A STATISTICAL INVESTIGATION TO DETERMINE THE PREDISPOSING RISK FACTORS FOR PREDICTING PREVALENCE OF ACTIVE TUBERCULOSIS IN NORTH POKOT SUB COUNTY, WEST POKOT COUNTY, KENYA

LYDIA NYABATE MBOGA (B.Ed. Science)

I56/CE/22360/2010

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIEMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE (BIOSTATISTICS) IN THE SCHOOL OF PURE AND APPLIED SCIENCES OF KENYATTA UNIVERSITY

OCTOBER, 2018

DECLARATION

I declare that this thesis is my original work and has not been presented for a degree in any university or any other award.

LYDIA NYABATE MBOGA

Signature …… Date……..

Approved by supervisors

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

Prof. Leo Odongo

Signature……. Date ……

Department of Statistics and Actuarial Science,

Kenyatta University.

Dr. John Maingi

Signature …… Date ……..

Department of Microbiology,

Kenyatta University.

DEDICATION

I dedicate this thesis with love to the following who were always there for me during my research work: My beloved mother Mary Asuma, my beloved father Peterson Mboga, my uncle James Bogonko, my sister Alice Nyarinda, my brother Caleb Kiriama, my beloved daughter Mary and my lovely son Isaac.

ACKNOWLEGMENT

I greatly express my special deep gratitude to my supervisor Pro. Leo Odongo of Kenyatta University for his invaluable and unlimited support and constructive advices that inspired me to prepare the thesis. I would like to acknowledge the support and guidance of my supervisor Dr. John Maingi of Kenyatta University. Special gratitude to Mr. Jackson Owiti for his support in this research. I greatly appreciate the support and assistance given me by the medical practitioners, the TB monitoring officer the late Mr. Samuel Bargottio of Kacheliba Sub-county hospital, my family and friends. I do acknowledge the National Commission for Science, Technology and Innovation (NACOSTI) through Kenyatta University Ethics Review Committee for granting me permission to undertake my research. I finally give my gratitude to my Heavenly Father for the strength, courage and endurance He granted me during the whole period of my work.

TABLE OF CONTENTS

DECLARATION ..... **i**

DEDICATION ..... **ii**

LIST OF TABLES ..... **viii**

DEFINITION OF TERMS ..... **xi**

ABSTRACT ..... **xii**

CHAPTER ONE **1**

1.1 Background 1

1.2 Statement of the Problem. 3

1.3 Justification 3

1.4 Research questions 5

1.5 Hypotheses. 5

1.6 Objectives 5

1.6.1 General objective 5

1.6.2 Specific objectives 6

1.7 Significance 6

1.8 Organization of the thesis. 6

CHAPTER TWO **8**

LITERATURE REVIEW 8

2.1 Introduction 8

2.2 Risk factors for tuberculosis infection 9

2.3 Tuberculosis transmission 10

2.4 Diagnosis of active pulmonary tuberculosis 11

2.5 Prevention of active pulmonary tuberculosis 12

2.6 Management of active tuberculosis.....12

2.7 The epidemiological burden of tuberculosis.....13

2.7.1 Tuberculosis situation in the world.....13

2.7.2 Tuberculosis situation in Kenya 15

CHAPTER THREE.....18

MATERIALS AND METHODS.....18

3.1 Study area and source of population 18

3.2 Study design.....18

3.3 Study sample.....19

3.3.1 Inclusion criteria 19

3.3.2 Determination of sample size 19

3.3.3 Sampling procedure. 20

3.4 Data collection 20

3.6 Methods for data analysis 21

3.6.1 Chi-square test for homogeneity.....21

3.6.2 Odds ratio. 23

3.6.3 Fisher's exact test 23

3.6.4 Difference of proportions 25

3.6.5 Logistic regression.....26

3.6.6 Goodness of fit 30

3.8 Ethical considerations.....32

CHAPTER FOUR 34

RESULTS AND DISCUSSIONS 34

4.1 Introduction.....34

4.2 Group demographics 34

4.3 Factors for predicting pulmonary tuberculosis 35

4.3.1 Pulmonary tuberculosis prevalence and gender 35

4.3.2 Pulmonary tuberculosis prevalence and smoking 36

4.3.3 Pulmonary tuberculosis prevalence and alcoholism.....38

4.3.4 Pulmonary tuberculosis prevalence and daily nutritionally balanced meals 40

4.3.5 Pulmonary tuberculosis prevalence and level of education 42

4.3.6 Pulmonary tuberculosis prevalence and infection history 43

4.3.8 Pulmonary tuberculosis prevalence and accessibility to health facility.....44

4.3.9 Pulmonary tuberculosis prevalence and occupation.....46

4.3.10. Pulmonary tuberculosis prevalence and residential ventilation 47

4.3.11. Pulmonary tuberculosis prevalence and residential house.....49

4.3.12. Pulmonary tuberculosis prevalence and infection knowledge 51

4.3.13. Pulmonary tuberculosis prevalence and house congestion 52

4.3.14 Logistic regression model of pulmonary tuberculosis 54

4.4 North Pokot Sub-county pulmonary tuberculosis prevalence 57

4.4.1 North Pokot Sub-county pulmonary tuberculosis prevalence and age group ...57

4.4.2 North Pokot Sub-county and National pulmonary tuberculosis prevalence 59

4.5 Divisional pulmonary tuberculosis prevalence. 61

5.2 Summary.....63

5.3 Conclusions.....63

5.4.1 Recommendations from the study.....64

APPENDICES 71

Appendix 1: Informed Consent. 71

Appendix 2 Questionnaire.....74

Appendix 3: Data coding.....78

Appendix 4: West Pokot County Map 80

Appendix 5: R program.....81

LIST OF TABLES

Table 4.1: Pulmonary tuberculosis prevalence by gender 35

Table 4.2: Fisher's exact test result comparing male and female PTB prevalence 36

Table 4.3: Pulmonary tuberculosis prevalence among smokers and non-smokers 37

Table 4.4: Fisher's exact test result comparing PTB prevalence among smokers and non-smokers 37

Table 4.5: Pulmonary tuberculosis prevalence among alcoholic and non-alcoholic patients 39

