UNIVERSITY FOR DEVELOPMENT STUDIES

BAYESIAN CONTINUOUS-TIME SURVIVAL ANALYSIS FOR RECOVERY OF TUBERCULOSIS PATIENTS IN BULUK

AKAN-NYAATEMI AKANLIKSIUK JOSEPH

2021

UNIVERSITY FOR DEVELOPMENT STUDIES

BAYESIAN CONTINUOUS-TIME SURVIVAL ANALYSIS FOR RECOVERY OF TUBERCULOSIS PATIENTS IN BULUK

By

AKAN-NYAATEMI AKANLIKSIUK JOSEPH

(BSc. Actuarial Science)

(UDS/MAS/0002/19)

Thesis submitted to the Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies in partial fulfillment of the requirements for the award of Master of Philosophy Degree in Statistics

September, 2021

DECLARATION

Student

I hereby declare that this thesis is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere:

Candidate's Signature: ……. Date: ……..……

Name: Akan-Nyaatemi Akanliksiuk Joseph

Supervisors'

We hereby declare that the preparation and presentation of the thesis was supervised in accordance with the guidelines on supervision of thesis laid down by the University for Development Studies.

Principal Supervisor's Signature: …… Date…….

Name: Dr. Dioggban Jakperik

Co-Supervisor's Signature ……..……... Date: ……..……

Name: Mr. Mathias Akamba

ABSTRACT

Tuberculosis is an infectious bacterial disease caused by the bacilli *Mycobacterium tuberculosis.* The study modeled the prognostic factors for TB patients in Buluk using Kaplan-Meier, Bayesian Continuous-time Survival models, log-logistic and logistic regression models. Age, treatment time, and smear results were found to be associated with treatment outcomes. Pulmonary positive TB was identified to be the most prevalent disease category. The study unveiled that the median recovery time for patients was 171 days with average recovery of 170 and 174 days for males and females respectively. Also, the percentage recoveries for males and females were found to be 79.32% and 87.32% respectively. In addition, recovery among children, adults and the aged were 100%, 81.28% and 80.28 % respectively. The patients who reported for treatment for the first time had lower recovery rates compared to relapsed and other forms of non-new patients. There was a high HIV testing rate of 97.73% with an alarming TB/HIV co-infection rate of 10.39%. TB/HIV Co-infection was found to be associated with TB related mortality with a fatality rate of 18.75%. Finally, it was established that a unit increase in age is associated with 1.2% decrease in the odds of recovery among TB patients whilst a unit increase in treatment time is associated with 3.9% increase in the odds of recovery.

ACKNOWLEDGEMENTS

My thanks go to Dr. Nasiru Suleman (Head of Department, Statistics), all lecturers of the Department of Statistics and all lecturers of the Faculty of Mathematical Sciences for their immerse support and shaping towards the conduct of a successful and credible research work.

Also, I will like to thank my parents, siblings and my coworkers for their support in diverse ways. I extend my regards to Mr. Joseph Naab Kologbire and the team at Builsa North District Health Directorate for the work they did for me during the data collection.

Finally, my thanks and appreciations will be incomplete if I fail to acknowledge my workaholic supervisors Dr. Dioggban Jakperik and Mr. Mathias Akamba for their time, comments and motivations that led to the successful compilation of this document. Despite their tight schedules, they still made time to support me technically and professionally sometimes just to favour me to the detriment of their own interest.

DEDICATION

This work is dedicated to my lovely parents Mr. and Mrs. Akan-Nyaatemi Akanvariba and my wife. It is their appetite for education that has carried me to this stage.

TABLE CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

A	Age
AFB	Acid Fast Bacilli
AFT	Accelerated Failure Time
AIC	Akaike's Information Criterion
AICc	Corrected Akaike's Information Criterion
AIDS	Acquired Immunodeficiency Syndrome
ART	anti-retroviral Treatment
BCG	Bacillus Calmette-Guerin
BIC	Bayesian Information Criteria
CD4	Cluster of Differentiation 4
cdf	Cumulative Density Function
CI	Confidence Interval
CS	Cycloserine
DC	Disease Category
DHD	District Health Directorate
DIC	Deviance Information Criterion
DOT	Directly Observed Treatment

E	Ethambutal
EPTB	Extra-Pulmonary Tuberculosis
ETH	Ethionamide
GHS	Ghana Health Service
H	HIV/AIDS Status
HFI	Household Food Insecurity
HIV	Human Immunodeficiency Virus
HL	Hosmer-Lemeshow
IAP	Indoor Air Pollution
IFN-y	Interferon-y
K-M	Kaplan-Meier
LOGLIK	Log-Likelihood
LTB	Latent Tuberculosis
MDR-TB	Multidrug-Resistant Tuberculosis
NTP	National Tuberculosis control Programme
OR	Odd Ratio
P	Pyrazinamide
PCR	Polymerase Chain Reaction

pdf	Probability Density Function
PH	Proportional Hazard
PTB	Pulmonary Tuberculosis
RD1	Region of Difference 1
ROC	Receiver Operating Characteristic
S	Streptomycin
SR	Smear Results
T	Treatment time
TB	Tuberculosis
TP	Type of Patient
TST	Tuberculin Skin Test
WHO	World Health Organization
XDR-TB	Extensively Drug Resistant Tuberculosis

CHAPTER ONE

INTRODUCTION

1.1 Background of Study

Tuberculosis (TB) is an infectious disease that mostly attacks the lungs which is caused by the *Mycobacterium tuberculosis* bacteria (Jurado et al., 2018). According to WHO Global Report (2020) on tuberculosis, TB is among the top 10 killer diseases worldwide and even rated higher ahead of HIV in terms of mortality. Tuberculosis appears to be one of the highest secret killer diseases (opportunistic diseases) as it confounds with some immune weakening diseases like HIV/AIDS to act and destroy if not quickly detected and treated. TB among HIV patients was estimated to be about 173 times more than non-HIV patients in Guangxi (Cui et al., 2017). This incidence as reported by Cui et al. (2017) is the picture of cases in most part of the world, especially in the under developed Regions.

According to the WHO (2012) report, HIV is a number one cause of death among TB patients that could be prevented if detected on time and the right care or treatment commences on the patient for both illnesses. The WHO (2013) report further stressed that, to be able to properly manage and reduce mortality among co-infected patients is to increase the screening and testing rate to identify the patients and put them on the appropriate medication such as Antiretroviral drugs if there is need. The major mode of transmission is when the bacterium is expelled from the lungs of an active patient through coughing, sneezing or

spitting into the air where the bacteria can be inhaled by another person who comes in contact (airborne). In the year 2018, there was an estimated number of 10 million people worldwide who suffered from TB with Africa alone recording about 20% of these numbers. According to the National TB control Program (NTP), TB prevalence in Ghana stood at 144 persons per every 100,000 people as at 2019 (Global TB report, 2020).

Several research woks have been conducted on TB screening, early detection, and treatment of people living with TB (Date et al., 2015; Eang et al., 2012).

According to research conducted in the Kassena Nankana West District of the Upper East Region of Ghana, smoking, alcohol consumption, contact with TB patients and Low household income were identified as risk factors associated with TB (Boah et al., 2020). Other researchers identified some works such as construction, gold mine, Indoor Air Pollution (IAP), Household TB Exposure, Household Food Insecurity (HFI), Diabetes, HIV/AIDS, harmful use of alcohol, undernourishment and Patients living with HIV/AIDS as well as relapsed patients as risk factors of TB (Jakperik et al., 2013; Jubulis et al., 2014; Jakperik et al., 2015; WHO, 2020).

Successful treatment (recovery) of TB refers to patients who are completely cured of the infection with at least a smear or cultural negative in the last month of treatment and on at least one previous occasion and patients who have completed the full anti-tuberculosis dose without bacteriological results (Biruk et al., 2016; Xie et al., 2020).

Study by Terefe et al. (2018) reported TB recovery of 75% with a median recovery time of 185 days and a 73.75% recovery in the Upper West Region (Jakperik et al., 2013). Similarly, the median recovery time of TB patients in Northern Region, Upper West Region and Northern Ghana are, 22 weeks, 25.43 weeks and 24.14 weeks respectively (Jakperik et al., 2013; Jakperik et al., 2011; Jakperik et al., 2012).

In addition, most studies in survival and epidemiological analysis concentrates much in classical analysis. However, Bayesian analysis has been proven to have high precision compared to the traditional classical approach (Mahanta et al., 2015; Ojo et al., 2017).

This study therefore to model factors that are associated with recovery in tuberculosis patients using Bayesian analysis.

1.2 Problem Statement

Tuberculosis is considered one of the top 10 infectious diseases with infection and mortality rates globally (WHO, 2018). It is estimated that, about 10 million people in the world get infected every year. In Ghana, the tuberculosis infection rate stood at 290 per every 100,000 people as at 2013 (WHO, 2018). The prevalence rate in Ghana has reduced from 290 per every 100,000 people in 2013 to a rate of 144 persons per every 100,000 population in the year 2019 (WHO, 2020).

Although the National Tuberculosis control Program (NTP) has made progress by improving on treatment outcomes to over 80%, TB case detection rate in

Ghana is still low to about 40.1% (WHO, 2018; Osei et al., 2020). TB is a curable bacterial infection if detected early; it can however become life threatening if left untreated after a period of time as it remains one of the major infectious causes of death in the world (Sulistiyani et al., 2010).

According to the WHO (2020) report, treatment successes for new and non-new cases are 84% and 87% respectively. TB/HIV co-infection (77%), Multi-Drug Resistant (63%) and previously treated (excluding relapse cases) of 87% were reported. TB successful treatment among TB/HIV co-infected patients in two urban hospitals in Ghana was reported to be 91.17% (Ogyiri et al., 2019).

According to Mahanta et al. (2015), it is reported that, estimations made using Bayesian approach gives better estimates than their corresponding classical approach in the estimation of a Markov based logistic model. Also, the Bayesian approach was seen to be more efficient in the selection of the most significant factors that are associated with the risk of TB in South Africa as against the classical methods (Ojo et al., 2017)

Though the treatment success is steadily increasing nationally, no research has been done to discover the factors that influence treatment outcomes of TB patients as well as provide estimates for prevalence of the disease. Also, most of the studies on recovery of TB patients are done using classical statistical approaches. Consequently, this study seeks to model factors that are associated with recovery of TB patients in Buluk using Bayesian Continues-Time Survival Model.

1.3 General Objective

The main objective of the study is to model the factors that are associated with recovery of Tuberculosis patients in Buluk using Bayesian Continuous-Time Survival Model.

1.4 Specific Objectives

The specific objectives of the study are to:

- Model factors associated with recovery of TB patients.
- Identify factors that influence treatment outcome of TB patients.
- Estimate the median recovery time of TB patients.

1.5 Research Questions

The research questions for the study are:

- What are the factors that influence the successful treatment of Tuberculosis?
- Which factors influence treatment outcomes in Buluk?
- What is the median recovery time for TB patients in Buluk?

1.6 Significance of the Study

This study provides a level knowledge on the impact of some factors associated with TB treatment outcomes. This will inform TB caregivers, health personnel and policy makers the measures to take when handling patients with comorbidities such as diabetes and HIV/AIDS which compromises the immune system of the patient.

The findings of this study could be used by the District Health Directorates (DHD), the Ghana Health Service (GHS) and the World Health Organization (WHO) to manage the prevalence of Tuberculosis in the Districts.

Also, the DHD, GHS and WHO can use the results of this study as a policy guide through education and policy formulation as part of strategies towards Treating TB infections. The Bayesian approach employed provides better estimates that will help identify the most significant prognostic factors that influence treatment outcome compared to the classical approach.

Finally, findings of this study could serve as basis for further research to identify the causes and management of tuberculosis.

1.7 Scope of the Study

The study relied solely on secondary data from the Builsa North District Health Directorate for the period January 25, 2012 to March 20, 2020 (an 8-year period). The district health directorate collates data from all health centers across Buluk. Also, people from nearby Districts such as Kassena-Nankana East and West, Mampurugu Moaduri and Sisala East visit facilities within Buluk for healthcare. The following variables were considered when collecting the data; age, sex, date of onset of symptoms, date patient reports to the health facility for registration, date of diagnosis, date treatment commences, date treatment completed, treatment outcomes, type of TB and the risk factors patients are exposed to.

The various treatment outcomes as identified by this study are: Death, cured, treatment complete, lost to follow-up, defaulted and transferred out.

The study analysed and computed the average recovery time using Kaplan-Meier estimator. Factors that influence TB treatment outcomes and the effect of these factors on recovery were also analysed using Accelerated Failure Time (AFT) Models and Binary logistic regression analyses. AIC, BIC, AICc, DIC, log-likelihood values, Hosmer-Lemeshow goodness of fit test and Receiver Operating Characteristic Curve were used to select the best fit models.

1.8 Limitations of the Study

Because the study used a secondary data, some key variables which were to be investigated in this study could not be found, hence led to variation of the scope of study. Also, causes of TB such as smoking and drinking among others as identified in literature were not available and their effects could not be studied.

1.9 Organization of the Study

This study is made up of five chapters. Chapter one comprises of the background, problem statement, objectives, significance, scope, limitations and the organization of the study. Chapter two contains literature review where past similar studies were reviewed, the detailed methodology is presented in chapter three. Results and discussion of the results are presented in chapter four while chapter five contains the summary, conclusions and recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.0 Tuberculosis Prevalence

Tuberculosis is an infectious disease caused by *mycobacterium tuberculosis* which is transmitted via droplets and a mutable period of latency of infections (Bhargava et al., 2018). In other words, tuberculosis is defined as a chronic infectious disease caused by *Mycobacterium tuberculosis* bacilli (Xie et al., 2020). There are several means of detecting and testing for the presence of the *mycobacterium.* These include the use of X-ray, the skin test, the culture test, molecular (GeneXpert) test and the Sputum Smear test among others.

The incidence of tuberculosis is estimated to be 10 million new cases worldwide with a global fatality of 1.24 million. The mortality rate of Tuberculosis is estimated to 16 per every 100,000 individuals in 2018 (Global tuberculosis Report, 2019). In Ghana TB incidence rate for the year 2019 stands at 144 patients per every 100,000 people out of a 30.4 million estimated population (Global tuberculosis report, 2020). This implies that out of an estimated population of 30.4 million people, about 43,776 will be at risk of being infected. Even though the global TB incidence rate has dropped by 9% between 2015 and 2019, its impact on high TB burden countries in Africa and South-East Asia cannot be underestimated. Also, the WHO target of reducing TB incidence rate by 20% by the year 2020 was not achieved. The WHO target of reducing TB related

deaths by 35% by 2020 was unachievable as the reduction in deaths between 2015 and 2019 was 14% (Global TB report, 2020).

2.1 Types of Tuberculosis

The types of tuberculosis of patients are classified based on the part of the body that the causative agent of tuberculosis resides. These are Pulmonary Tuberculosis (PTB) and Extra-Pulmonary Tuberculosis (EPTB). When the Bacilli resides in the lungs parenchyma of its victim, it is called pulmonary tuberculosis (PTB). Military tuberculosis which is a form of TB that is widely disseminated into the lungs by tiny side of the lesions is also classified as PTB.

However, Extra-Pulmonary Tuberculosis refers to TB where the bacilli invade other parts of the human body other than the lungs. For example, TB of the pleura, lymph nodes, skin, abdomen, joints and bones are all forms of EPTB (Xie et al., 2020). EPTB can either be diagnosed clinically or bacteriologically.

