-<25	Ref		Ref	
BMI Cat-<18.5	1.383(0.564,3.393)	0.478	1.65(0.634,4.294)	0.305
BMI Cat-25+	0.574(0.131,2.51)	0.461	0.861(0.188,3.948)	0.847
Level of education (lower).	0.952(0.914,0.992)	0.019	0.931(0.833,1.041)	0.209
Sex: Male	1.633(1.239,2.152)	<.0001	1.884(0.731,4.855)	0.19
ARV perfect adherence: Yes	0.668(0.506,0.882)	0.004	1.393(0.534,3.633)	0.498

Ref=Reference category. **Key: C.I** Confidence Interval **Cat** Category

4.5.3 The Unadjusted Hazard Ratios

Table (4.5) shows the unadjusted hazard ratio, confidence interval and p-value for each explanatory variable that were discussed as follows: -

Results of the unadjusted Cox model showed that patients who were HIV/AIDS severe had less risk of death or associated with improved survival since the UHR was less than 1.00 i.e. [UHR=0.729, 95% C.I = 0.547, 0.972 p-value = 0.032]. This association was statistically significant since its 95% C.I does not contain the null value (1) and the p-value was less than 0.05.

HIV/AIDS patients with age category between 30 and 40 years had decreased risk of death or associated with improved survival compared to the reference age category less or equal to 30 years, with [UHR=0.824, 95%C.I = 0.565-1.201, p-value = 0.314] whereas patients with age category 40 years and above had 11% higher risk of death compared to the same reference age category with [UHR=1.11, 95% C.I = 0.77-1.609, p-value = 0.567]. Both associations were not statistically significant since their 95% C.I contained the null value (1) and their p-values were greater than 0.05.

Patients who had under body weight defined by BMI less than18.5kg/m^2 had 38% higher risk of death or decreased survival rate compared to the reference BMI category of normal body weight of 18.5-<25kg/m^2, with [UHR=1.383,95% C.I = 0.564-3.393, p-value = 0.478], whereas patients with over body weight given by BMI $\geq$ 25kg/m^2 had increased chances of survival or less risk of death as compared with the same reference category with [UHR= 0.574,95% C.I = 0.131-2.51, p-value = 0.461]. Both associations were not statistically significant since their 95% C.I contained the null value (1) and their p-values were greater than 0.05.

Level of education had protective effect on the risk of death (improved survival) since the UHR was less than (1.00) marked by lower level of education [UHR= 0.952, 95% C.I =0.914-0.992, p-value= 0.019]. This association was statistically significant in that the 95% C.I does not contain the null value (1) and its p-value was less than 0.05. However, patients with lower level of education had higher risk of death (95.2%) compared to those who had higher levels of education, with 4.8% risk of death.

Male patients had 63% higher risk of death compared to female patients on ART in MTRH with [UHR= 1.633, 95% C.I = 1.239-2.152, p-value < 0.0001]. The association was highly statistically significant in that the 95% C.I does not contain the null value (1) and its p-value was approximately 0.00.

Most of the patients who perfectly adhered to ARV had decreased risk of death or had enhanced survival compared to those who did not adhered, with unadjusted hazard rate ratio [UHR = 0.668, 95% C.I = 0.506-0.882, p-value 0.004]. The association was highly statistically significant in that the 95% C.I does not contain the null value (1) and its p-value was approximately 0.00.

4.5.4 Adjusted Hazard Ratios

From Cox model (Table 4.5), the adjusted hazard ratio, confidence interval and respective p-values for each explanatory variable considered while adjusting for the other variables were as follows: -

HIV/AIDS patients' with severe cases had 18% higher risk of death as compared with non-severe patients' when age, BMI, school years completed, gender and ARV Adherence was adjusted for with adjusted hazard ratio [AHR =1.18, 95% C.I = 0.451-3.085, p-value = 0.735]. Though the association was statistically insignificant

since the 95% C.I for AHR contained the null value (1) and the p-value was greater than 0.05.

Patients' with age category of between 30 to 40 years had less risk of death compared to reference age category less or equal to 30 years when other variables were adjusted for with [AHR= 0.459, 95% C.I = 0.148-1.424), p-value 0.178]. However, HIV/AIDS patients' with age category 40 years above had higher risk of death compared to the same reference age category with [AHR= 0.644, 95% C.I = 0.207- 2.003, p-value = 0.447]. Both associations were not statistically significant in that their 95% C.I contained the null value (1) and their p-values were greater than 0.05.

Underweight patient's with BMI less than18.5kg/m^2 had 65% higher risk of death compared to the reference normal body weight marked by BMI of 18.5 - < 25 when other covariates in the study were adjusted for in the model, with [AHR=1.65,95% C.I = 0.634-4.294, p-value = 0.305]. However, for patients with overweight defined by B.M.I $\geq$ 25kg/m^2 had increased chances of survival as compared to the same BMI reference category with [AHR= 0.861, 95% C.I = 0.188-3.948, p-value 0.847]. However, both associations were not statistically significant since their 95% C.I contained the null value (1) and their p-values were greater than 0.05.

Level of education had protective effect on the risk of death (improved survival) when other variables considered in the model were adjusted for, since the AHR for lower level of education was less than (1.00) i.e., its adjusted hazard ratio was [AHR= 0.931, 95% C.I = 0.833-1.041, p-value 0.209]. However, Patients' who had lower level of education had higher risk of death (decreased survival) i.e., 93.1% compared to those who had completed higher levels of education when other covariates were

adjusted for, although the relationship was insignificant since the 95% C.I contained the null value (1) and its p-values was greater than 0.05.

When HIV severity, age, BMI, school years completed and perfect adherence to ARV were adjusted for in the model, the male patients had 88% higher risk of death (decreased chances of survival) compared to the female counterpart with [AHR= 1.884, 95% C.I = 0.731- 4.855, p-value=0.19]. However, the association was statistically insignificant since the 95% C.I contained the null value (1) and their p-value was greater than 0.05.

Patients' who perfectly adhered to ARV had 39.3% higher risk of death compared to those who did not perfectly adhered to the ARV drugs, with adjusted hazard ratio [AHR= 1.393, 95% C.I = 0.534- 3.633), p-value 0.498] when other covariates are adjusted for in the Cox model. However, the association was statistically insignificant since the 95% C.I contained the null value (1) and its p-value was greater than 0.05.

4.6 Results Discussion

In this study, Cox proportional hazard model was applied to estimate the predictor variables of survival time among HIV/AIDS patients under ART in Moi Teaching and Referral Hospital (MTRH).

The ten years retrospective adult study of HIV/AIDS patients' in MTRH (ART) Center gave an insight into survival time patterns of patients and its determinants in a hospital setting in Kenya.

In this study 2% of the patients' died and most of the deaths occurred after six years of initiation of ART (Table 4.2). The overall death rate was lower compared to other similar studies in Africa, for instance it was 29.7% in Tanzania (Johannessen et al.,

2008), 14.5% in Uganda (Bamulangeyo, 2018) and 23.0% in Cameroon (Sieleunou, Souleymanou, Schönenberger, Menten, & Boelaert, 2009). There was no mortality reported during the first one year which was in contrast to other studies in Africa where patients' died within the first one year of initiation of ART (Johannessen et al., 2008), and (Alemu, & Sebastián, 2010). This was mainly due to perfect adherence to ARV by most of the subjects investigated in this study.

