UNIVERSITY FOR DEVELOPMENT STUDIES

SURVIVAL ANALYSIS OF AVERAGE RECOVERY TIME OF TUBERCULOSIS

PATIENTS IN NORTHERN REGION, GHANA

JAKPERIK DIOGGBAN

Thesis submitted to the Department of Statistics, Faculty of Mathematical Science,

University for Development Studies in partial fulfillment of the requirements for the award

of Master of Science Degree in Applied Statistics

2011

UNIVERSITY FOR DEVELOPMENT STUDIES

UNIVERSITY FOR DEVELOPMENT STUDIES

SURVIVAL ANALYSIS OF AVERAGE RECOVERY TIME OF TUBERCULOSIS

PATIENTS IN NORTHERN REGION, GHANA

BY

JAKPERIK DIOGGBAN (BSc. MATHEMATICAL SCIENCE (STATISTICS OPTION))
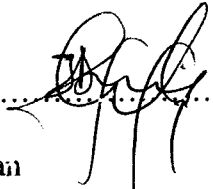
(UDS/MAS/0008/09)

Thesis submitted to the Department of Statistics, Faculty of Mathematical Science, University

for Development Studies in partial fulfillment of the requirements for the award of Master of

Science Degree in Applied Statistics

November, 2011

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...12 – 11 – 2011

Name: Jakperik Dioggban

Supervisor

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

Supervisor's Signature   Date : 14th November, 2011

Name: Prof. M. O. Ozoje

ABSTRACT

In this study, the average recovery time of Tuberculosis patients and the associated risk of treatment failure/death were examined based on a retrospective moving cohort of sixty-one patients. Mainly, four models: Cox regression, Kaplan-Meier estimator, Log-Pearson III, and the generalized gamma distributions were employed in the analysis to explore all useful information that may be of help to policy makers and stakeholders in their quest to improve service delivery to patients. The generalized gamma distribution produced results that were closer to that of the semi-parametric model, Cox regression model. It was observed that the survival rates for males and females were 85.71% and 88.46% respectively. Evidently, pulmonary Tuberculosis was most prevalent. It was discovered that, not only the age of patient at diagnosis matter, but also the category, and type of the patient. The study also reported a much shorter median recovery time of 22 weeks in the Region compared to those reported in other parts of the world. The risk of relapse and death were found to be related to age. High percentage of success rate of recovery was recorded in patients with new diagnosis. Although, it is generally reported that the levels of drug resistance in Africa are lower than in other parts of the world, measures to provide controlled application of second-line drugs, supervision of drug distribution and compliance, enforcement of Directly Observed Therapy – Short course protocols, and sustained training of all personnel involved in TB management should be enforced for effective combat of the tuberculosis disease.

ACKNOWLEDGEMENT

DEDICATION

This work is dedicated to the loving memory of my late father, Mr. Jatuat Liyialib.

TABLE OF CONTENTS

UNIVERSITY FOR DEVELOPMENT STUDIES

ix

LIST OF TABLES

UNIVERSITY FOR DEVELOPMENT STUDIES

x

LIST OF FIGURES

UNIVERSITY FOR DEVELOPMENT STUDIES

ORGANIZATION OF THE THESIS

The thesis consists of five chapters. Chapter one contains the Introduction of the thesis; Chapter two presents Literature Review whiles Chapter Three gives the Methodology. Analysis is presented in Chapter Four with Conclusions and Recommendations in Chapter Five.

UNIVERSITY FOR DEVELOPMENT STUDIES

CHAPTER ONE

1.0   Introduction

Tuberculosis (TB) is a common and often deadly infectious disease caused majorly by *Mycobacterium tuberculosis* in humans (Kumar et al., 2007). It was first isolated in 1882 by Robert Koch, a German physician who received a Nobel Prize for his discovery. Tuberculosis attacks the lungs but can also affect other parts of the body. It is an air borne disease which spread through coughing, sneezing, and spitting by infected persons (Konstatinos, 2010). Most infections in humans result in an asymptomatic, latent infection. Nevertheless, about one in every ten latent infection cases eventually progresses to active disease, which, if left untreated, kills more than 50% of its victims (http://en.wikipedia.org/wiki/Tuberculosis, 2009). Many years ago, TB was referred to as "consumption" disease because patients often waste away. But today, tuberculosis can be treated successfully with antibiotics.

Although, TB is an age long disease, health workers all over the world are finding hard to control. The World Health Organization in 1993 declared TB a global emergency in recognition of the growing importance of TB as a public health problem. In the year 2000 alone, about 8.7million new cases of TB with 1.9million deaths were reported worldwide. *Mycobacterium tuberculosis*, TB causing bacteria kills more people than any other known single infectious disease agent. Deaths from tuberculosis accounts for 25% of all avoidable deaths in developing countries (http://en.wikipedia.org/wiki/Tuberculosis, 2009).

Tuberculosis is an immense medical, social, and economic wasteful disease of great magnitude which has however received substantial attention from the general public health system and the scientific communities (Sharma and Mohan, 2006).

1

According to WHO (1996), *Mycobacterium tuberculosis* and Human Immunodeficiency Virus are the main agents of deaths resulting from infectious diseases in the world. In 2007, they estimated that 32% of the world population are carriers of *mycobacterium tuberculosis*, the causative agent of TB. According to WHO report in 2008, an estimated 9.2million new cases of TB and 1.7million deaths from TB was recorded in 2006 alone.

The Control of TB remains one of the most serious challenges to global health till date. A potentially devastating threat to TB control is the emergence of drug resistant strains. These drug resistant strains of M. tuberculosis generally arise through the selection of mutated strains by inadequate chemotherapy. A strain resistant to at least two major anti-tubercular drugs Isoniazid and Rifampicin has been reported and its termed multidrug resistant tuberculosis, MDR-TB (Zager and McNerney, 2008). Since the introduction of the first effective anti-TB drug – streptomycin (SM), in the late 1940s, resistance strains of *mycobacterium tuberculosis* to chemotherapeutic agents has been identified as a major problem in the management of TB disease. Clinical relapse after three to six months of improvement has been reported (Fox et al., 1999). Randomized control trials carried out by British researchers upon the introduction of Aminosalicylic Acid(PAS) in 1948 revealed that patients receiving combined therapy (PAS and SM) had lower rates of relapse than those receiving either the PAS or SM alone (Wolinksy et al., 1948). As the tuberculosis chemotherapy era evolved, increasing cases of drug resistance strains continue to appear mainly as a result of inadequate regimens and non-adherence to therapy (Jassal and Bishai, 2009).

