# UNIVERSITY OF NAIROBI


## FACULTY OF SCIENCE AND TECHNOLOGY

# Modelling Survival among Tuberculosis Patients in Kenya.

BY

**SHEILA AWUOR ONYANGO**

**I56/37320/2020**

A Thesis Submitted to the Department of Mathematics for Examination in Partial Fulfillment of the Requirements for the Award of Degree of Master of Science in Biometry of the University of Nairobi.

November,2022

**Master Project in Biometry** **University of Nairobi**

**November 2022**

# Modelling Survival among Tuberculosis Patients in Kenya

**Research Report in Biometry, 2022**

Sheila Onyango I56/37320/2020

Department of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Thesis

Submitted to the Department of Mathematics in partial fulfilment for a degree in Master of Science in Biometry

Submitted to: The Graduate School, University of Nairobi, Kenya

# Abstract

The airborne illness tuberculosis (TB), which mostly affects the lungs, is a substantial source of mortality when the right controls are not implemented. Tb is preventable and treatable, but despite this, close to 10 million individuals contract it annually throughout the world. This study examined factors related to survival time in tuberculosis patients from Kenya. This study compared survival rates across Treatment phases and several sub-counties. The study also sought to determine factors that predict survival among Kenyan patients with tuberculosis.

The study employed secondary population-based tuberculosis data from electronic database Treatment Information from Basic Unit (TIBU). To compare survival function for Treatment stages and the different sub-counties, the Kaplan-Meier estimation method was utilized. Independent variables against survival time were modeled using the Cox proportional hazards regression model .

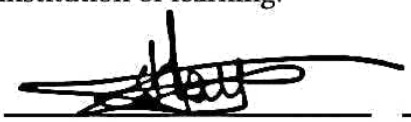
The results of the Cox proportional hazards regression examining the relationship between predictors and survival time were as follows: Age, HR = 1.0241 (CI: 1.0028 – 1.046). Consequently, the risk of death for someone with tuberculosis is 2.4% higher for every additional year. HIV, HR = 3.0812(CI: 1.6508 – 5.751). When diagnosed with tuberculosis, HIV positive people had a 3 times higher mortality rate than HIV negative individuals. TB type and Alcohol consumption were significant predictors of death in TB patients. Kaplan-Meier estimation method shows that survival between Intensive and Continuation Phase were significantly different as it had a p-value of <0.0001. Results from Kaplan-Meier data, however, did not reveal any significant variations in survival in the four sub-counties'.

**Master Thesis in Mathematics at the University of Nairobi, Kenya.**
**ISSN 2410-1397: Research Report in Mathematics**

**DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya**

# Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

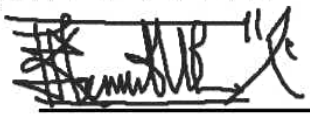


01/31/2023

Signature Date

SHEILA ONYANGO
Reg No. I56/37320/2020

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.



06/02/2023

Signature Date

Dr Isaac Kipchirchir
Department of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: kipchirchir@uonbi.ac.ke

# Dedication

My wholehearted gratitude to my research supervisor: Dr. Isaac Kipchirchir for his guidance and professional support during the entire course of the project. I sincerely appreciate the support from my family; Dad, Mum and Kuyuti thank you very much for your selfless support throughout my entire course, may God abundantly bless you. Special credit to Wekunda et al for allowing me to use the data he had collected for a study on Determinants of Tuberculosis Treatment Interruption among patients.

# Contents

# List of Tables and Figures

## Figures



## Tables

# CHAPTER ONE

This chapter gives background information on Tuberculosis, the objectives of the study, problem statement and significance of the study.

## 1.1 Background Information

Tuberculosis (TB) remains a major global health problem. Fatal outcomes including TB recurrence, death and Multidrug resistance TB (MDR-TB) which occur during and after treatment is a big challenge to tuberculosis control. There is a 10% risk of getting infected with TB but in the presence of weakened immunity (this include people living with HIV, malnutrition or diabetes, or people who smoke) the risk of TB increases. According to the World Health Organization (WHO) 2020 annual Tuberculosis report, there were 1.5 million deaths occurring yearly. This makes Tuberculosis the top killer infectious disease in the world. Despite the fact that Tuberculosis is curable and preventable, nearly 10 million people get it yearly (Harding, 2020).

TB is second only to HIV/AIDS as the leading killer in the world because of a single infectious agent (WHO, 2015). In 1993, the World Health Organization (WHO) declared tuberculosis a global emergency. Bettering TB treatment results was part of the Millennium Development Goals. There has still been high frequency of TB mortality despite productive treatment. 95% of TB deaths occur in low and middle income countries (WHO, Global Tuberculosis Report, 2015). In 2020, an estimated 10 million people fell ill with tuberculosis (TB) worldwide. 3.2 million females, 7 million males and 1.2 million children. TB is present in all countries and age groups but TB is curable and preventable.

In 2013, the Division of Leprosy, Tuberculosis and Lung Disease (DLTLD) published guidelines for the management of TB and leprosy in the country. Tuberculosis treatment involves the use of several drugs taken in combination. The recommended regimen for treatment for new SM+ adults entails two months of intensive phase treatment which is then followed by a continuation phase for four months.

Kenya is ranked 15th up from 13th position among the 22 high TB burden countries worldwide which contributes 80% of the global TB (World Health Organization (WHO), 2013). Tuberculosis remains a major cause of morbidity and mortality in Kenya. It affects all age groups, but has it's greatest toll in the most productive age group of 15 to 44 years. The major factor responsible for the large TB disease burden in Kenya is the concurrent HIV epidemic. Other factors that have contributed to this large TB disease burden in-

clude poverty and social deprivation that have led to mushrooming of peri-urban slums, congestion and limited access to general health services. Recently, there have been increasing concerns about the emergence of drug resistant TB, a threat that would pose major challenges in the fight against TB in resource limited countries like Kenya.

The Government of Kenya (GOK) is implementing all the six components of STOP TB strategy which include: Pursue high-quality DOTS expansion and enhancement, Address TB-HIV, MDR- TB, and the needs of poor and vulnerable populations, Contribute to health system strengthening based on primary health care, Engage all care providers, Empower people with TB, and communities through partnership and finally enable and promote research. The target of Sustainable Development Goals (SDGs) is to attain 80% decrease in TB incidence, 90% decrease in TB related deaths by 2030. For there to be a realisation of vision 2030, Tuberculosis (TB) monitoring will be key to that achievement.