The overall mean and median survival times for patients who adhered to ART services in MTRH was 9.06 and 10 years respectively. There was no difference in the median survival time for both gender, this was in agreement with the study done by Guerreiro *et al.,* (2002) but male recorded a mean survival time of 9.347years against 8.939 years for the female counterpart as shown in (Table 4.2). This was higher in comparison to similar studies in Ethiopia which established a mean survival time of 65.22 months (Abebe,.*et al* (2014), 63.7 months and 77 months (Tsehaineh, 2010) and (Kebebew & Wencheko, 2012) respectively and 5.7 months by (Akessa *et al.,* 2015). This was attributed to the longer follow up time considered in this study, quality of ART service, perfect adherence to ARV and improved socio demographic characteristics of the patients.

There was significant difference between male and female survivorship function (figure 4.2) which was contrary to the previous study carried out by (Shebeshi, 2011) where the study established that there was no significant difference in survivorship. This was attributed to fact that male patients were often diagnosed at the later WHO stages of HIV/AIDS and in turn led to delayed intervention of AIDS. In addition, male patients had tendency of drug abuse.

The Cox proportional hazard regression model fitted to the data. HIV severity, age, BMI, level of education, gender and perfect adherence to the ARV were explanatory variables and survival time was dependent variable in this study. This was similar to all past studies that considered fitting the Cox model to medical human data (Maposa, 2016; Shebeshi, 2011). However, for this study Patients' data were classified into severe and non-severe cases of the HIV/AIDS disease based on the level of CD4 cell count and W.H.O staging, where patients with CD4 count less than 350cells /mm^3 (stage III and IV) have been classified to be severe and those with CD4 count at least 350 cells/mm^3 (stage I and II) classified as non-severe cases of HIV disease. The reason for this classification is to avoid the multi-collinearity effects of the two variables i.e., WHO staging and CD4 count clustering.

HIV/AIDS severity was statistically significant predictor of survival time when other factors were held constant i.e. It was associated with improved survival (protective variable) with [UHR = 0.729, 95% C.I = 0.547- 0.972, p-value = 0.032]. This means that patients with severe cases of HIV/AIDS would die at the rate of 0.729 times or would be 1.371 years death free from HIV/AIDS related cases, and statistically significant in the sense that its confidence interval does not contain the null value i.e. (1) and its p-value less than 0.05. The reason for the above observation was that patients with severe cases would tend to strictly adhere to the ARV drugs and ART guidelines compared to the non-severe cases in order for them to improve their health conditions.

However, adjusting for other covariates included in the model, increases the risk of death by 18% among patients with severe cases as compared to those with non-severe disease cases or decreases the years of death free time to 0.85 years. Although the

association was not statistically significant due to interrelationships of the considered covariates in the model. This results was consistent with most of the past studies in Kenya (Sika, Wools-Kaloustian, Mwangi, Kimaiyo, Diero, Ayuo, & Musick, 2010) established that WHO stage IV was found the main predictor of death i.e. patients' in stage IV had lesser chance of survival this was supported by most of the African nations studies (Maposa, 2016; Shebeshi, 2011).

Age of the patients in this study was statistically insignificant predictor of survival time in both univariate and multivariate Cox models since the C.I for both unadjusted and adjusted hazard ratios contain the null value (1) and their p-values were greater than 0.05 i.e., UHR for age category 30-40 and above 40 years were [UHR = 0.824, 95% C.I = 0.565-1.201, p-value=0.314] and [UHR = 1.11, 95% C.I =0.77-1.609, p-value= 0.567] respectively. Also, AHR for the respective age category were [AHR= 0.459, 95% C.I = 0.148-1.424), p-value 0.178] and [AHR= 0.644, 95% C.I = 0.207-2.003, p-value = 0.447]. This was in agreement to comparable study in Singapore (Hentrich, Maretta, Chow, Bogner, Schürmann, Neuhoff, & Mitrou, (2006) and other research papers done before (Getahun, 2010; Birtukan, 2010; Balcha, Jeppsson, & Bekele, 2011) Nevertheless, when survivorship of categorized aged group was done it was established that the rate of failure increases with increase in the age of the patients for both unadjusted and adjusted Cox models, Table (4.5), i.e. the failure rate for age brackets 30-40 years and 40 years were 82.4% and 111.4% for unadjusted and 45.9% and 64.4% for the adjusted models respectively. The value above 100% signifies decreased survival time or increased rate of failure in reference to $\leq$30 years age category. The above results could also be interpreted to mean, patients who were 30-40 age category would be 1.21 years and 2.18 years death free from HIV/AIDS defining illness for unadjusted and adjusted Cox models respectively. Conversely

patients with age bracket $\geq$ 40 would be 0.90 years and 1.55 years death free from AIDS defining illness for unadjusted and adjusted models respectively. These results indicated that there was significant difference in the survival rates within age category. This was consistent with the studies done in Ethiopia (Shebesi, 2011; Assefa, & Wencheko, 2012) who found out that HIV/AIDS patients who were above 60 years had shortest survival time as compared to other age categories below that age group. This was attributed to decrease in body physiological functions for instance decreases in cell differentiation rendering the body to have incompetent immunity status and thus would be at higher risk of complications and respond poorly to ART.

Furthermore, there was decreased risk of death within age categories when other factors were adjusted for in the model. This was due to multicollinearity among protective univariate variables (HIV severity, level of education and perfect adherence to ARV) with age of the patients.

Body Mass Index was statistically insignificant predictor of survival in both univariate and multivariate Cox models in (Table 4.5) for the reason that their 95% C.I contained the null value (1) and their p-values were greater than 0.05. This was marked by the UHR of [UHR= 1.383,95%C.I = 0.564-3.393,p-value=0.478] for BMI< $18.5kg/m^2$ and [UHR= 0.574,95% C.I = 0.131-2.51, p-value = 0.461] for BMI$\geq 25kg/m^2$ in addition the adjusted hazard ratios of [AHR=1.65,95% C.I = 0.634-4.294, p-value = 0.305] for BMI< $18.5\text{kg}/m^2$ and [AHR= 0.861, 95%C.I = 0.188-3.948, p-value 0.847] for BMI$\geq 25kg/m^2$ .This was not in agreement with the study done in Ethiopia, which found out that BMI was statistically significant predictor of survival (Tadesse, Haile, & Hiruy, 2014). However, some of the studies

uses weight as a proxy measure of BMI because most hospitals do not record heights of the patients (Assefa, & Wencheko, 2012). One study Similar to this was done in Tanzania yielded similar results that patients who were underweight marked by BMI less than 18.5kg/m^2 had higher mortalities in comparison to those with normal and over body weights marked by BMI of $18.5 \leq 25kg/m^2$ and $\geq 25kg/m^2$(Li, N., Spiegelman, D., Drain, P., Mwiru, R. S., Mugusi, F., Chalamilla, G., & Fawzi, W. W. 2012). This was evident by the failure rate of 138.3% and 57.4% for BMI < 18.5kg/m^2 and $\geq$ 25kg/m^2 in reference to BMI category of $\geq 18.5\text{-}\leq 25$ kg/m^2 respectively in univariate Cox models (table 4.5). Similarly, there was decreased risk of death with increase in BMI among HIV/AIDS positive patients on ART when other factors considered in the multivariate Cox model table (4.5) were adjusted This observation was validated by failure rate of 165% and 86.1% for BMI < 18.5kg/m^2 and $\geq$ 25kg/m^2 Respectively.