Table 4.6: Fisher's exact test result comparing PTB prevalence between alcoholic and non-alcoholic patients 39

Table 4.7: Pulmonary tuberculosis prevalence and daily nutritionally balanced meals 41

Table 4.8: Pulmonary tuberculosis prevalence and level of education 42

Table 4.9: Pulmonary tuberculosis among patients with and without infection history ...43

Table 4.10: Fisher's exact test result comparing PTB prevalence between patients with and without infection history 44

Table 4.11: Pulmonary tuberculosis prevalence by hospital accessibility 45

Table 4.12: Fisher's exact test result comparing PTB prevalence between patients with and without accessibility 46

Table 4.13: Pulmonary tuberculosis prevalence and patient occupation 47

Table 4.14: Pulmonary tuberculosis prevalence by residential ventilation 48

Table 4.15: Fisher's exact test result comparing PTB prevalence between patients residing in good and poorly ventilated houses 48

Table 4.16: Pulmonary tuberculosis prevalence by type of residential house 50

Table 4.17: Fisher's exact test result comparing PTB prevalence between patients residing in cemented and mud walled houses 50

Table 4.18: Pulmonary tuberculosis prevalence and patient knowledge.....51

Table 4.19: Fisher's exact test result comparing PTB prevalence between patients with and without infection knowledge 52

Table 4.20: Pulmonary tuberculosis prevalence and house congestion 52

Table 4.21: Fisher's exact test result comparing PTB prevalence between patients residing in congested and non-congested houses 53

Table 4.22: Result of logistic regression model of PTB prevalence - I.....55

Table 4.23: Result of logistic regression model of PTB prevalence - II 56

Table 4.24: Pulmonary tuberculosis prevalence by age group 58

Table 4.25: District and national pulmonary tuberculosis prevalence 59

Table 4.26: A Chi square test result comparing Pokot District and national PTB prevalence.....60

Table 4.27: Divisional pulmonary tuberculosis prevalence.....61

ABBREVIATIONS AND ACRONYMS

AIDS: Acquired Immune Deficiency Syndrome.

AFB: Acid Fast Bacilli

BCG: Bacilli Calmette Guerin Vaccine

CDC: Centers for Disease Control

DOT: Directly Observed Treatment

DLTDLD: Division of Leprosy, Tuberculosis and Lung disease

HIV: Human Immune Deficiency Virus

LTBI: Latent TB infection

MDR TB: Multi Drug resistant Tuberculosis

PTB: Pulmonary Tuberculosis

TB: tuberculosis

WHO: World Health Organization

XDR TB: Extensive Drug Resistance Tuberculosis

DEFINITION OF TERMS

Tuberculosis Incidence: This is a measure of the number of new cases who develops active tuberculosis in a given population during a specified period of time.

Tuberculosis prevalence: This is the proportion of people in a given population living with active TB at a given time.

Risk factor: This is a person's characteristic, environment, or a situation, which increases the chances of a person to develop a disorder.

TB suspects: Individuals who present with signs and symptoms suggestive of TB.

TB cases: Individuals who are diagnosed with TB and are smear sputum positive.

Epidemiology: It is the study that deals with the spread and control of diseases in a population.

ABSTRACT

Tuberculosis is a major threat to world health. Finding out the TB cases and treatment of the disease are the tenet means of controlling TB transmission and reducing its incidence. In many industrialized countries, the prevalence of tuberculosis has declined significantly in the last decade, and elimination of TB has come back as a foreseeable goal, based on efficient treatment of overt TB cases and treatment of latent TB infection to prevent development of the disease. In developing countries, however, the number of TB cases is reported to increase steadily, especially in Africa South of Sahara, where TB is a leading cause of mortality. TB disease burden in Kenya is large and rising. Kenya is ranked 13th out of the 22 countries which collectively bring about 80% of TB cases in the World. It affects all age groups but more so the economically productive age group of 15 and 44 years. Few studies have considered the risk factors for tuberculosis at community level in highly resource-poor countries. The study was conducted in Kacheliba Sub-County hospital in West Pokot County. The aim of the study was to investigate and determine the predisposing risk factors for predicting the prevalence of active tuberculosis in the region. The study design which was employed was analytical cross sectional design targeting all persons above 15 years and children of 0-14 years. Hospital records and questionnaires were used in collecting data. Chi-square and Fisher's exact tests at 5% level of significance were used in comparing pulmonary tuberculosis prevalence between subgroups in r × c tables. A Logistic regression model was used to determine the significant factors for predicting tuberculosis in North Pokot Sub-county. The study findings indicated that tuberculosis prevalence in North Pokot Sub-county was higher than the national prevalence; 9/1,000 and 2/1000 respectively. The logistic regression model indicated that, alcoholism, smoking and congestion as being statistically significant factors in predicting tuberculosis in North Pokot Sub-county. The study findings will be used by community health care workers in creating awareness among community members on predisposing factors for tuberculosis and on the need to sought medication early to avoid complication of the disease and further transmission.

CHAPTER ONE

INTRODUCTION

1.1 Background

Tuberculosis is caused by *Mycobacterium tuberculosis.* The highly infectious disease was discovered by Robert Koch a great physician on March 24, 1882. TB affects almost all organs of the body but mostly it affects the lungs. It spreads when an individual in close contact to a TB patient breaths in infected air. Tuberculosis may remain in its inactive state for many years without infecting other people or causing symptoms (Schiffman, 2012).

Tuberculosis develops in two stages. First a susceptible person exposed to a TB case becomes infected and later may develop the disease depending on various factors. A condition which changes the stability established between the immune defenses and the *Mycobacterium tuberculosis* in infected individuals can impact the risk to develop the TB disease. Factors which influence this stability have been investigated for a very long period of time. The factors were found to be both 'innate' and 'external' to the host (Lienhardt *et al*., 2005). Most studies have considered host-related and environmental risk factors for TB separately. The studies used different designs. This has made it difficult to assess the respective effects of these factors (Lienhardt, 2001). In addition, there are few studies which have been carried in resource-poor countries (Rieder, 1999). The re-emergence of tuberculosis disease which mainly affects developing countries, posed the need to investigate the predisposing risk factors for predicting prevalence of active tuberculosis in order to improve control policies for tuberculosis.