Tuberculosis drug resistance can be either primary (transmission of resistant organisms) or secondary (resistance acquired in the host related to inadequate treatment). Four broad categories of mechanisms of acquiring resistance to drugs by *M*. tuberculosis have been identified. These

2

are; 1) the creation of a lipid-rich cell wall that can reduce the permeability of drugs (and arrest phagasomematuration); 2) the production of enzymes that degrade or modify compounds, rendering them useless; 3) the efflux of drugs through protein pumps, described for Isoniazid and Ethambutol; and 4) spontaneous chromosomal mutations that affect key drug targets (Rich, 2006). Among these, the fourth mechanism is considered to be the most important. Mobile or horizontal transmission of resistance, such as plasmid mediated resistance, does not occur in *M. tuberculosis*. Random genetic mutations occur with low but predictable frequencies in the range of one mutation per 106 to 109 organisms. The frequency of mutations conferring resistance to particular agents varies from the range of 103 for many second line drugs (Thiacetazone, Ethionamide, Capreomycin, Cycloserine, and Viomycin) to an intermediate level (around 106) for some first and second line drugs (Isoniazid, Streptomycin, Ethambutol, Kanamycin, and Pamino salicylic acid) to the lowest levels for Rifampicin, on the order of 108 to 1010. When large populations of *M. tuberculosis* are formed in a host and selective pressure is placed by a chemotherapeutic agent, the small population of *M. tuberculosis* that has acquired resistance will continue to multiply, while the susceptible *M. tuberculosis* is suppressed. This enables the drug resistant organism to become the dominant organism in hosts (Warner and Mizrahi, 2006).

Shortage of drugs is one of the most common reasons for the inadequacy of the initial anti-TB regimen, especially in resource poor settings. Other major issues significantly contributing to the higher complexity of the treatment of MDR-TB is the increasing cost of treatment.

The fight against TB in the last two decades has been further challenged with the emergence and pandemic spread of the Human Immunodeficiency Virus (HIV). Recent surveillance data revealed that the prevalence of drug resistant tuberculosis have risen to its highest rate ever in the history of the disease (Marahatha, 2010).

1.1   Justification of the Study

TB is one of the leading infectious diseases in the globe which is almost becoming a pandemic especially in the developing countries of the world. Its devastating effects has attracted global interest and led many scientists to devote research time to checking its incidence and reduce its menacing effects in the socio-economic life of the people. Currently, there are more cases of TB on the planet than have ever existed in the entire history of mankind and most agree that, there is an urgent need for a newer, more effective vaccine that would prevent all forms of TB infectious agents—including drug resistant strains—in all age groups and among people with HIV (Sadoff, and Jerry, 2006).

The story in Ghana is not different as many concerted efforts have been made towards remedying its effects since the implementation of the Directly Observed Therapy – Short course (DOTS) programme in 1994. It is estimated that Ghana has 123 smear.positive pulmonary TB cases per 100,000 population and 281 of all types TB cases per 100,000 population per year. This means that, Ghana with a population of about 20 million should expect about 25,000 smear positive pulmonary TB cases and 55,000 new TB cases of all types every year with about 12,000 deaths.

The 2003 report of the Ghanaian National Tuberculosis Programme (NTP) stated a TB incidence of 281/100,000 people. On the 9th July, 2010, the Upper East Regional TB Coordinator reported that, the incidence of TB in 2009 was 75%, whiles it increases to 85% in 2010, indicating that it was on the ascendancy. The incidence of TB cases for the years, 2007, 2008 and 2009 in Northern Region were **545, 677, and 556** respectively. With this observation, much attention need be paid to the incidence of TB, which therefore informs the need for this research study.

1.2 Objectives of the Study

1.2.1 Main Objective

The main objective of this research is to model the average recovery time of tuberculosis patients in Northern Region, Ghana using a survival function.

1.2.2 The Specific Objectives

- To estimate the average time to recovery of TB patients

- To determine the hazard ratio of TB patients

- To assess the survival rate of TB patients

1.3 Research Questions

The open questions which called for investigation in this research are:

- Is the hazard ratio of TB patients in Northern Region alarming?

- Is there any difference in the expected recovery time between the patients in Ghana and other countries?

- Does the prevalence rate relate to age or sex?

CHAPTER TWO

2.0 Literature Review

2.1 History of Tuberculosis

The earliest unambiguous detection of *Mycobacterium tuberculosis* is in the remains of bison dated back to about 18,000 years ago (Rothschild, et al., 2001). Whether tuberculosis originated in cattle and then transferred to humans, or diverged from a common ancestor infecting a different species, is however unclear (Pearce-Duvet, 2006). Nevertheless, it is clear that *M. tuberculosis* is not directly descended from *M. bovis*, which seems to have relatively evolved recently (Ernest et al., 2007).

Skeletal remains from a Neolithic settlement in Eastern Mediterranean show prehistoric humans (7000 BC) with TB (Hershkovitz et al., 2008). In addition tubercular decay had been found in the spines of mummies of 3000–2400 BC (Zink et al., 2003). Phthisis is a Greek term for tuberculosis. Around 460 BC, Hippocrates identified phthisis as the most widespread disease of the times involving coughing of blood and fever, which was almost always fatal. In South America, the earliest evidence of tuberculosis is associated with the Paracas-Caverna culture (circa 750 BC to circa 100 AD) (Konomi et al., 2002).

A third of the world's population is thought to be infected with *M. tuberculosis* (Jasmer et al., 2002), and new infections occur at a rate of about one per second (WHO, 2007). However, the proportion of people who become sick with tuberculosis each year is stable / falling worldwide, but because of population growth, the absolute number of new cases is still increasing (WHO, 2007). In 2007 alone, there were an estimated 13.7 million chronic active cases, 9.3 million new

6

cases, and 1.8 million deaths, mostly in developing countries (WHO, 2009). More people in the developed world are contracting tuberculosis because their immune systems are compromised by immunosuppressive drugs, substance abuse, or HIV/AIDS. Nevertheless, the distribution of tuberculosis is not uniform across the globe; about 80% of the population in many Asian and African countries test positive to tuberculin tests, while only 5-10% of the US population test positive (Kumar et al., 2007).

Tuberculosis was once described as "**consumption disease**", because it appeared to consume from within, with bloody cough, fever, pallor, and long relentless wasting. At other times, it was called *Phthisis* (Greek for consumption) or *Phthisis pulmonalis*; scrofula (in adults), affecting the lymphatic system and resulting in swollen neck glands; *tabes mesenterica*, - TB of the abdomen and *lupus vulgaris*, TB of the skin; wasting disease; white plague -because sufferers appear markedly pale; king's evil- because it was believed that a king's touch would heal scrofula; and Pott's disease, or gibbus of the spine and joints (Encyclopedia Britannica).