Alternative inference on the results was that patients with BMI < 18.5kg/m^2 were 0.72 and 0.61 years death free for unadjusted and adjusted models respectively. Moreover, patients with BMI $\geq 25kg/m^2$ were 1.74 and 1.16 years death free for unadjusted and adjusted models respectively from AIDS defining illnesses.

From above results it was observed that adjusting for other covariates in multivariate model increase the risk of death within categories of BMI, due to interrelationship between HIV severities, male gender, non-adherence to ARV and BMI.

BMI contribute to drug metabolism and would therefore have an effect to the efficacy of HAART. Underweight Patients with BMI $\leq$ 18.5kg/m^2 tend to have higher viral loads than those with normal and overweight patients marked by BMI 18.5kg/m^2 $\leq 25kg/m^2$ and $\geq 25kg/m^2$ respectively. This was due to emaciated CD4, CD8 and T

lymphocyte cells caused by malnutrition hence increases the risk of death (Li, X., Ding, H., Geng, W. *et al, 2019).*

Level of education was significant predictor of improved survival time in univariate Cox model with lower level of education having unadjusted hazard ratio of [UHR= 0.952, 95% C.I = 0.914-0.992, p-value = 0.019] since the 95% C.I does not contain the null value (1) and its p-value was less than 0.05. However, the same level of education was not statistically significant in the multivariate model when other covariates were adjusted for i.e., the adjusted hazard ratio was [AHR= 0.931, 95% C.I 0.833-1.041, p-value 0.209] contained the null value (1) and its p-value was greater than 0.05.

However, patients with lower level of education had 95.2% risk of death compared to those with higher levels (tertiary level) patients with 4.8% risk of death. These analogous results would also be interpreted to mean, patients with lower level of education would be 1.05 years death free from AIDS defining illness compared to 2.08 years of death free among patients with higher level of education. The results were statistically significant in univariate Cox model which was consistent with other comparable studies (Fausta, 2013; Tadesse, & Hiruy, 2014). The observation was attributed to better know-how by patients with higher level of education about the importance of drug compliance, their understanding on the need for earlier presentation for hospitalization (ART) and rightful medication procedures. In addition, higher education reduces stigmatization attitude and the small sample size of patients with tertiary education was an indicator of lower infection rate within this group.

Adjusting for other covariates in the study decreases the risk of death among patients with lower level of education from 95.2% to 93.1% or increases the time of death free from 1.05 years to 1.07 years this was due to interrelationship between variables with protective effects (non-severity and perfect adherence to ARV) and level of education.

Gender (male) was the most statistically significant predictor of death in univariate Cox model with UHR of [UHR= 1.633, 95% C.I = 1.239-2.152, p-value < 0.0001] reason being its 95% C.I does not contain the null value (1) and the p- value was less than 0.05 This output was consistent to most of the early studies for instance (Tadesse, Haile & Hiruy, 2014), established that most men were on higher risk of death in comparison to women patients.

Conversely gender (male) was not statistically significant predictor of survival time in multivariate Cox model with [AHR= 1.884, 95% CI 0.731- 4.855, p-value=0.19], the reason was that its 95% C.I contained the null value (1) and its p- value was greater than 0.05.

From Table (4.5) male patients had 63.3% and 88.4% higher risk of death compared to the female patients in univariate and multivariate Cox models respectively This discrepancy was accounted for by late diagnosis among male patients i.e., living in denial for a long time, while female patients undergo compulsory testing during both pre-natal and post-natal clinics visits.

The above results (gender variable) were also interpreted to mean; male patients were 0.61 and 0.53 years death free from AIDS defining illnesses in univariate and multivariate models respectively.

Adjusting for other covariates in multivariate model increases the risk of death among the male patients from 63.3% to 88.4% due to multicollinearity between variables associated with worst prognosis (HIV Severe cases, lower BMI, lower level of education, old age and non-adherence to ARV drugs) with gender variable.

Adherence of ART pertains intake of all ARV pills in correctly prescribed doses at the right time in the right way observing any dietary restrictions (Shebeshi, 2011). Successful antiretroviral therapy is dependent on sustaining high rate of adherence. The minimum level of adherence required for ARV to work effectively is 95 % (Dombrowski *et al.*, 2013).

In this study Patients 'perfect adherence to ART was statistically significant predictor of survival time i.e., was associated with improved survival time in the univariate Cox model with [UHR = 0.668, 95% C.I = 0.506 - 0.882, p-value 0.004] since its 95% C.I does not contain the null value (1) and the p- value was less than 0.05. This result was in tandem to similar comparable studies in Ethiopia, (Birtukan, 2010), and (Ibrahim, 2007). However, analogous comparable study in south Africa (Mzileni, M. O., Longo-Mbenza, B., & Chephe, 2008) showed that estimated survivorship was statistically significant different. Also a Kampala study in Uganda (Abaasa *et al.,* 2008) established that non-adherence was found a significant predictor of survival time.

Patients' perfect adherence to ART was not statistically significant predictor of survival time in multivariate Cox model since the 95% CI contained the null value and its p-value was greater than 0.05 given by [AHR= 1.393, 95% C.I = 0.534-3.633), p-value 0.498].

Similarly, this result would be interpreted to mean patients with perfect adherence to ART were 1.50 years and 0.72 years death free from HIV/AIDS defining illness in unadjusted and adjusted models respectively.

Adjusting for other covariates in model (Table 4.5) increases the risk of death by 39.3% among patients who had perfectly adhered to ART compared to the non-adherence due to the confounding effects of variables considered jointly in the multivariate model. Consequently non-adherence to antiretroviral drugs results in treatment failure by increasing the chances of mutation that could lead to drug resistant virus and finally death.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the conclusion and recommendations from the study, section 5.2 gives conclusions derived from the results and 5.3 puts forward recommendations from the study.

5.2 Conclusions

The estimated mean and median survival times of 9.06 years and 10 years respectively confirmed that ART adherence enhances longer survival time among adult HIV/AIDS patients in MTRH. The Cox proportional hazard regression model fitted well to the HIV/AIDS data hence was able to describe the data optimally. HIV severity, level of education, gender and ART adherence level were independent significant predictors of survival time among adult HIV/AIDS patients in MTRH whereas BMI and age were not at univariate level. In multivariate CPHRM none of the covariates was jointly significant predictor of survival time among this group of patients due to confounding effects of the considered covariates. However, adjusting for other variables in the model, the study established that the following adult HIV/AIDS patients' characteristics: - severity (low CD4 count) at the start of the therapy, underweight, late diagnosis and delayed presentation for medication by male gender and perfect adherence of ART at later WHO stage (III&IV) increased the risk of death. Both unadjusted and adjusted model showed that the risk of death among HIV/AIDS patients on ART during this period increase with the rise in age of the patients. Conversely an increase in BMI decreases the risk of death. The rate of survival increases with rise of literacy level. Start of ART and perfect adherence at a latter disease stage (terminal) increases the risk of death.