Tuberculosis among other diseases has caused public health as well as economic crises in developing countries. The principle means of controlling tuberculosis and reducing its incidence is finding out the cases and treating the disease. Tuberculosis has been found to have declined greatly in most of the industrialized countries in a period of the last ten years. This has brought forth a foreseeable goal of the elimination of TB based on effective treatment of active TB cases and treatment of inactive TB infection so as to prevent the development of the disease. (Barner, 2001). This is not the case in developing countries since TB is reported to be a leading cause of death rate due to great and increasing number of TB cases especially in Africa South of Sahara (Dye *et al*., 1999).

Kenya is positioned 13th out of the 22 nations which together contribute to nearly 80% of the TB cases in the world. A report indicates that the notified rate of TB cases between 1987 to 2009 rose from 51 to 326 per 100,000 populations respectively (WHO Report, 2009). The disease was also found to cause public health crisis among pastoralists within North Pokot Sub-county, West Pokot County. The region was found to be semi-arid and therefore contributes to 10% of the country's TB burden (Sitienei *et al*., 2010). There was no reliable information on predisposing risk factors for predicting the prevalence of tuberculosis in the region. This study endeavored to statistically investigate and determine the predisposing risk factors for predicting the prevalence of tuberculosis using hospital records and questionnaires in Kacheliba Sub County hospital. The study findings will be useful in promoting TB awareness in the community and in provision of appropriate health services and resources by clinicians and community health workers. This in turn will considerably reduce the time lapse between infection and diagnosis then treatment, thus controlling the spread of TB.

1.2 Statement of the Problem.

Tuberculosis has caused public health as well as economic crises in developing countries (Barner, 2001). According to (Dye *et al*., 1999), TB is a leading cause of death in developing countries.

Tuberculosis develops in two stages. First a susceptible person exposed to a TB case becomes infected and later may develop the disease depending on various factors. A condition which changes the stability established between the immune defenses and the *Mycobacterium tuberculosis* in infected individuals can impact the risk to develop the TB disease. Factors which influence this stability have been investigated for a very long period of time. The factors were found to be both 'innate' and 'external' to the host (Lienhardt *et al.,* 2005). Most studies have investigated host-related and environmental risk factors for TB separately. The studies used different designs. This has made it difficult to assess the respective effects of these factors (Lienhardt, 2001). In addition, there are few studies which have been carried out in resource-poor countries (Rieder, 1999). The re-emergence of tuberculosis disease which mainly affects developing countries, posed the need for a statistical investigation to determine the predisposing significant risk factors for predicting the prevalence of tuberculosis using logistic regression model at community level in highly resource poor country in order to improve control policies for tuberculosis.

1.3 Justification

The Kenyan population has individuals who are predisposed greatly to tuberculosis infection than the general population. This is due to the fact that 80% of the land of

Kenya is categorized as desert or semi-desert which constitutes 10% of the of the tuberculosis burden of the country. The communities that live in these areas carry approximately 10 percent of the country's TB burden. The community under study was pastoral-nomadic dwelling in a semi-desert area. The demographic and environmental predisposing risk factors to tuberculosis in the community is a great concern. Information on significant predisposing risk factors for predicting the prevalence of tuberculosis North Pokot Sub County is not known hence the need for updated information. The study findings will provide a snapshot of the prevalence of tuberculosis in the period under study and its burden to National TB prevalence. In this case the government will develop initiatives to control TB disease like employing more health practitioners to cub with the disease and create awareness to the community on the burden of TB in the region so as to be keen on preventive measures. The government can construct more TB clinics due to vast nature of the topography of the region under study for accessibility by the patients from the onset of the symptoms. This will enhance treatment process and prevent further spread to health individuals. The study findings will also provide information on the predisposing risk factors which will be used to predict TB in the region under study. This will be used by community health workers in educating the community on risk factors predisposing one to TB infection and development of the disease. This in turn will help the community to observe health living free of TB so as to remain economically active. The government will also use the obtained information in creating proper measures to control some risk factors such alcoholism and smoking which enhance development of TB disease. Housing and nutrition factors will be investigated and necessary support be

provided by the government. All these measures will help meet the vision 2030 goal of a TB free community.

1.4 Research questions

i. What would be the significant predisposing risk factor for predicting tuberculosis in North Pokot Sub-county, West Pokot County?

ii. What is the contribution of TB prevalence in North Pokot Sub-county, West Pokot County to the National TB prevalence?

iii. Among the divisions in North Pokot Sub-county which one has significantly highest TB prevalence?

1.5 Hypotheses.

H_{01}: There is no significant predisposing risk factor for predicting tuberculosis in North Pokot Sub-county, West Pokot County.

H_{02}: The prevalence of tuberculosis in North Pokot Sub-county, West Pokot County is not different from that of the country.

$H_{03:}$ TB prevalence among divisions in North Pokot Sub-county is the same.

1.6 Objectives

1.6.1 General objective

To investigate the predisposing risk factors for predicting the prevalence of active pulmonary TB in North Pokot Sub-county, West Pokot County.

1.6.2 Specific objectives

i. To investigate the significant predisposing factors for predicting tuberculosis in North Pokot Sub-county, West Pokot county.

ii. To investigate the prevalence of PTB in North Pokot Sub-county, West Pokot County.

iii. To compare PTB prevalence among divisions in North Pokot Sub-county, West Pokot County using appropriate statistical tools.

1.7 Significance

The study will contribute to TB knowledge in the community in that individuals will be enlightened on the significant predicting factors for TB infection in the region. The county government will obtain results on TB prevalence in North Pokot Sub-county compared to National TB prevalence. The study findings will be useful in promoting TB awareness in the community and in provision of appropriate health services and resources by the government, clinicians and community health workers. This in turn will considerably reduce the time lapse between infection and diagnosis then treatment, thus controlling the spread of TB.