2.2 Diagnosis and treatment of tuberculosis.

The first step to eliminating the menace of tuberculosis is diagnosing for the presence of the *mycobacterium tuberculosis* bacteria using the appropriate scientific and clinical methods and tools early. The TB infection can only be managed appropriately and properly if it is accurately diagnosed. Therefore, proper identification of the bacilli and the part of the body it can be located is the only sure way to fighting the disease. Nahid et al. (2006) acknowledged that, Tuberculin Skin Test (TST) was the only tool used in the detection of Latent tuberculosis (LTB) which estimated sensitivity (“in patients with active TB”) in the measurement of immunocompromised populations is about 75% to 90% lower

with a corresponding estimated specificity ("in patients with no known TB infection") of 70% to 95% lower in Bacillus Calmette-Guerin (BCG) vaccinated until the emergence of the RD1-based-IFN-y assays with 80% to 95% estimated sensitivity in patients with active TB in immunocompromised populations and an estimated specificity of 95% to 100%. Drug resistance TB detection is one of the most difficult diagnoses to undertake since it requires very highly sensitive kits. The PCR assay which employs the use of Molecular beacons with Oligonucleotides which usually emit light whenever a chemical reaction is detected has proving to be one of the tests that have a very high sensitivity in the detection of rifampicin resistance tuberculosis (El-Hajj et al., 2001). However, the availability and the affordability of this technology are lacking in most health facilities especially in the underdeveloped countries since it requires certain sophisticated technology. The detection of the *Mycobacterium tuberculosis* in a multi-Drug resistance TB like the rifampicin resistant TB can also be done fast and accurate using the INNO-LiPA-Rif.TB (LiPA) which has a sensitivity greater than 95% with a specificity of 100% in all cases (Morgan et al., 2005).

Tuberculosis is a bacterial infection that can be cured with the right prescription. However, this cannot be done effectively if patients are not classified properly in regiments to reduce the risk of wrong prescriptions and foster easy monitoring and outcome evaluations. The first step towards the treatment of TB is in the hands of the patient. A successful treatment outcome cannot be achieved without the commitment and adherence of the patient to therapeutic measures and instructions (Nahid et al., 2006). The Directly Observed Treatment (DOT)

approach is one of the most efficient and effective means of ensuring patient adherence.

The treatment of tuberculosis can be done in two phases; the intensive which usually takes two (2) months for newly diagnosed patients and three (3) months for previously treated and the continuation phase which also takes four months and five months for the new and previously treated patients respectively. According to Bhargava et al. (2018), treatment of TB can be done based on WHO recommendations using first-line anti-TB drugs (Isoniazid (I), Rifampicin (R), Pyrazinamide (P), Ethambutol (E)) and Streptomycin (S)) for the treatment of newly diagnosed patients, Fixed-Dose combination for the treatment of Drug Resistant TB ((R150/H75/E400/E275)mg for the initial phase and (R150 and /H75)mg for the continuation phase) for drug resistant TB patients and pediatric TB treatment for children (R60/H30/Z150)mg and E100mg for first phase and R60/30 for the continuation phase). Despite the presence of curable medications and management measures, certain challenges that are still lingering around the treatment of TB keep suppressing the expected treatment success desired. These challenges include the continues increase in the number of drug resistance, multi-drug resistance and extreme-drug resistance TB as well as the presence of comorbidities such as diabetes, TB/HIV co-infection and so on. The limited supplies of first-line anti-TB drugs, irregular treatment, overcrowding, poor ventilation and malnutrition is a major challenge in the management and treatment of tuberculosis (Bhargava et al., 2018).

2.3 Delays in Detection and Treatments.

The term "Delay" is used to describe the interval between two events for all TB cases (Golub et al., 2006). Total treatment delay is the period between the time a patient first experiences the symptom of the infection and the time the diagnosed patient is put on treatment at a health facility (Getnet et al., 2017).

Treatment delay is a major concern in our health care system and a hinderance towards the achievement of the WHO Goal of Global End of TB since it gives room for the Delayed Patients to continuously transmit the disease while in their delayed period. Treatment delay has a strong correlation with the transmission of TB to contacts exposed to smear positive patients and the magnitude of delay also has a proportional impact in the transmission of the infection. People in contact with a patient who has a treatment delay above 90 days are at greater risk of being infected with TB compared to contacts to patients with relatively shorter period of treatment delays (Golub et al., 2006).

Though the correlation between treatment delay and contact infection is valid, the study could not establish a distinction between delay in TB detection and delay in treatment and for that matter could not do any comparative analysis on the rate of contact infections between patient delay and Health care system delay.

Patient Delay is the time interval between time a patient first detects a symptom to the time the patient reports to a health facility to register for treatment. A study in Addis Ababa revealed that "Patient Delay" accounted for about 42.1% while Health care system delay which is the period between the time a patient visits a health facility for registration and the time of commencement of treatment

recorded 54.8%. A Total Treatment delay was 61.6% delay for more than 28 days before the start of treatment which is an indication that, there is significant delay in treatment of patients living with TB (Getnet et al., 2017).

2.4 **Treatment Outcomes of Tuberculosis**

Tuberculosis as one of the leading causes of mortality is a threat to public health and hence global concern. Treatment of TB is a priority of the WHO and for that matter efforts have been made globally to curb the menace through funding and logistic support worldwide to health facilities and organizations responsible for the screening, treatment and cure of TB patients (Dräger et al., 2006). According to Biruk et al. (2016), treatment outcomes were categorized into five. These are; cured, treatment complete, died, defaulted and transferred out. This was however regrouped into successful treatment and unsuccessful treatment or treatment failure.

However, researchers and policy makers ought to understand that, tackling the disease from treatment of TB patients is a failure in disguise. Thus, more new cases will continue to be recorded and even patients who abscond and goes back to the community to continue their old life styles might even develop drug resistant or Multidrug-Resistant Tuberculosis (MDR-TB) and will be transmitting the disease unknowingly. Multidrug Resistant TB is defined as TB caused by some strains of the *Mycobacterium tuberculosis bacilli* that are resistant to at least isoniazid and rifampicin (Amita et al., 2008).

On the other hand, survival outcome refers to the end result or treatment outcome of a patient under treatment after being diagnosed of TB. Some of the survival outcomes or treatment outcomes according to literature are as follows;

Relapse: this is a situation where the health condition of a patient under treatment deteriorates after a period of temporal treatment and stability or "A patient previously treated for TB who had been declared cured or treatment completed, and is again diagnosed with bacteriologically positive (smear or culture)" (WHO, 2019). Relapse patients have a relatively greater risk of treatment failure compared to newly diagnosed TB cases (Jakperik et al., 2015).

Lost to follow-up: this refers to a situation where a patient on TB regimen cannot be found to complete the treatment probably due to migration or patient refusal to continue medication. A patient might not be in touch with health personnel and hence cannot be traced.

Treatment failure: a situation whereby a patient under treatment remains positive for TB after treatment has been completed due to ineffectiveness of treatment combinations to cure the disease.

According to Ifebunandu et al. (2012), treatment outcomes can be categorized into six principal headings: "cured, treatment completed, defaulted, died, treatment failure (patient remains positive for TB after a treatment has been completed) and transferred (a situation where a patient has been transferred from one health facility to another to complete treatment for the purposes of

proximity)". Where cured in this instance means patient has been treated and confirmed TB negative for (*smear+*), treatment completed seeks to say that, the patient has successfully gone through all the required treatment process and has completed taking all the required dose of the anti-tuberculosis drugs. A recent analysis of the cascade of TB care in India reported that in 2013, 72% of the estimated 2.7 million cases presented to government TB diagnostic services, only 45% completed treatment (Padayatchi et al., 2019).

Also, Naidoo et al. (2017) reported that about 53% of TB patients who were under treatment successfully completed their treatments. Their study further revealed that, of the estimated TB cases of 507,533 patients, about 17% of patients who are under treatment could not complete their treatments, 5% could not access diagnostic facility, 13% could not be traced during the diagnostic process and about 12% of patients did not initiate the treatment process at all.

However, this particular study though has acknowledged the fact that, a group of patients were not traceable after the commencement of treatments, the researcher failed to do further investigations to determine the exact cause of loss of follow-up. Could it be that the patient migrated? What exactly is the cause of loss of follow-up? Ifebunandu et al. (2012) did not also look beyond completion of treatment in the categorization of treatment outcomes.

The act of wrongly diagnosing a patient negative and subsequently confirming the patient cured can lead to the patient relapsing. Research revealed that, about 14.6% of treated tuberculosis cases do relapse (Merle et al., 2014). Even though the cause of recurrence is not emphatically stated, what however is certain and

more alarming is that these patients are in the community with a lot of TB susceptible around them with the risk of contracting the infection.

In addition to the continuous transmission of the *Mycobacterium* bacteria to others, the relapsed patients are at risk of developing MDR-TB. This is a growing concern because, Gandhi et al. (2012) in their study to unearth some risk factors associated with mortality among MDR-TB and XDR-TB patients in a very high HIV prevalence settings did indicate that, at least one country in every region of the world has reported a high prevalence of MDR-TB (greater than 3% of all new TB cases), and XDR-TB has been diagnosed in at least 57 countries. They went further to suggest that Drug resistant TB cases are harder to treat because they require about 18–24 months of treatment with second-line anti-tuberculosis medications (For instance aminoglycosides, fluoroquinolones, ethionamide [ETH], cycloserine [CS], and so on) which are less potent, more toxic, and more expensive than drug-susceptible TB medications.

A tubercle whose bacilli is resistant to flouroquinolone (ofloxacin, maxifloxacin, levofloxacin and gatifloxacin) and second-line drugs such as capreomycin, kanamycin and amikamycin are said to be suffering from Extreme drug resistance tuberculosis (XDR-TB) (Salil Bhargava et al., 2018). Drug resistance TB can further be categorized based on drug susceptibility as; rifampicin resistance, multi-drug resistance, excessive drug resistance, polydrug resistance and monoresistance. Age is one of the factors that are associated with TB treatment outcomes. Studies that looked at treatment outcomes of TB patients in Ghana and China reported that, an increase in age will result in a corresponding decrease in

treatment success (Abuaku et al., 2010; Ogyiri et al., 2019). It is not surprising to know that an opportunistic disease like tuberculosis will always take advantage of a person under treatment with a weak immune system.

2.5 Risk factors associated with Tuberculosis

In other to help achieve the WHO aim of Global End of TB strategy, several Researchers have done a lot of great researches and publications on some risk factors associated with TB, TB screening among HIV patients, TB treatment outcomes, factors and causes of mortality of TB patients and even factors associated with delays in diagnosis among TB patients (Jubulis et al., 2014; Boah et al., 2020; Cui et al., 2017; Gandhi et al., 2012; Date et al., 2015).

Risk factors associated with TB are basically factors or life styles that pre-expose an individual to the infection or they are factors that increase the chance of an individual getting the TB infection. Research indicates that, household income, smoking, alcohol consumption and household exposure to a known TB patient are risk factors associated with TB (Boah et al., 2020).

The study however did not find any association between history of diabetes and TB which is contrary to the WHO report that “Diabetes triples a person’s risk of developing TB”. Literature further revealed that, mining activities, construction works, drug abuse, exposure to dust and tar, and the injection of illicit drugs are all risk factors associated with TB (Jakperik et al., 2013). Low CD4 counts together with high degree of drug resistance are risk factors attributed to the high

volumes of MDR-TB and XDR-TB related deaths with the highest being drug resistance related deaths (Gandhi et al., 2012).

This therefore implies that we are fighting a lose battle if people do not understand the risk factors that pre-exposes them to tuberculosis.

2.6 Tuberculosis and HIV co-infection.

Tuberculosis is a bacterial infection caused by the bacilli *mycobacterium tuberculosis* that mostly attacks the lungs of an infected person though it can affect other parts of the body as well (WHO, 2013). HIV/AIDS on the other hand is a viral disease caused by the human immunodeficiency virus which point of attack is the human immune system. HIV is a viral disease that weakens the immune system when infected and hence making patients prone to TB and many other illnesses.

The chances of getting TB infection among HIV positive patients are higher compared to non-HIV patients. Similarly, the probability of a previously treated HIV positive TB patient relapsing is higher than that of non-co-infected TB patients (Bhargava et al., 2018). The WHO (2008) report suggests that, HIV/TB co-infection is prevalent in Africa and South-East of Asia.

According to Ahmed et al. (2004), the increase in the prevalence of HIV patients of 39% among TB patients could be the reason for the continuous increase in the number of TB cases in the Southern Region of Ethiopia. It is revealing to know that HIV/TB co-infection is quite high in Ghana especially among children who recorded a co-infection of 52.2% (Antwi et al., 2017). A study conducted in five

Districts of the Volta Region came out with an HIV/TB co-infection of 18.2% with the Kpando District recording the highest co-infection of 26%. The study indicated that women in the Volta Region recorded the highest co-infected cases than the male population within the period of 2012 - 2015 (Osei et al., 2017).

According to the WHO (2012) report, HIV is a leading cause of deaths among TB patients that could be prevented if detected early and the appropriate care or treatment commences on the patient for both illnesses. Sharma et al. (2005) reported TB related fatalities as higher among HIV positive patients even though patients co-infected are responding very well to treatment. This is clear indication of a correlation between HIV and TB related deaths. In addition, the treatment of tuberculosis among HIV patients could be complicated due to the activities of TB drug interaction with antiretroviral drugs. Jiang et al. (2008) in their research did confirm the impact of HIV in TB related mortalities.

The WHO (2013) report further stressed that, to be able to properly manage and reduce mortality among co-infected patients is to increase the screening and testing rate to identify the patients and put them on the appropriate medication such as Antiretroviral drugs if there is need.

2.7 Tuberculosis Profile of Ghana and Africa

According to the Global TB report (2020), analysis was made on 2019 data comparing the situation in Ghana to Africa as a whole. The estimated population was 30 million and 1 billion for Ghana and Africa respectively. Ghana had an incidence of 144 per 100,000 as against the 226 per every 100,000 population in

Africa with decrease rates of 10% and 16% respectively between 2015 and 2019. TB/HIV Co-infection rates were 30 and 54 per every 100,000, MDR recorded was 4 and 7 per 100,000 populations, and HIV positive TB mortality was 16 per 100,000 for both Africa and Ghana. The percentage decrease in TB mortality in Ghana and Africa are 4.3% and 19%. This implies that Ghana is not on track in preventing TB related deaths compared to the regional average. Also, the percentage of co-infection in Ghana (75%) that is on ART is quite low compared to the 92% in the region. This further translated into a relatively low treatment success of 77% in Ghana with a corresponding success of 79% for Africa.

The percentages of HIV positive patients that are on ART in Ghana for the year 2019 is 4.1% while Africa as a whole recorded 53% (Global TB report, 2020).

Finally, the risk factors that were found to be associated with TB in Ghana and Africa are; Alcohol, diabetes, smoking and undernourishment.

2.8 Bayesian Analysis

Bayesian analysis is a statistical technique that estimates unknown parameters using probability statements. In Bayesian inference, the most important are the Prior and posterior distributions. Unlike classical analysis where the basis of statistical inference is only the likelihood, in Bayesian approach, inference is drawn from the posterior distribution of the parameter which is a combination of both the prior distribution and the likelihood.

It is evidential as used in the estimation of the Markov based Logistic model that, Bayesian approach gives better model estimates compared to their alternative

classical approach (Mahanta et al., 2015). In addition, according to Ojo et al. (2017), the classical approach in selecting the most significant factors that are associated with the risk of TB in South Africa was discovered to be inferior compared to the Bayesian analytical approach. The Bayesian approach was confirmed to be more efficient over the classical approach as it is more flexible when modeling, and also provides individualized risk assessment in a simpler manner compared to the classical (Onisko et al., 2016).

2.8 Survival Analysis

Survival analysis is a statistical technique employed in estimating an event period. Survival analysis sometimes called failure-time analysis refers to a collection of statistical methods or procedures used to assist analyze time-to-event data (Prentice et al., 2001). Survival is one of the primary statistical techniques for analyzing data on time to an event such as death, heart or device failure and other communicable diseases (Tolley et al., 2016; Park et al., 2012; O'Connor et al., 1999). As applied by Ajagbe et al. (2014), survival analysis can be used in the analysis of tuberculosis data. Also, survival analysis can be employed to study event history such as time until death, recovery, employment or time to divorce (Browne et al., 2009).

The chance of an event occurring before a censoring time is called a discrete-time hazard (Browne et al., 2009). Discrete-time survival models provide avenues for determining the relationship that exist between the risk factors and the hazard. According to Efron (1988), discrete-time survival analysis can be used in examining the predictors of the hazard.