5.3 Recommendations

The main objective of establishment of ART service was to improve health of HIV infected patients hence prolonging their lives and reducing HIV related mortality. However, there is continued mortality within the ART. Based on the findings of the study, some of the recommended feasible approaches of improving HIV/AIDS patient survival time include: - HIV patients should be advised to begin ART when their CD4 count is at least 350mm^3 (non-severe status) and when they have normal body weight during medication. Kenyan population should be sensitized on the adverse effect of HIV at old age. MTRH to enhance HIV/AIDS education on importance of perfect adherence to ARV regimens at initial stages of the disease infection. The Government of Kenya through ministry of should spearhead development of policy framework for provision of regular compulsory screening services for male population to avoid late diagnosis and intervention of the HIV/AIDS disease. The intervention points to include; school opening days for primary, secondary and college children and male adults coming to seek treatment should begin with the HIV test first. Finally, Future research to be done using this study as the baseline while considering a longer follow up period say 20-30 years in order to establish ART efficacy on reduction of mortality within treatment.

REFERENCES

Abaasa, A. M., Todd, J., Ekoru, K., Kalyango, J. N., Levin, J., Odeke, E., & Karamagi, C. A. S. (2008). Good adherence to HAART and improved survival in a community HIV/AIDS treatment and care programme: The experience of the AIDS Support Organization (TASO), Kampala, Uganda. *BMC Health Services Research, 8*. https://doi.org/10.1186/1472-6963-8-241

Abebe, N., Alemu, K., Asfaw, T., & Abajobir, A. A. (2014). *Survival status of HIV positive adults on antiretroviral treatment in Debre Markos Referral Hospital, Northwest Ethiopia*: retrospective cohort study.

Alemu, A. W., & Sebastián, M. S. (2010). Determinants of survival in adult HIV patients on antiretroviral therapy in Oromiyaa, Ethiopia. *Global health action, 3(1), 5398*. https://doi.org/10.3402/gha.v3i0.5398

Alomepe, J. L. (2011). *Illness experience of HIV-infected women in Cameroon, Central Africa.* Milwaukee, US*:* The University of Wisconsin-Milwaukee.

Andrade, H. B., Shinotsuka, C. R., Ho, L., Carvalho, F. B. D., Bozza, F. A., & Japiassú, A. M. (2017). Systematic review of short-term prognostic factors of HIV patients admitted to an intensive care unit. *Journal Brasileiro de Doenças Sexualmente Transmissíveis,* 29(2), 36-43. https://www.arca.fiocruz.br/handle/icict/34181

Andrinopoulou, E.-R., Rizopoulos, D., Takkenberg, J. J. M., & Lesaffre, E. (2014). Joint modeling of two longitudinal outcomes and competing risk data. *Statistics in Medicine, 33*(18), 3167–3178. https://doi.org/10.1002/sim.6158

Assefa, T., & Wencheko, E. (2012). Survival analysis of patients under chronic HIV-care and antiretroviral treatment at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia. *Ethiopian Journal of Health Development, 26*(1), 22-29.

Bamulangeyo, M. O. (2018). *Determinants of survival time among HIV positive patients after being initiated on Combivir. The case of Gulu Regional Referral Hospital in Uganda (Doctoral dissertation*). Kampala, Uganda: Makerere University. http://196.43.133.114/handle/10570/7026

Balcha, T. T., Jeppsson, A., & Bekele, A. (2011). Barriers to antiretroviral treatment in Ethiopia: a qualitative study. *Journal of the International Association of Physicians in AIDS Care,* 10(2), 119-125.

Birtukan, T. (2010). *Predictors of Early Death in HIV Positive Individuals Enrolled for Chronic HIV Care in Jimma Zone, South Western Ethiopia*. Addis Abeba, Ethiopia: University.

Burkey, Matthew D., Sheri D. Weiser, Desiree Fehmie, Stella Alamo-Talisuna, Pamella Sunday, Joy Nannyunja, Steven J. Reynolds, and Larry W. Chang. "Socioeconomic determinants of mortality in HIV: evidence from a clinical cohort in Uganda." *Journal of acquired immune deficiency syndromes* (1999) 66, no. 1 (2014): 41.

Chaisson, R. E., Keruly, J. C., & Moore, R. D. (1995). *Race, sex, drug use, and progression of human immunodeficiency virus disease*. New England Journal of Medicine, 333(12), 751–756. https://doi.org/10.1056/NEJM199509213331202

Choi, B. Y., Choi, J. Y., Han, S. H., Kim, S. I., Kee, M. K., Kim, M. J., ... & Kim, J. (2018). Korea HIV/AIDS Cohort Study: study design and baseline characteristics. *Epidemiology and health*, *40*.

Ciaranello, A. L., Myer, L., Kelly, K., Christensen, S., Daskilewicz, K., Doherty, K., & Wools-Kaloustian, K. (2015). Point-of-care CD4 testing to inform selection of antiretroviral medications in South African antenatal clinics: A cost-effectiveness analysis. *PLoS One*, *10*(3), 1-19. https://doi.org/10.1371/journal.pone.0117751

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: SeriesB (Methodological)*, *34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Croxford, S., Kitching, A., Desai, S., Kall, M., Edelstein, M., Skingsley, A., Burns, F., Copas, A., Brown, A. E., Sullivan, A. K., & Delpech, V. (2017). Mortality and causes of death in people diagnosed with HIV in the era of highly active antiretroviral therapy compared with the general population: an analysis of a national observational cohort. *The Lancet Public Health*, *2*(1), 35-46. https://doi.org/10.1016/S2468-2667(16)30020-2

De Castro, M., Cancho, V. G., & Rodrigues, J. (2010). A hands-on approach for fitting long-term survival models under the GAMLSS framework. *Computer methods and programs in biomedicine*, 97(2), 168-177.

Dombrowski, J. C., Kitahata, M. M., Van Rompaey, S. E., Crane, H. M., Mugavero, M. J., Eron, J. J., Boswell, S. L., Rodriguez, B., Mathews, W. C., Martin, J. N., Moore, R. D., & Golden, M. R. (2013). High levels of antiretroviral use and viral suppression among persons in HIV care in the United States, 2010. *Journal of Acquired Immune Deficiency Syndromes*, 63(3), 299–306. https://doi.org/10.1097/QAI.0b013e3182945bc7

Getahun, Z. (2010). *Statistical Modelling of HIV/AIDS Progression and Survival of AIDS Patients. A Case Study of Bahir-Dar, Feleg-Hiwot Referral Hospital.* Hawassa, Ethiopia: Hawassa University,

Guerreiro, M. F., Kerr-Pontes, L. R. S., Mota, R. S., Franạ, M. C., FTávora, F., & Caminha, I. (2002). Survival of adult AIDS patients in a reference hospital of a metropolitan area in Brazil. *Revista de Saude Publica*, 36(3), 278–284. https://doi.org/10.1590/S0034-89102002000300004

Hendricks, K. M., Dong, K. R., Tang, A. M., Ding, B., Spiegelman, D., Woods, M. N., & Wanke, C. A. (2003). High-fiber diet in HIV-positive men is associated with lower risk of developing fat deposition. *The American Journal of Clinical Nutrition*, *78*(4), 790–795. https://doi.org/10.1093/ajcn/78.4.790

Hentrich, M., Maretta, L., Chow, K. U., Bogner, J. R., Schürmann, D., Neuhoff, P., & Mitrou, P. (2006). Highly active antiretroviral therapy (HAART) improves survival in HIV-associated Hodgkin's disease: results of a multicenter study. *Annals of oncology*, *17*(6), 914-919.