## 1.2 Problem Statement

Tuberculosis (TB) comes second to HIV/AIDS as one of the top killer disease. In 2012, 8.6 million people were infected with TB and 1.3 million deaths occurred as a result of TB and over 95% of the deaths occurred in third world countries. It is among the top three causes of death for women aged 15 to 44. Tuberculosis disease in pregnant women is known to cause unfavourable outcomes in pregnancy. These outcomes include six-fold increase in perinatal mortality mainly in females co-infected with HIV, premature birth and low birth weight. TB is a leading killer of people living with HIV causing one fifth of all deaths.(World Health Organization (WHO), 2012). HIV enables the quick progression of Mycobacterium tuberculosis infection to active TB disease, among both people who have acquired infection recently and also in individuals with latent infection. The higher morbidity and mortality due to TB and HIV co-infection emphasizes the need for prevention, early diagnosis and treatment of TB among People Living with HIV/AIDS.

Successful treatment of TB involves taking anti-TB drugs for 2 months during intensive phase and 4-6 months in the continuation phase but many find it difficult to complete their course of treatment and this serves as a major constraint to eradicating the disease (Cuneo, 1989). Development of MDR TB is sometimes caused by default from treatment, that is quite costly to manage and can quickly progress into X-DR TB which is fatal. It costs an estimate of KSh. 2 million to manage MDR TB in an individual. This increases the burden to the already stretched health services in Kenya, thus there is need for the country to understand and address factors contributing to default and mortality.

## 1.3 Objectives of the Study

### Main Objective

The broad objective of this study is to find out factors associated with survival time among patients with Tuberculosis in Kenya.

### Specific Objectives

The specific objectives are:

i) To compare survival among TB patients in different sub-counties Kenya.

ii) To compare survival among TB patients between different Treatment stages.

iii) To investigate the predictors of survival in Tuberculosis patients in Kenya.

## 1.4 Significance of the study

This research attempts to understand the survival pattern of Tuberculosis over time. Tb treatment takes 6 months now compared to 8 months previously hence understanding the survival pattern over time is key for the different counties in Kenya which have implemented various measures in Tuberculosis management. Understanding the predictors and survival pattern is very important as this will inform counties and the country in general on implementation TB control policies in Kenya. Further, towards the actualisation of the national strategy plan which was launched in 2018. The plan seeks to control TB by successful treating at the very least 597,000 cases by the year 2023. It will inform the attentive monitoring of patients at the county level to fasten treatment successes.

# CHAPTER 2: LITERATURE REVIEW

A number of studies have been carried out to model survival among TB patients. This chapter provides a summary of studies that have been done previously.

Pardeshi (2009) conducted a study whose main objective was to determine the pattern of survival among tuberculosis patients undergoing treatment categorized by age and sex of patients in Yavatmal, India. The analysis utilized Kaplan Meier plots and log rank tests. Multivariate analysis was performed using Cox proportional hazards model. A sample of 716 patients registered in the Tb clinic was used for analysis. At the conclusion of the Intensive phase the survival rates were 96%, 93% and 99% in respectively in the three categories. Log rank statistic was 7.26 with a p-value of less than 0.05 showing that there was a significant difference in the survival among the three categories and also among the different age groups where Log rank statistic was 8.76 with a p-value of less than 0.05. There was no difference in the survival curves of the sex of patients. Cox proportional hazard analysis showed that patients with age greater than 40 years had a high risk of death among patients with Tuberculosis.

A retrospective cohort study which used secondary data collected on hospital registers was done in Yaounde, Cameroon. To investigate demographic and predictors of mortality Cox regression model was utilized. 337 (20%) defaulted from treatment, 86 (5.1%) died, treatment failure in 6 (0.4%) and 104 (6.2%) were transferred to a different health center among the 1188 patients. Median time to treatment default was 90 days and 62% of treatment default happened in the continuation phase. Admission during the intensive phase (adjusted HR 0.69; 14 95% CI 0.54 to 0.89) and lack of consent for HIV screening (1.65;1.24 to 2.21) were the main predictors of TB mortality in multi-variable analysis. (Yone et al., 2011).

A retrospective cohort study was done by Birlie et al. (2015) in Addis Ababa in 2014 in six health centers during the month of March. The objective of the study was to determine the time of deaths reported and factors associated with Tuberculosis among patients. 872 cases were registered in the hospital log book in the TB clinic. A sample of 810 cases was used in the study, 7.4% of the patients died while receiving treatment. 56.7% of the deaths were experienced in the Intensive phase of treatment compared to continuous phase of treatment. The study found that the overall rate of mortality was 12.8 per 1000 person months of observation and two months of treatment was established as the median time

of death. Age, body weight at baseline and HIV status were found to be the independent variables associated with death while in treatment.

Onyango et al. (2018) conducted a study where Bivariate and Cox proportional hazard modelling was used to investigate childhood Tuberculosis in Kenya. Survival analysis was done and death was the main outcome. A sample of 23,753 children who were aged 15 years or less were incorporated in the analysis. Data for registered patients during the years 2012 to 2015 was obtained from the Kenyan National Tuberculosis Program. Their results indicated that HIV infected but not on ARV, HIV infected and on ARV, Having pulmonary disease and being a child of 5 years and below with tuberculosis were the risk factors. Findings showed that children with TB accounted for 9% of the overall TB patients. The death rate for children with TB was at 4% and Tb treatment success for children with TB was at 90%. Mortality Predictor variables included, not taking ART while having HIV (4.84 is the adjusted hazard ratio; 95 percent confidence interval [CI], 3.59-6.51), being HIV positive and taking ART (3.14-4.35 at 95% CI and aHR, 3.69), having children under the age of five (aHR, 1.25; 95 percent CI, 1.08-1.44), and having pulmonary disease that is smear-negative (aHR, 1.68; 95 percent CI, 1.27-2.24).

Limenih and Workie (2019) examined the time to cure and the rates of survival for patients with tuberculosis that is multi-drug resistant in Ethiopia's Amhara region. Retrospective research was conducted at multi-drug resistant-TB treatment facilities between November 2015 and March 2018. The shared frailty and accelerated failure time models were used in the investigation. The study's goal was to look into how long it takes for MDR-TB patients to recover. As per the study, Patients with MDR-TB need 21 months to recover. 64.6% of the MDR-TB patients were cured while 35.4% were censored. According to the study, patients with extra pulmonary MDR-TB recovered more slowly than patients with seamier pulmonary MDR-TB. Additionally, this research showed that male MDR-TB patients, those having co-morbid conditions and those with clinical complications recovered more slowly than the control groups.

Abdullahi et al. (2019) performed a survival analysis of electronic TB monitoring data from Kilifi County, spanning five years (2012–2016). The outcome was an all-cause death within 180 days of starting treatment. The risk factors that were looked at were clinical characteristics at the time anti-TB medication was initiated. They carried out survival analysis with time at risk defined from the day that TB treatment began till the day of death, lost-to-follow-up or finishing dosage. To take into account the competing risk of "lost-to-follow-up", survival analysis was employed to look at mortality risk factors. The investigation discovered that majority of deaths happened within the initial 3 months and the rate of mortality increased in those receiving treatment from TB. The study compared the survival rates of those still living and perished using the log-rank test.