The most widely used methods in the estimation of survival data are;

I. **The parametric methods:** methods that rely on fixed number of parameters to build models usually for a small data set. An example of parametric model is the accelerated failure time (AFT) model based on distributions (exponential, lognormal, log-logistic, Weibull and Gamma).

II. **Nonparametric methods.** These are models that unlike the parametric do not assume a data set is finite but are usually defined on an infinite data set.

An example of a nonparametric regression model that is widely used in analyzing survival data by means of hazard and survival functions estimation is the Kaplan-Meier (K-M) estimator (Kaplan and Meier 1958). According to Dzimah et al. (2016), the Kaplan-Meier (product limit) estimator is the most widely used standard nonparametric model in the estimation of survival functions.

III. **Semi Parametric.** This is usually a modeling process that combines both the parametric and nonparametric models in performing a task the researcher thinks only parametric or nonparametric models will not be functional enough to yield desired results. A typical example is the Cox Proportional hazard model.

2.8.1 Accelerated Failure Time model (AFT).

The accelerated failure time model describes the linkage between failure time and independent variables.

In medical and most epidemiological studies, the two models that are widely used are the Proportional Hazard (PH) and the Accelerated Failure Time (AFT) models (Singogo et al., 2016). Also, models that are considered appropriate for fitting and analyzing Medical, biomedical and survival data is the Logistic regression model and the accelerated failure time model (Kalhori. et al., 2010; Biruk. et al., 2016; Barman. et al., 2017).

An AFT model is a parametric model in statistics that serves as an option to the widely used proportional hazard since it has better predictions with more reliable and consistent estimates than the PH model (Nardi and Schemper, 2003). The AFT models comprise the Weibull, exponential, log-logistic, Gamma and the log-normal distributions.

CHAPTER THREE

METHODOLOGY

3.0 Introduction

This chapter presents the study area, the statistical techniques, tools and methods that were used to conduct this research. It contains the data collection processes, statistical models used, the model assumptions and the data analysis procedures.

3.1 Study Area

Buluk is a traditional area made up of two administrative Districts namely the Builsa North and South Districts with Sandema and Fumbisi as their administrative capitals respectively. Builsa South is the youngest District that was carved out of the then Builsa district on 7th June 2012 as one of the four newly created Districts in the Upper East Region at the time.

Buluk is situated south-west of the Upper East Region and can be found between longitudes 1^0 05'' West and 1^0 35'' West and the latitudes 10^0 20'' North and 10^0 50'' North. Buluk shares boundaries with Mamprugu Moagduri District of the North East Region to the South, the Sissala East District of the Upper West Region to the West and the Kassena-Nankana west and Kassena-Nankana Municipal of the Upper East Region to the North and East respectively (Nyanwura et al., 2013). According to the Ghana Health Service Annual Report (2004) of the Upper East Regional Health Administration, the Builsa traditional area has a land coverage of about 2,205km which is equivalent to 24.94% of the total land mass of the Upper East Region (8,842km).

The major occupation of the Builsas (people of Buluk) is farming. And the major crops cultivated by the people are cereals with few of them especially those settled along the banks of the Sisili River in recent years venturing heavily into beans cultivation.

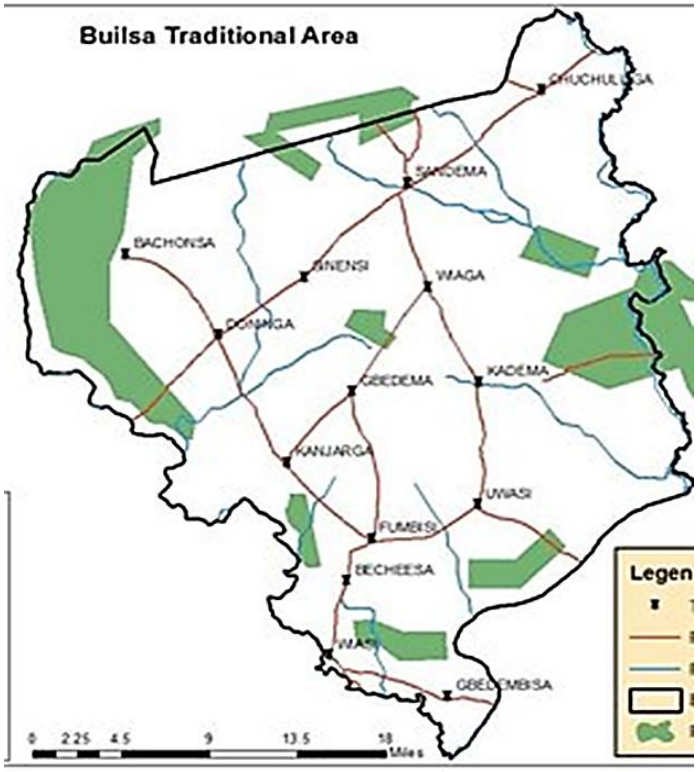


Figure 3.1 Map of Buluk

3.2 Data and Source

The study relied solely on secondary data from the Builsa North District Health Directorate in Sandema. Data was collected on every individual patient enrolled for TB treatment from January 25, 2012 to March 20, 2020. Some of the variables that were collected on each patient from the TB registers were as follows; the date of registration, date treatment commenced, date treatment completed, age, gender,

disease category, type of patient, smear results, treatment outcome and the HIV history of the patient. The study did not directly involved patients; hence ethical clearance was not necessary.

A total of 308 patients registered for treatment were studied and patients who died, transferred out, defaulted, lost to follow-up and patients whose treatment time exceeded the six months of complete dose were considered censored

3.3 Description of Covariates

The description, codes and the interpretation of the covariates or prognostic factors used in this study are presented below.

3.3.1 Prognostic factors

- Age: this refers to a patient's age at commencement of treatment.
- Gender: this represents the sex of a patient and was coded as 0 = Male and 1= Female
- HIV Status: This presents the HIV testing status of patients. These are; HIV positive, P=1, HIV negative, N=0 and No HIV test done, DN=2
- Treatment time**:** this was determined as date treatment completed minus date treatment commenced, TT.
- Type of patient (TP): The types of patient as found in the register were new coded as '0' and non-new, coded as '1'.

a. New patients**.** This refers to patients who reported and are registered for anti-tuberculosis treatment for the first time or these are patient who have

never had treatment for tuberculosis or who have taken anti tuberculosis drug for less than one month and was coded as '0'

b. Non-new patients. These are patients who are not receiving the ant-tuberculosis medication for the first time. These are mostly relapsed patients (A patient previously treated for TB, declared cured or treatment completed and is again diagnosed with bacteriological (+) TB (smear or culture)) and was coded as '1'

- Disease category: This refers to the categorization of patients into pulmonary and extra pulmonary tuberculosis patients.

a. Pulmonary tuberculosis. This is a type of tuberculosis where the *mycobacterium tuberculosis* bacteria recede in only the lungs, though it can be spread from the lungs to other parts of the body in which case it is no more pulmonary tuberculosis.

b. Extra pulmonary tuberculosis. This refers to tuberculosis in which the infection can be detected in other parts of the human body like the bone, the skin or the abdomen other than the lugs.

- Smear results

a. Smear positive. This refers to the detection of Acid-Fast Bacilli (AFB+) in the sputum sample of a TB patient and is coded as '1'.

b. Smear negative. The absence of AFB in the sputum of a patient during TB screening at the laboratory and is coded as '0'.

3.3.2 Treatment Outcomes

- Cured: a patient whose sputum smear test or culture results were positive before the commencement of treatment but who was smear negative or culture negative at the last month of treatment and on at least one previous occasion (source is TB register, DHD)

- Treatment completed: a patient who completed the treatment but did not have a negative sputum smear or culture result in the last month of treatment and on at least one previous occasion.

- Died: a patient who dies for any reason in the course of treatment.

- Defaulted**:** a patient whose treatment was interrupted for two consecutive months or more.

- Lost-to-follow-up: patients who could not be traced by TB care givers for monitoring and proper drug medication.

- Treatment failure**:** a patient whose sputum smear or culture is positive at five months or later during treatment. Also, included in this definition are patients found to harbor Rifampicin Resistance or multi-Drug Resistance (MDR) strain at any point in time during the treatment.

- Transfer-out: a patient who has been transferred from one facility to another outside the catchment area of the District Health Directorate to continue treatment where the end results of the patient (treatment outcome) cannot be traced.

According to Biruk et al. (2016) which is in line with WHO categorization, treatment outcomes were further grouped into Successful treatment outcome and unsuccessful treatment outcome.

- Successful treatment outcome: this refers to patients that were completely cured of the infection with corresponding bacteriological negative results after five months of anti-tuberculosis treatment and patients who have completed the full anti-tuberculosis dose without a bacteriological results.
- Unsuccessful treatment outcome: this refers to patients that were lost-to-follow-up by TB care givers, those who defaulted and patients who die for any cause while on treatment.

i. Censoring: this basically is the situation where observations under study failed to experience the event of interest during the follow-up period. In censoring, information about an object's event time is known but the actual event time is not known. In most times, an incomplete time is observed due to;

- Individual not experiencing the event by the end of the study.
- Individual lost-to-follow-up.
- Individual voluntarily withdrawing from the study due to reasons known to him or her alone.

Censoring can be classified into three main categories. That is; right censoring, left censoring and interval censoring.

- Right censoring: according to Dzimah et al. (2016), this type of censoring exists if an individual experienced the event of interest after the given follow-up period. Right censoring is further divided into type1 where the individual will not experience the event before the event time expires, a type 2 censoring where subjects does not experience the event since the process will be terminated after the required number is reached and a random censoring where the entire event termination process is not under the control of the researcher since individuals have different censoring times aside the study or event time set by the investigator.

- Left Censoring: this is a type of censoring where subjects under study experience the event before the time set by the investigator. A typical example is TB patient recovering before the 168 days set for this study.

- Interval Censoring**:** this is a type of censoring where the only information available to the investigator is that, the event of interest has occurred say between time A and B but as to when exactly this event occurred, the researcher cannot tell. For example, a TB patient who tested positive after three months of treatment and again tested negative after the fifth month is said to recovered between the third and firth months but the exact day of recovery cannot be determined by the investigator is said to be interval censored.

ii. Censored: a complete dose for the treatment of tuberculosis is a six months medication (28 days in a month). This therefore gives an expected recovery by 168 days (24weeks) of treatment which coincides with

literature of 24.14weeks (168.98 days) by Jakperik et al. (2012). Therefore, all observations that had their treatment time exceeding 168 days together with patients who died while on treatment, patients who were lost-to-follow-up, those transferred out and could not be traced for information on their treatment outcome and those that defaulted were classified as observations that are censored.

3.4 Bayesian Analysis

3.4.1 Bayes' Rule

The Baye's rule which drives its bases from the conditional probability is given as;

$$P(y/x) = \frac{P(x/y)P(y)}{P(x)} \tag{3.1}$$

The Bayes' rule in equation (3.1) above in a discrete form can be written as

$$P(y/x) = \frac{P(x/y)P(y)}{\sum_y P(x/y)P(y)} \tag{3.2}$$

However, continues form of the Bayes' rule in equation (3.1) is given as:

$$P(y/x) = \frac{P(x/y)P(y)}{\int dy P(x/y)P(y)} \tag{3.3}$$

Where $P(y/x)$, $P(y)$ and $P(x/y)$ is the posterior, the prior and the likelihood respectively. (Kruschke et al., 2010).

3.5 Survival Analysis

This is a collection of statistical procedures used in analyzing time-to-event data. The term "failure" is mostly used to describe the occurrence of an event of interest to the researcher even though the event could actually be a success like the recovery of TB patients (Stevenson et al., 2009). The duration between the inception of monitoring and the time of occurrence of a failure is called survival time. This time can be days, weeks, months and even years for an individual until the event of interest occurs (Dzimah et al., 2016). Typical examples of time-to-event data are; from the onset of symptoms of TB to time patient recovers, time of treatment commencement to end of treatment, time of start of a football match to time of scoring the first goal in the game and finally, the commencement of a surgery to time of surgery completion.

3.5.1 Survival Function

This gives the probability that an event of interest has not yet occurred by time *t*.

In other words, it gives the probability that the survival of the event of interest has gone past a given time *t*

The survival function $S(t)$ is given as:

$$S(t) = \Pr\{T \geq t\} = 1 - F(t) = \int_t^{\infty} f(t)dt \,. \qquad (3.4)$$

Where $F(t)$ is the cumulative density function (cdf) and $f(t)$ is the probability density function (pdf). Given that *t* is bounded by 0 and ∞, $S(t)$ is non-increasing with $S(t) = 1$ at $t=0$ and $S(t) = 0$ at $t = \infty$ (Dzimah., 2016; Jakperik., 2011).

3.5.2 The Hazard Function

This refers to the sudden rate at which events occur given no previous events (Dzimah, 2016).

It is simply the probability of a subject experiencing an event within a small (limited) time interval (Chechile et al., 2003). The risk of failure at any point in time can only be explained by the hazard function

According to Dzimah, (2016), the hazard rate function $h(t)$ is given as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}, \tag{3.5}$$

within the interval of $[t,\ t + dt]$.

Also $h(t)$ can be written as:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} InS(t). \tag{3.6}$$

3.5.3 Relationship between the Hazard and Survival functions

Given the hazard rate function $h(t)$ in equation (3.3) above, there exist a relationship between hazard and survival as follows:

$$h(t) = -\frac{d}{dt} InS(t). \tag{3.7}$$

Taking an integral of the hazard $h(t)$ from 0 to t gives:

$$S(t) = \exp\{-\int_0^t h(x)dx\}. \tag{3.8}$$

3.5.4 Kaplan-Meier estimator

The K-M estimator $S(t)$ is defined as:

$$S(t)=\left\{\prod_{j:t_j\le t}[1-\frac{d_j}{n_j}]\right\}. \tag{3.9}$$

Where $t_1<t<t_k$

(Allison, 2010).

Where n_j represents number of observations that are yet to experience the event of interest and are not also censored by time t. j is the premier failure time and d_j is the number of individuals who experienced the event at time t.

Jackknife was used to validate the K-M estimate of the average survival time.

According to Kaplan and Meier (1958), the K-M estimator has the following assumptions

1. Irrespective of the time a subject is selected, the survival probability is the same.
2. Censoring is independent of the prognosis of a patient.
3. Events actually occurred at the event time.

3.5.5 Bayesian Accelerated Failure Time (AFT) Model

According to Prabhash et al. (2016), the AFT is given as

$$\log\left(t_{ij}\right)=\alpha+\beta x_{ij}+\sigma\varepsilon_{ij}\,. \tag{3.10}$$

Where $\alpha + \beta x_{ij}$ is the linear predictor and ε_{ij} the error term.

The density, survival and the hazard are given as follow

$$f\left(t_0 / \lambda_{ij}\right) = \frac{1}{\sigma t_0} f_0\left(\frac{\log\left(t_0\right) - \lambda_0}{\sigma}\right) \tag{3.11}$$

$$s\left(t_0 / \lambda_0\right) = s_0\left(\frac{\log\left(t_{ij}\right) - \lambda_{ij}}{\sigma}\right) \tag{3.12}$$

And

$$h\left(t_0 / \lambda_{ij}\right) = \frac{1}{\sigma t_{ij}} h_0\left(\frac{\log\left(t_{ij}\right) - \lambda_{ij}}{\sigma}\right) \tag{3.13}$$

The likelihood function is given by

$$L = \prod_{l=1}^{n} \prod_{j=1}^{n_i} \left[\frac{1}{\sigma t_{ij}} f_0\left(\frac{\log\left(t_{ij}\right) - v_{ij}}{\sigma}\right)\right]^{v_{ij}} s_0\left(\frac{\log\left(t_{ij}\right) - v_{ij}}{\sigma}\right)^{1-c_{ij}} \tag{3.14}$$

The AFT model parameters were estimated by posterior sampling based on a Markov Chain Monte Carlo (MCMC) simulation method

If p(a) and p(b) are the prior distributions for the parameters a and b, the posterior distribution will be given as

$$p\left(b, v, a / t\right) \propto L\left(t / b, v\right) p\left(w / a\right) p\left(b\right) p\left(a\right) \tag{3.15}$$

The Bayesian results for the Accelerated Failure Time models and the Binary logistic regression was done using the SAS Bayesian Built-in procedure (GENMOD). The Built-in Bayesian procedure makes use of default prior distributions depending on the model specified in the SAS coding.