Ibrahim, N. (2007). *Education of factor affecting the chance of survival/death status among hiv positive people under the antiretroviral treatment program: The Case of Adama Referral Hospital (Doctoral dissertation, M. Sc. Thesis)*. Addis Ababa, Ethiopia: Addis Ababa University

Jerene, D., Endale, A., Hailu, Y., & Lindtjøorn, B. (2006). Predictors of early death in a cohort of Ethiopian patients treated with HAART. *BMC Infectious Diseases*, *6*. https://doi.org/10.1186/1471-2334-6-136

Johannessen, A., Naman, E., Ngowi, B. J., Sandvik, L., Matee, M. I., Aglen, H. E., Gundersen, S. G., & Bruun, J. N. (2008). Predictors of mortality in HIV-infected patients starting antiretroviral therapy in a rural hospital in Tanzania. *BMC Infectious Diseases*, *8*. https://doi.org/10.1186/1471-2334-8-52

Kebebew, K., & Wencheko, E. (2012). *Survival analysis of HIV-infected patients under antiretroviral treatment at the Armed Forces General Teaching Hospital, Addis Ababa, Ethiopia*. AIDS Behav, 12, 86.

Kee, M. K., Lee, J. H., Kim, E. J., Lee, J., Nam, J. G., Yoo, B. H., & Kim, S. S. (2009). Improvement in survival among HIV-infected individuals in the Republic of Korea: Need for an early HIV diagnosis. *BMC Infectious Diseases*, *9*. https://doi.org/10.1186/1471-2334-9-128

Lee, M., & Rotheram-Borus, M. J. (2001). Challenges Associated With Increased Survival Among Parents Living With HIV. *American Journal of Public Health*, *91*(8), 1303–1309. https://doi.org/10.2105/AJPH.91.8.1303

Li, N., Spiegelman, D., Drain, P., Mwiru, R. S., Mugusi, F., Chalamilla, G., & Fawzi, W. W. (2012). Predictors of weight loss after HAART initiation among HIV-infected adults in Tanzania. *Aids*, *26*(5), 577-585.

Li, X., Ding, H., Geng, W. *et al.* (2019). Predictive effects of body mass index on immune reconstitution among HIV-infected HAART users in China. *BMC Infect Dis* **19,** 373 (2019). https://doi.org/10.1186/s12879-019-3991-6.

Maposa, I. (2016). *Survival Modelling and Analysis of Hiv/Aids Patients on Hiv Care And Antiretroviral Treatment to Determine Longevity Prognostic Factors*. Cape Town, South Africa: University of the Western Cape. http://hdl.handle.net/11394/5444

Mecha, J. O., Kubo, E. N., Nganga, L. W., Muiruri, P. N., Njagi, L. N., Mutisya, I. N., Odionyi, J. J., Ilovi, S. C., Wambui, M., Githu, C., Ngethe, R., Obimbo, E. M., & Ngumi, Z. W. (2016). Trends in clinical characteristics and outcomes of Pre-ART care at a large HIV clinic in Nairobi, Kenya: a retrospective cohort study. *AIDS Research and Therapy*, *13*(1), 38. https://doi.org/10.1186/s12981-016-0122-y

Mzileni, M. O., Longo-Mbenza, B., & Chephe, T. J. (2008). Mortality and causes of death in HIV-positive patients receiving antiretroviral therapy at Tshepang Clinic in Doctor George Mukhari Hospital. *Pol Arch Med Wewn, , 118*(10), 548-554.

NASCOP. (2016). *Kenya AIDS response progress report*. Nairobi, Kenya: Ministry of Health.

Nischang, M., Sutmuller, R., Gers-Huber, G., Audigé, A., Li, D., Rochat, M. A., ... & Amssoms, K. (2012). Humanized mice recapitulate key features of HIV-1 infection: a novel concept using long-acting anti-retroviral drugs for treating HIV-1. *PloS one*, *7*(6), 38853.

Palombi, L., Mancinelli, S., Liotta, G., Narciso, P., & Marazzi, M. C. (1997). The impact of socio-economic factors, mental health and functional status on survival in a sample of AIDS patients. *Aids Care*, *9*(6), 671-680.

Kelkile, D. P. (2016). Statistical analysis of adult HIV/AIDS patients and modelling of AIDS disease progression. *Science Journal of Applied Mathematics and Statistics*, 4(5), 189.

Piulachs Lozada-Benavente, X., Alemany Leira, R., Guillén, M., & Rizopoulos, D. (2017). Joint models for longitudinal counts and left-truncated time-to-event data with applications to health insurance. *Sort (Statistics and Operations Research Transactions), 2017, vol. 41, num. 2, p. 347-372*.

Rai, S., Mahapatra, B., Sircar, S., Raj, P. Y., Venkatesh, S., Shaukat, M., & Rewari, B. B. (2013). Adherence to antiretroviral therapy and its effect on survival of HIV infected individuals in Jharkhand, India. *PLoS One*, *8*(6), 66860.

Shebeshi, D. (2011). *Survival analysis of adult HIV/aids patients and stochastic modelling of aids disease progression*: a case study of jimma. Addis Ababa, Ethiopia: jimma University Specialized Hospital.

Sieleunou, I., Souleymanou, M., Schönenberger, A. M., Menten, J., & Boelaert, M. (2009). Determinants of survival in AIDS patients on antiretroviral therapy in a rural centre in the Far-North Province, Cameroon. *Tropical Medicine & International Health, 14*(1), 36-43.

Sika, A. M., Wools-Kaloustian, K., Mwangi, A. W., Kimaiyo, S. N., Diero, L. O., Ayuo, P. O., ... & Musick, B. (2010). Risk factors for death in HIV-infected adult african patients recieving anti-retroviral therapy. *East African Medical Journa, 87(11).*

Stewart, A., Chan Carusone, S., To, K., Schaefer-McDaniel, N., Halman, M., & Grimes, R. (2012). Causes of death in HIV patients and the evolution of an AIDS hospice: 1988–2008. *AIDS research and treatment*, *2012*.

Sungkanuparph, S., Eampokalap, B., Chottanapund, S., Thongyen, S., & Manosuthi, W. (2007). Impact of drug-resistant tuberculosis on the survival of HIV-infected patients. *International Journal of Tuberculosis and Lung Disease, 11*(3), 325–330.

Tadesse, K., Haile, F., & Hiruy, N. (2014). Predictors of mortality among patients enrolled on antiretroviral therapy in Aksum hospital, northern Ethiopia: a retrospective cohort study. *PloS one, 9*(1), e87392.

Tsegaye, E., & Worku, A. (2011). Assessment of antiretroviral treatment outcome in public hospitals, South Nations Nationalities and Peoples Region, Ethiopia. *Ethiopian Journal of Health Development*, *25*(2), 251911–471276.

Tsehaineh, B. (2010). *Assessment of Factors Associated With High Risk of Mortality of HIV Patients Treated With Highly Active Antiretroviral Therapy in Jimma Zone, South Western Ethiopia: Application of Survival Analysis Methods.* Adiss ababa, Ethiopia: Addis Abeba university.