1.8 Organization of the thesis.

The rest of the thesis is organized as follows: chapter two is the literature review which gives information on other studies on the epidemiology of TB disease and its risk factors. Chapter three discusses all the statistical methods, formulae, statistical model, theoretical properties of the statistics including their manipulation in testing hypotheses, construction of confidence intervals, and interpretation of results to be obtained in chapter four based on the study objectives. Chapter four provides the analysed results of the study and

discusses them in relation to other study findings. Chapter five provides a conclusion summary of the study findings in terms of the study objectives. Recommendations on what to be done to solve the problem under study and suggestions on further study are also contained in chapter five.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Tuberculosis is a highly infectious disease. It is widely spread and fatal in many cases. TB is caused by *Mycobacterium tuberculosis (*Kumar *et al., 2007).*TB affects all parts of the body but mostly it attacks the lungs. It is an airborne disease spread from individuals with active TB when they transmit respiratory fluids as they cough or sneeze. When infections do not have symptoms it is known as latent tuberculosis. The latent infections can eventually move forward to active TB disease, which can lead to death in the infected individuals if left untreated (Konstantinos, 2010).

Tuberculosis which infects the lungs is known as pulmonary tuberculosis. When it affects other parts of the body other than the lungs it is known as extra pulmonary TB. Generally TB symptoms include: lack of appetite, weight loss, fever, night sweats, fatigue and chills (Dolin *et al.,* 2010). 90% of active tuberculosis infection is Pulmonary TB. PTB symptoms include prolonged cough of more than two weeks producing sputum, chest pain, in some cases individuals cough small amounts of blood. The PTB infection may progress to pulmonary artery and lead to intensive bleeding (Halezeroglu *et al., 2013*). 15-20% of active cases develop extra pulmonary tuberculosis. Extra PTB mostly affects young children and immunosuppressed individuals. 50% of the cases are those with HIV. The parts of the body mostly infected with extra pulmonary tuberculosis include the central nervous system, the genitourinary system, the lymphatic system, the pleura, the joints and bones (Jindal, 2011). This study has researched on active pulmonary

tuberculosis, its epidemiology and predisposing risk factors at community level, however there is need to study more on extra pulmonary tuberculosis.

2.2 Risk factors for tuberculosis infection

A number of factors make people more susceptible to TB infections. According to (WHO, 2011) HIV was noted globally to be the greatest risk to TB infection where 13% of TB cases were also HIV positive. According to (Gibson *et al.,* 2005) in sub-Saharan Africa HIV infection rates are very high. Out of the people with HIV 30% develop active tuberculosis while 5-10% of those without HIV infection develop TB active disease. According to (Lawn *et al.,* 2011) malnutrition and overcrowding are other risk factors for tuberculosis infection. These places TB as the core disease of poverty. This study has investigated the extent to which malnutrition and overcrowding factors predisposes individuals to pulmonary tuberculosis infection at community level. However, the contribution of HIV to TB infection was not investigated.

There are individuals who are at a high risk to TB infection. These include children who are in close contact with TB patients, TB health care workers, resource poor communities, and people in prisons, high risk ethnic minorities and employees in areas with vulnerable people such as prisons. Diseases such as silicosis and chronic lung disease are also significant risk factors to TB infection. Other disease states can also increase the risk of developing tuberculosis, these include alcoholism and diabetes (CDC, 2010). According to (Van *et al.,* 2010) cigarette smoking is another risk factor for TB infection. It has been noted that individuals who smoke are twice likely to develop TB compared to those who do not smoke. There is a genetic element to TB infection of which its overall importance has not been defined, this is according to (Lawn *et al.,*

2011). This study has researched on the extent of TB infection in children from 0-14 years as well as individuals above 15 years and the contribution of cigarette smoking and alcoholism to PTB infection in a resource poor community.

Demographic factors are also contributing risk factors to TB infection. For instance the male gender is highly susceptible to TB infection than the counterpart females especially those who are not married. This is according to a study which was done in West Africa (Lienhardt *et al., 2005*).This study has considered the male and female gender to determine the risk group to TB infection in a resource poor community.

Housing conditions are also contributing factors to TB infection. According to a study conducted in West Africa TB infection increased if houses were mud walled with soiled floors compared to cemented houses (Lienhardt *et al., 2005*). Poor ventilation is a contributing factor to TB infection. This is according to a study carried out in hospitals and health care facilities (Menzies *et al*., 2000).The current study has investigated the contribution of ventilation, mud and cemented houses to TB infection in the study area.

2.3 Tuberculosis transmission

TB is transmitted from people with active pulmonary TB. This happens when they release an infectious aerosol droplets as they sneeze, spit, sing, speak or cough. A single cough or sneeze releases a maximum of 40,000 droplets of which each one may transmit the disease, because a dose for tuberculosis infection is very small (CDC, 2011).

Tuberculosis transmission depends on certain factors like the number of infectious droplets expelled from TB active individual not latent TB individuals since they are not contagious (Kumar *et al.,* 2007), the exposure duration, the ventilation condition of the

house, the virulence of the *Mycobacterium tuberculosis* strain and the immunity level of the infected individual CDC,2011). A newly infected individual takes three to four weeks to become infectious so as to transmit the TB disease to other people (Mayo, 2007).

The main environmental factors for the transmission of TB are the extent and persistence of contact with an infected person. Prolonged contact in enclosed environments with an infectious person thus leads to TB transmission. People with prolonged, frequent, or close contact with TB are particularly at high risk of becoming infected with tuberculosis (Ahmed *et al.*, 2011). A person with untreated active tuberculosis may infect more than 15 people per year (WHO, 2010). People who are at a greatest risk to TB exposure are those who live in the same household with an infected person and also sleep together (Singh *et al.*, 2005). Persons who share a house with an infected person are therefore at a risk of TB infection. A relationship has been found out between overnight cough and increased TB transmission among household contacts (Fennelly *et al*., 2006).

This study did not investigate household contact with an infected person as a predisposing factor for tuberculosis.

2.4 Diagnosis of active pulmonary tuberculosis

TB is diagnosed by using a clinical sample of sputum to identify *Mycobacterium tuberculosis*. The sputum samples are produced for examination early in the morning for at least three times. Blood or sputum culture are also used to identify the bacterium which takes two to six weeks since the organism grows slowly. Treatment is usually begun early before cultures are confirmed (Pai *et al.*, 2008). The study considered only sputum sample diagnosed results.