In this case, the prior distribution selected for the log-logistic is Gamma.

Model assumptions

1. The model accepts censored observations.
2. It requires that the predictors are multiplicative (proportional) with respect to survival time
3. It also assumes that the estimated time ratios are constant.

The most frequently used accelerated failure time distributions are; exponential distribution, Weibull distribution, gamma distribution, lognormal and log-logistic distribution.

3.5.6 Exponential Distribution

The exponential distribution of the AFT model relies on the assumption that, the hazard ($h(t)$) depends solely on the coefficients and dependent variables which are time invariant. It is a memory less distribution (the occurrences of future events are independent of past events) with just a single parameter (λ).

The probability density function, hazard function, cumulative hazard and the survival functions of the exponential distribution are shown below.

The probability density function is given as;

$$f(t) = \lambda \exp(-\lambda t) \quad t > 0, \lambda > 0. \tag{3.16}$$

The hazard function is given as;

$$h(t) = \frac{f(t)}{S(t)} = \lambda ; \quad \lambda > 0 . \tag{3.17}$$

Therefore;

$$H(t) = \lambda t . \tag{3.18}$$

The cumulative density function of the exponential distribution will now be given as;

$$F(t) = 1 - \exp(-\lambda t) \quad 0 \le t < \infty . \tag{3.19}$$

This gives rise to a survival function as follows

$$S(t) = 1 - F(t) = 1 - (1 - \exp(-\lambda t)) = \exp(-\lambda t) . \tag{3.20}$$

The main disadvantage of the exponential distribution is the assumption of a constant hazard over time.

3.5.7 Weibull Distribution

Unlike the exponential distribution, the Weibull distribution is a two-parameter monotonic decreasing or increasing distribution with a constant hazard. The two parameters are the shape parameter (α) and the scale parameter (λ). Where α show how the behavior of the failure is. That is α <1 and α >1 means a decrease and an increase in the failure rates respectively.

The probability density function, hazard function, cumulative hazard and the survival functions of the Weibull distributions are shown as follows.

The probability density function

$$f(t) = \lambda\alpha(\lambda t)^{\alpha-1} * \exp(-(\lambda t)^{\alpha}) . \quad (3.21)$$

The Hazard function is given as:

$$h(t) = \lambda\alpha(\lambda t)^{\alpha-1} . \quad (3.22)$$

The cumulative hazard function is shown below

$$H(t) = (\lambda t)^{\alpha} . \quad (3.23)$$

The survival function is as follows:

$$S(t) = \exp-(\lambda t)^{\alpha} . \quad (3.24)$$

3.5.8 Gamma Distribution

The gamma which has the Weibull imbedded also have both the scale (λ) and shape (κ) parameters.

The probability density function of the Gamma distribution is given as:

$$f(t) = \frac{\lambda^{\kappa} t^{\kappa-1}}{\Gamma(\kappa)} \exp(-\lambda t); \quad t > 0, \lambda > 0 \text{ and } \kappa > 0 . \quad (3.25)$$

The hazard function is

$$h(t) = \frac{\lambda^{\kappa} t^{\kappa-1} \exp(-\lambda t)}{(1 - I_{\kappa}(\lambda x))\Gamma(\kappa)} . \quad (3.26)$$

The survival function is shown below

$$S(t) = 1 - I_{\kappa}(\lambda t). \tag{3.27}$$

3.5.9 Log-Normal Distribution

The log-normal is an AFT distribution which follows the normal distribution with the parameters μ and σ^2 if logarithms of its variables are taken.

The probability density function, hazard and survival functions of the log-normal distribution are presented below

The probability density function is given as:

$$f(t) = (2\pi)^{-\frac{1}{2}} \alpha t^{-1} \exp\left(\frac{-\alpha^2 (\log(\lambda t))^2}{2}\right). \tag{3.28}$$

The survival function is

$$S(t) = 1 - \Phi(\alpha \log(\lambda t)). \tag{3.29}$$

The hazard is also given as:

$$h(t) = \frac{f(t)}{S(t)} = \frac{(2\pi)^{-\frac{1}{2}} \alpha t^{-1} \exp\left(\frac{-\alpha^2 (\log(\lambda t))^2}{2}\right)}{1 - \Phi(\alpha \log(\lambda t))}. \tag{3.30}$$

Where $\alpha > 0$ and $\lambda > 0$ and Φ is the cumulative distribution function of the standard normal.

At $t = 0$, the hazard of the log-normal distribution is 0. Though not perfect as compared to the log-logistic, the log-normal can also be used to analyze survival

data of TB patients since the probability of dying initially increases and later decrease as time goes up.

3.5.10 Log-Logistic Distribution

The log-logistic AFT distribution is an alternative to the Weibull distribution that takes the form of the standard logistic distribution but with an error term.

This distribution can fit non-monotonic hazards where the hazards can rise and fall. This distribution is appropriate for fitting models for diseases such as Tuberculosis where the hazard can decrease, increase or hump-shaped. The merit of this distribution is that, the coefficients can be translated as time or odd ratios.

The pdf, hazard and the survival functions are stated below.

The probability density function is

$$f(t) = \lambda\alpha(\lambda t)^{\lambda-1}(1+(\alpha t)^{\lambda})^{-2}. \tag{3.31}$$

The hazard function is also given as;

$$h(t) = \frac{\alpha\lambda(\alpha t)^{\lambda-1}}{1+(\alpha t)^{\lambda}}. \tag{3.32}$$

The survival function is given as:

$$S(t) = \frac{1}{1+(\lambda t)^{\alpha}}. \tag{3.33}$$

Where $\alpha > 0, \lambda > 0$

The log-logistic is very simple compared to the Weibull and the lognormal when dealing with censored observations.

3.5.11 Model Selection

The Accelerated Failure Time models that were tested to ascertain their fitness are; the Weibull distribution, the Gamma distribution, the Exponential distribution, the log-normal distribution and the log-logistic distribution.

The model selection criteria were based on accuracy measures such as the Akaike's Information Criterion (AIC), the corrected Akaike's Information Criterion (AICc), the Bayesian Information Criteria (BIC), the Deviance Information Criterion (DIC) and the Log-likelihood values to determine the AFT model that best fit the TB data.

The model with the lowest AIC, AICc, BIC, DIC and highest log-likelihood values is selected as the model or distribution that best fit the tuberculosis data.

Deviance Information Criteria (DIC)

$$DIC = 2\overline{D(\theta)} - D\left(\hat{\theta}\right) \tag{3.34}$$

Where $2\overline{D(\theta)}$ is the mean deviance for the posterior distribution while $D\left(\hat{\theta}\right)$ is the estimate for the effective number of parameters in the model.

(Gao et al., 2011)

Bayesian Information criteria (BIC)

$$BIC = -2(\log - likelihood) + (p + k) * \log(n) \tag{3.35}$$

Where p, k, and log(n) are the number of parameters, coefficients and observations respectively

(Barman et al., 2017)

3.5.12 Logistic Regression Analysis

According to Ibrahim (2019), logistic function is the inverse cumulative distribution function of the logistic distribution which creates a map of probability values from [0, 1] to $[-\infty, +\infty]$. Given q as the probability, the odds ratio will be $\frac{q}{1-q}$.

The logarithm of the odds gives the logit of the probabilities as shown below

$$\log it(q) = \log\left(\frac{q}{1-q}\right) = \log q - \log(1-q). \quad (3.36)$$

The logistic of any function say α is the inverse of the logit function given as:

$$\text{logit}(\alpha) = \text{logit}^{-1}(\alpha) = \frac{1}{1+e^{-\alpha}}. \quad (3.37)$$

(Ibrahim, 2019).

Therefore, the binary logistic regression for recovery is formulated as follows:

$$P(R) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_p x_p+\varepsilon)}} \quad (3.38)$$

Where:

P(R) = the probability or chance of predicting the outcome variable (in this study the probability of recovering).

$\beta_1, \beta_2, ..., \beta_p$ = The model parameters,

β_0 = A constant.

$x_1, x_2, x_3, ..., x_p$ = The independent variables

ε = The error term.

Model Assumptions

1. A binary or ordinal logistic regression requires that the dependent variable is binary or dichotomous
2. Logistic regression requires that the data set is relatively large
3. There must be no or little multi-collinearity among covariates
4. There exists a linear relationship between the dependent and independent variables.

The fitness of the binary logistic regression model to the data was tested using the Hosmer-Lemeshow (HL) goodness of fit test. A small Chi-Square value with a large probability value (P-Value greater than 0.05) of HL indicates that, there exists no evidence of poor model fit.

And finally, the discrimination ability of the model was also verified using the Receiver Operating Characteristic (ROC) Curve.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.0 Introduction

The results of the study emanating from the application of the methods described under methodology and discussion of results are presented in this chapter

4.1 Preliminary analysis

A total of 308 TB patients are registered for treatment in Buluk from January 25, 2012 to March 20, 2020. Out of this number, 250(81.16%) of the patients recovered. Also, 200(64.94%) patients where censored while 27(8.77%) out of the 308 patients registered for treatment died.

4.1.1 Descriptive Statistics based on Gender

From Table 4.3 (Appendix I), about 237(76.95%) of patients registered for TB treatment are males while the female patients are 71 representing 23.05% of the total registered patients. It could be observed that, mortality for males stands at 21 out of a total of 237 patients representing 8.86% of the patients.

Also, out of the 71 female patients, about 6 (8.45%) patients lost their lives to the disease while on treatment. From the mortality statistics presented, it is clear and evidential that TB related mortality among males in Buluk from 2012 to 2020 was relatively higher than females.

The recoveries of males and females is 188(79.32%) and 62(87.32%) respectively. The deaths and recoveries in the gender statistics indicated a higher

percentage of TB deaths and lower recovery among males. However, the percentage death among females is lower, with a corresponding higher recovery of 87.32%. Also, it was noticed that, a total of 152 representing 64.14% of the 237 male patients who suffered the sickness within the period got censored whereas female patients that were censored are 48 out of the 71 female patients. A censored percentage of about 67.61% among females is relatively high compared to the male censoring.

4.1.2 Descriptive Statistics based on Age

The results presented in Table 4.2 showed that the minimum age for patients registered for treatment within the period was 17 years while the maximum age recorded was 87 years. The mean age was seen to be 47.08 with a standard deviation of 17.363. However, the median age was discovered as 43 years.

Table 4.2 Descriptive Statistics for Some Covariates

Parameter	Min.	Max.	Mean	Median	Std.Dev
Age	17	87	47.08	43	17.363
Dead	27	85	53.07	49	17.095

4.1.3 Descriptive Statistics for a Type of Patient

New cases registered for treatment within the period were 265(86%), out of which 174(65.66%) patients were censored. A total of 24 of the 265 new cases died while on treatment. Also, the number of non-new cases that were registered for treatment is 43(13.96%). Non-new cases that were censored are 26(60.47%) while

mortality for non-new cases stood at 3(6.98%). Recovery is high among non-new 36(83.72%) than new cases 214(80.75%).

4.1.4 Descriptive Statistics for Disease Categories

Tuberculosis Patients who tested pulmonary positive were 217 (70.45%) while extra-pulmonary TB and Pulmonary negative patients were 2(0.65%) and 89 (28.90%) respectively. The number of deaths recorded for each of pulmonary positive and pulmonary negative were 20(9.22%) and 7(7.87%) respectively. However, there was no single death recorded among extra-pulmonary TB patients. The recoveries for pulmonary positive patients is lower (78.80%) compared to the pulmonary negative and the extra-pulmonary TB which recoveries were 86.52% and 100% respectively. The censored observations for pulmonary positive, pulmonary negative and extra-pulmonary TB are 139(64.06%), 59(66.29) and 2(100%) respectively as shown in table 4.3.

4.1.5 Descriptive Statistics for Smear Results

The number that tested smear positive and smear negative out of the 308 are 196 (63.6%) and 112 (36.3%) respectively. Deaths for smear positive and smear negative are 17(8.67%) and 10(8.93%). Recoveries stood at 155(79.08%) and 95(84.82%) for smear positive and smear negative while the number of censored observations for both smear positive and smear negative are 127(64.80%) and 73(65.18%).

4.1.6 Descriptive Statistics for HIV Testing Results

A total of 301 representing 97.73% of TB patients under treatment within the study period were also tested for HIV/AIDS which reveal an HIV/TB co-infection of 32(10.39%). This is an indication of a strong interest and a high HIV testing rate among TB patients. HIV/TB co-infected patients recorded the highest percentage of deaths (18.75%) followed by a 14.29% for TB patients who were not tested for HIV, and finally a 7.43% of deaths being HIV negative TB patients.

4.1.7 Descriptive Statistics for Treatment Outcomes

Table 4.4 revealed that out of the 308 patients, a total of 170 patients representing 55.19% were completely cured and discharged, 80(25.97%) who actually tested smear negative initially before the commencement of the TB treatment successfully completed the mandatory 24 weeks full dose. Just as it was done by Biruk et al. (2016), the various treatment outcomes were further divided into two main categories. That is, successful treatment (which can also be referred to as patients who have completely recovered from the disease) and unsuccessful treatment. Since both patients who are cured and patients who completed treatment tested negative after the six months of taking the complete TB dose, they are together classified as patients who have actually recovered after treatment or patients that experienced treatment success. This therefore gave rise to a total recovery of 250(81.16%). Also, a total of 23(7.47%) patients could not be traced by TB care givers to assist them complete their medications (Lost to follow up), 2(0.65%) defaulted in the drug administration along the line, 6(1.95%) were referred out of the districts whereas a total of 27(8.77%) lost their lives

while receiving treatment. The patients who died, defaulted, transferred out, or lost to follow-up were classified as unsuccessful treatment (18.84%).

Table 4.4: Frequency of Treatment Outcomes

Treatment Outcome	Frequency	Percent
Treatment completed	80	25.97
Cured	170	55.19
Defaulted	2	0.65
Died	27	8.77
Lost to follow	23	7.47
Referred	6	1.95
Total	**308**	**100**

4.2 FURTHER ANALYSIS

4.2.1 Median Recovery Time

The median recovery time was found to be 171 days with a confidence interval of (169 – 175) days. This translated into approximately 6 months of treatment. These estimates were validated using Jackniffe estimates which yielded a median recovery time of 171 (169,174).