Another study conducted in Tanzania Leveri et al. (2019) which investigated predictors of death among Tuberculosis patients in Tanzania. The retrospective cohort study was done in Tanzania at Kibong'oto Infectious Diseases Hospital (KIDH). Patients' data which consisted of parameters that are demographic and clinical were obtained from the MDRTB registry and hospital files. Investigation of the factors influencing MDRTB treatment results was done using logistic regression. During the year 2019 to the year 2014 about 332 patients were found to have MDRTB and were started on treatment. Males were 67% and 95.48% were aged 18 years and above. More than 50% of the patients were residing in Dar es Salaam. There was an increase of 106 MDRTB patients from 2009 to 2014. Close to 75.7% of the patients had positive treatment outcomes. It was also established that low BMI, presence of cavities in the chest were significantly associated with death in Tuberculosis patients. It was observed that patients who are smokers and those on Ethambutol had a high risk of death due to Tuberculosis.

Ranzani et al. (2020) carried out a study by describing a five year survival for patients who were newly diagnosed with TB in Sao Paulo, Brazil. Population was selected from electronic system TB web. The main objective was to estimate the excess mortality from 2010 to 2015 and within the group who survived the first year in comparison with the Sao Paulo state population. Data on social vulnerability, drug usage and comorbidities was collected to estimate association with all-cause and cause-specific mortality. The association between the six exposures and the cohort's all-cause mortality was examined using the Cox proportional hazards model. In this population, it was observed that 17% of the patients died from tuberculosis within 5 years after diagnosis. Deaths was observed to be six times higher among tuberculosis patients compared to the original source population, and four-times greater in patients who survived the first year. Infection was established to be the cause for most deaths. Other predictors of lower survival was drug abuse, diabetes and homelessness.

Adeboye et al. (2020) conducted a study in Eastern Cape, South Africa, they used joint modelling which included Cox proportional hazard and a generalized logit model. A total of 729,463 teenagers and kids were enrolled in TB therapy; 2.5% passed away while receiving it. Between 2007 and 2016, the case fatality ratio dropped from 3.3% to 1.9%. It was discovered that mortality was lower in the age groups of 5 to 9 years compared to 0 to 4, 10 to 14, and 15 to 19 years. The mortality rate was shown to be higher in people who were HIV positive, had previously undergone TB treatment, and had extrapulmonary involvement. SMRs grew with time in the group of HIV-positive individuals and children aged 10 to 14. The findings revealed that age, smoking status, Diabetes were significant factors and also in Time to event data in adults with TB smoking status and Diabetes were significant factors. Better prediction power was shown by the joint model for Tuberculosis prognostic factors.

(Osman et al., 2021) undertook a 13-year study to determine the excess TB-related mortality among South African kids and teenagers. 729 463 kids and teenagers were documented receiving TB therapy between 2004 and 2016; While 2.5 percent (18 539) of patients who had TB treatment died, 84.0 percent had successful treatment outcomes. The case fatality ratio decreased from 3.3% to 1.9% between 2007 and 2016. In the multivariable Cox regression model, a higher risk of mortality was associated with ages 0-4, 10-14 and 15-19 (vs. ages 5 to 9), along with HIV infection, prior TB therapy, and extrapulmonary involvement.

Another study conducted in Nairobi Kenya, Gichuki and Mategula (2021) investigated using logistic regression the risk of dying as a result of tuberculosis. 930 (14.96%) of the 6218 deaths that occurred during the analysis period were TB-related. There were 62 TB fatalities on average each year (SD 23.9). Deaths from TB decreased from 21.2 percent in 2005 to 1.7 percent in 2016. The odds of dying from TB were 1.39 times greater in men than in women (AOR 1.39; 95 % CI 1.18-1.64; p-value 0.001). The risk of dying from TB was 42 percent lower in those over 50 than it was in people aged 30-39 (AOR 0.57; 95 percent confidence interval [CI] 0.47-0.73; p-value 0.001). 1.39 percent more people who died at home from TB than those who died in a hospital had this disease as their cause of death.(AOR: 1.93; 95% CI: 1.17-1.64; p = 0.001).

# CHAPTER 3: METHODOLOGY

## 3.1 Setting

The study was carried out in Vihiga County, Kenya. Vihiga County has four Tuberculosis treatment zones; Hamisi, Sabatia, Vihiga and Emuhaya. The observation unit for this study was the 4 sub-counties in Vihiga County. This was a retrospective cohort study of Tuberculosis cases diagnosed between 2015 to 2019 in twenty selected health care facilities in Vihiga County. The county comprises of seventy-one health care facilities located within four sub-counties.

## 3.2 Data Source

The study used population based data obtained from Treatment Information from Basic Unit (TIBU) electronic database. All cases are recorded in the relevant TB registers at hospital level and later entered into TIBU.

## 3.3 Data Management and analysis

Microsoft Excel, SAS 9.4 and R will be used to do statistical analysis. 95 percent confidence interval will be used for inferences.

## 3.4 Survival analysis

### 3.4.1 Introduction

In this study, survival analysis will be used. It is employed in time to happening of a specific event. The occurrence of death is the study's event. (Persson and Glimelius, 2002). It is also able to handle censoring which occur in our study due to below reasons:

1. Before the study's conclusion a participant is lost to follow-up
2. Throughout the study time, a study participant does not experience the study's event.

### 3.4.2 Survival Function, $S(t)$

$S(t)$ is the survival function that provides the likelihood of surviving past a specified period, is defined as

$$S(t) = P(T > t) = 1 - F(t) \tag{3.1}$$

where the lifetime random variable T's cumulative distribution function is denoted by F(t).

The properties of the survival function are as follows.

1. $S(t)$ properties are non-increasing.
2. The probability of remaining alive after time $t = 0$ is one.
3. $S(t) = 0$ at time $t = \infty$

### 3.4.3 The Hazard Function, $h(t)$

The function, indicated by $h(t)$, is the rate of instantaneous events occurring at time $t$ given a person has lived up until time $t$. It is given by

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \tag{3.2}$$

and is expressible as

$$h(t) = \frac{f(t)}{S(t)} \tag{3.3}$$

where $f(t)$ is the probability density function (pdf) of T.

### 3.4.4 Cumulative Hazard Function, $H(t)$

$h(t)$ is integrated to obtain it.

$$H(t) = \int_0^t h(u)du \tag{3.4}$$

This implies that

$$\frac{d}{dt}H(t) = h(t) \tag{3.5}$$

### 3.4.5 Kaplan Meir Estimator

It is a non-parametric estimator of the survival function. To compare survival in various sub-counties, the Kaplan-Meir survival function estimation method will be utilized.