Miliary tuberculosis—now commonly known as disseminated TB—occurs when the infection invades the circulatory system, resulting in millet-like seeding of TB bacilli in the lungs (Encyclopedia Britannica). TB is also called Koch's disease, after the scientist Robert Koch (Bhansali, 1977).

Before the industrial revolution, tuberculosis was sometimes regarded as vampirism. When one member of a family died from it, the other members that were infected would lose their health slowly. People believed that this was caused by the original victim draining the life from the other family members. Furthermore, people who had TB exhibited symptoms similar to what people considered to be vampire traits. People with TB often have symptoms such as red,

swollen eyes (which also creates a sensitivity to bright light), pale skin, extremely low body heat, a weak heart bloody cough, suggesting the idea that the only way for the afflicted to replenish this loss of blood was by sucking blood (Sledzik and Bellantoni,1994).

TB was romanticized in the nineteenth century. Many people believed TB produced feelings of euphoria referred to as *Spes phthisica* ("hope of the consumptive"). It was believed that TB sufferers who were artists had bursts of creativity as the disease progressed. It was also believed that TB sufferers acquired a final burst of energy just before they died that made women more beautiful and men more creative (Lawlor and Clark, 2003). In the early 20th century, some believed TB to be caused by masturbation (Laumann et al., 1994).

2.1.1. Mycobacterium Tuberculosis - Causative Agent of Tubarculosis

The bacillus bacteria causing tuberculosis, *Mycobacterium tuberculosis*, was identified and first described in 1882 by Robert Koch for which he received the Nobel Prize for physiology and medicine in 1905. Koch did not believe that bovine (cattle) and human tuberculosis were similar, which delayed the recognition of infected milk as a source of infection. Later, this source was eliminated by the pasteurization process. Koch announced a glycerine extract of the tubercle bacilli as a remedy for tuberculosis in 1890, calling it "tuberculin". It was not effective, but was later adapted as a test for pre-symptomatic tuberculosis (Waddington, 2004).

The first genuine success in immunization against tuberculosis was developed from attenuated bovine-strain tuberculosis by Albert Calmette and Camille Guérin in 1906. It was called "BCG" (Bacillus of Calmette and Guérin). The BCG vaccine was first used on humans in 1921 in France

(Bonah, 2005), but it was not until after World War II that BCG received widespread acceptance in the USA, Great Britain, and Germany (Comstock., 1994).

In 1815, one in four deaths in England was reported to be caused by the consumption disease. By 1918, one in six deaths in France was caused by TB. In the 20th century, tuberculosis killed an estimated 100 million people (Torrey and Yolken, 2005). After the establishment in the 1880s that the disease was contagious, TB was made a Notifiable disease in Britain which elicited campaigns to stop people from spitting in public places, while the infected poor were pressured to enter sanatoria that resembled prisons. The sanatoria for the middle and upper classes offered excellent care and constant medical attention. Whatever the purported benefits of the fresh air and labor in the sanatoria, even under the best conditions, 50% of those who entered were dead within five years (McCarthy, 2001).

It was not until 1946, with the development of the antibiotic streptomycin that effective treatment and cure became possible. Prior to the introduction of this drug, the only treatment besides sanatoria were surgical interventions, including the pneumothorax or plombage technique — collapsing an infected lung to "rest" it and allow lesions to heal — a technique that was of little benefit and was mostly discontinued by the 1950s (Wolfart, 1990). The emergence of multidrug-resistant TB again introduced surgery as part of the treatment for these infections. Here, surgical removal of chest cavities reduces the number of bacteria in the lungs, as well as increasing the exposure of the remaining bacteria to drugs in the bloodstream. It was therefore thought to increase the effectiveness of the chemotherapy (Lalloo and Ambaram, 2005).

Hopes that the disease could be completely eliminated have been dashed since the increase of drug- resistant strains in the 1980s. Due to the elimination of public health facilities and the

emergence of HIV, there was a resurgence of TB in the late 1980s (Paolo and Nosanchuk, 2004). The number of patients failing to complete their course of drugs increased and New York had to cope with more than 20,000 TB patients with multidrug-resistant strains. The resurgence of tuberculosis resulted in the declaration of a global health emergency by the World Health Organization (WHO) in 1993.

2.2   Types and Causes of Tuberculosis

2.2.1   Causal Agents of Tuberculosis

Tuberculosis is an infectious disease caused by several species of the mycobacteria. These organisms are also known as tubercle bacilli.

The primary cause of TB, *Mycobacterium tuberculosis* (MTB), is a small aerobic non-motile bacillus. High lipid content of this pathogen accounts for many of its unique clinical characteristics (Southwick and Frederick, 2007). It divides every 16 to 20 hours, an extremely slow rate compared with most other bacteria, which usually divide in less than an hour (Cox, 2004). MTB has a cell wall but lacks aphospholipid outer membrane, it is therefore classified as a Gram-positive bacterium. However, if a Gram stain is performed, MTB stains very weakly either as Gram-positive or retain no dye because of the high lipid & mycolic acid content of its cell wall (Madison, 2001). Since MTB retains certain stains after treatment with acidic solution, it is classified as an acid-fast bacillus (AFB) (Kumar et al., 2007, Parish and Stoker, 1999). The most common acid-fast staining technique, the Ziehl-Neelsen stain, dyes AFBs a bright red that stands out clearly against a blue background.

10

MTB can withstand weak disinfectants and survive in a dry state for weeks. In nature, the bacterium can grow only within the cells of a host organism, but *M. tuberculosis* can be cultured *in vitro* (Parish and Stoker, 1999). Using histological stains on expectorate samples from phlegm (also called sputum), scientists can identify MTB under a regular microscope.

The *M. tuberculosis* complex includes four other TB-causing mycobacteria: *M. bovis*, *M. africanum*, *M. canetti* and *M. microti* (Van et al., 1997). In West Africa, the commonest types are *Mycobacterium tuberculosis* and *Mycobacterium africanum. M. africanum* is not widespread, but in parts of Africa, it is a significant cause of tuberculosis (Niemann. et al, 2002. Niobe-Eyangoh et al., 2003). *M. bovis* was once a common cause of tuberculosis, but the introduction of pasteurized milk has largely eliminated this as a public health problem in developed countries (Kumar et al., 2007, Thoen et al., 2006). *M. canetti* is rare and seems to be limited to Africa, although a few cases have been seen in African emigrants (Pfyffer. et al., 1998). *M. microti* is mostly seen in immune-deficient people, although it is possible that the prevalence of this pathogen has been underestimated (Niemann et al., 2000).