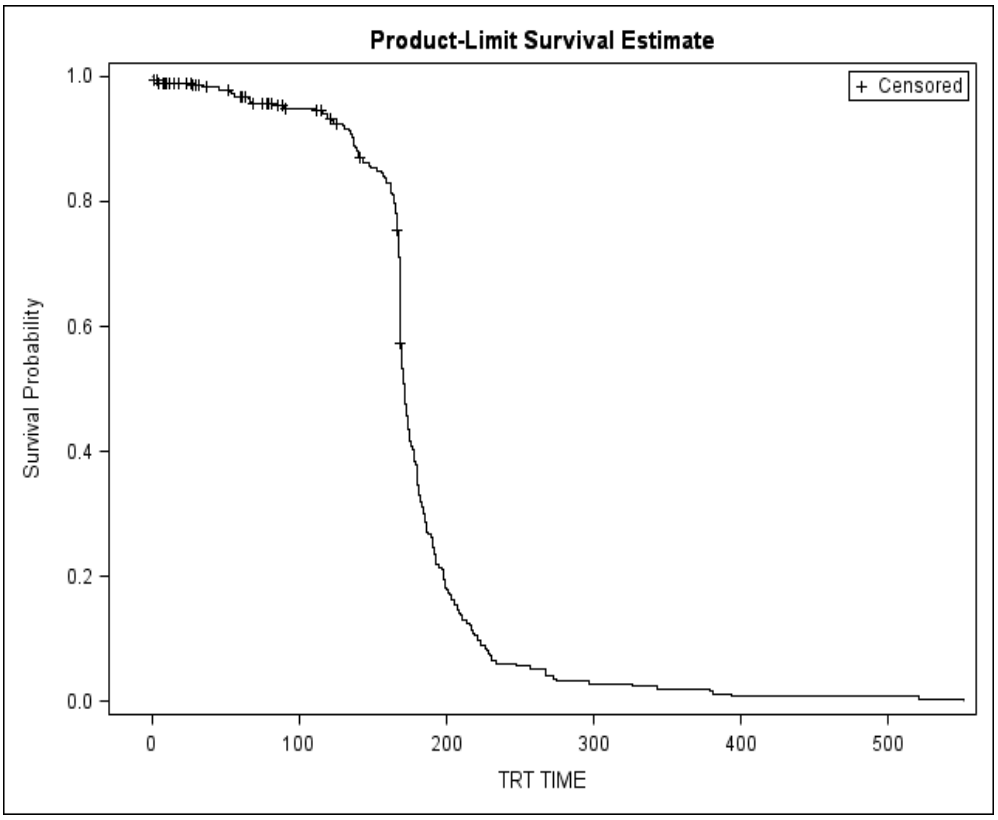


Figure 4.1 Kaplan-Meier plots for treatment time

Table 4.6 Median Recovery Time Estimates

			95% Confidence	
Percent	**Point Estimate**	**Transform**	**Interval [Lower Upper]**	
75	190.000	LOGLOG	184.000	198.000
50	**171.000**	**LOGLOG**	**169.000**	**175.000**
25	167.000	LOGLOG	164.000	168.000

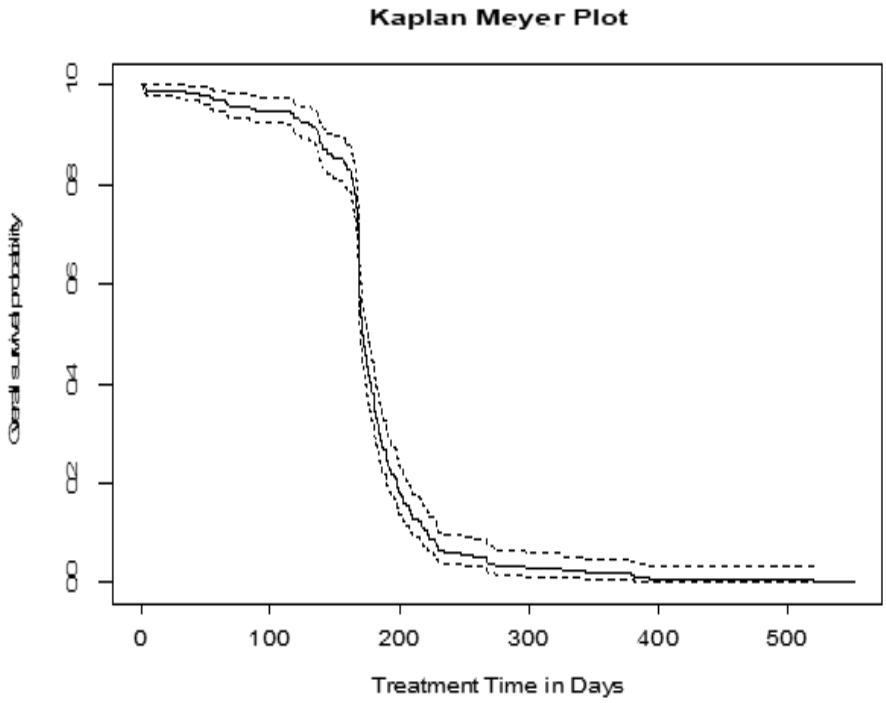


Figure 4.2 Kaplan-Meier plots for treatment time

The Kaplan-Meier approach was used in estimating the median recovery times of TB patients based on gender as well as the median recovery times for the children, adult and elderly age categories.

A probe uncovered that; the median recovery time based on Gender varied significantly from the overall average recovery time of patients in Buluk. In table 4.7, the median recovery time for males and females are 170(168.000-174.000) and 174(168-180) days respectively. The median recovery time for children, adults and the elderly as presented in table 4.7 (appendix I) are 192(169.000 - 216.000), 170(168 - 173) and 180(169.000- 186.000) days.

4.2.2 Accelerated Failure Time Model

4.2.2.1 Goodness of Fit Test

The estimated AIC, AICc BIC, DIC and Log-likelihood values of fitted AFT distributions are shown in Table 4.8.

Table 4.8: AFT Models Goodness of Fit Test Results.

Model	AIC	AICc	BIC	DIC	LOGLIK
Weibull	264.046	264.720	296.662	264.186	-123.023
Log-logistic	**98.733**	**99.729**	**138.597**	**96.636**	**-38.366**
Exponential	561.119	561.793	593.735	560.563	-271.559
Gamma	445.001	445.742	482.302	445.324	-212.501
Log-Normal	300.676	301.503	336.916	300.793	-140.338

According to the goodness of fit test results presented above, all the statistical measures (AIC, AICc, DIC and BIC) had relatively lower estimates for the log-logistic distribution with a corresponding high Log-Likelihood. The relatively smaller estimated values in favour of the Log-logistic places it as the best fit accelerated failure time distribution for the tuberculosis data. Also, convergence of the log-logistic distribution was studied and parameters were duly converged as can be seen in the trace plots (Apendix-III)

4.2.2.2 Log-Logistic Distribution

Log-logistic is identified to be the most appropriate distribution that best fitted the TB data. From the posterior summary statistics of the log-logistic below, it could be seen from the mean parameter estimates and the standard deviations that, the only factors that were significantly associated with treatment outcome are age, smear results and treatment time (table 4.12). However, parameters like gender, TP, DC and HIV seems not to have any significant influence in the recovery of a TB patient in Buluk

Table 4.12 Posterior summary statistics for Log-Logistic Regression (without interactions)

Parameter	DF	Mean Estimate	Standard Error	95% Confidence Limits	
Intercept	1	13.3045	61113.21	-119766	119793.0
AGE	**1**	**0.0131**	**0.0063**	**-0.0013**	**0.0235**
GENDER 0	1	-0.0235	0.2369	-0.4878	0.4408
GENDER 1	0	0.0000			
DC 0	1	-4.8067	60476.11	-118536	118526.2
DC 1	1	-5.5611	60476.11	-118537	118525.4
DC 2	0	0.0000			
TP 0	1	-0.3296	0.2932	-0.9043	0.2451
TP 1	0	0.0000			
SR 0	**1**	**-0.8479**	**0.4052**	**-1.6422**	**-0.0537**
SR 1	0	0.0000			
HIV0	1	-5.9124	8800.315	-17254.2	17242.39

HIV 1	1	-5.8431	8800.315	-17254.1	17242.46
HIV 2	0	0.0000			
TIME	**1**	**0.0324**	**0.0029**	**0.0267**	**0.0382**
Scale	**1**	**0.3402**	**0.0512**	**0.2533**	**0.4568**

Further probe by way of interacting the covariates in the log-logistic distribution uncovered age, HIV and the interaction between the two as factors that significantly affect treatment outcome (Table 4.9; appendix 1).

4.2.3 Logistic Regression Analysis

A small Chi-Square value of 4.6750 and a relatively large P-value of 0.7917 as can be found in Table 4.14 is an indication that, the binary logistic regression is an appropriate model for fitting the TB data. Also, the discrimination ability of the model was determined using the ROC curve. According to figure 4.5 (appendix III), there is no discrimination found for the model and that further confirmed the fitness of the logistic model for the data.

Table 4.14 Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
4.6750	8	0.7917

The multivariate logistic regression results as shown in Table 4.11(Appendix I) revealed four parameters that are significantly associated with treatment outcome. These parameters are age, disease category, HIV and treatment time. This insinuated that, there is a relationship between age, disease category, HIV status

and treatment time with the recovery of a TB patient. The model for TB recovery is presented below.

$$P(TO=1)=\frac{1}{1+e^{-[48.5965-0.0419A-19.4654\,DC-26.6675\,H+0.0399\,T]}}$$

Where A = age, DC=Disease Category, H = HIV/AIDS status and T= Treatment Time

The odd ratio estimates in Table 4.10 (Appendix I) suggests that, the odds for the treatment time is (OR= 1.039; 95% CI: 1.027 – 1.050).

The odds of age, disease category, Smear Results, Type of patient, gender and HIV are (OR: 0.988; CI: 0.954 - 1.023), (OR: 0.190; CI: 0.020 - 1.770), (OR: 3.632; CI: 0.469 - 28.121), (OR: 2.918; CI: 0.437 - 19.477), (OR: 1.009; CI: 0.213 - 4.768) and (OR: 1.136; CI: 0.239 - 4.768) respectively.

4.3 DISCUSSION

Out of the 308 patients in the study, 200(64.94%) patients are censored. The case fatality is 27(8.77%) and the treatment success which is technically the recovery is 250(81.16%). The percentage of males and females are 76.95% and 23.05% respectively. A male dominated infection in this research agreed with the findings of a study in Cape Coast Metropolis by Tetteh et al. (2018) where the percentage of males and females were 74.6% and 25.4%. A treatment outcome of 81.16% though impressive, it is still below the target of 85% by the WHO and the achieved treatment success of 90.2% by the central region of Ghana within the period of 2012 to 2016 (Tetteh et al., 2018).

However, it is also better than the 75% recovery of adult TB patients at South-West Ethiopia (Terefe et al., 2018) and the 73.75% recorded in the Upper West Region (Jakperik et al., 2013). Tuberculosis related mortality recorded was 27(8.77%). This again is seen to be higher than the percentage of TB deaths in South-West Ethiopia (4.4%) and the 8.5% recorded in the Central Region but lower than the 17.7% at an Ethiopia university hospital (Tetteh et al., 2018; Biruk et al., 2018; Terefe et al., 2018). The recoveries by gender were found to be 79.32% and 87.32% while the respective gender-based deaths reported during the analysis were 8.86% and 8.45% for males and females respectively. This is in accordance with the findings of Jakperik et al. (2013) in the Upper West Region where the probability of recovery is higher in females than in males. Literature further revealed that, smoking, alcohol intake and works such as mining, construction and farming which are male dominated activities are risk factors

associated with tuberculosis (Ariyothai et al., 2004; Rehm et al., 2009). This could be the reason why males recorded a low percentage recovery compared to their female counterparts.

It could be seen from the analysis that, the percentage of males contracting the disease (76.95%) and the percentage of death (8.86%) among males are relatively high compared to the percentage of contraction (23.0%) and death (8.45%) among females. Residents of Buluk and its surrounding communities are predominantly farmers with some residents actively involved in small scale mining (galamsey) activities at communities like Kadema in Buluk and the Nyanguruma galamsey site in the Mampurugu Moaduri District of the North East Region. These two main occupational activities of the people produce dust. The heavy use of agrochemicals and cyanide for farming and mineral purification are all risk factors that are associated with TB. The most actively involved in the use of these chemicals are mostly the male population which probably is the cause of the huge number of males contracting the disease than females. Literature has it that, the use of hard drugs, high alcohol intake and smoking are TB risk factors (Ariyothai et al., 2004; Rehm et al., 2009) and all these activities are mostly practiced by males.

The percentage of patients cured as identified in this study was 55.19% while those who completed treatment stood at 25.97% which together formed a treatment success of 81.16%. Patients who were lost to follow up were 7.47%, transfer-out was 1.95% and the least is those who defaulted (0.65%). A default of 0.65% is considered a target achieved since it is far less than the WHO target of

5% for defaulters and the default of 0.9% reported for the Central Region (Tetteh et al., 2018).

The study revealed that, HIV testing among TB patients is 97.73% with an HIV/AIDS and TB co-infection of 10.39%. The percentage of deaths recorded for co-infected patients was 18.75% while their recovery is 71.88%. The high number of deaths recorded among HIV positive TB patients is an indication that seems to confirm the suspicion that the co-infection of TB and HIV can be very fatal and should be treated or handled with caution.

According to a research conducted to determine the "trends of tuberculosis case detection, mortality and co-infection with HIV in Ghana", the HIV testing among TB patients was 84.2% with HIV positivity of 22.6%. The fatality among co-infected patients was reported to be 21.8% as against the overall fatality of 13% (Osei et al., 2020). A co-infection of 10.39% is though quite high, but still remains one of the lowest in the country as against the 14.7% and 22.6% recorded earlier (Addo et al., 2018; Osei et al., 2020). Comparatively, the rate of testing for HIV among TB patients in Buluk is very high. This probably is due to prior knowledge about the impact of HIV/AIDS on TB patients. Research has it that, the best way to handle and deliver quality TB and HIV service is to screen for the two diseases or the early identification of both HIV and TB and linking the patients to the appropriate health institutions for effective monitoring and proper heath care to achieve desirable results (Naidoo et al., 2019). Even though the fatalities recorded among co-infected patients is still high, an early detection of co-infection cannot be ruled out as one of the factors that contributed to the high

recoveries recorded. Since TB is an opportunistic disease, a rigorous testing and early detection of HIV and any other immune weakening illness is in the right direction to fighting towards the reduction of TB prevalence.

From the three disease categories, it was found that, pulmonary positive, pulmonary negative and extra-pulmonary tuberculosis patients' fatalities were 9.22%, 7.87% and 0% respectively. The treatment success for pulmonary positive, pulmonary negative and extra-pulmonary cases were 78.80%, 86.52% and 100% respectively. A report on the extra pulmonary tuberculosis patients in Accra indicates that, extra pulmonary TB constituted 21.8% of the study population while pulmonary TB was 78.2 % (Ohene et al., 2019). According to Chennaveerappa et al. (2011), both pulmonary and extra pulmonary in south India was 65% and 35% respectively.

This study revealed a high percentage of pulmonary TB of 70.45% as against relatively low reports of extra pulmonary cases of 0.65%. That notwithstanding, the percentage of pulmonary tuberculosis is closely related to the findings in the case of Accra (Ogyiri et al., 2019).

Also, new cases constituted about 86% with deaths constituting 9.06% as against non-new deaths of 6.98%. The recoveries for both new and non-new cases for this study were 80.75% and 83.72% respectively. The recovery or treatment success of this study is in line with the report by Chennaveerappa et al. (2011) on data from a Teaching hospital in South India which suggests that cure (in this case the report is using the word cure for treatment success) for new and non-new TB cases were 84% and 83%. This is in contrast with the findings of this study which

reveal that, patients who were previously treated for tuberculosis had a higher percentage recovery from the disease than patients who contracted the disease and are receiving treatment for the first time. This could be attributed to the fact that, relapsed or defaulters (non-new patients) report early for treatment and are also committed to the medication due to past experience.

The Sputum smear test approach which is widely used in Ghana and beyond is the predominantly used in Buluk with few cases being detected by means of X-ray. Smear positive constituted 64.80% of cases under study out of which the fatality among smear positive patients detected was 8.67% while 79.08% of them had a treatment success.

Again, the study revealed TB prevalence to be high among the adult population with a percentage of 76.30% followed by the aged which also constituted 23.05% of the study population with the least and seemly negligible being children (0.65%). This is consistent with the age specific rates of the Cape Coast Metropolis where the adult population which in their case is people within the age bracket of 20 to 59 constituted 79.8% followed by the aged (that is persons of age 60 years and above) who were 14.1% and finally the children just like the case of this study having the least infection rate of 6.1%. A high rate of infection among this age group could be as a result of the fact that they are obviously the active workforces who are normally engaged in activities that exposed them to the risk of infection. Some of these risks are; farming activities that exposes them to dust and agrochemical inhalation, Galamsey operations which also expose them to the risk of cyanide and other poisonous chemical intake and probably the cohort that

have the tendencies of being under the addiction of high alcohol intake, smoking and drug abuse. An attack or infection on this category of people in society is basically an attack on the food supply chain or in general terms is an attack on the economy of the nation since they are the workforce.