2.5 Prevention of active pulmonary tuberculosis

Prevention and control of TB is basically done on the vaccination of infants. The vaccine used is known as Bacillus Calmette-Guerin (BCG) which decreases the risk of infection by 20% and development of TB disease by 60%. The immunity built normally decrease after a period of ten years. The principle means to control the spread of TB in active TB cases is early detection and provision of appropriate treatment (Lawn *et al*., 2011).This study did find out the extent of BCG vaccination in children who participated in the study.

2.6 Management of active tuberculosis

Antibiotics are used in the treatment of TB. Many antibiotics used in TB treatment are ineffective due to the unusual structure and chemical composition of the cell wall of *Mycobacterium tuberculosis* which prevents drugs to pass through (Manzies *et al.,* 2011). Isoniazid and rifampicin are the two common antibiotics used in the treatment of TB. Treatment normally takes several months. Treatment of latent TB employs only a single antibiotic. The treatment of active TB disease involves a combination of several antibiotics. This reduces the risk of the bacterium to develop resistance to the antibiotic (CDC, 2011). WHO recommends directly observed therapy (DOT).This is where a health care provider watch the person take their medication. This helps to reduce the number of people who do not take the antibiotics appropriately (Arch *et al.,* 2010). Methods to remind people of importance of treatment appear effective than DOT (Leu *et al.,* 2008).

WHO recommends treatment of new-onset of tuberculosis to take six months. This involves a combination of antibiotics such as Isoniazid, rifampicin, ethambulol and

pyrazimamide. Ethambulol may be added for the last four months as an alternative in case there is resistance to Isoniazid (WHO, 2009).

After treatment of TB in case it recurs testing is done to determine the antibiotic one is resistant to before treatment is begun. If an individual has multi-drug resistance TB, treatment for a period of 18 to 24 months is recommended. This treatment involves at least four effective antibiotics (Lawn *et al.*, 2011).

Drug resistance TB is a serious public health issue in many developing countries, as its treatment is longer and requires more expensive drugs. MDR-TB is defined as resistance to two more effective first-line TB drugs; rifampicin and Isoniazid. Extensively drug-resistant TB is resistant to three or more of the six classes of second-line drugs .This occurs due to inadequate treatment, inappropriate use of prescribed drugs, or using low quality medicine (Maryn, 2012). This study did not investigate the prevalence of MDR-TB and extensive drug resistance TB.

2.7 The epidemiological burden of tuberculosis

2.7.1 Tuberculosis situation in the world

Mycobacterium *tuberculosis* has infected about one-third of population in the world (WHO, 2010). New infections occur in nearly 1% of the population per year. According to (CDC, 2011) most infections with *Mycobacterium tuberculosis* do not cause TB disease and 90 to 95% of infections remain asymptomatic. 8.6 million chronic active TB cases were notified in 2012 (WHO, 2012). According to WHO report in 2010, 8.8 million new cases of TB were diagnosed and 1.3-1.45 million deaths occurred. Most of these

deaths occurred in developing countries (WHO, 2011). About 0.35 million from these 1.45 million deaths are also those infected with HIV (WHO, 2012).

Tuberculosis follows HIV/AIDS as being the most common infectious diseases which cause death. Since the year 2005 prevalence of tuberculosis has been declining while since the year 2002 TB incidence has also been decreasing (WHO, 2011). The tuberculosis mortality rate in China from 1990 to 2010 had reduced drastically by about 80% while the number of new cases declined by 17% between 2009 and 2014 (Kiesitra *et al.*, 2014). Tuberculosis mostly affects developing countries than the developed ones. This is seen whereby about 80% of the population in many African and Asian countries test positive in tuberculin tests comparing to the US population where only 5-10% test positive. The efforts towards controlling tuberculosis disease completely have been dramatically depressed due to the following factors such as; the expensive and time-consuming diagnostic process, difficulty of developing an effective vaccine, the necessity of many months of treatment, the emergence of drug-resistant cases in the 1980s and the increase in HIV associated tuberculosis (Lawn *et al.*,2011).

Switzerland had the highest TB incidence rate of 1,200 cases per 100,000 people in the year 2007 while India had an estimated 2.0 million new cases (WHO, 2009). Tuberculosis is less common in developed countries where it is mainly found in urban areas. In different areas of the world rates per 100,000 people in the year 2010 were: globally 178, the Americans 36, Africa 332, Eastern Mediterranean 173, Southeast Asia 278, Europe 63, and Western Pacific 139 (WHO, 2011).

In other countries such as the United States tuberculosis incidence rates have decreased dramatically. In many areas TB mainly affects the immunocompromised and the older people. Generally in the whole World the 22 high TB burden countries experience together 80% of cases and 83% of deaths (Kumar *et al*., 2007).

2.7.2 Tuberculosis situation in Kenya

Kenya is ranked 13th out of the 22 high TB burden countries. This is according to the 2010 WHO Global Tuberculosis Report. The report indicated that TB cases increased from 11,625 in 1990 to 110,065 cases in 2009. Over the past ten years the average annual increase rate for all forms of TB was 7%.In the year 2009 the case notification rate increased from 52 per 100,000 population for all forms of TB to 280 per 100,000 and form 32 per 100,000 population for sputum smear-positive pulmonary tuberculosis cases to 95 per 100,000 population. The proportion of sputum smear-positive PTB cases increased by 4% in 2009 compared to 2008.In 2009 the age group with highest TB notification rate as well with high HIV prevalence was 25-34 in both males and females. This has been the trend in the last ten years. TB notification rate in males start to exceed that of females after the age of 24 (Sitienei, *et al*., 2010). In the year 2010, the total number of all forms of TB patients notified were 106,083 which represented a 4% decrease compared to the 110,065 cases who reported in year 2009.The high HIV prevalence which is estimated at 7.1% in the general population has contributed to the high tuberculosis burden. This is according to 2010 data, where 41% of tuberculosis patients were infected with HIV (Sitienei, *et al*., 2013).