2.2.2 Types of Tuberculosis

The site of the disease in the body defines the type of tuberculosis illness. Eighty percent of tuberculosis cases occurs in the actual lung tissue and is called pulmonary TB (PTB). The remaining 20% of tuberculosis cases can occur anywhere in the body, for example, in the lymph nodes, the abdomen and in other tissues. This is known as extra pulmonary TB (EPTB).

There are two types of Pulmonary Tuberculosis:

1. Sputum Smear-Positive pulmonary tuberculosis (Sm+ PTB). The sputum of patients suffering from Sm +PTB contain mycobacterium when viewed under microscopy.

2. Sputum Smear-negative Pulmonary Tuberculosis (Sm- PTB). In this case patient sputum smears test negative to mycobacterium on microscopy, but X-rays evidence consistently show active tuberculosis, which does not clear with ordinary antibiotics. In some cases even though sputum smears are all negative for mycobacterium under microscopy, the culture of the sputum test positive for mycobacterium tuberculosis.

Extra-Pulmonary tuberculosis disease occurs anywhere other than the actual lung tissue. It includes TB inside the chest but outside the lung. This means that TB of the lymph nodes of the chest and TB of the pleura are classified as extra-pulmonary TB. Different types of extra-pulmonary TB includes Pleural, Glandular, Intestinal, Miliary, Meningitis, Bone, Urogenital, Skin, and Eye TB. Extra-pulmonary TB is relatively more common in HIV positive patients than in HIV negative patients. Diagnosis of extra-pulmonary TB is difficult and an experienced medical officer must confirm this as the correct diagnosis.

Extra-pulmonary TB is not infectious, but many patients with extra-pulmonary TB may also have pulmonary TB.

2.2.3 Signs and Symptoms of Tuberculosis

Symptoms of variants and stages of tuberculosis (George, 2009) overlap with each other. Multiple variants may be present simultaneously. When the disease becomes active, 75% of the cases are pulmonary TB, that is, TB in the lungs show symptoms that include chest pain, coughing up blood, and a productive, prolonged cough for more than three weeks. Systemic symptoms include fever, chills, night sweats, appetite loss, weight loss, pallor, and often a tendency to fatigue with ease (WHO, 2007).

This occurs more commonly in immune suppressed persons and young children. Extra-pulmonary infection sites include the pleura in tuberculosis pleurisy, the central nervous system in meningitis, the lymphatic system in scrofula of the neck, the genitourinary system in urogenital_ tuberculosis, and bones and joints in Pott's disease of the spine. An especially serious form of TB is disseminated TB, more commonly known as miliary tuberculosis. Extra-pulmonary TB may co-exist with pulmonary TB as well (CDC, 2003).

2.2.4 Diagnosis

Tuberculosis is diagnosed definitively by identifying the causative organism (*Mycobacterium tuberculosis*) in a clinical sample (for example, sputum orpus). When this is not possible, a probable - although sometimes inconclusive (Konstantinos, 2010) - diagnosis may be made using imaging (X-rays or scans) and/or a tuberculin skin test (Mantoux test).

The main problem with tuberculosis diagnosis is the difficulty in culturing this slow-growing organism in the laboratory (it may take 4 to 12 weeks for blood or sputum culture). A complete medical evaluation for TB must include a medical history, a physical examination, a chest X-ray, microbiological smears, and cultures. It may also include a tuberculin skin test, a serological test. The interpretation of the tuberculin skin test depends upon the person's risk factors for infection and progression to TB disease, such as exposure to other cases of TB or immune suppression (CDC, 2000).

Currently, latent infection is diagnosed in a non-immunized person by a tuberculin skin test, which yields a delayed hypersensitivity type response to an extract made from *M. Tuberculosis* (Kumar et al., 2007). Those immunized for TB or with past-cleared infection will respond with

delayed hypersensitivity parallel to those currently in a state of infection, so the test must be used with caution, particularly with regard to persons from countries where TB immunization is common (Rothel and Andersen, 2005). Tuberculin tests have the disadvantage of producing false negatives, especially when the patient is co-morbid with sarcoidosis, Hodgkins lymphoma, malnutrition, or most notably active tuberculosis disease (Kumar et al., 2007). The newer interferon release assays (IGRAs) overcome many of these problems. IGRAs are *in vitro* blood tests that are more specific than the skin test. IGRAs detect the release of interferon gamma in response to mycobacterial proteins such as ESAT-6 (Nahid et al., 2006). These are not affected by immunization or environmental mycobacteria, so generate fewer false positive results (Pai et al., 2008). There is also evidence that the T-SPOT.*TB* IGRA is more sensitive than the skin test (Lalvani et al., 2005). Diagnosis of TB has also been done with use of various radiotracers using nuclear medicine methods, which not only detects but also locates tubercular infection (Singh, 2009, and Onsel et al., 1998).

2.2.5. Treatment

2.2.5.1.   Vaccines

Many countries use Bacillus Calmette-Guérin (BCG) vaccine as part of their TB control programmes, especially for infants. According to the WHO (1995), this is the most often used vaccine worldwide, with 85% of infants in 172 countries immunized in 1993. This was the first vaccine for TB and developed at the Pasteur Institute in France between 1905 and 1921 (Bonah, 2005). However, mass vaccination with BCG did not start until after World War II (Comstock, 1994). The protective efficacy of BCG for preventing serious forms of TB (e.g. meningitis) in

children is greater than 80%; its protective efficacy for preventing pulmonary TB in adolescents and adults is variable, ranging from 0 to 80% (Bannon and Finn, 1999).

In South Africa, the country with the highest prevalence of TB, BCG is given to all children under age three (WHO, 2006).

BCG provides some protection against severe forms of pediatric TB, but has been shown to be unreliable against adult pulmonary TB, which accounts for most of the disease burden worldwide. Currently, there are more cases of TB on the planet than at any other time in history and most agree that, there is an urgent need for a newer, more effective vaccine that would prevent all forms of TB—including drug resistant strains—in all age groups and among people with HIV (Sadoff and Jerry, 2006).