According to table 4.2, the youngest TB related death was 27 years old and the oldest to have died while receiving treatment is 85 years. The mean and the median ages to have died are 53.07 and 49 respectively with a standard deviation of 17.095. It is shown in the table that, the hard-hit age group (age group with the highest fatality) in terms of TB related Mortality is patients who are more than 60 years. This is in conformity with the situation of Addis Ababa, Ethiopia as reported by Tolosie et al. (2014).

The age specific fatality of this study is 0%, 7.82% and 12.68% for children, adults and the aged. Though, prevalence is very high among adults, the aged turn to die more than any other age group when they contract the infection. This may be due to the fact that tuberculosis is an opportunistic disease that kills faster in populations that have their immune systems compromised by either other illnesses or age. And the aged are such patients who normally suffer from comorbidities such as diabetics. TB mortality of 0% and 7.82% among children and adults is not too different from the results of Tetteh et al. (2018) who have not recorded a single death for children who are below 19 years and a 7.5% for adults (persons within the ages of 20 to 59). The data further showed that, the recovery of TB patients decreases with an increase in age where the recoveries of children, adults and the aged are 100%, 81.28% and 80.28% respectively.

In this study, the median recovery time (50th percentile) which is the most efficient measure of central tendency in survival analysis (since the mean can be affected by extreme values) is 171 days with a 95% confidence interval of 169 and 175 days (this is 5 months 21 days which is approximately six months) or 24.43 weeks with the 25th and 75th percentiles being 167 and 190 respectively. This is similar to findings in the upper West region of Ghana where median recovery time of 25.43 weeks was recorded (Jakperik et al., 2013) and a median recovery in Northern Ghana of 24.14 weeks (Jakperik et al., 2011). However, it is quite higher than an average treatment time of 22 weeks as reported in Northern region, Ghana (Jakperik et al., 2012) but far lower than a median recovery time of 185 days (Terefe et al., 2018). A lower average recovery time in this case could be attributed to early detection and treatment, low comorbidity as well as high DOTS measures adherence by patients.

Also, the median recovery time for males and females are 170(168.000 - 174.000) and 174(168 - 180) days respectively. The median recovery time for children, adults and the elderly as presented in table 4.7(appendix 1) are 192(169.000 - 216.000), 170(168 - 173) and 180(169.000 - 186.000) days respectively. In this regard, the adult who are usually active and probably have relatively stronger immune systems seems to recover faster than the aged and children.

But surprisingly, it was uncovered that persons above age 60 have the chance of recovering from tuberculosis faster than children. Though age is instrumental in the recovery process of a TB patient, a possibility of children having a longer recovery period than the other age groups could be attributed to comorbidities

such as malnutrition, HIV and other infections that compromised their immune system fighting ability. Despite the impact of the immune system on the recovery of TB patients due to age; strict adherence to the DOTS protocols is a key determinant of treatment outcomes and time (Terefe et al., 2018) which possibly is the reason for a lowered treatment time in the case of the aged.

In an attempt to identify the specific distribution of the accelerated failure time model that best fits the tuberculosis data under study, the Weibull, exponential, Gamma, log-Normal, and the Log-Logistic distributions were tested with the data using the AIC, AICc, BIC, DIC and the Log-likelihood values. The relatively smaller AIC AICc, BIC, DIC values and a high LOG-LIKELIHOOD value in table 4.8 suggests the log-logistic regression is the best fit model for the tuberculosis data. The log-logistic model suggests that, the covariates; age, smear results and treatment time are the significant parameters. A probe by interacting some of the covariates further revealed the age, HIV status of a patient and the interaction between the two independent variables to be significant. A decreasing treatment success with an increase in age is in line with literature where an increase in age results in a decrease in treatment success in both Ghana and china (Abuaku et al., 2010; Ogyiri et al., 2019). Although recoveries among non-new case patients were seen to be higher than new cases in percentage terms, the type of a TB patient is not significant in this study. The chances are that, both new and non-new patients were committed to the medications and none probably is a drug resistant TB patient.

Just like the log-logistic, a logistic regression analysis revealed that, the most significant variables that were associated with treatment outcome are age, disease category, HIV and the duration of treatment. The age of a patient has an impact on the treatment of tuberculosis. Age being identified as a contributing factor in the treatment of tuberculosis is consistent with findings of other researchers (Jakperik et al., 2015; Jakperik et al., 2012; Pardeshi et al., 2007). This will mean that the success of tuberculosis treatment turns to decrease as the age of a person increases. Therefore, the probability of recovery is higher among younger patients than the aged and the disease category of a patient also have a great impact on the treatment success of a patient.

According to the odd ratio estimates in table 4.10, a unit increase in the age of a patient has a negative effect on the odds of recovery. This is an indication that a one-point increase in the age of a patient is associated with 1.2% decrease in the odds of recovery. Tuberculosis is an opportunistic disease that takes advantage of the weakness of one's immune system. The decrease in the rate of recovery of a patient as one gets older could be due to the fact that, the aged have relatively weaker immune systems and cannot actively fight enough to aid faster recovery. Also, there exists an effect of the treatment time with recovery. This could be translated into about 3.9% increase in the odds of recovery with a corresponding one unit increase in treatment time. Interestingly, as there is no effect of type of tuberculosis, gender and smear results, disease category has proven to be strongly associated with recovery but with a negative effect. Finally, there is 13.6% increase in the odds of recovery among non-co-infected patients than HIV

positive tuberculosis patients. HIV is an immune compromising disease that weakens and reduces the power and ability of the immune system from fighting the *mycobacteria* and any other infection. This probably is the reason why HIV negative patients have about 13.6% chance of recovery over patients co-infected with TB/HIV. However, recovery among co-infected patients can be improved if such patients are put on immune boosters such as antiretroviral drugs or are advised to eat foods that can boost the immune system.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the summary of the findings of the study, conclusions and recommendations emanating from the study

5.2 Summary

The prime purpose of this study was to estimate the median recovery time of TB patients and model the recovery of patients with its associated prognostic factors.

The preliminary analyses unveil that out of the 308 patients, 273 were males and 71 were females. The overall mortality was found to be 8.77% whereas patients who recovered from the infection recorded 81.16%. Treatment success was higher among females (87.32%) than males (79.32%). However, males recorded high percentage deaths (8.86%) than females (8.45%). Treatment outcomes that were uncovered from the study were treatment complete, cured, defaulted, died, lost to follow-up and transferred out with corresponding percentages of 25.97%, 55.19%, 0.65%, 8.77%, 7.47% and 1.95% respectively. A high HIV testing rate of 97.73% revealed a co-infection of 10.39% with 18.75% mortality among co-infected patients.

The most prevalent type of TB was discovered to be pulmonary positive (217) followed by pulmonary negative (89) and 2 cases of extra pulmonary TB. Recoveries were found to be higher among extra pulmonary patients than

pulmonary negative and positive. However, pulmonary positive recorded the highest percentage of deaths.

Recovery among new cases was seen to be lower than recovery among non-new patients. It was further revealed that, recovery was high among the adult population than both the aged and children. The youngest and the oldest patients who died while on treatment within the study period were 27 years and 85 years with mean and median ages to have died being 53.07 and 49 respectively.

The median recovery time for patients registered for treatment was found to be 171 days which was validated using the Jackkniffe. The Kaplan-Meier estimates based on gender revealed an early recovery times for males. Also, average recovery was seen to be least among adults followed by the aged whilst children recorded the highest average recovery time.

The log-logistic regression model was found to be the best fit accelerated failure time model since it recorded the least AIC, AICc, BIC, DIC and a high log-likelihood values. The further analysis using the log-logistic initially uncovered age, smear results and treatment time to be the only prognostic factors that are statistically significant. However, an interaction of the parameters indicated that HIV is significantly related to recovery of TB patient whilst disease category and type of patient showed no sign of significance. But when recovery as the dependent variable was tested using the binary logistic regression model, again age, treatment time, HIV and disease category were seen to be statistically significant.

5.3 Conclusions

The average recovery time for TB patients in Buluk is 171 days (that is 24.43 weeks). In addition, the log-logistic distribution provided a better fit to the data. The posterior summary statistics for mean estimates and standard deviation suggest that, the parameters; age of patient, treatment time, HIV status and disease category are factors that are associated with recovery. Furthermore, a unit increase in the age of a Patient is linked to a decrease in the patient's recovery (a point increase in age will result in 1.2% decrease in the odds of recovery) while a unit increase in treatment time provides a corresponding 3.9% increase in the chances of recovering from the infection.

5.4 Recommendations

The following are recommendations emanating from the study:

1. Health care workers should pay much attention to the aged since ageing is linked to decrease in a patient's recovery.
2. A research to determine the relationship between patient delay and treatment outcome is needed.
3. Finally, further research is needed to assess spatial effect of TB in Buluk.

REFERENCES

Abuaku, B. K., Tan, H., Li, X., Chen, M., and Huang, X. (2010). A comparative analysis of tuberculosis treatment success between Hunan Province of China and Eastern Ghana. *Medical Principles and Practice*, *19*(6), 451-456.

Addo, K. K., Ampofo, W. K., Owusu, R., Bonsu, C., Nartey, N., Mensah, G. I., and Bonsu, F. A. (2018). First Nationwide Survey of the Prevalence of TB/HIV Co-Infection in Ghana. *Journal of Tuberculosis Research*, *6*(2), 135-147.

Adenager, G. S., Alemseged, F., Asefa, H., and Gebremedhin, A. T. (2017). Factors associated with treatment delay among pulmonary tuberculosis patients in public and private health facilities in Addis Ababa, Ethiopia. *Tuberculosis research and treatment*, *2017*.

Ahmed Yassin, M., Takele, L., Gebresenbet, S., Girma, E., Lera, M., Lendebo, E., and Cuevas, L. E. (2004). HIV and tuberculosis coinfection in the southern region of Ethiopia: a prospective epidemiological study. *Scandinavian journal of infectious diseases*, *36*(9), 670-673.

Ajagbe, O. B., Kabair, Z., & O'Connor, T. (2014). Survival analysis of adult tuberculosis disease. *PloS one*, *9*(11), e112838.

Allison, P. D. (2010). *Survival analysis using SAS: a practical guide*. Sas Institute.

Antwi, S., Yang, H., Enimil, A., Sarfo, A. M., Gillani, F. S., Ansong, D., ..and Kwara, A. (2017). Pharmacokinetics of the first-line antituberculosis drugs in Ghanaian children with tuberculosis with or without HIV coinfection. *Antimicrobial agents and chemotherapy*, *61*(2).

Ariyothai, N., Podhipak, A., Akarasewi, P., Tornee, S., Smithtikarn, S., and Thongprathum, P. (2004). Cigarette smoking and its relation to pulmonary tuberculosis in adults.

Barman, R. S. M. P. (2017). Comparing accelerated failure time models with its specific distributions in the analysis of esophagus cancer patients data. *International Journal of Computational and Applied Mathematics*, *12*(2), 411-424.

Bhargava, S., Unissa, A. N., Range, N. S., Tiwari, D., Aadil, S. C., and Khan, A. H. (2018). Diagnosis and Management of Tuberculosis, Vol. 2, pp. 1-94. *Diagnosis and Management of Tuberculosis*, *2*, 4.

Biruk, M., Yimam, B., Abrha, H., Biruk, S., and Amdie, F. Z. (2016). Treatment outcomes of tuberculosis and associated factors in an Ethiopian University Hospital. *Advances in Public Health*, *2016*.

Boah, M., Adampah, T., Jin, B., Wang, W., and Wang, K. (2020). Trend of tuberculosis casenotifications and their determinants in Africa and South-East Asia during 2000–2018: a longitudinal analysis of national data from 58 countries. *Infectious Diseases*, *52*(8), 538-546.

Boah, M., Amporfro, D. A., Adampah, T., Bordotsiah, S., Jin, B., Coffie, E. S., and Akanpabadae, E. (2020). The Risk Factors for Pulmonary Tuberculosis Incidence in Ghana: A Small Matched Case-control Study. *Asian Journal of Research in Infectious Diseases*, 14-23.

Browne, W. J., Steele, F., Golalizadeh, M., and Green, M. J. (2009). The use of simple reparameterizations to improve the efficiency of Markov chain Monte Carlo

estimation for multilevel models with applications to discrete time survival models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(3), 579-598.

Castro, K. G. (1995). Tuberculosis as an opportunistic disease in persons infected with human immunodeficiency virus. *Clinical infectious diseases*, *21*(Supplement_1), S66-S71.

Chechile, R. A. (2003). Mathematical tools for hazard function analysis. *Journal of Mathematical Psychology*, *47*(5-6), 478-494.

Chennaveerappa, P. K., Siddharam, S. M., Halesha, B. R., Vittal, B. G., and Jayashree, N. (2011). Treatment outcome of tuberculosis patients registered at dots centre in a teaching hospital, South India. *Int J Biol Med Res.*, *2*(2), 487-489.

Chin, D. P.,and Hanson, C. L. (2017). Finding the missing tuberculosis patients. *The Journal of infectious diseases*, *216*(suppl_7), S675-S678.

Cleves, M., Gould, W., Gould, W. W., Gutierrez, R., & Marchenko, Y. (2008). *An introduction to survival analysis using Stata*. Stata press.

Cui, Z., Lin, M., Nie, S., & Lan, R. (2017). Risk factors associated with Tuberculosis (TB) among people living with HIV/AIDS: A pair-matched case-control study in Guangxi, China. *PLoS One*, *12*(3), e0173976.Date, A., and Modi, S. (2015). TB screening among people living with HIV/AIDS in resourcelimited settings. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *68*, S270-S273.

Date, A., & Modi, S. (2015). TB screening among people living with HIV/AIDS in resource- limited settings. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *68*, S270-S273.

DiSaia, P. J., Creasman, W. T., Mannel, R. S., McMeekin, D. Sand Mutch, D. G. (2017). *Clinical gynecologic oncology e-book*. Elsevier Health Sciences.

Dräger, S., Gedik, G., & Dal Poz, M. R. (2006). Health workforce issues and the Global Fund to fight AIDS, Tuberculosis and Malaria: an analytical review. *Human Resources for Health*, *4*(1), 1-12.

Dzimah, D. A. (2016). *Modelling the Risk Factors of Neonatal Mortality Using Survival Analysis* (Doctoral dissertation, University of Ghana). Eang, M.T., Satha, P.,Yadav, R. P.,

Dzimah, D. A. (2016). *Modelling the Risk Factors of Neonatal Mortality Using Survival Analysis* (Doctoral dissertation, University of Ghana).

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association*, *83*(402), 414-425.

El-Hajj, H. H., Marras, S. A., Tyagi, S., Kramer, F. R., and Alland, D. (2001). Detection of Rifampin Resistance inMycobacterium tuberculosis in a Single Tube with Molecular Beacons. *Journal of clinical microbiology*, *39*(11), 4131-4137.

Gandhi, N. R., Andrews, J. R., Brust, J. C., Montreuil, R., Weissman, D., Heo, M., and Shah,N. S. (2012). Risk factors for mortality among MDR-and XDR-TB patients in a high HIV prevalence setting. *The International Journal of Tuberculosis and*

Lung Disease, *16*(1), 90-97. Geneva: World Health Organization 2020. Licence: CC BY- NC-SA 3.0 IGO.

Gao, H., Bryc, K., & Bustamante, C. D. (2011). On identifying the optimal number of population clusters via the deviance information criterion. *PloS one*, *6*(6), e21014.

George, B., Seals, S., & Aban, I. (2014). Survival analysis and regression models. *Journal of Nuclear Cardiology*, *21*(4), 686-694.

Getnet, F., Demissie, M., Worku, A., Gobena, T., Seyoum, B., Tschopp, R., and Andersen, C. T. (2019). Determinants of patient delay in diagnosis of pulmonary tuberculosis in Somali Pastoralist Setting of Ethiopia: a matched case-control study. *International journal of environmental research and public health*, *16*(18), 3391.

Golub, J. E., Bur, S., Cronin, W. A., Gange, S., Baruch, N., Comstock, G. W., and Chaisson, R. E. (2006). Delayed tuberculosis diagnosis and tuberculosis transmission. *The international journal of tuberculosis and lung disease*, *10*(1), 24-30.