In year 2010, WHO was notified with a total of 112 MDR TB cases. By the end of 2010, the country had in total initiated 180 patients on treatment with 70 of them being initiated

within the year. DR TB is projected to rise over the years (USAID, TB CARE-1 2013). Tuberculosis treatment success rate for the 2009 cohort is 85.5% for new smear positive pulmonary TB cases in a total of 37,402 patients. The case detection rate for all forms of TB was reported as 85% by WHO 2010 Global report.

Kenya attained 70 out of 85 of the global TB control targets in the year 2007. These successes were attributed to the well-developed health care infrastructure such as the laboratory network with 1,538 AFB microscopy centers translating to one microscopy center per 26,000 persons and 2,818 treatment centers, trained health workers with integrated tuberculosis management and case detection, quarterly review meeting and monthly supportive supervisions. The private and public health sectors have incorporated to provide all levels of health care to TB patients (WHO, 2010).In the year 2012 the total TB notified cases were 89,568. In 2011 the successful rate of treatment was 88% for new smear-positive and culture-positive (USAID and TB CARE- 1, 2013).

An epidemiological study on the tuberculosis prevalence in north Eastern Province indicates that out of 400 subjects screened for tuberculosis 9.2% were smear sputum positive an indication that the community has a high prevalence of tuberculosis (Abdikadir, 2004).

Homa-Bay district in rural Kenya is a region with intermediate TB prevalence. A study carried out from 2005-2007 indicates that about 17% (1122) of the number of patients visiting a TB laboratory at Homa-Bay district hospital were positive for TB. (Shitandi *et al.*, 2010). There is no available study on the predisposing risk factors for predicting the

prevalence of tuberculosis in North Pokot sub –County, West Pokot County hence the purpose of this study.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study area and source of population

North Pokot Sub-county is situated in West Pokot County, Great rift region, Kenya. The area is divided into four divisions namely: Kacheliba, Konyao, Kasei and Alale. It hosts the sub County hospital, a TB manyatta, mission clinics, private clinics and public clinics. The Sub-county borders with the Republic of Uganda. The weather in the area is generally hot and dry. The indigenous people in the study area are called the Pokot. Generally most of the people are semi-illiterate and speak the Pokot language. The region is a lowland experiencing long rainfall between April and July but not as much as in the highland regions. The economic activity of the area is agro-pastoralism. Small scale farming is practiced by people who live along the river.

3.2 Study design

The research employed analytical cross sectional study design in gathering information about the prevalence of active pulmonary tuberculosis and the possible predisposing risk factors. The targeted population was that of the patients attending Kacheliba Sub county hospital. Out of this population patients who presented themselves to the hospital with symptoms of tuberculosis were chosen randomly upon leaving the doctor's room from the period of October 2014 up to March 2015. The patients after undergoing laboratory sputum test the results were recorded in the TB register. The patients were divided into two groups those with positive sputum test and those with negative sputum test. These patients were subjected to a questionnaire upon signing a consent form to determine PTB

prevalence and investigate further on the predisposing risk factors. The factors under investigation included both demographic and socio economic factors. The following were the factors which were investigated: age and sex to determine the highly susceptible age group and gender to TB infection, education level, occupation, TB knowledge, TB medication history, nature of the residence house whether mud or cemented and its ventilation, alcoholism, nutrition, accessibility to a health facility, smoking, and congestion in the residential house .

3.3 Study sample

3.3.1 Inclusion criteria.

The targeted sample was of the patients presenting themselves to the hospital with symptoms of tuberculosis. The patients must have stayed in the study area for at least six months. They must be present in the TB register and consent to participate in the study. The age of the participants was children of 0-14 years and individuals from 15years up to 80 years. Both the male and female gender were included in the study.

3.3.2 Determination of sample size.

The formula which was employed in determining the sample size was the Slovin's formula (1960). This is the formula:

$n = \frac{N}{1+Ne^2}$ where:

n represents the sample size, N represents the population size and e represents margin of error.

Substituting:

$$n = \frac{23,679}{1+23,679(0.05^2)}$$

n= 393

3.3.3 Sampling procedure.

Simple random sampling method was employed in selecting the required sample for the study. A random number table was used to select participants who met the inclusion criteria during the study period of October 2014 up to March 2015. The participants were obtained from the patients attending Kacheliba Sub County Hospital presenting with symptoms of TB in the TB clinic register. The participants must have stayed in the study area for at least six months and consent to take part in the study. The age bracket of the participants was those above 15 years and children of 0 to14 year age group. At the end of the study period only 300 participants were sampled and this was the sample size that was used due to the convenience of the study time and expenditure.

3.4 Data collection

Data was collected from patients attending Kacheliba Sub County hospital presenting with symptoms of tuberculosis from October 2014 to March 2015. The patients upon leaving the doctor's room were tested for tuberculosis using sputum sample in a laboratory. The results were recorded in a TB register. Hospital registers contained data on the name, age, sex, address, date of registration , alongside clinical details such as final diagnosis whether PTB sputum positive or PTB sputum negative. The patients who were selected randomly to participate in the study were provided with questionnaires by research assistants. The questionnaires contained guided questions on demographic and socio economic predisposing factors for TB infection. An interview schedule was also

employed in collecting more information from the participants. For the younger children from 0-14 years, the parents consented on their behalf and were interviewed then information filled in the questionnaire.

3.6 Methods for data analysis

The data captured in the questionnaires was grouped and examined for errors, then entered in a summary sheet and analysed using R-software version 2.3.0 a statistical package. Hypotheses on individual risk factors were tested using Fisher's exact test and Chi-square test at 5% level of significance. A logistic regression model was fitted to determine the predisposing risk factors which were significant in predicting tuberculosis in North Pokot Sub-county, West Pokot County.