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i} \tag{3.6}$$

$$\hat{V}ar(\hat{S}(t)) = (\hat{S}(t))^2 \sum_{t_i \leq t} \frac{d_i}{n_i(n_i - d_i)} \tag{3.7}$$

where
$n_i$ shows number of people at risk at time $t_i$
$d_i$ denotes number of individuals experiencing death at time $t_i$

Kaplan Meier terms are produced by multiplying the conditional probability that a person will survive past time $t_i$ given that a person has survived until time $t_i$

### 3.4.6 Survival Function; Test of Hypothesis

To check for equality of survival functions between groups we use either Wilcoxon or Log-rank test. Wilcoxon and Log-rank tests will be used to look for significant variations in survival across various sub-counties and treatment periods. The test statistic is given below:

$$\chi^2 = \frac{(\sum_i w_j(d_{ij} - \hat{e_{ij}}))^2}{\sum_i w_j^2 \hat{v_{ij}}} \tag{3.8}$$

where
$d_{ij}$ denotes the sum total of deaths that occurred in group $i$ at time $j$
$\hat{e_{ij}}$ denotes the deaths forecasted in group $i$ at time $j$
$\hat{v_{ij}}$ denotes the approximated variance for actual deaths $d_{ij}$
$w_j$ denotes the weighted difference at time $j$

For Wilcoxon text weighting is by the sum total of persons at risk at every interval therefore $w_j$ =$n_j$ while for the Log-rank test similar weighting is done at every interval hence $w_j = 1$.

### 3.4.7 Cox Proportional Hazard Regression Model

**Assumptions of Cox Proportional Regression Model**

1. There exists a multiplicative relationship between Hazard function and the predictor variables.
2. For different persons the hazards plots need to be proportionate and should never crisscross.

Cox Proportional Hazard Regression Model is used to model association between independent variables and survival. It will be used to determine predictors of survival among TB patients in Kenya It can be expressed as follows:

$$\log \frac{h(t)}{h_o(t)} = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \tag{3.9}$$

where
$t$ denotes survival time.
$h(t)$ denotes hazard function.
$h_o(t)$ denotes baseline hazard function.
$x_i$ denotes predictor variables.
$\beta_i$ denotes coefficients used to measure effect of the independent variables

Interpretation
$\beta_1$ > 0 implies the hazard ratio increases with an increase in the explanatory variable. This means there is a reduction in length of survival time.

$\beta_1$ < 0 implies the hazard ratio reduces as predictor variable reduces. This means that there is an increase in length of survival time.

$\beta_1$ = 0 implies that that the predictor has no impact on hazard.

**Model adequacy and Selection**
AIC is used to select a good fit model.

$$AIC = -2log(L) + 2(k + c + 1) \tag{3.10}$$

where
L denotes the likeihood value
k denotes the number of predictors in the model
c denotes the number of model distributional

The model that has the smallest AIC is the better fit model compared to the model with the largest AIC.

**Checking for Validity of Cox Proportional Regression Model**
Assessing the model's assumptions is a crucial step in model building since it is important to ensure that the model is valid.

**Propotional Hazards Assumption**
Schoenfield residuals is used in checking this assumption. The residuals are not time dependent hence for this assumption to be met the plot aof residuals against time displays a random pattern. In addition, a non-significant association of residuals with time backs the assumption.

# CHAPTER 4: DATA ANALYSIS AND RESULTS

## 4.1 Descriptive Statistics

From the 291 patients in this study, the gender proportions was 209 (71.8%) male and 82 (28.2%) female. Male patients showed higher percentage of death (10.91%) compared to female (4.47%).

**Table 4.1. Tuberculosis Patients Characteristics**

| Patient Characteristics | Total | % of the total deaths | No and %from total deaths |
|---|---|---|---|
| **Gender** | | | |
| Male | 209 (71.8%) | 10.91% | 32 (71.1%) |
| Female | 82 (28.2%) | 4.47% | 13 (28.9%) |
| Total | **291** | | 45 |
| **HIV Status** | | | |
| Positive | 101 (34.7%) | 9.28% | 27 (60%) |
| Negative | 190 (65.3%) | 6.91% | 18 (40%) |
| Total | **291** | | 45 |
| **TB Type** | | | |
| EPTB | 269 (92.8%) | 2.41% | 7 (15.6%) |
| PRB | 22 (7.2%) | 13.1% | 38 (84.4%) |
| Total | **291** | | 45 |
| **Alcohol Consumption** | | | |
| No | 138 (47.4%) | 4.81% | 14 (31.1%) |
| Yes | 153 (52.6%) | 10.7% | 31 (68.9%) |
| Total | **291** | | 45 |
| **Education Level** | | | |
| Post High School | 24 (8.25%) | 0.69% | 2 (4.44%) |
| Primary | 172 (59.1%) | 10.7% | 31 (68.9%) |
| High School | 95 (32.65%) | 4.12% | 12 (26.7%) |
| Total | **291** | | 45 |

A total of 24 (8.25%) patients had Post High School education, 172 (59.1%) had primary level of education and 95 (32.65%) had reached High School level of education. As shown in Table 4.1, 138 (47.4%) patients took alcohol while 153 (52.6%) did not take alcohol.

**Table 4.2. Count of Participants by different Treatment Phases**

| Outcomes | n | % |
|---|---|---|
| Complete Treatment | 212 | 72.9 |
| LTFU | 32 | 10.9 |
| Death | 45 | 15.5 |
| Treatment Failure | 2 | 0.7 |

**Table 4.3. Count of Participants by different Treatment Phases**

| Treatment Phase | n | % |
|---|---|---|
| Intensive Phase | 53 | 18.1 |
| Continuous Phase | 238 | 81.9 |

As shown on Table 4.2 and Table 4.3, A total of 212 (72.9%) completed treatment, 32 (10.9%) were lost to follow up, 45 (15.5%) died and 2(0.7%) failed treatment as shown in table (4.2). Most of the patients 238 (81.9%) were in the Continuation Phase while the remaining 53 (18.1%) were in the Intensive Phase.

**Table 4.4. Count of Participants in different Sub-Counties**

| Sub-counties | n | % |
|---|---|---|
| Emuhaya | 87 | 29.9 |
| Hamisi | 63 | 21.6 |
| Sabatia | 40 | 13.7 |
| Vihiga | 101 | 34.8 |

Participants were sampled from four different sub-counties within Vihiga County. Vihiga had the highest number of patients 101(34.8%) while Sabatia had the lowest percentage of patients at 40(13.7%).