Hashim, D., and Weiderpass, E. (2018). Cancer survival and survivorship. In *Encyclopedia of Cancer* (pp. 250-259). Elsevier.

Ibrahim, I. (2019). *DIFFERENCES IN SOME DIABETIC INDICATORS AMONG PATIENTS WITH SEXUAL DYSFUNCTION* (Doctoral dissertation).

Ifebunandu, N. A., and Ukwaja, K. N. (2012). Tuberculosis treatment default in a large tertiary care hospital in urban Nigeria: prevalence, trend, timing and predictors. *Journal of infection and public health*, *5*(5), 340-345.

Jain, A., and Mondal, R. (2008). Extensively drug-resistant tuberculosis: current challenges and threats. *FEMS Immunology and Medical Microbiology*, *53*(2), 145-150.

Jakperik D., and Lea, A. Characterization of Prognostic Factors for Recovery in Tuberculosis Patients in Northern Ghana.

Jakperik, D., and Acquaye, B. K. (2013). Assessing the effects of prognostic factors in recovery of tuberculosis patients in the upper west region. *Mathematical and Modeling*, *3*(11).

Jakperik, D., and Ozoje, M. (2012). Survival analysis of average recovery time of tuberculosis patients in Northern region, Ghana. International Journal of Current Research, 4(9), 123-125.

Jiang, Xueyan, Hongzhou Lu, Yuexin Zhang, Zengquan Zhou, Hanhui Ye, Qingxia Zhao, Hui Wang et al. "A cross-sectional study of HIV and tuberculosis coinfection cases in mainland China." *Southern medical journal* 101, no. 9 (2008): 914-917.

Jubulis, J., Kinikar, A., Ithape, M., Khandave, M., Dixit, S., Hotalkar, S., ... and Jain, S. (2014). Modifiable risk factors associated with tuberculosis disease in children in Pune, India. *The International journal of tuberculosis and lung disease*, *18*(2), 198-204.

Jurado, L. F., & Palacios, D. M. (2018). Tuberculosis: A Risk Factor Approach.*Tuberculosis*, 37. Kalhori, S. R. N., Nasehi, M., and Zeng, X. J. (2010). A Logistic Regression Model to Predict High Risk Patients to Fail in Tuberculosis Treatment Course Completion. *International Journal of Applied Mathematics*, *40*(2).

Kalhori, S. R. N., Nasehi, M., and Zeng, X. J. (2010). A Logistic Regression Model to Predict High Risk Patients to Fail in Tuberculosis Treatment Course Completion. *International Journal of Applied Mathematics*, *40*(2).

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457-481.

Kruschke, J. K. (2010). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. Academic Press /Elsevier

Mahanta, J., Biswas, S. C., Roy, M. K., & Islam, M. A. (2015). A comparison of bayesian and classical approach for estimating Markov based logistic model. *American Journal of Mathematics and Statistics*, *5*(4), 178-183.

Merle, C. S., Fielding, K., Sow, O. B., Gninafon, M., Lo, M. B., Mthiyane, T., ...and N'Diaye, (2014). A four-month gatifloxacin-containing regimen for treating tuberculosis. *New England Journal of Medicine*, *371*(17), 1588-1598.

Morgan, M., Kalantri, S., Flores, L., and Pai, M. (2005). A commercial line probe assay for the rapid detection of rifampicin resistance in Mycobacterium tuberculosis: a systematic review and meta-analysis. *BMC infectious diseases*, *5*(1), 1-9.

Morishita, F., Nishikiori, N., van-Maaren, P., and Lambregts-van Weezenbeek, C. (2012). Early detection of tuberculosis through communitybased active case finding in Cambodia. *BMC Public Health*, *12*(1), 469.

Nahid, P., Pai, M., and Hopewell, P. C. (2006). Advances in the diagnosis and treatment of tuberculosis. *Proceedings of the American Thoracic Society*, *3*(1), 103-110.

Naidoo, K., Gengiah, S., Singh, S., Stillo, J., & Padayatchi, N. (2019). Quality of TB care among people living with HIV: gaps and solutions. *Journal of clinical tuberculosis and other mycobacterial diseases*, *17*, 100122.

Naidoo, K., Gengiah, S., Singh, S., Stillo, J., & Padayatchi, N. (2019). Quality of TB care among people living with HIV: gaps and solutions. *Journal of clinical tuberculosis and other mycobacterial diseases*, *17*, 100122.

Naidoo, P., Theron, G., Rangaka, M. X., Chihota, V. N., Vaughan, L., Brey, Z. O., and Pillay, Y. (2017). The South African tuberculosis care cascade: estimated losses and methodological challenges. *The Journal of infectious diseases*, *216*(suppl_7), S702-S713.

Nardi, A., and Schemper, M. (2003). Comparing Cox and parametric models in clinical studies.*Statistics in medicine*, *22*(23), 3597-3610.

Nardi, A., and Schemper, M. (2003). Comparing Cox and parametric models in clinical studies. *Statistics in medicine*, *22*(23), 3597-3610.

Nyanwura, E. M., and Esena, R. K. (2013). Essential medicines availability and affordability: a case study of the top ten registered diseases in Builsa District of Ghana. *Int J Sci Technol Res*, *2*(8), 1-12.

O'Connor, C. M., Gattis, W. A., Uretsky, B. F., Adams Jr, K. F., McNulty, S. E., Grossman, S. H., ... and FIRST Investigators. (1999). Continuous intravenous dobutamine is associated with an increased risk of death in patients with advanced heart failure: insights from the Flolan International Randomized Survival Trial (FIRST). *American heart journal*, *138*(1), 78-86.

Ogyiri, L., Lartey, M., Ojewale, O., Adjei, A. A., Kwara, A., Adanu, R. M., & Torpey, K. (2019). Effect of HIV infection on TB treatment outcomes and time to mortality in two urban hospitals in Ghana-a retrospective cohort study. *The Pan African Medical Journal*, *32*.

Ojo, O. B., Lougue, S., & Woldegerima, W. A. (2017). Bayesian generalized linear mixed modeling of Tuberculosis using informative priors. *PloS one*, *12*(3), e0172580.

Ohene, S. A., Bakker, M. I., Ojo, J., Toonstra, A., Awudi, D., and Klatser, P. (2019). Extra pulmonary tuberculosis: A retrospective study of patients in Accra, Ghana. *PLoS One*, *14*(1), e0209650.

Onisko, A., Druzdzel, M. J., & Austin, R. M. (2016). How to interpret the results of medical time series data analysis: classical statistical approaches versus dynamic Bayesian network modeling. *Journal of pathology informatics*, *7*.

Osei, E., Der, J., Owusu, R., Kofie, P., and Axame, W. K. (2017). The burden of HIV on Tuberculosis patients in the Volta region of Ghana from 2012 to 2015: implication for Tuberculosis control. *BMC infectious diseases*, *17*(1), 1-9. Global tuberculosis report 2020.

Osei, E., Oppong, S., and Der, J. (2020). Trends of tuberculosis case detection, mortality and co infection with HIV in Ghana: A retrospective cohort study. *Plos one*, *15*(6), e0234878.

Padayatchi, N., Daftary, A., Naidu, N., Naidoo, K., and Pai, M. (2019). Tuberculosis: treatment failure, or failure to treat? Lessons from India and South Africa. *BMJ global health*, *4*(1).

Pardeshi, G., and Deshmukh, D. (2007). Disease characteristics and treatment outcome in elderly tuberculosis patients on DOTS. *Indian Journal of Community Medicine*, *32*(4), 292.

Park, S. J., Milano, C. A., Tatooles, A. J., Rogers, J. G., Adamson, R. M., Steidley, D. E., ... and HeartMate II Clinical Investigators†. (2012). Outcomes in advanced heart failure patients with left ventricular assist devices for destination therapy. *Circulation: Heart Failure*, *5*(2), 241-248.

Prabhash, K., Patil, V. M., Noronha, V., Joshi, A., & Bhattacharjee, A. (2016). Bayesian Accelerated Failure Time and its Application in Chemotherapy Drug Treatment Trial. *STATISTICS*, *671*.

Rehm, J., Samokhvalov, A. V., Neuman, M. G., Room, R., Parry, C., Lönnroth, K., ... and Popova, S. (2009). The association between alcohol use, alcohol disorders and tuberculosis (TB). A systematic review. *BMC public health*, *9*(1), 1-12.

Samokhvalov, A. V., Neuman, M. G., Room, R., Parry, C., Lönnroth, K., ... and Popova, S. (2009). The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review. *BMC public health*, *9*(1), 1-12.

Sharma, S. K., Mohan, A., and Kadhiravan, T. (2005). HIV-TB co-infection: epidemiology, diagnosis and management. *Indian Journal of Medical Research*, *121*(4), 550-567. Singogo, E. (2016). *Modelling survival in HIV cohorts with applications to data from Zomba, Malawi* (Doctoral dissertation, Lancaster University).

Singogo, E. (2016). *Modelling survival in HIV cohorts with applications to data from Zomba, Malawi* (Doctoral dissertation, Lancaster University).

Smelser, N. J., and Baltes, P. B. (Eds.). (2001). *International encyclopedia of the social and behavioral sciences* (Vol. 11). Amsterdam: Elsevier.

Stevenson, M., and EpiCentre, I. V. A. B. S. (2009). An introduction to survival analysis. *EpiCentre, IVABS, Massey University*.

Sulistiyani, S., Nugraheni, S. A., Radjasa, O. K., Sabdono, A., and Khoeri, M. M. (2010). Antibacterial activities of bacterial symbionts of soft coral Sinularia sp. against tuberculosis bacteria. *Journal of Coastal Development*, *14*(1), 45-50.

Terefe, A. N.,and Gebrewold, L. A. (2018). Modeling Time to Recovery of Adult Tuberculosis (Tb) Patients in Mizan-Tepi University Teaching Hospital, South-West Ethiopia. *Mycobact Dis*, *8*(258), 2161-1068.

Tetteh, A. K., Agyarko, E., Otchere, J., Bimi, L., and Ayi, I. (2018). An Evaluation of Treatment Outcomes in a Cohort of Clients on the DOTS Strategy, 2012–2016. *Tuberculosis research and treatment*, *2018*.

Tolley, H. D., Barnes, J. M., and Freeman, M. D. (2016). Survival analysis. In *Forensic Epidemiology* (pp. 261-284). Academic Press.

Tolosie, K., and Sharma, M. K. (2014). Application of Cox proportional hazards model in case of tuberculosis patients in selected Addis Ababa health centers, Ethiopia. *Tuberculosis research and treatment*, *2014*.

Trinh, Q. M., Nguyen, H. L., Nguyen, V. N., Nguyen, T. V. A., Sintchenko, V., & Marais, B. J. (2015). Tuberculosis and HIV co-infection—focus on the Asia-Pacific region. *International Journal of Infectious Diseases*, *32*, 170-178.

Wand, H., Whitaker, C., and Ramjee, G. (2011). Geoadditive models to assess spatial variation of HIV infections among women in Local communities of Durban, South Africa. *International Journal of Health Geographics*, *10*(1), 28.

World Health Organization. (2012). WHO policy on collaborative TB/HIV activities: guidelines for national programmes and other stakeholders.

World Health Organization. (2013). *Global tuberculosis report 2013*. World Health Organization.

World Health organisation, Geneva(2018). Global Tuberculosis Report 2018.

Xie, Y., Han, J., Yu, W., Wu, J., Li, X., and Chen, H. (2020). Survival Analysis of Risk Factors for Mortality in a Cohort of Patients with Tuberculosis. *Canadian Respiratory Journal*, *2020*.

APPENDIX I - TABLES

Table 4.3: Descriptive Statistics

Parameters	Total	Percent	Censored	Percent	Died	Percent	Recovered	Percent
Sex								
Male	237	76.95	152	64.14	21	8.86	188	79.32
Female	71	23.05	48	67.61	6	8.45	62	87.32
Age Category								
0-17	2	0.65	2	100	0	0.00	2	100
18-60	235	76.30	147	62.55	18	7.82	191	81.166
>60	71	23.05	51	71.83	9	12.68	57	80.28
Type of Patient								
New	265	86.04	174	65.66	24	9.06	214	80.75
Non-new	43	13.19	26	60. 47	3	6.98	36	83.72
Smear Results								
Positive	196	63.4	127	64.80	17	8.67	155	79.08
Negative	112	36.3	73	65.18	10	8.93	95	84.82
HIV Test								
Positive	32	10.39	22	68.75	6	18.75	23	71.88
Negative	269	87.39	1733	64.32	20	7.43	221	82.16
Not Tested	7	2.27	5	71.42	1	14.29	6	85.71
Disease Category								
Pulmonary positive	217	70.45	139	64.06	20	9.22	171	78.80
Pulmonary	89	28.90	59	66.29	7	7.87	77	86.52
Negative								
Extra Pulmonary	2	0.65	2	100	0	0.00	2	100

Table 4.7: Quartile Estimates for Gender and Age

Percent	Point Estimate	Transform	95% Confidence Interval [Lower	Upper)
Males				
75	190.000	LOGLOG	183.000	199.000
50	**170.000**	**LOGLOG**	**168.000**	**174.000**
25	166.000	LOGLOG	164.000	168.000
Females				
75	192.000	LOGLOG	180.000	210.000
50	**174.000**	**LOGLOG**	**168.000**	**180.000**
25	167.000	LOGLOG	143.000	168.000
0-17 Years				
75	216.000	LOGLOG	169.000	216.000
50	**192.500**	**LOGLOG**	**169.000**	**216.000**
25	169.000	LOGLOG	169.000	216.000
18-60 Years				
75	186.000	LOGLOG	180.000	198.000
50	**170.000**	**LOGLOG**	**168.000**	**173.000**
25	166.000	LOGLOG	163.000	167.000
>60 Years				
75	193.000	LOGLOG	186.000	217.000
50	**180.000**	**LOGLOG**	**169.000**	**186.000**
25	168.000	LOGLOG	136.000	169.000

Table 4.9: **Posterior summary statistics for Log-Logistic Regression (with interactions)**

Parameter	DF	Mean Estimate	Standard Error	95% Confidence Limits	
Intercept	**1**	**13.7734**	**2.0371**	**9.7808**	**17.7660**
AGE	**1**	**-0.1536**	**0.0297**	**-0.2119**	**-0.0954**
GENDER 0	1	-0.1727	0.1914	-0.5478	0.2023
GENDER 1	0	0.0000			
DC	1	-0.0385	0.2009	-0.4322	0.3553
TP 0	1	0.1607	0.1476	-0.1287	0.4500
TP 1	0	0.0000			
SR 0	1	0.2316	0.2730	-0.3035	0.7666
SR 1	0	0.0000			
HIV 0	**1**	**-8.6377**	**2.0541**	**-12.6637**	**-4.6117**
HIV1	1	-9.5312	2.1324	-13.7105	-5.3518
HIV2	**0**	**0.0000**			
AGE*HIV0	**1**	**0.1584**	**0.0298**	**0.0999**	**0.2168**
AGE*HIV1	**1**	**0.1811**	**0.0337**	**0.1150**	**0.2473**
AGE*HIV2	0	0.0000			
SR*TP0 0	1	-0.1633	0.2823	-0.7167	0.3900
SR*TP0 1	0	0.0000			
SR*TP1 0	0	0.0000			
SR*TP1 1	0	0.0000			
DC*GENDER 0	1	0.1447	0.2222	-0.2908	0.5801
DC*GENDER 1	0	0.0000			
Scale	**1**	**0.3576**	**0.0311**	**0.3016**	**0.4240**

Table 4.10: Odds Ratio Estimates for Logistic Regression

Effect	Point Estimate	95% Wald Confidence	Limits
AGE_	0.988	0.954	1.023
DC	0.190	0.020	1.770
TP	2.918	0.437	19.477
SR	3.632	0.469	28.121
HIV	1.136	0.239	5.393
GENDER	1.009	0.213	4.768
TREATMENT TIME	1.039	1.027	1.050

Table 4.11 Analysis of Maximum Likelihood Parameter Estimates for Logistic Regression

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	
Intercept	1	48.5965	1.5973	45.4658	51.7272
AGE_	**1**	**-0.0419**	**0.0186**	**-0.0583**	**0.0144**
GENDER 0	1	-0.4006	0.8806	-2.1265	1.3252
GENDER 1	0	0.0000	0.0000	0.0000	0.0000
DC 0	**1**	**-19.4654**	**1.3919**	**-22.1934**	**-16.7374**
DC1	0	-21.6084	0.0000	-21.6084	-21.6084
DC 2	0	0.0000	0.0000	0.0000	0.0000

TP	0	1	-1.1775	0.9690	-3.0768	0.7218
TP	1	0	0.0000	0.0000	0.0000	0.0000
SR	0	1	-1.8276	1.1990	-4.1776	0.5224
SR	1	0	0.0000	0.0000	0.0000	0.0000
HIV	**0**	**1**	**-26.6675**	**0.9341**	**-28.4982**	**-24.8367**
HIV	1	0	-27.8888	0.0000	-27.8888	-27.8888
HIV	2	0	0.0000	0.0000	0.0000	0.0000
TRT_TIME		**1**	**0.0399**	**0.0062**	**0.0278**	**0.0520**
Scale		0	1.0000	0.0000	1.0000	1.0000

NOTE: The scale parameter was held fixed.