3.6.1 Chi-square test for homogeneity

In the general situation of a $r \times$ c table obtained from a sample survey, let X_1 and X_2 represent two categories of variables, X_1 with *r* levels and X_2 with *c* levels; there are *r c* combinations of classes (*r c* cells). Considering $\{H_0\}$ the null hypothesis that the cell probabilities is equivalent to certain fixed values$\{\pi_{ij}\}$. The expected frequencies are the values $\{e_{ij} = n\pi_{ij}\}$which represent the values of the expectations $\{E(n_{ij})\}$ for a sample of size *n* with cell counts $\{n_{ij}\}$ when H_0 is true (Agresti, 2007).

The chi-square test for homogeneity finds out whether or not the *r* independent samples (represented by *r* rows) are homogeneous according to the proportion of observations in each of the *c* categories. If the data are homogeneous, the proportion of the observations in the j^{th} category will be the same throughout the *r* populations (Sheskin, 2000). The observed frequencies $\{n_{ij}\}$ and the expected frequencies $\{e_{ij}\}$ are compared to find out

whether the data varies with the H_0. Suppose we want to test for homogeneity, the null hypothesis and alternative hypothesis can be written as:

H_0: In the underlying population the samples represent, all of the proportions in the same column of the $r \times$ c table are equal

H_1: In the underlying population the samples represent, all of the proportions in the same column of the $r \times$ c table are not equal for at least one column

The Pearson chi-squared statistic for testing H_0 is given as:

$$\chi^2 = \sum \frac{\left(n_{ij} - e_{ij}\right)^2}{e_{ij}}$$

The χ^2 statistic has approximately a chi-squared distribution, for large *n*. For *r* rows and *c* columns, the degrees of freedom is

$$df = (r-1)(c-1)$$

This statistic takes its minimum value of zero when $n_{ij} = e_{ij}$ for all i and j. For a fixed sample size, greater differences $\{n_{ij} - e_{ij}\}$ produce larger values of χ^2 hence stronger evidence against H_0.The chi-squared approximation improves as e_{ij} increase, and $e_{ij} \geq 5$ is usually sufficient for a decent approximation.

The chi-square test for homogeneity will be used to compare proportions of patients with pulmonary tuberculosis infection in $r \times$ c tables $(c > 2)$.The chi-square test for homogeneity assumes that the sums of the *r* rows determined by the researcher prior to the data collection phase of the study and is based on the following assumptions:

a) Mutually exclusive categories are employed in the analysis
b) The data represent a random sample comprising of *n* independent observations
c) The expected frequency of each cell in the contingency table is equal to or greater than 5

3.6.2 Odds ratio.

Odds ratio θ is a measure of strength of association for 2×2 contingency tables. A sample odds ratio is given by:

$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

Values of $\hat{\theta}$ further from 1.0 in a given direction represent a stronger association. When $\widehat{\theta} > 1$, the odds of success are higher in row 1 than in row 2. A large sample $100(1-\alpha)\%$ confidence interval for the estimated odds ratio $\hat{\theta}$ is given as:

$$exp\left[\log\hat{\theta} \pm Z_{\alpha/2}\{se\ (\hat{\theta})\}\right], \quad se\ (\hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

Where $se(\ \hat{\theta}\ )$ is the standard error for $\log\hat{\theta}$.

3.6.3 Fisher's exact test

For 2×2 contingency tables (Table 3.1), homogeneity or independence corresponds to an odds ratio $\theta = 1$. Suppose n_{ij} cell counts results from two independent binomial samples over the four cells. A small sample null probability distribution for the cell counts that does not depend on any unknown parameter results from considering the set of tables having the same row and column totals as the observed data (Agresti, 2007).

3.1: A 2 × 2 contingency table

	Column 1	Column 2	Total
Row1	n_{11}	n_{12}	n_{1+}
Row2	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	n

The cell counts have the hypergeometric distribution once we condition on this restricted set of tables. For given row and column marginal totals, n_{11} determines the other three cell counts. The hypergeometric formula, therefore, expresses probabilities for the four cell counts in terms of n_{11} alone (Agresti, 2007). When $\theta = 1$, the probability of a particular value n_{11} is given as;

$$P(n_{11}) = \frac{\binom{n_{1+}}{n_{11}}\binom{n_{2+}}{n_{+1}-n_{11}}}{\binom{n}{n_{+1}}}$$

Where n_{i+} and n_{+j} are the i^{th} row and j^{th} column totals respectively.

To test H_0: homogeneity, the P-value is the sum of hypergeometric probabilities for outcomes at least as favourable to H_a as the observed outcome. Contingency tables having larger n_{11}also have larger sample odds ratios $\hat{\theta} = {}^{n_{11}n_{22}}/_{n_{12}n_{21}}$ providing stronger evidence in favour of the alternative hypothesis, $H_a: \theta > 1$. The P-value equals the right-tail hypergeometric probability that n_{11} is at least as large as the observed value (Agresti, 2007). Fisher's exact test for homogeneity will be used to compare proportions

of patients with pulmonary tuberculosis infection in 2×2 tables. The test shares the same assumptions as those noted for chi-square test for homogeneity, with the exception of the assumption regarding small expected frequencies.

3.6.4 Difference of proportions

For subjects in row 1 and row 2, denote success probabilities π_1 and π_2 respectively. The difference of proportions $\pi_1 - \pi_2$ compares the success probabilities in the two rows. The sample proportions of successes are denoted P_1 and P_2 respectively. The sample success proportions in row 1 and row 2 are given respectively as:

$$P_1 = \frac{n_{11}}{n_{1+}} \text{ and } P_2 = \frac{n_{21}}{n_{2+}}$$

The sample proportion difference $P_1 - P_2$ estimates $\pi_1 - \pi_2$. When the counts in the two rows are independent binomial samples, the estimated standard error of proportion difference $P_1 - P_2$ is:

$$se = \sqrt{\frac{P_1(1-P_1)}{n_{1+}} + \frac{P_2(1-P_2)}{n_{2+}}}$$

A large sample $100(1-\alpha)\%$ confidential interval for the proportion difference $\pi_1 - \pi_2$ is given as:

$$(P_1 - P_2) \pm Z_{\alpha/2}(se)$$

Proportion difference has been applied in determining the difference in pulmonary tuberculosis prevalence among divisions in North Pokot sub-County and between National PTB prevalence and North Pokot sub- County PTB prevalence.