Several new vaccines to prevent TB infection are being developed. The first recombinant tuberculosis vaccine rBCG30, entered clinical trials in the United States in 2004, sponsored by the National Institute of Allergy and Infectious Diseases (NIAID). A 2005 study showed that a DNA TB vaccine given with conventional chemotherapy can accelerate the disappearance of bacteria as well as protect against re-infection in mice; it may take four to five years to be available in humans (Ha. et al., 2005). Many other strategies are also being used to develop novel vaccines (Dohert and Andersen, 2005) including both subunit vaccines (fusion molecules composed of two recombinant proteins delivered in an adjuvant) such as Hybrid-1, HyVac4 or M72, and recombinant adenoviruses such as Ad35 (http://www.crucell.com/R, 2009). Some of these vaccines can be effectively administered without needles, making them preferable for areas where HIV is very common (Dietrich et al., 2006). All of these vaccines have been successfully tested in humans and are now in extended testing in TB-endemic regions.

2.2.5.2 Orthodox Treatment methods

Treatment for TB uses antibiotics to kill the bacteria. It consists of the initial phase which last two or three months. The risk of drug resistance developing is higher during the early stages of treatment when there are more TB bacteria present. Three or more anti-TB drugs are therefore given to kill the TB bacteria rapidly. The continuation phase lasts for four or five months. Fewer drugs are required to eliminate the remaining TB bacteria in this phase. Effective TB treatment is difficult, due to the unusual structure and chemical composition of the mycobacterial cell wall, which makes many antibiotics ineffective and hinders the entry of drugs (Brennan and Nikaide, 1995). The two antibiotics most commonly used are rifampicin and isoniazid. However, instead of the short course of antibiotics typically used to cure other bacterial infections, TB requires much longer periods of treatment (around 6 to 24 months) to entirely eliminate mycobacteria from the body (CDC, 2000). Latent TB treatment usually uses a single antibiotic, while active TB disease is best treated with combinations of several antibiotics, to reduce the risk of the bacteria developing antibiotic resistance (O'Brien, 1994). People with latent infections are treated to prevent them from progressing to active TB disease later in life.

TB drugs are effective if they are giving in the correct dosage and according to the weight group of the patient. If the prescribed dosage is less than the recommended dosage, the TB bacteria will not be killed and they may become resistant to the drugs. If the dose is higher, it can cause severe toxic effects. The prescribed dosages are based on four weight bands, for adults these are: 30–39kg, 40 – 54kg, 55 – 70kg, and over 70kg. While for children they are: less than 9kg, 10–14kg, 15 -19kg, 20 -29kg. The dose of TB drugs is always based on pre-treatment weight of the

patient and should not be change even if the patient's weight changes. Every TB patient must take the full course of treatment or they will not be cured and may develop drug resistant.

Some of the drugs come in fixed drug combination formulation with the following codes: HR - Isoniazid + Rifampicin, HT- Isoniazid + Thiacetazone, HRZ - Isoniazid + Rifampicin + pyrazinamide, HRZE – Isoniazid + Rifampicin + Pyrazinamide + Ethambutol.

An example of a TB treatment regimen written as a code is $2(HRZE)|4(HR)_3$. The code shows the two phases of TB treatment by writing information about intensive phase in front of a slash and about continuation phase after the slash. A number before the letters indicates the duration of treatment in months – 2months in this example. Each drug is represented by its abbreviation (letter). When these appear in brackets, it indicates that the drugs come in combination tablet. When these appear out of brackets it indicates that single drug preparations are used. A number as a subscript after the drugs indicates the doses are given intermittently, not daily. A subscript of 3 means give three times a week. If no subscript, then the drugs are given daily (NTP training course, 2006).

2.2.6. Relationship between Drug Resistance and Treatment

For new cases of TB, treatment is usually a 4-drug regimen of isoniazid, rifampicin, pyrazinamide, and ethambutol for 2 months (intensive phase), followed by 4-6 months of only isoniazid and rifampicin or isoniazid and ethambutol (continuation phase). The duration of this standardized regimen may vary, depending on a number of factors, including economic considerations and availability of culture-based diagnosis and monitoring. For example, United States guidelines, in contrast to those of most resource-constrained and developing countries,

emphasize mycobacterial culture for TB diagnosis and recommend extending TB treatment for patients with cavities visible on chest radiograph and persistence of positive sputum cultures after 2 months of treatment (Chansuda et.al., 2010).

To prevent inadequate drug ingestion and thereby resistance, staff in TB programs often directly observe patients ingesting their medications. Because directly observed therapy requires substantial human resources, WHO recommends that directly observed therapy should be used any time that rifampicin is administered (Chansuda et. al., 2010). Many countries began using a combination of isoniazid and ethambutol, rather than isoniazid and rifampicin, in the continuation phase, because of the cost of rifampicin and an inability to provide directly observed therapy for the entire duration of this phase. The most common reason to vary the treatment regimen is documented drug resistance or a history of previous TB treatment, which is a risk factor for resistance development (Chansuda et. al., 2010). In most developing countries, drug resistance data are sparse because confirmation of infection with *Mycobacterium tuberculosis* followed by drug susceptibility testing requires use of advanced molecular diagnostics and/or slower and more laborious culture methods. Therefore, patients are usually treated on the assumption that they are infected with a drug-susceptible strain. The drug regimen is usually changed only if the patient's condition does not clinically improve, including having persistently positive sputum smears, after months of treatment. Retreatment protocols in most countries require prolonged therapy with essentially the same basic drugs before the patient is eligible to receive a drug regimen containing second-line drugs specifically for treatment of MDR TB (Chansuda et. al., 2010).

Although multidrug regimens to prevent and treat drug-resistant TB were first evaluated in the 1950s, use of true combination therapy for malaria, the simultaneous use of ≥ 2 drugs (with independent modes of action and different chemical targets) to kill asexual blood-stage parasites, did not arise until much later. However, during the past decade, combination therapy has become the norm, intended to improve effectiveness and reduce the spread of resistance (Chansuda et al., 2008)

In uncomplicated malaria, an outpatient is usually treated with the first-line antimalarial drugs recommended by the local health authority for the malaria-endemic region in which the patient became infected. For example, a patient infected with *P. falciparum* on the eastern Thailand–Laos border would be treated with an artesunate–mefloquine combination plus primaquine at a government malaria clinic on the Thailand side of the border, or with an artemether–lumefantrine combination (Coartem; Novartis AG, Basel, Switzerland) on the Laos side of the border. For travelers returning to their home country outside a malaria-endemic area, different drugs may be prescribed. In none of these situations would a physician expect any laboratory tests to determine drug susceptibility before making a treatment decision. As an acute, potentially fatal disease, falciparum malaria requires effective treatment promptly (Chansuda et. al., 2010).