UNAIDS. (2019). Report on the Global HIV. AIDS epidemic. *AIDS, 33, 203-211.*

Zekan, Š., Novotny, T. E., & Begovac, J. (2008). Unsafe sexual behavior among HIV-infected patients in Croatia, 2006: *Prevalence and associated factors. AIDS and Behavior, 12*(1), 86. https://doi.org/10.1007/s10461-008-9420-2

APPENDICES

Appendix 1: IREC Approval Letter

MU/MTRH-INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE (IREC)

MOI TEACHING AND REFERRAL HOSPITAL
P.O. BOX 3
ELDORET
Tel: 33471//2/3

MOI UNIVERSITY
COLLEGE OF HEALTH SCIENCES
P.O. BOX 4606
ELDORET

Reference: IREC/2018/147
Approval Number: 0003206

31st January, 2019

INSTITUTIONAL RESEARCH & ETHICS COMMITTEE
31 JAN 2019
APPROVED
P. O. Box 4606-30100 ELDORET

Robert Kibichii Mengich,
Moi University,
School of Medicine,
P.O. Box 4606-30100
ELDORET-KENYA.

Dear Mr. Mengich,

RE: FORMAL APPROVAL

The MU/MTRH- Institutional Research and Ethics Committee has reviewed your research proposal titled: -

"Survival Analysis of Adult HIV/AIDS Patients: A Case Study of Moi Teaching and Referral Hospital, Kenya".

Your proposal has been granted a Formal Approval Number: ***FAN: IREC 3206*** on 31st January, 2019. You are therefore permitted to begin your investigations.

Note that this approval is for 1 year; hence will expire on 30th January, 2020. If it is necessary to continue with this research beyond the expiry date, a request for continuation should be made in writing to IREC Secretariat two months prior to the expiry date. You will be required to submit progress report(s) on application for continuation, at the end of the study and any other times as may be recommended by the Committee.

Furthermore, you must notify the Committee of any proposal change (s) or amendment (s), serious or unexpected outcomes related to the conduct of the study, or study termination for any reason. You will also be required to seek further clearance from any other regulatory body/authority that may be appropriate and applicable to the conduct of this study.

Sincerely,

DR. S. NYABERA
DEPUTY-CHAIRMAN
INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE

cc	CEO	-	MTRH	Dean	-	SOP	Dean	-	SOM
	Principal	-	CHS	Dean	-	SON	Dean	-	SOD

Appendix 2: MTRH approval letter.

An ISO 9001:2015 Certified Hospital

MOI TEACHING AND REFERRAL HOSPITAL

Telephone :(+254)053-2033471/2/3/4
Mobile: 722-201277/0722-209795/0734-600461/0734-683361
Fax: 053-2061749
Email: ceo@mtrh.go.ke/directorsofficemtrh@gmail.com

Nandi Road
P.O. Box 3 – 30100
ELDORET, KENYA

Ref: ELD/MTRH/R&P/10/2/V.2/2010

6th February, 2019

Robert Kibichii Mengich,
Moi University,
School of Medicine,
P.O. Box 4606-30100,
ELDORET-KENYA.

APPROVAL TO CONDUCT RESEARCH AT MTRH

Upon obtaining approval from the Institutional Research and Ethics Committee (IREC) to conduct your research proposal titled:-

"Survival Analysis of Adult HIV/AIDS Patients: A Case Study of Moi Teaching and Referral Hospital, Kenya".

You are hereby permitted to commence your investigation at Moi Teaching and Referral Hospital.

06/02/2019
DR. WILSON K. ARUASA, *MBS*
CHIEF EXECUTIVE OFFICER
MOI TEACHING AND REFERRAL HOSPITAL

MOI TEACHING AND REFERRAL HOSPITAL
CEO
APPROVED
06 FEB 2019
SIGN.....
P. O. Box 3 - 30100, ELDORET

cc - Senior Director, (CS)
- Director of Nursing Services (DNS)
- HOD, HRISM

Appendix 3: SAS codes

```
libname men 'D:\Mengich\Data';

PROC IMPORT OUT= WORK.men
    DATAFILE= "D:\Mengich\Data\cleaned data.xlsx"
    DBMS=XLSX REPLACE;
  GETNAMES=YES;
  DATAROW=2;
RUN;

PROC IMPORT OUT= WORK.kib
    DATAFILE= "D:\Mengich\Data\kib data.csv"
    DBMS=CSV REPLACE;
  GETNAMES=YES;
  DATAROW=2;
RUN;

*Check for duplicates;
proc sql; select count (distinct patient_identifier) into: num from men;quit;
*Format dates;
/*data men0;set men;format Date_of_last_clinic_encounter date9.;run;*/

*Explore dataset;
proc freq data = men;tables male_gender ARV_perfect_adherence Death
WHOStage___enrolment /missing;run;

proc means data = men n nmiss min max mean std median p25 p75 maxdec= 2;
var School_Years_Completed CD4___enrolment age_at_therapeutic_ARV_start
age_at_enrollment Weight___enrolment;
run;

proc sql; select count (patient_identifier) into: num from men where
School_Years_Completed = 0;quit;

proc sql; select count (patient_identifier) into: num from men where
WHOStage___enrolment gt 3 | 0 < CD4___enrolment <200;quit;
proc sql; select count (patient_identifier) into: num from men where
WHOStage___enrolment = 4 ;quit;
proc sql; select count (patient_identifier) into: num from men where 0 <
CD4___enrolment <200 ;quit;
proc print data = men (obs=10);var patient_identifier WHOStage___enrolment
CD4___enrolment;where WHOStage___enrolment gt 3 | 0 < CD4___enrolment
<200;run;

*Derive analysis variables;
data men0;
set men;

*Create age categories (@ art therapeutic);
```

```
if age_at_therapeutic_ARV_start ge 18 and age_at_therapeutic_ARV_start lt 30 then agecat_therap = 1;
if age_at_therapeutic_ARV_start ge 30 and age_at_therapeutic_ARV_start lt 40 then agecat_therap = 2;
if age_at_therapeutic_ARV_start ge 40 and age_at_therapeutic_ARV_start lt 55 then agecat_therap = 3;
if age_at_therapeutic_ARV_start ge 55 then agecat_therap = 4;

*Create age categories (@ enrollment);
if age_at_enrollment ge 18 and age_at_enrollment lt 30 then agecat_enroll = 1;
if age_at_enrollment ge 30 and age_at_enrollment lt 40 then agecat_enroll = 2;
if age_at_enrollment ge 40 and age_at_enrollment lt 50 then agecat_enroll = 3;
if age_at_enrollment ge 50 then agecat_enroll = 4;

if age_at_enrollment ge 18 and age_at_enrollment lt 30 then agecat_enroll0 = 1;
if age_at_enrollment ge 30 and age_at_enrollment lt 45 then agecat_enroll0 = 2;
if age_at_enrollment ge 45 then agecat_enroll0 = 3;

*exlcude patients with WHOStage greater than 3 or cd4 less than 200;
if WHOStage___enrolment eq 4 or (CD4___enrolment ne . & CD4___enrolment lt 200) then delete;*339 with WHO = 4 & 4218 cd4 total = 4303(overlap);

*create the HIV severity variable;
if WHOStage___enrolment in (2,3) | (0 < CD4___enrolment <= 350) then H = 1;
if WHOStage___enrolment =1 | CD4___enrolment gt 350      then H = 0;

*create education level variables;
if 0 < School_Years_Completed <=8 then educ_level = 1;
if 8 < School_Years_Completed <=12 then educ_level = 2;
if School_Years_Completed gt 12   then educ_level = 3;
run;

*Explore dataset;
proc freq data = men0;tables male_gender ARV_perfect_adherence Death WHOStage___enrolment /missing;run;
proc means data = men0 n nmiss min max mean std median p25 p75 maxdec= 2;
var School_Years_Completed CD4___enrolment age_at_therapeutic_ARV_start age_at_enrollment Weight___enrolment __days_prior_to_therapeutic_ARV __days_post_therapeutic_ARV_star;
run;
proc freq data = men0;tables agecat_therap agecat_enroll agecat_enroll0 H/missing;run;

*Create bmi at enrollment dataset;
*check visit date status,min and max - 01JAN2005 to 09FEB2018;
proc tabulate data = kib;var encounter_date;table encounter_date,n nmiss (min max median)*f=date9. range;run;

proc sort data = kib;by Patient_Identifier encounter_date;run;
```

```
data bmi(rename=Body_Mass_Index = bmi_b drop = encounter_date);
set kib (keep = Patient_Identifier encounter_date Body_Mass_Index);
by Patient_Identifier encounter_date;
if Body_Mass_Index ne .;
if first.Patient_Identifier;

*create bmi_cat;
if Body_Mass_Index ne . & (Body_Mass_Index lt 15 or Body_Mass_Index gt 40) then delete;
if Body_Mass_Index ne . & Body_Mass_Index lt 18.5 then bmi_cat = 2;
if Body_Mass_Index ge 18.5 and Body_Mass_Index lt 25 then bmi_cat = 1;
if Body_Mass_Index ge 25 and Body_Mass_Index lt 30 then bmi_cat = 3;
if Body_Mass_Index ge 30 then bmi_cat = 4;
run;

proc freq data = men0;tables agecat_therap agecat_enroll H/missing;run;
proc means data = bmi n nmiss min max mean std median p25 p75 maxdec= 2;var bmi_b ;run;
proc freq data = bmi;tables bmi_cat/missing;run;

*Merge data with the rest of the datasets and keep the analysis variables;
data men1 (rename=(__days_post_therapeutic_ARV_star=t_days WHOStage___enrolment = WHOStage_b CD4___enrolment = CD4_b));
merge men0(keep = Patient_Identifier male_gender ARV_perfect_adherence agecat_therap agecat_enroll H __days_post_therapeutic_ARV_star WHOStage___enrolment CD4___enrolment School_Years_Completed Death agecat_enroll0 educ_level age_at_therapeutic_ARV_start age_at_enrollment in=a) bmi;
by Patient_Identifier;
if a;

*create time object in years;
if __days_post_therapeutic_ARV_star eq . then delete;
t_yrs = __days_post_therapeutic_ARV_star/365.25;
if __days_post_therapeutic_ARV_star = 0 then delete;*27 patients;

*Create CD4 category;
if 200 <= CD4___enrolment <= 350 then cd4_cat = 0;
if CD4___enrolment gt 350 then cd4_cat = 1;
run;

*Check for duplicates;
proc sql; select count (distinct patient_identifier) into: num from men1;quit;
proc means data = men1 n nmiss min max mean std median p25 p75 maxdec= 2;var School_Years_Completed t_days t_yrs ;run;
proc sql; select count (patient_identifier) into: num from men1 where t_yrs = 0;quit;
proc sql; select count (patient_identifier) into: num from men1 where CD4_b = 200;quit;
proc sql; select count (patient_identifier) into: num from men1 where School_Years_Completed = 0;quit;
```

```
proc freq data = men1;tables Death H male_gender educ_level cd4_cat/missing;run;

*Export Analysis Dataset;
PROC EXPORT DATA= work.men1
     OUTFILE= "D:\Mengich\Data\men.csv"
     DBMS=CSV REPLACE;
   PUTNAMES=YES;
RUN;
```