## 4.2 Comparing survival between sub-Counties

Comparison of Survival Function between different sub-counties in Vihiga County

**Table 4.5. Kaplan Meier Estimates for Survival functions in sub-counties in Vihiga**

| Time | $n_t$ | $d_t$ | $\hat{S}(t)$ | $s.e(\hat{S}(t))$ | 95% CI; LL | 95% CI; UL |
|---|---|---|---|---|---|---|
| **E**muhaya | | | | | | |
| 2 | 87 | 1 | 0.989 | 0.0114 | 0.966 | 1.000 |
| 4 | 86 | 2 | 0.966 | 0.0196 | 0.928 | 1.000 |
| 5 | 84 | 1 | 0.954 | 0.0225 | 0.911 | 0.999 |
| 19 | 81 | 1 | 0.942 | 0.0251 | 0.894 | 0.993 |
| 25 | 80 | 1 | 0.930 | 0.0274 | 0.878 | 0.986 |
| 28 | 78 | 1 | 0.919 | 0.0295 | 0.862 | 0.978 |
| 41 | 72 | 2 | 0.893 | 0.0338 | 0.829 | 0.962 |
| 56 | 70 | 1 | 0.880 | 0.0356 | 0.813 | 0.953 |
| 58 | 69 | 1 | 0.868 | 0.0373 | 0.797 | 0.944 |
| 69 | 68 | 1 | 0.855 | 0.0389 | 0.782 | 0.934 |
| 139 | 65 | 1 | 0.842 | 0.0404 | 0.766 | 0.925 |
| **H**amisi | | | | | | |
| 2 | 63 | 2 | 0.968 | 0.0221 | 0.926 | 1.000 |
| 3 | 61 | 1 | 0.952 | 0.0268 | 0.901 | 1.000 |
| 11 | 60 | 1 | 0.937 | 0.0307 | 0.878 | 0.999 |
| 25 | 59 | 1 | 0.921 | 0.0341 | 0.856 | 0.990 |
| 30 | 57 | 1 | 0.904 | 0.0371 | 0.835 | 0.980 |
| 46 | 55 | 1 | 0.888 | 0.0399 | 0.813 | 0.970 |
| 64 | 50 | 1 | 0.870 | 0.0429 | 0.790 | 0.958 |
| 88 | 48 | 1 | 0.852 | 0.0456 | 0.767 | 0.946 |
| **S**abatia | | | | | | |
| 6 | 40 | 1 | 0.975 | 0.0247 | 0.928 | 1.000 |
| 25 | 38 | 1 | 0.949 | 0.0349 | 0.883 | 1.000 |
| 29 | 37 | 1 | 0.924 | 0.0424 | 0.844 | 1.000 |
| 54 | 35 | 1 | 0.897 | 0.0487 | 0.807 | 0.998 |
| 77 | 34 | 1 | 0.871 | 0.0539 | 0.771 | 0.98 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **V**ihiga | | | | | | |
| 2 | 101 | 2 | 0.980 | 0.0139 | 0.953 | 1.000 |
| 3 | 99 | 1 | 0.970 | 0.0169 | 0.938 | 1.000 |
| 5 | 98 | 1 | 0.960 | 0.0194 | 0.923 | 0.999 |
| 8 | 97 | 1 | 0.950 | 0.0216 | 0.909 | 0.994 |
| 17 | 94 | 2 | 0.930 | 0.0254 | 0.882 | 0.981 |
| 28 | 91 | 1 | 0.920 | 0.0271 | 0.868 | 0.975 |
| 35 | 90 | 1 | 0.910 | 0.0287 | 0.855 | 0.968 |
| 38 | 89 | 1 | 0.900 | 0.0301 | 0.842 | 0.961 |
| 39 | 88 | 1 | 0.889 | 0.0315 | 0.830 | 0.953 |
| 55 | 85 | 1 | 0.879 | 0.0328 | 0.817 | 0.946 |
| 71 | 83 | 1 | 0.868 | 0.0341 | 0.804 | 0.938 |
| 75 | 82 | 2 | 0.847 | 0.0364 | 0.779 | 0.922 |
| 124 | 77 | 2 | 0.825 | 0.0386 | 0.753 | 0.904 |
| 153 | 75 | 1 | 0.814 | 0.0396 | 0.740 | 0.896 |

Lost to follow-up observations or those that had already received treatment were subjected to right censoring. The survival table 4.5 is a descriptive table that shows the time till death. Incorporating both censored and uncensored data, as determined by observed survival and censored times, is the Kaplan-Meier estimator. Figure 4.1 compares the four sub-counties' survival probabilities using KM Curves. The Wilcoxon and Logrank test has a p-value of 0.88. According to our data, there aren't any significant variations among the four sub-counties' survival rates.