Table 4.13 Partition for the Hosmer and Lemeshow Test

		TO = 0		TO = 1	
Group	Total	Observed	Expected	Observed	Expected
1	31	0	0.05	31	30.95
2	31	0	0.18	31	30.82
3	31	0	0.32	31	30.68
4	31	1	0.51	30	30.49
5	31	0	0.69	31	30.31
6	31	0	0.93	31	30.07
7	31	2	1.36	29	29.64
8	31	6	4.73	25	26.27
9	31	23	21.92	8	9.08
10	29	26	27.31	3	1.69

APPENDIX II - PROGRAMME CODES

Codes for odd ratio

```
proc logistic data=aa;
model to= age_ dc tp sr hiv gender trt_time;
run;
quit;
```

SAS codes for Binary logistic regress model

```
ods graphics on;
  proc genmod data=aa;
  class dc hiv tp sr gender;
  model TO = AGE_ GENDER TIME DC TP SR HIV/ dist=bin link=logit;
  bayes seed=1 OutPost=PostSurg;
  run;
  ods graphics off;
```

SAS codes for Hosmer-Lemeshow Goodness of Fit Test

```
ods select lackfitpartition lackfitchisq;
proc logistic data=aa;
model TO = age_ gender dc tp sr hiv trt_time /  lackfit;
run;
```

SAS codes for Receiver Operating Characteristic (ROC) Curve

```
ods graphics on;
proc logistic data=ab plots(only)=roc;
logisticModel: model TO(event='1') = age_  gender dc tp sr hiv trt_time;
```

output out=LogiOut predicted=LogiPred;

run;

ods graphics off;

R program codes for median recovery time plots

KM_fit <- survfit(Surv(TIME, TO) ~ 1, data = aa)

KM_fit

plot(survfit(Surv(TIME, TO) ~ 1, data = aa),

xlab = "Treatment Time in Days", main = 'Kaplan Meyer Plot',

ylab = "Overall survival probability")

SAS programme codes for median treatment time

ods graphics on;

proc lifetest data=aa;

*time time*to(0);*

run;

ods graphics off;

SAS programme codes for Median Treatment time based on Gender

ods graphics on;
proc lifetest data=aa;

*time time * to(0);*

strata gender;

run;

ods graphics off;

SAS programme codes for Median Treatment time based on Gender

ods graphics on;

proc lifetest data=aa;

*time time * to(0);*

strata age_cat;

run;

ods graphics off;

SAS codes for log-logistic regression

ods graphics on;
proc lifereg data=aa;

class gender hiv sr tp;

*model time*to(0)=age_ gender dc tp sr hiv hiv*age_ tp*sr gender*dc /*

dist=llogistic;

bayes;

run;

ods graphics off;

APPENDIX III - GRAPHS

Kaplan – Meier estimates for gender and Age

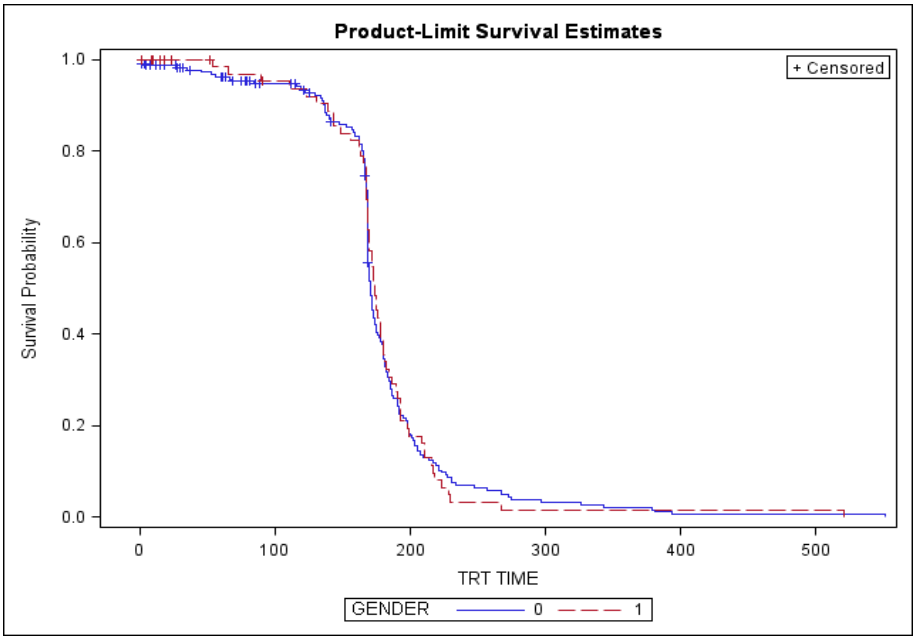


Figure 4.3 Survival plots for Gender

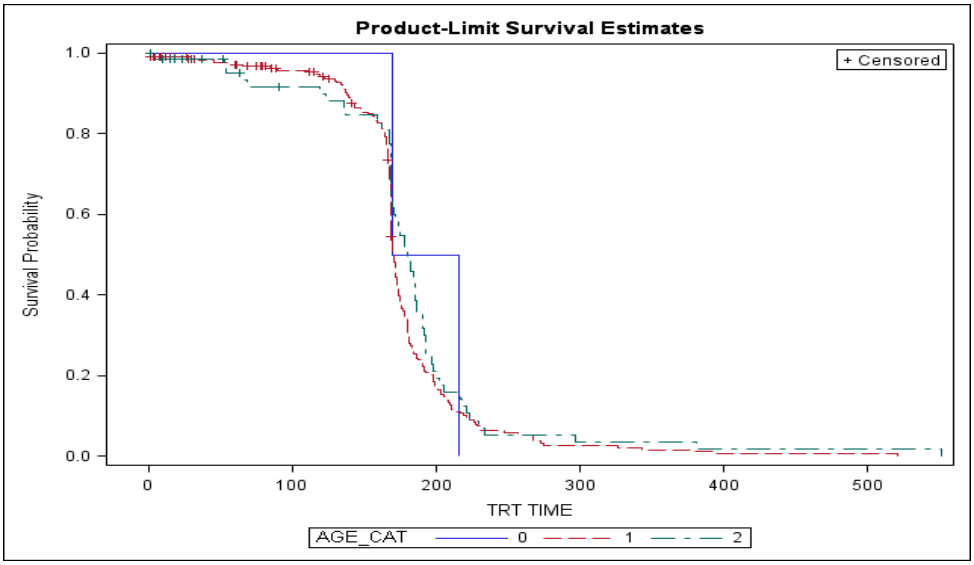


Figure 4.4 Survival plots for age

Receiver Operating Characteristic (ROC) Curve

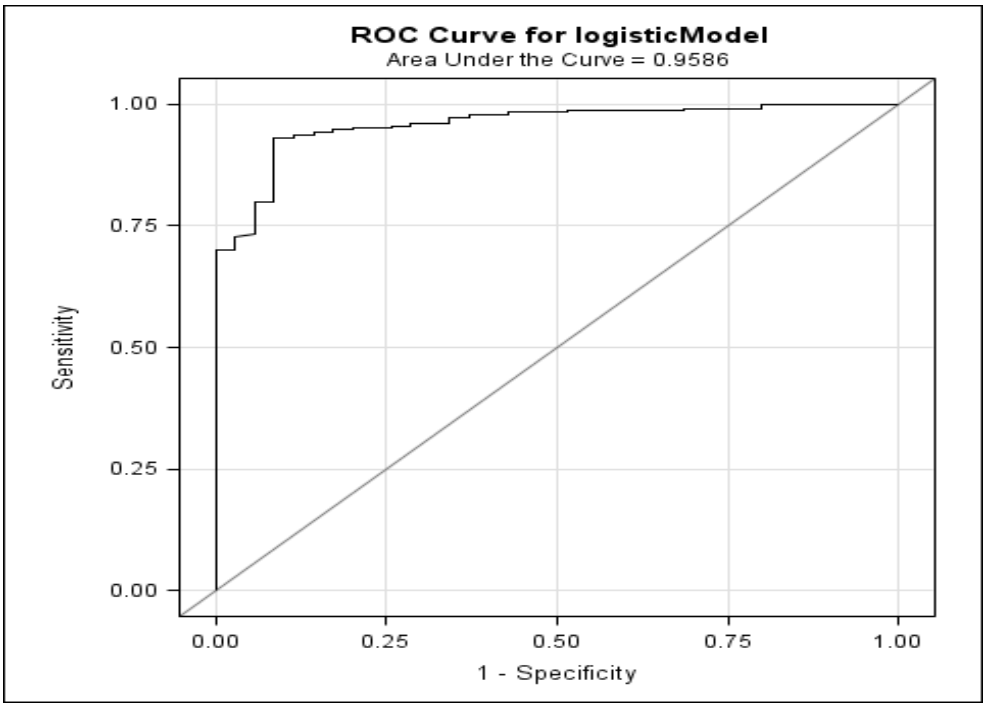
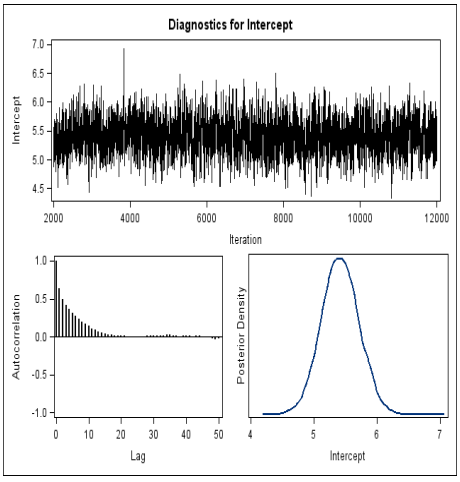
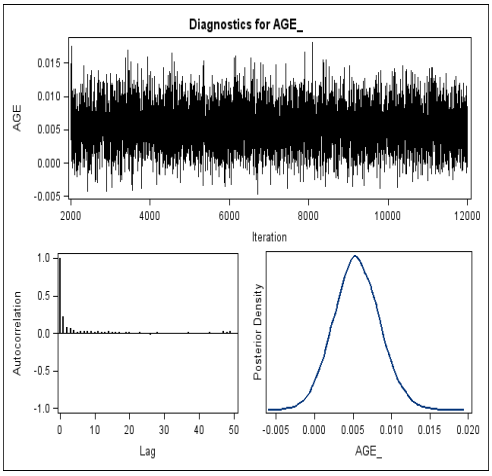
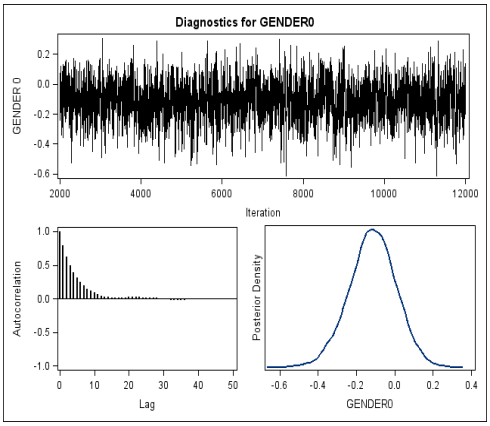


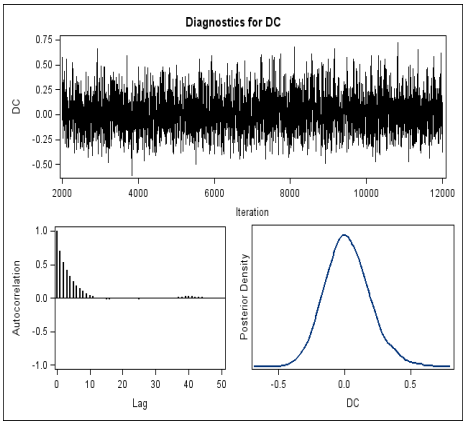
Figure 4.5: ROC Curve

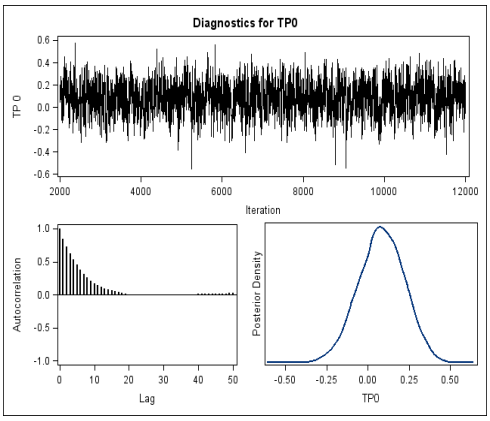
Trace plots for Log-logistic Regression analysis

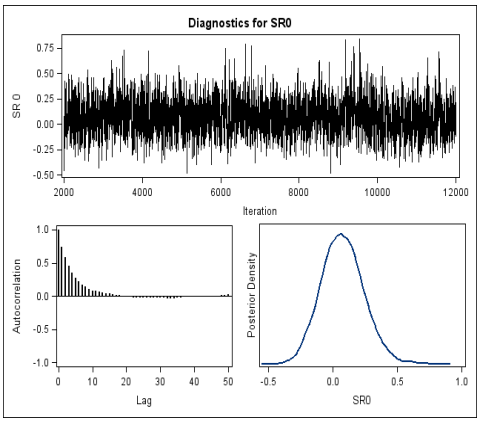


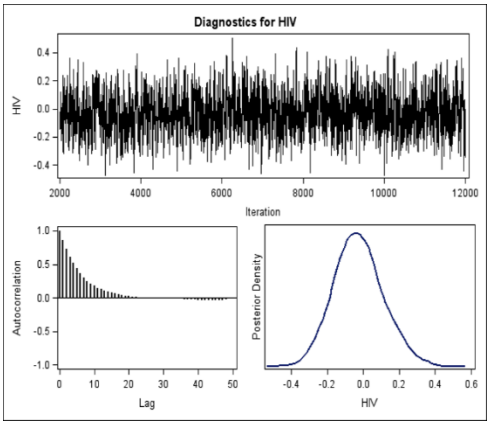


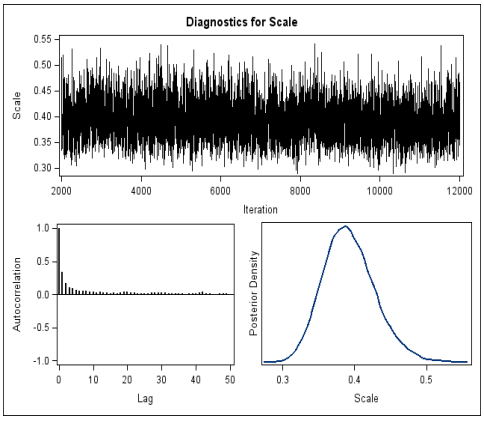












LOGISTIC REGRESSION TRACE PLOTS