Appendix 4: R-codes.

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

```{r,warning=FALSE, message=FALSE,echo=F}
#LOAD THE NECESSARY PACKAGES
library (plyr)
library (readr)
library(Hmisc)
library(dplyr)
library(broom)
library(tidyr)
library(knitr)
library(kableExtra)
library(MASS)
library(car)
library(boot)
library(MatchIt)
library(ggplot2)
library(arm)
library(epiDisplay)
library(png)
library(tidyverse)
library(grid)
library(gridExtra)
library(survival)
```

```{r,warning=FALSE,message=FALSE,echo=F}
rm(list = ls())
#Import analysis dataset
m = read_csv("D:/Mengich/Data/men25Jul19.csv")
#table(m$male_gender)
#combine tail 4th bmicat with the 3rd bmicat
```

```
m$bmicat_n[m$bmi_cat == 4] = 3
m$bmicat_n[m$bmi_cat == 3] = 3
m$bmicat_n[m$bmi_cat == 2] = 2
m$bmicat_n[m$bmi_cat == 1] = 1
#combine tail 4th agecat_enroll with the 3rd agecat_enroll
m$agecat_enroll_n[m$agecat_enroll == 4] = 3
m$agecat_enroll_n[m$agecat_enroll == 3] = 3
m$agecat_enroll_n[m$agecat_enroll == 2] = 2
m$agecat_enroll_n[m$agecat_enroll == 1] = 1
#combine tail 4th agecat_enroll with the 3rd agecat_enroll
m$educ_level_n[m$educ_level == 3] = 2
m$educ_level_n[m$educ_level == 2] = 2
m$educ_level_n[m$educ_level == 1] = 1
#Fit cox regression
m$agecat_enroll = as.factor(m$agecat_enroll)
m$agecat_enroll0 = as.factor(m$agecat_enroll0)
m$agecat_enroll_n = as.factor(m$agecat_enroll_n)
m$bmicat_n = as.factor(m$bmicat_n)
m$educ_level = as.factor(m$educ_level)
m$educ_level_n = as.factor(m$educ_level_n)
```

```{r,warning=FALSE,message=FALSE,echo=F}
#By Gender
Male  = summary(m$t_yrs_[m$male_gender==1])
Female = summary(m$t_yrs_[m$male_gender==0])
Overall = summary(m$t_yrs_)
Summ_survtime = data.frame(round(rbind(Male,Female,Overall),3))
colnames(Summ_survtime) = c("Min","1st Quart.","Median","Mean","3rd Quart.","Max")
kable(Summ_survtime,caption = "Summary of Survival Time")
p1 = ggplot(m, aes(x = male_gender , y = t_yrs_)) + geom_boxplot() + labs(x="Gender",y="Survival time (years)") + ggtitle("Box plot for Survival time by Gender") + facet_wrap(~Death)
```

```{r,warning=FALSE,message=FALSE,echo=F}
library(survival)
m$T = Surv(m$t_yrs_,event = m$Death)
km <- survfit(T ~ 1, data=m)
print(km)
plot( km,
main="Overall Kaplan Meier Curve",
ylab="Survival Probability", xlab="Time(Years)" )
# legend("topright", c("AG-", "AG+"), col=c("red", "blue"),
#        lty=1, lwd=2, bty="n")
abline(h=0.5, lty=2)
```

```{r,warning=FALSE,message=FALSE,echo=F}
m$T = Surv(m$t_yrs_,event = m$Death)
gsurv = survfit(T~male_gender,data = m)
#summary(gsurv)
# plot of cumulative survival probabilities
pdfind = FALSE
if (pdfind) { pdf(file = "figure1A.pdf", width = 10.2, height = 5) }
 par(oma = c(2, 2, 0.5, 0.5), mar = c(2, 2, 0, 0))
 plot(gsurv, lty = 1:2, las = 1, lwd = 2)
 mtext(side = 1, line = 2.5, text = "Time (years)", cex = 1.2)
 mtext(side = 2, line = 3, text = "Survival Probability", cex = 1.2)
 #mtext("mtext", side=3, at=1.5, line=1, cex=2)
 legend("bottomright", title = "Gender", legend = c("Female", "Male"),
    lty = 1:2, inset = 0.05, bty = "n", cex = 1.0)
  abline(h=0.5, lty=2)
if (pdfind) { dev.off() }
```

```{r,warning=FALSE,message=FALSE,echo=F}
#Are the survival curves different?
#Log rank test ;
#H0: No difference between the survival curves vs H1: The survival curves are not statistically equivalent
```

```
#survdiff(T~male_gender,data = m)
# The p-value = <0.0001 is significant therefore we reject the null hypothesis - The survival curves are not statistically equivalent
```
```
```{r,warning=FALSE,message=FALSE,echo=F}
#Univariate Cox Regression
#HIV Severity
H = coxph(T~H,data = m,ties = "breslow")
#summary(H)
beta.H = round(summary(H)$coefficients[,1],3)
H.pval = round(summary(H)$coefficients[,5],3)
H.summ = (round(summary(H)$conf.int,3))[1:4][-2]
H.summ_ = c(beta.H,H.summ,H.pval)