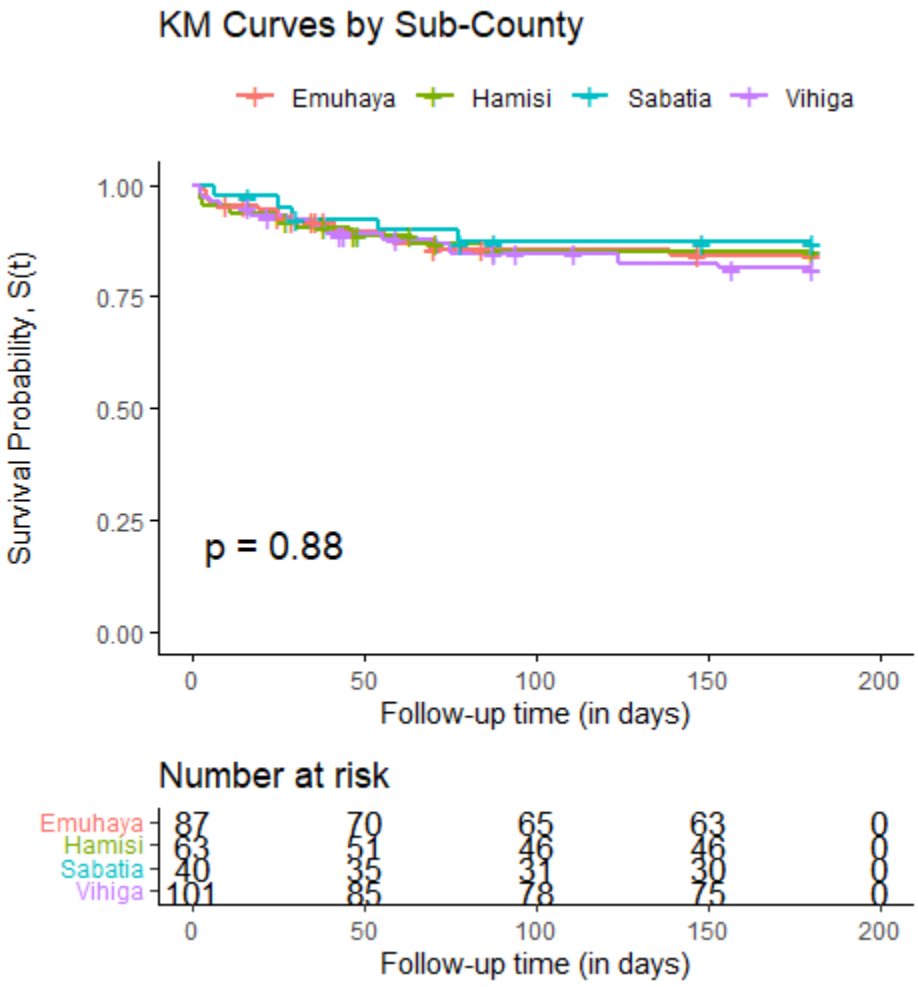



**Figure 4.1. KM Curves by Sub-County**

## 4.3 Comparing survival between Patients in Intensive Phase and Continuation Treatment Phase

Rifampicin (R), ethambutol (E) (RHZE), pyrazinamide (Z) and isoniazid (H) are used to treat TB patients. The duration of the treatment's intensive phase is 8 weeks and the maintenance phase lasts for 16 weeks. Comparison was therefore made between Intensive and Continuation Phase.

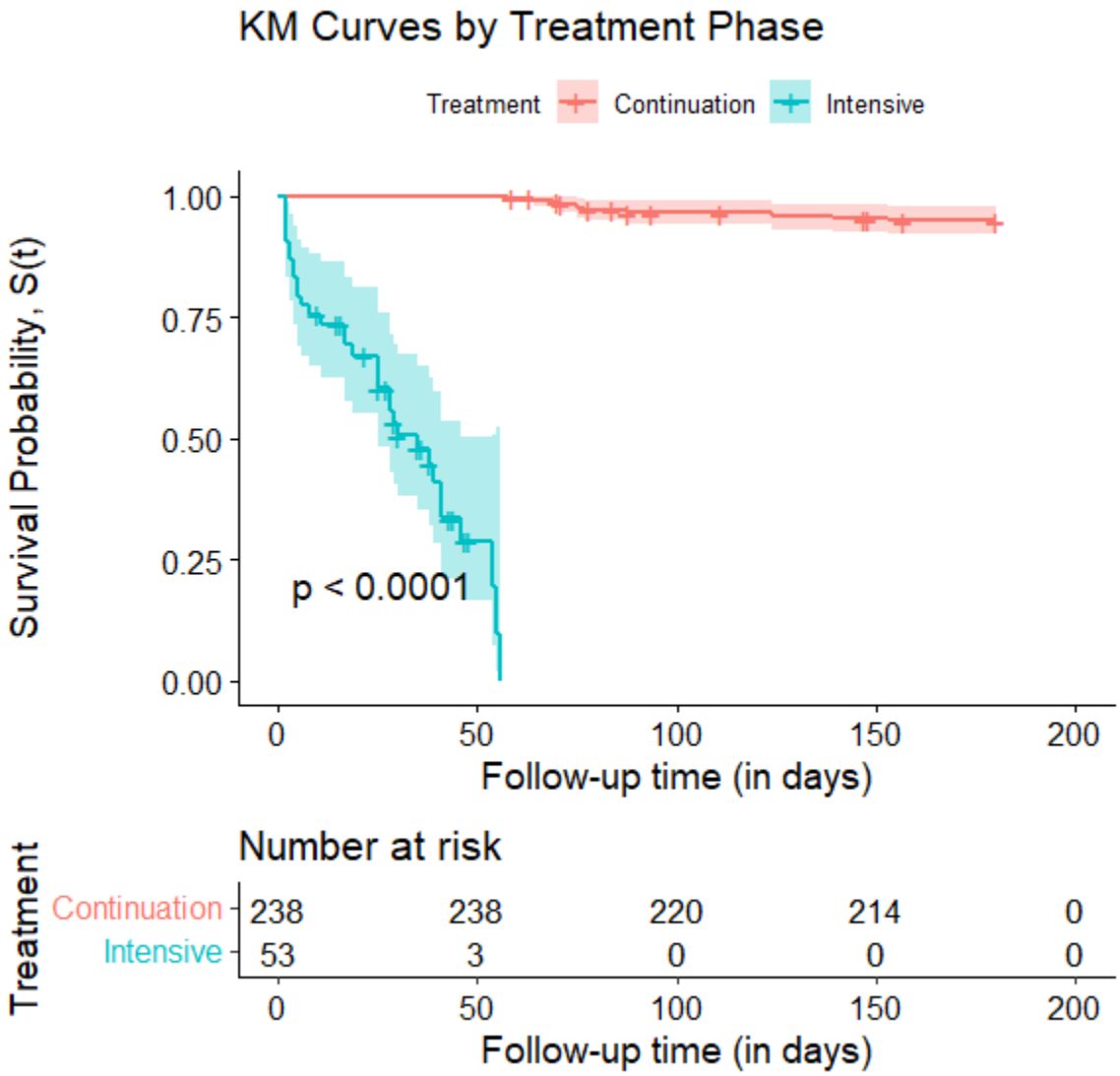



**Figure 4.2. KM Curves by Treatment Phase**

P-value is <0.0001 in both Log-rank and Wilcoxon Test, this shows that survival between Intensive and Continuation Phase were significantly different. The number at risk is the total number of living individuals, whereas the survival rate is the percentage of participants who were alive from the start of the table until the relevant period. The probability of surviving beyond the maximum 180 days in the continuous phase is 0.948 while probability of surviving beyond 55 days in the intensive phase is 0.0958 as seen from Table 4.6. Figure 4.2 illustrates a substantial decline in patient survival within the first few days of the intensive phase, showing that the majority of patients experienced the event earlier.