For malaria, the geographic location in which infection is acquired is the primary determinant of the risk for a drug-resistant infection. Unlike MDR TB, the decision to treat and the treatment of MDR malaria do not require complex clinical and laboratory assessment of an individual patient's isolate, except for severe malaria, which requires critical care capacity. For TB, the geographic area in which infection is acquired is not as reliable a determinant of the treatment choice. Despite wide differences in MDR TB prevalence across countries, the most reliable

predictors of MDR risk for a TB patient are a history of prior treatment or known exposure to another case-patient (i.e., contact with an index MDR TB case-patient), not geography (Granich . et al., 1994).

2.2.7. Epidemiology and Public health Implications of Tuberculosis

According to the report of WHO (2007), roughly a third of the world's population has been infected with *M. tuberculosis*, and new infections occur at a rate of one per second. However, not all infections of *M. tuberculosis* cause TB disease and many infections are asymptomatic (CDC, 2005). In 2007, an estimated 13.7 million people had active TB disease, with 9.3 million new cases and 1.8 million deaths; the annual incidence rate varied from 363 per 100,000 in Africa to 32 per 100,000 in the Americas (WHO, 2009). Tuberculosis is the world's greatest infectious killer of women of reproductive age and the leading cause of death among people with HIV/AIDS (Stop TB partnership, 2002). In the year 2000, one-third of HIV-infected persons worldwide (about 13million people) were also infected with *M. tuberculosis*. Of people infected with both HIV and *M. tuberculosis*, 50% will become sick with TB during their lifetime; 10% will become sick per year. Thus, the prevalence of HIV in a community has an important effect on the incidence of TB.

The rise in HIV infections and the neglect of TB control programs have enabled a resurgence of tuberculosis (Stop TB partnership, 2002). The emergence of drug-resistant strains has also contributed to this new epidemic with, from 2000 to 2004, 20% of TB cases being resistant to standard treatments and 2% resistant to second-line drugs (CDC, 2006). The rate at which new TB cases occur varies widely, even in neighboring countries, apparently because of differences in health care systems (Iademacro and Castro, 2003).

In 2007, the country with the highest estimated incidence rate of TB was Swaziland, with 1200 cases per 100,000 people. India had the largest total incidence, with an estimated 2.0 million new cases (WHO, 2009). The Philippines ranks fourth in the world for the number of cases of tuberculosis and has the highest number of cases per head in Southeast Asia. Almost two thirds of Filipinos have tuberculosis, and up to an additional five million people are infected yearly (http://www.doh.gov.ph/programs/tb, 2009). In developed countries, tuberculosis is less common and is mainly an urban disease. In the United Kingdom, the national average was 15 per 100,000 in 2007, and the highest incidence rates in Western Europe were 30 per 100,000 in Portugal and Spain. These rates compared with 98 per 100,000 in China and 48 per 100,000 in Brazil. In the United States, the overall tuberculosis case rate was 4 per 100,000 persons in 2007 (WHO, 2009). The incidence of TB varies with age. In Africa, TB primarily affects adolescents and young adults (WHO, 2009). However, in countries where TB has gone from high to low incidence, such as the United States, TB is mainly a disease of older people, or of the immunocompromised (Kumar et al., 2007).

Prisoners, especially in poor countries, are particularly vulnerable to infectious diseases such as HIV/AIDS and TB. Prisons provide conditions that allow TB to spread rapidly, due to overcrowding, poor nutrition and a lack of health services. Since the early 1990s, TB outbreaks have been reported in prisons in many countries in Eastern Europe. The prevalence of TB in prisons is much higher than among the general population – in some countries as much as 40 times higher (Larouze et al., 2008).

Knowledge of the drug susceptibility pattern of a TB strain helps not only with therapeutic decision making, but also with predicting of patient's prognosis. From the public health

perspective, information on drug resistance is useful for strategic planning. The proportion of TB case-patients infected with MDR strains, when stratified by previous treatment status, helps public health officials evaluate the intensity of community transmission and the strength of the TB program in curing patients. Unfortunately, the absence of continuous, systematic, representative, and timely drug susceptibility data, especially for second-line anti-TB drugs, is a major obstacle for the control of drug-resistant TB. Consequently, a large number of infectious MDR- and XDR TB cases globally may go undiagnosed.

For TB, the term XDR was created to describe not only TB strains that are resistant to more of the available drugs but also infections that are substantially more difficult to cure. For example, for patients co-infected with HIV, XDR TB is often fatal (Granich et al., 2005). An equivalent term for malaria does not yet exist, although infections with similar characteristics—resistance developed successively to more drugs and the lack of alternative drug choices—will represent identical challenges to control programs.

2.3   Tuberculosis Profile in Ghana

Ghana is not among the WHO's 22 high-burden tuberculosis countries, yet the disease is a major health problem in the country. With an estimated 47,632 new TB cases in 2007, Ghana ranks 19th in Africa for the highest estimated number of new cases per year. Nine percent of the 7,786 TB patients registered in 2007 died before completing TB treatment (WHO, 2009). Ghana's National TB Control Program (NTCP), which is based on DOTS (the internationally recommended strategy for TB control), was established in 1994. DOTS was phased in over a number of years, with an initial strategy to roll out DOTS to all public health facilities in all regions and districts,

and reached 100 percent coverage in 2000. However, the quality of DOTS in public health facilities is still below expectations.

A recent study supported by the U.S. Government showed that some districts had not incorporated DOTS into their routine activities. The case detection rate of sputum smear-positive (SS+) TB has slowly declined over the past five years from 40 percent in 2003 to 36 percent in 2007, well below WHO's target of 70 percent. The treatment success rate has improved from 60 percent since 2002 to 76 percent in 2006; but this is still below WHO's target of 85 percent. The HIV/AIDS epidemic, a generalized epidemic with a prevalence of 2.3 percent, is fueling TB incidence. According to WHO, approximately 15.6 percent of new TB cases are HIV positive, and an additional 30,000 new TB cases annually could be attributable to HIV/AIDS by 2015. Though accurate data are limited, there is a growing concern about multidrug-resistant (MDR) TB. WHO estimated that 1.9 percent of new TB cases were MDR in 2007.