#Age Category at Enrollment
agecat = coxph(T~agecat_enroll_n,data = m,ties = "breslow")
#summary(agecat)
beta.agecat = round(summary(agecat)$coefficients[1,1],3)
agecat.pval = round(summary(agecat)$coefficients[1,5],3)
agecat.summ = (round(summary(agecat)$conf.int[1,],3))[1:4][-2]
agecat.summ_ = c(beta.agecat,agecat.summ,agecat.pval)

beta.agecat0 = round(summary(agecat)$coefficients[2,1],3)
agecat.pval0 = round(summary(agecat)$coefficients[2,5],3)
agecat.summ0 = (round(summary(agecat)$conf.int[2,],3))[1:4][-2]
agecat.summ0_ = c(beta.agecat0,agecat.summ0,agecat.pval0)

#BMI at Enrollment
bmi_cat = coxph(T~bmicat_n,data = m,ties = "breslow")
#summary(bmi_cat)
beta.bmi_cat = round(summary(bmi_cat)$coefficients[1,1],3)
bmi_cat.pval = round(summary(bmi_cat)$coefficients[1,5],3)
bmi_cat.summ = (round(summary(bmi_cat)$conf.int[1,],3))[1:4][-2]
bmi_cat.summ_ = c(beta.bmi_cat,bmi_cat.summ,bmi_cat.pval)
```
```

```
beta.bmi_cat0 = round(summary(bmi_cat)$coefficients[2,1],3)
bmi_cat.pval0 = round(summary(bmi_cat)$coefficients[2,5],3)
bmi_cat.summ0 = (round(summary(bmi_cat)$conf.int[2,],3))[1:4][-2]
bmi_cat.summ0_ = c(beta.bmi_cat0,bmi_cat.summ0,bmi_cat.pval0)

#School_Years_Completed at Enrollment
sch = coxph(T~School_Years_Completed,data = m,ties = "breslow")
#summary(sch)
beta.sch = round(summary(sch)$coefficients[1,1],3)
sch.pval = round(summary(sch)$coefficients[1,5],3)
sch.summ = (round(summary(sch)$conf.int[1,],3))[1:4][-2]
sch.summ_ = c(beta.sch,sch.summ,sch.pval)

#male_gender at Enrollment
sex = coxph(T~male_gender,data = m,ties = "breslow")
#summary(sex)
beta.sex = round(summary(sex)$coefficients[1,1],3)
sex.pval = round(summary(sex)$coefficients[1,5],3)
sex.summ = (round(summary(sex)$conf.int[1,],3))[1:4][-2]
sex.summ_ = c(beta.sex,sex.summ,sex.pval)

#ARV_perfect_adherence
adh = coxph(T~ARV_perfect_adherence,data = m,ties = "breslow")
#summary(sex)
beta.adh = round(summary(adh)$coefficients[1,1],3)
adh.pval = round(summary(adh)$coefficients[1,5],3)
adh.summ = (round(summary(adh)$conf.int[1,],3))[1:4][-2]
adh.summ_ = c(beta.adh,adh.summ,adh.pval)
```

```{r,warning=FALSE,message=FALSE,echo=F}
#summary of undjusted HRs
#create vector for reference groups
age_ref = c(rep('Ref',5))
```

```
bmi_ref = c(rep('Ref',5))
m.mat =
rbind(H.summ_,age_ref,agecat.summ_,agecat.summ0_,bmi_ref,bmi_cat.summ_,bmi_
cat.summ0_,sch.summ_,sex.summ_,adh.summ_)
rownames(m.mat) = c("HIV Severity:Yes","Age Cat-","Age Cat-","Age Cat-","BMI
Cat-","BMI Cat-","BMI Cat-","Sch Years
Comp.","Sex:Male","ARV_perfect_adherence")
undj.f = paste((m.mat[,2]), '(', m.mat[,3],',',m.mat[,4], ') ',sep = "" )
```
```
```{r,warning=FALSE,message=FALSE,echo=F}
#Adjusted Cox Regression
#r2 = coxph(T ~ H + agecat_enroll_n + bmicat_n + School_Years_Completed +
male_gender,data = m,ties = "breslow")
r2 = coxph(T ~ H + agecat_enroll_n + bmicat_n + School_Years_Completed +
male_gender + ARV_perfect_adherence,data = m,ties = "breslow")
#summary(r2)
beta.r2 = round(summary(r2)$coefficients[,1],3)
r2.pval = round(summary(r2)$coefficients[,5],3)
r2.summ = (round(summary(r2)$conf.int,3))[,1:4][,-2]
r2.summ_ = cbind(beta.r2,r2.summ,r2.pval)
r2.mat  = matrix(r2.summ_,8,5)
ref_mat = matrix("Ref",nrow=10,ncol=5)
ref_vec = c(2,5)
ref_mat[-ref_vec,] = r2.mat
#ref_mat
colnames(ref_mat) = c("Beta","Hazard Ratio","2.5% C.I", "97.5% C.I","p.value")
rownames(ref_mat) = c("HIV Severity:Yes","Age Cat-","Age Cat-","Age Cat-","BMI
Cat-","BMI Cat-","BMI Cat-","Sch Years
Comp.","Sex:Male","ARV_perfect_adherence:Yes")
#summary of adjusted HRs
adj.f = paste((ref_mat[,2]), '(', ref_mat[,3],',',ref_mat[,4], ') ',sep = "" )

#combine undjusted and adjusted HRs
hr = cbind(undj.f,m.mat[,5],adj.f,ref_mat[,5])
```
```

```
rownames(hr) = c("HIV Severity:Yes","Age Cat-<30","Age Cat-30-40","Age Cat-40+","BMI Cat-18.5-<25","BMI Cat-<18.5","BMI Cat-25+","Sch Years Comp.","Sex:Male","ARV_perfect_adherence:Yes")
colnames(hr) = c("Unadjusted HR (95% CI)","P-value","Adjusted HR (95% CI)","P-value")
hr[,2][hr[,2] == 0] = "<.0001"
kable(hr,caption = "Unadjusted/Adjusted Hazard Ratios")
```

```{r,warning=FALSE,message=FALSE,echo=F}
```

```{r,warning=FALSE,message=FALSE,echo=F}
```