**Table 4.6. Kaplan Meier Estimates for Survival functions in Treatment Phases**

| Time | $n_t$ | $d_t$ | $\hat{S}(t)$ | $s.e(\hat{S}(t))$ | 95% CI; LL | 95% CI; UL |
|---|---|---|---|---|---|---|
| **C**ontinuation | | | | | | |
| 58 | 238 | 1 | 0.996 | 0.00419 | 0.988 | 1.000 |
| 64 | 235 | 1 | 0.992 | 0.00594 | 0.980 | 1.000 |
| 69 | 234 | 1 | 0.987 | 0.00727 | 0.973 | 1.000 |
| 71 | 232 | 1 | 0.983 | 0.00839 | 0.967 | 1.000 |
| 75 | 230 | 2 | 0.975 | 0.01027 | 0.955 | 0.995 |
| 77 | 228 | 1 | 0.970 | 0.01108 | 0.949 | 0.992 |
| 88 | 225 | 1 | 0.966 | 0.01184 | 0.943 | 0.989 |
| 124 | 219 | 2 | 0.957 | 0.01327 | 0.931 | 0.983 |
| 139 | 217 | 1 | 0.953 | 0.01392 | 0.926 | 0.980 |
| 153 | 214 | 1 | 0.948 | 0.01455 | 0.920 | 0.977 |
| **I**ntensive | | | | | | |
| 2 | 53 | 5 | 0.9057 | 0.0402 | 0.8303 | 0.988 |
| 3 | 48 | 2 | 0.8679 | 0.0465 | 0.7814 | 0.964 |
| 4 | 46 | 2 | 0.8302 | 0.0516 | 0.7350 | 0.938 |
| 5 | 44 | 2 | 0.7925 | 0.0557 | 0.6905 | 0.910 |
| 6 | 42 | 1 | 0.7736 | 0.0575 | 0.6687 | 0.895 |
| 8 | 41 | 1 | 0.7547 | 0.0591 | 0.6473 | 0.880 |
| 11 | 39 | 1 | 0.7354 | 0.0607 | 0.6256 | 0.864 |
| 17 | 34 | 2 | 0.6921 | 0.0644 | 0.5768 | 0.830 |
| 19 | 32 | 1 | 0.6705 | 0.0659 | 0.5530 | 0.813 |
| 25 | 30 | 3 | 0.6034 | 0.0697 | 0.4811 | 0.757 |
| 28 | 25 | 2 | 0.5552 | 0.0720 | 0.4305 | 0.716 |
| 29 | 23 | 1 | 0.5310 | 0.0728 | 0.4058 | 0.695 |
| 30 | 21 | 1 | 0.5057 | 0.0736 | 0.3802 | 0.673 |
| 35 | 19 | 1 | 0.4791 | 0.0744 | 0.3534 | 0.650 |
| 38 | 15 | 1 | 0.4472 | 0.0760 | 0.3205 | 0.624 |
| 39 | 12 | 1 | 0.4099 | 0.0783 | 0.2819 | 0.596 |
| 41 | 11 | 2 | 0.3354 | 0.0798 | 0.2103 | 0.535 |
| 46 | 7 | 1 | 0.2875 | 0.0815 | 0.1649 | 0.501 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 54 | 3 | 1 | 0.1916 | 0.0953 | 0.0723 | 0.508 |
| 55 | 2 | 1 | 0.0958 | 0.0828 | 0.0176 | 0.521 |
| 56 | 1 | 1 | 0.0000 | NaN | NA | NA |

## 4.4 Application of Cox Proportional Hazards regression in significant predictors of death identification among Tuberculosis patients.

**Model Selection**

Using Stepwise Regression we will check if all the seven variables should be included in the final model. A low AIC value was got when the predictor variables Treatment Supporter, Education Level, Sex were eliminated from the model.

**Table 4.7. Results of Stepwise**

| Variable | DF | AIC |
|---|---|---|
| None | | 480.33 |
| TB Type | 1 | 481.00 |
| Alcohol Consumption | 1 | 482.65 |
| Age in years | 1 | 483.12 |
| HIV Status | 1 | 491.03 |

The table 4.8 shows the final model with significant covariates. The study found that the variables Age, HIV status, Tb type and Alcohol consumption were significant predictors of death for TB patients. Patients who were HIV positive had a higher chance of dying 3.08 times in contrast to those who are HIV negative. Patients whose TB type was PTB were less at risk of 0.446 compared to those whose TB type was EPTB. There was evidence that patients who take alcohol had a higher risk of death, 1.931 times in comparison to those who do not take alcohol.

**Table 4.8. Results of Final Model**

| Covariate | $\hat{\beta}$ | $s.e(\hat{\beta})$ | HR(CI;95%) | P |
|---|---|---|---|---|
| Age | 0.02385 | 0.01073 | 1.0241(1.0028-1.046) | 0.026283 |
| HIV Status Positive | 1.12533 | 0.31840 | 3.0812(1.6508-5.751) | 0.000409 |
| TB Type PTB | -0.76345 | 0.42034 | 0.4661(0.2045-1.062) | 0.049326 |
| Alcohol Consumption Yes | 0.65817 | 0.32722 | 1.9313(1.0170-3.668) | 0.034284 |

### 4.4.1 Cox Proportional Hazards model Diagnostics

The model makes the assumption that any person's hazard is a fixed percentage of everyone else's hazard. A p-value of 0.68 is given by the global test, implying that assumption of proportional Hazard is supported.

**Table 4.9. Test For The Assumption of Proportional Hazard Ratio**

| | chi square | df | P |
|---|---|---|---|
| Sex | 0.297 | 1 | 0.59 |
| Age.in.years | 1.076 | 1 | 0.30 |
| HIV.Status | 1.517 | 1 | 0.22 |
| TB.Type | 0.954 | 1 | 0.33 |
| Alcohol.consumption | 0.694 | 1 | 0.40 |
| Education.level | 0.355 | 2 | 0.84 |
| GLOBAL | 4.855 | 7 | 0.68 |

From figure 4.3, There is no pattern in residuals against time further supporting proportional hazards assumption.