2.3.1   National and Northern Ghana Picture of Tuberculosis Cases

It is estimated that Ghana has 123 smear positive pulmonary TB cases per 100,000 population and 281 of all types TB cases per 100,000 of its population per year. This means that with a population of about 20 million, we should expect about 25,000 smear positive pulmonary TB cases and 56,000 new TB cases of all types every year with about 12,000 deaths (Ghana NTP, 2006). However, only about a fifth of cases (12,000) are reported in this country every year. This is because:

- Many people with TB do not report to health facilities

- Those who report to our health facilities are not diagnosed as having TB(missed diagnosis)

- Not all diagnosed cases at the health facilities are captured by our disease surveillance system (Ghana NTP, 2006).

Most TB patients are adults in the most productive age group and this adversely affects the national and family economy (Ghana NTP, 2006).

The Tamale Metropolis recorded the highest incidence of TB in Northern Region from 2007 to 2010 with Bole District in second place. TB incidence for the region for 2007, 2008, and 2009 are **545, 677, and 556** persons respectively (Northern Regional TB control Unit).

2.4.   Survival Analysis

The origin of survival analysis goes back to mortality tables from centuries ago. However, it was not until World War II that a new era of survival analysis emerged (Allison, 2009). This new era was stimulated by interest in reliability (or failure time) of military equipment. At the end of the war these newly developed statistical methods emerging from strict mortality data research to failure time research, quickly spread through private industry as customers became more demanding of safer, more reliable products (Allison, 2009).

As the uses of survival analysis grew, parametric models gave way to nonparametric and semi parametric approaches for their appeal in dealing with the ever-growing field of clinical trials in medical research. Survival analysis was well suited for such work because medical intervention follow-up studies could start without all experimental units enrolled at start of observation time and could end before all experimental units had experienced an event (Tyler and Besa, 2000).

24

Pardeshi (2009) used survival analysis on TB patients and revealed that TB was a disease with a high case fatality of 4.65% in India. There was no difference in the survival curves of male and female patients.

2.4.1 Applications of Survival Analysis

Survival analysis encompasses a wide variety of methods for analyzing the timing of events. It was originally developed for studying time from commencement of treatment of any disease until death which was commonly used for evaluating treatment efficacy in fatal conditions like cancer (Fox, 2006). The prototypical event is death, which accounts for the name given to these methods. Survival analysis is also appropriate for many other kinds of events, such as criminal recidivism, divorce, child-bearing, unemployment, and graduation from school (Fox, 2006). Survival analysis typically focuses on time to event data from a known origin. Generally, it comprises of techniques for positive-valued random variables such as time to death, time to onset (or relapse) of a disease, length of stay in a hospital, duration of a strike, money paid by health insurance, viral load measurements, time to finishing a doctoral dissertation and so on. Thus, the outcome of interest is time to an event. The best way to define such events is simply to realize that these events are a transition from one discrete state to another at an instantaneous moment in time. Of course, the term "instantaneous", which may be years, months, days, minutes, or seconds, is relative and has only the boundaries set by the researcher (Allison, 2009).

In India, survival analysis and risk factors for death in TB patients revealed that TB is a disease with a high case fatality of 4.65%. There was however, no difference in the survival curves of male and female patients (Pardeshi, 2009). Survival analysis has been used extensively by researchers in studying various diseases. De La et al. (1995) evaluated the duration of treatment

and the reasons for discontinuing therapy with disease modifying antirheumatic drugs in Spanish rheumatoid arthritis patients. That study recommended that methotrexate should be a preferred drug for second line treatment of rheumatoid arthritis patients.

Currently, the most widely used method in the disease management industry for evaluating program effectiveness is the total population approach. This model which is a pretest-posttest design carries the most basic limitation of no control group, which may be sources of bias and/or competing extraneous confounding factors that offer plausible rationale explaining the change from baseline. Cole et al. (2010) presented survival analysis as an alternative, and more appropriate approach to evaluating Disease Management program effectiveness. They had also studied the survival of patients after AIDS diagnosis using survival analysis.

Diagnosis of chronic disease, of either infectious or non-infectious causes, is generally delayed depending on the duration of the latency period: the interval between induction and detection of disease. Consequently, to make inferences on the risk of disease induction based on data of disease detection, appropriate methods must be used to adjust for the latency period. Survival analysis methods which adjust for the latency period are therefore most suitable for the study of chronic disease epidemiology (Kostoulas et al., 2011). An illustrative example was given for the case of *Mycobacterium avium* subsp. *paratuberculosis* (MAP) infection in naturally infected Danish dairy cattle result in reduced milk yield, diarrhea and death (Kostoulas et al., 2011). Most importantly, it has been shown that cows infected earlier in their lives were more likely to subsequently shed detectable levels of MAP and therefore, be a liability to herd-mates (Kostoulas et al., 2011).

CHAPTER THREE

3.0   Methodology

3.1   Research Area

The research considered tuberculosis cases from the entire Northern Region of Ghana because patients are admitted from all over the northern region and detained for treatment at the Baptist Medical Center, Nalerigu where the data used for this study were obtained.

The Northern Region is the largest of the ten regions in Ghana in terms of landmass. It occupies about 70,384 square kilometers accounting for 29.5 per cent of the total land area of Ghana. It is made up of 20 Districts with average population of about 1,820,806, representing 9.6 per cent of the total population of the country (2000 Census).

3.1.1   Source and Data Collection

The data used were secondary data obtained for patients who enrolled at the centre between January and April, 2010. For the sake of this study, a patient was considered cured, if and only if he/she is negative to Smear+ Test. The prognostic factors in the data include: age, sex, date treatment started, date treatment completed, disease classification, category of patient, outcome of treatment, sub-district, and type of patient. The study did not have direct patient involvement, therefore, ethical clearance was not needed.

In all, a retrospective moving cohort of 61 patients was involved in the study over seven months. Patients who die, or got treatment failure, or whose recovery time exceeded the seven months in the course of the treatment were considered censored.

3.2.1 Description of Covariates

The interpretation and definitions of the various prognostic factors employed in the study are presented below:

- Sex: This represents the gender of the respondent coded as 1=female and 0=male.

- Age: It is the age of the patient at the time of admission into the treatment/study

- Disease Classification (dc):   classifies the patients into whether they are Pulmonary positive, pp coded as 1, 0=otherwise.

- Type of Patient: defines the status of the patient as 'new' or 'relapse'. New diagnosis is coded as '1' and '0' otherwise.

- Category: indicates whether a patient belongs to category I, II, or III. Category I (Cat I) patients refers to patients who are diagnosed for the first time, referred to as 'new cases' and coded as '1'. Category II (Cat II) denotes patients who have a treatment failure (relapse or retreatment) and coded as '0'. Category III (Cat III) refers to children who are diagnosed positive of TB and coded as '-1'.

- Treatment Duration: this is the time in weeks it takes a patient to experience an event.

- Outcome: is the result of treatment, if a patient is either cured or completed the treatment, the code '1' is given and '0' otherwise.

3.2.2 Censoring

Censoring comes in many forms and occurs for many different reasons. The most basic distinction is between *left censoring* and *right censoring*. An observation on a variable T is right censored if all you know about T is that it is greater than some value c. In survival analysis, T is typically the time of occurrence for some event, and cases are right censored because observation

is terminated before the event occurs. Thus, if T is the number of months it takes a TB patient to recover, you may only know that T> 7 in which case the person's time of recovery is right censored at 7months. By symmetry, an observation on the variable T is left censored if all you know about T is that it is less than c (T<c). Thus the event occurs before observation starts.

Singly Type I censoring arises when all the observations have the same and fixed censoring time which is under the control of the investigator.

Type II censoring occurs when observation is terminated after a pre-specified number of events have occurred.

Random censoring occurs when observations are terminated for reasons that are not under the control of the investigator. It can also be produced when there is a single termination time but entry times vary randomly across individuals. It is random because entry times are not under the control of the investigator.

3.3.1 Description of Survival Distributions

All the standard approaches to survival analysis are probabilistic or stochastic. Thus, the times at which events occur are assumed to be realizations of some random process. It follows that the event time for some particular individual; T is a random variable having a probability distribution (Allision, 2009). The distribution of these survival times can be characterized by three functions as survivor function, the probability density function, and the hazard function. These functions are equivalent ways of describing a continuous probability distribution, owing to the fact that, given any one of them, we can rediscover the other two.

29

3.3.2   Survival function

The survival function gives the probability of surviving or been event-free beyond time t. it is given by;

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

Where $F(t)$ is the cumulative density function (*c.d.f.*) of the variable T, thus the *c.d.f.* gives the probability that the variable is less than or equal to any value t that we choose. Mathematically the *c.d.f.* is given as;

$$F(t) = \int_{-\infty}^{t} f(t)dt \tag{3.2}$$

Since $S(t)$ is a probability, it is bounded between 0 and 1, and since t can not be negative, we know $S(0) = 1$ implying $f(t) = 0, \, for \, t < 0$.

3.3.3   Probability density function

For continuous variables, one common way of describing their probability distribution is the *probability density function* (*c.d.f.*) denoted by $f(t)$. This function is defined as;

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}. \tag{3.3}$$

Thus, the *p.d.f* is the derivative or slope of the *c.d.f.* and it corresponds to the intuitive notion of distributional shape.

3.3.4   Estimation of survival functions

Prior to 1970, the estimation of survivor function was the predominant method of survival analysis, nowadays, Cox regression is the workhorse of survival analysis. Nevertheless survival

curves are still useful for preliminary examination of the data, for computing derived quantities from regression models like the median survival time and for evaluating the fit of regression models (Allison, 2009). The two methods for estimating survivor functions are the life table and the Kaplan-Meier methods.

PROC LIFETEST produces estimates of survivor functions using either of the two methods.

The **life-table method** or actuarial method may be tractable for large data sets especially when the measurement of event times is crude.

Kaplan-Meier (KM) method or the product-limit estimator is the most widely used method for estimating survivor functions and is more suitable for smaller data sets. The Kaplan-Meier estimator incorporates information from all of the observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. The survival curve describes the relationship between the probability of survival and time.

When there are no censored times or the observations are single right censored (thus for $T \leq t$), the KM estimator is just the sample proportion of observations with event times greater than t. thus if 75% of the observations have event times greater than 5, we have $\hat{S}(5) = 0.75$. However for left censored data (thus, for $T > t$), $\hat{S}(t)$ is undefined.

Supposed there are k distinct event times, $t_1 < t_2 < ... < t_k$. At each time t_j, there are n_j individuals who are said to be at risk of an event. At risk means they have not experienced an event nor have been censored prior to time with censoring time t_j. If any cases are censored at exactly t_j, they are also considered to be at risk at t_j. Let d_j be the number of individuals who die at time t_j. The KM estimator is then defined as;

$$\hat{S}(t) = \prod_{j:t_j \le t} \left[1 - \frac{d_j}{n_j} \right] \text{ for } t_1 \le t \le t_k \tag{3.4}$$

In words, the formula says that for a given time t, take all the event times that are less than or equal to t. For each of those event times, compute the quantity in the brackets, which can be interpreted as the conditional probability of surviving up to time t_{j+1} given that one has survived to time t_j. Then multiply all of these conditional probabilities together (Allison, 2009).

The Kaplan–Meier estimator is defined for any time between 0 and the largest event or censoring time, just that it changes only at an observed event time. When there are tied values the KM estimate is reported only for the last of the tied cases.

3.3.5   Cox Proportional Hazard Model

The model is usually written as $h_i(t) = \lambda_o \exp\{\beta_1 X_{i1} + ... + \beta_k X_{ik}\}$ $\qquad$ (3.5)

This equation says that the hazard for individual i at time t is the product of two factors:

- A baseline hazard function λ_0 that is left unspecified, except that it cannot be negative

- A baseline function of a set of k fixed covariates, which is then exponentiated

The function λ_o can be regarded as the hazard function for an individual whose covariates all have values 0.

Taking the logarithms of both sides, we can re-write the model as

$\log h_i(t) = \alpha(t) + \beta_1 X_{i1} + ... + \beta_k X_{ik}$ where $\alpha(t) = \log \lambda_o(t)$ $\qquad$ (3.6)

If we further specify $\alpha(t) = \alpha$, we get the exponential model. If we specify $\alpha(t) = \alpha t$, we get the

Gompetz model. Finally, if we specify $\alpha(t) = \alpha \log t$, we have the Weibull model. However, the great attraction of Cox regression is that such choices are unnecessary. The function $\alpha(t)$ can take any form whatever, even that of a step function.

It is called a proportional hazards model because the hazard for any individual is a fixed proportion of the hazard for any other individual. To see this, take the ratio of the hazards for two individuals i and j , and apply equation (3.5):

$$\frac{h_i(t)}{h_j(t)} = \exp\left\{\beta_1(X_{i1} - X_{j1}) + ... + \beta_k(X_{ik} - X_{jk})\right\} \tag{3.7}$$

What's important about this equation is that $\lambda_0(t)$ cancels out of the numerator and denominator. As a result, the ratio of the hazards is a constant over time.

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

3.3.5.1 Relative Risk

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