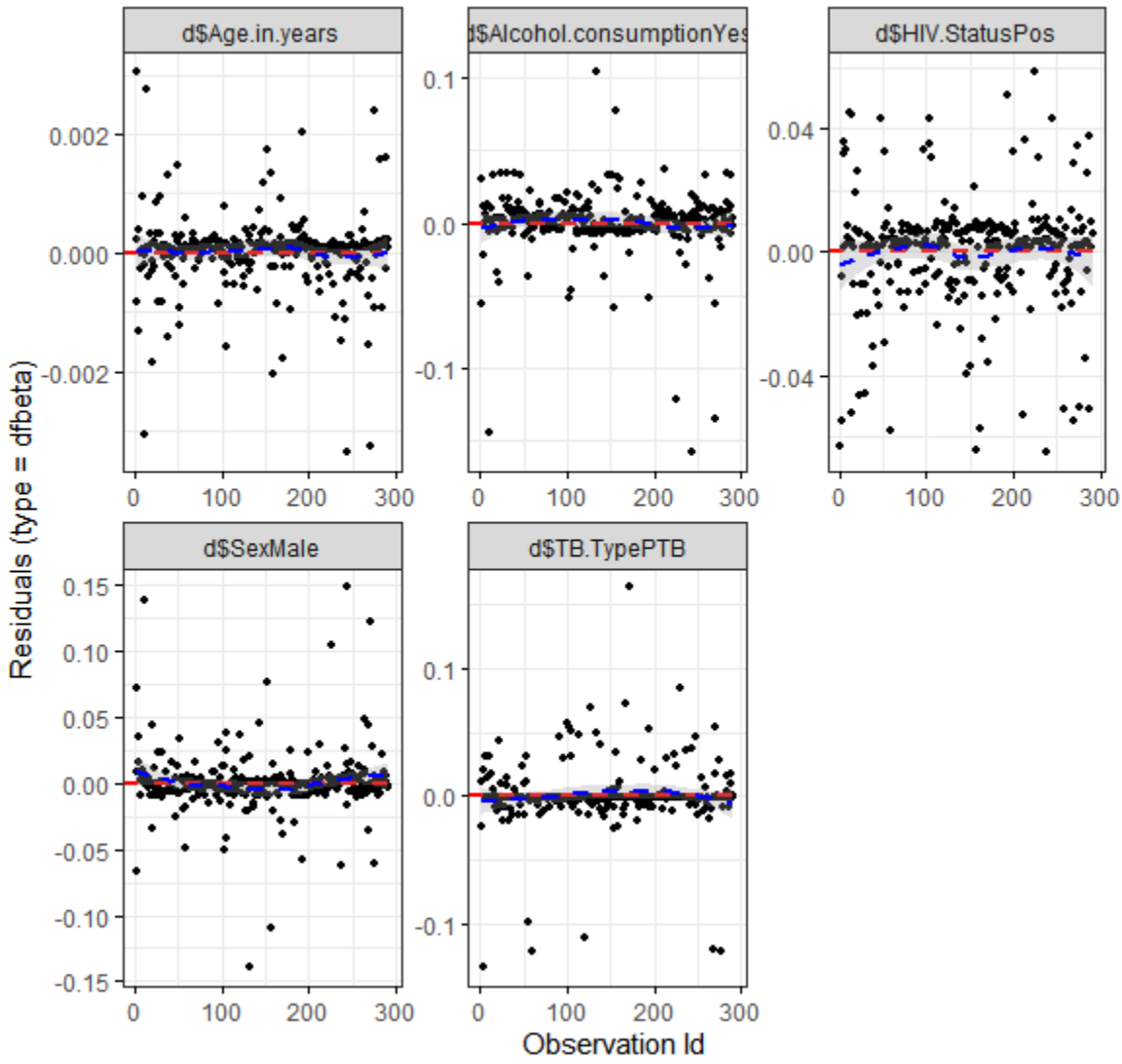



**Figure 4.3. df Beta of Residuals**

# CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

## 5.1 Conclusion

Findings from the study provide essential details which will help Vihiga county health care in making decisions. The study utilized Kaplan Meier to compare survival time between different sub-counties in Vihiga County. It was also utilized to compare survival between different phases of Tuberculosis treatment. The findings show that survival of patients is very poor in the Intensive treatment phase compared to the Continuation treatment phase. Wilcoxon and Log-rank tests were then used to test on whether there was significant differences between the categories.

The relationship between predictors variables and survival among the tuberculosis patients was further examined using a Cox proportional hazards regression model. Age, HIV Status, TB type and Alcohol consumption were significant predictors affecting the survival of patients. The findings also supported the notion that aging increases the risk of dying from tuberculosis..

## 5.2 Recommendations

As indicated in the results section, the likelihood of death in tuberculosis increases with age, Having Extra Pulmonary TB type, Alcohol consumption and HIV status. Previous studies have shown similar results which have been confirmed by the present study in Vihiga setting. The government is therefore encouraged to do public education on the importance of treatment completion to prevent developing of EPTB. Encourage patients to avoid alcohol during TB treatment.

Further research on the time to relapse, and the effect of treatment on the disease should be carried out to give more insight into the burden of the disease and disease management.

# Bibliography

Abdullahi, O. A., Ngari, M. M., Sanga, D., Katana, G., and Willetts, A. (2019). Mortality during treatment for tuberculosis; a review of surveillance data in a rural county in kenya. *PLoS One*, 14(7):e0219191.

Adeboye, A., Ruffin, M., James, N., and Yongsong, Q. (2020). Joint modelling of longitudinal-time-to-event with categorical variable indicators of latent classes: Application to tuberculosis data.

Birlie, A., Tesfaw, G., Dejene, T., and Woldemichael, K. (2015). Time to death and associated factors among tuberculosis patients in dangila woreda, northwest ethiopia. *PloS one*, 10(12):e0144244.

Gichuki, J. and Mategula, D. (2021). Characterisation of tuberculosis mortality in informal settlements in nairobi, kenya: analysis of data between 2002 and 2016. *BMC Infectious Diseases*, 21(1):1–8.

Harding, E. (2020). Who global progress report on tuberculosis elimination. *The Lancet Respiratory Medicine*, 8(1):19.

Leveri, T. H., Lekule, I., Mollel, E., Lyamuya, F., and Kilonzo, K. (2019). Predictors of treatment outcomes among multidrug resistant tuberculosis patients in tanzania. *Tuberculosis research and treatment*, 2019.

Limenih, Y. A. and Workie, D. L. (2019). Survival analysis of time to cure on multi-drug resistance tuberculosis patients in amhara region, ethiopia. *BMC public health*, 19(1):1–11.

Onyango, D. O., Yuen, C. M., Masini, E., and Borgdorff, M. W. (2018). Epidemiology of pediatric tuberculosis in kenya and risk factors for mortality during treatment: a national retrospective cohort study. *The Journal of Pediatrics*, 201:115–121.

Osman, M., du Preez, K., Seddon, J. A., Claassens, M. M., Dunbar, R., Dlamini, S. S., Welte, A., Naidoo, P., and Hesseling, A. C. (2021). Mortality in south african children and adolescents routinely treated for tuberculosis. *Pediatrics*, 147(4).

Pardeshi, G. (2009). Survival analysis and risk factors for death in tuberculosis patients on directly observed treatment-short course.

Persson, C. and Glimelius, B. (2002). The relevance of weight loss for survival and quality of life in patients with advanced gastrointestinal cancer treated with palliative chemotherapy. *Anticancer research*, 22(6B):3661–3668.

Ranzani, O. T., Rodrigues, L. C., Bombarda, S., Minto, C. M., Waldman, E. A., and Carvalho, C. R. (2020). Long-term survival and cause-specific mortality of patients newly diagnosed with tuberculosis in são paulo state, brazil, 2010–15: a population-based, longitudinal study. *The Lancet Infectious Diseases*, 20(1):123–132.

Yone, E. W. P., Kengne, A. P., and Kuaban, C. (2011). Incidence, time and determinants of tuberculosis treatment deaths in yaounde, cameroon: a retrospective hospital register-based cohort study. *BMJ open*, 1(2):e000289.