3.6.5 Logistic regression

Logistic regression is widely used to model binary response variables. The primary goal of a logistic regression is to describe the effects of changes in a set of covariates X. It has proven to be one of the most versatile techniques in the class of generalized linear models (GLM) .The logistic regression model has been used extensively and successfully in health sciences to describe probability (or risk) of developing a condition (disease) over a specified time period as a function of certain risk factors (Chap, 2003). Since the response variable in this study is binary, logistic regression will be used for modeling predisposing factors for predicting pulmonary tuberculosis in North Pokot sub-County.

A generalized linear model for *Y* has the following three components:

a) **Random component** of a GLM identifies the response variable *Y* and selects a probability distribution for it.

b) **Systematic component** of a GLM specifies the explanatory variables in the formula:

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

c) **Link function** specifies a function g(.) that relates $E(Y) = \mu$ to the linear predictor as:

$$g(\mu) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k, \;\; \mu = E(Y)$$

The generalized linear models equate the linear component to some function of the probability of a given outcome. The link function $g(\mu) = \{\mu/1 - \mu\}$ models the log of the odds that some event will occur. This is called logit link and is appropriate when μ is

between 0 and 1, such as probability. A generalized linear model that uses the logit link is called a logistic regression model.

3.6.5.1 Multiple logistic regression

The effect of some factor on a response variable may be influenced by the presence of other factors through effect modifications (Chap, 2003). It is desirable to consider a large number of factors and sort out which ones are most closely related to the dependent variable (Chap, 2003). Therefore, to provide a more comprehensive analysis, multiple logistic regression is used for risk determination. Multiple logistic regression involves a linear combination of explanatory variables. A covariate or independent variable, such as patient characteristic, may be dichotomous, polytomous or continuous. Categorical variables are represented by dummy variables (Chap, 2003). Denote the k predictors for a binary response *Y* by $x_1, x_2, \dots, x_k$. The multiple logistic regression model has the form:

$$\pi(x) = \frac{1}{1 + exp\{-(\beta_0 + \sum_{j=1}^{k} \beta_j x_{ji})\}}, \qquad y_i = 0,1;\ i = 1, 2, \dots, n$$

For a binary response variable *Y*, $E(Y) = \pi(x)$ denotes the "success" probability at value *x* and is the parameter for binomial distribution (Agresti, 2007). The logistic regression model has a linear form:

$$log\left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ji}\,, \qquad y_i = 0,1;\ i = 1, 2, \dots, n$$

The parameters β_0 and β_j are the intercept and partial regression coefficients respectively. It is assumed that:

(i) The dependent variable *Y* is a binomial random variable

(ii) The independent variables *X* are not random (fixed)

(iii) The relationship between logit and *X* is linear

The likelihood function is of the form:

$$L = \prod_{i=1}^{n} \left\{ \frac{\exp\left(\beta_0 + \sum_{j=1}^{k} \beta_j x_{ji}\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{k} \beta_j x_{ji}\right)} \right\}^{yi} , y_i = 0,1$$

The parameters β_0 and $\beta_j (j = 1,2, \dots, k)$ are estimated iteratively using a computer program. The parameter β_j refers to the effect of a covariate on the log odds that $Y = 1$, controlling the other predictors. The exponential of the regression coefficients e^{β_j} are the associated adjusted odds ratios.

3.6.5.2 Inference for multiple logistic regression

A large sample Wald $100(1-\alpha)\%$ confidence interval for the parameter $\beta_j (j = 1,2, \dots, k)$ is given by:

$$\hat{\beta}_j \pm Z_{\alpha/2}\{se(\hat{\beta}_j)\}$$

The $100(1-\alpha)\%$ confidence interval for the odds ratio e^{β_j} is given as:

$$exp\left[\hat{\beta}_j \pm Z_{\alpha/2}\{se(\hat{\beta}_j)\}\right]$$

For logistic regression model, $H_0: \beta_j = 0$, states that the probability of success is independent of the covariate X_j under investigation. The Wald test statistic for large samples,

$Z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$ has a standard normal distribution when $\beta_j = 0$. Equivalently, for the two-sided alternative $H_a: \beta_j \neq 0$, Z^2 has a large sample chi-squared null distribution with 1 degree of freedom (Agresti, 2007).

The global null hypothesis is:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

The likelihood ratio test (LRT) can be used to test the global hypothesis. The likelihood ratio test statistic is given as:

$$\chi^2_{LR} = 2\left[logL\left(\hat{\beta}\right) - logL(0)\right]$$

$where\ L(\hat{\beta})$ is the maximum likelihood of a full model (model with all predictors included) and L(0) is the null model (model with the intercept only).

It has asymptotic chi − square distribution with k degrees of freedom under H_0. If H_0 is rejected, it suggests that taken collectively, the entire set of independent variables contribute significantly to the prediction of the response.

3.6.6.3 Model comparison

We now consider two logistic regression models, denoted by M_0(simpler model) and M_1 (complex model), such that M_0 is a special case of M_1. Given that the more complex model holds, the likelihood ratio statistic for testing that the simpler model holds is given as:

$$-2[L_0 - L_1] = -2[L_0 - L_s] - \{-2[L_1 - L_s]\}$$

$$= Deviance_0 - Deviance_1$$

Where L_0, L_1 and L_s are the maximized loglikelihood of simpler, complex and saturated models respectively.

We can compare models by comparing their deviances (Agresti, 2007). The statistic has an approximate chi-squared distribution, with degrees of freedom equal to the difference between residual degrees of freedom for the separate models. This degree of freedom equals to the number of additional parameters in M_1 but not M_0 (Agresti, 2007). Large values of test statistic and a small P-value suggest that M_0 fits poorly than M_1.

3.6.6 Goodness of fit

The test compares the model fit with the data. It regards the data as representing the fit of the saturated model, the most complex model possible. The saturated model has a separate parameter for each observation (Agresti, 2007). We test whether all the parameters that are in the saturated model but not in the model *M* equal zero to detect lack of fit. The likelihood ratio test statistic is the deviance of the model. The deviance statistic is given by:

$$Deviance = -2\{L_m - L_s\}$$

Where L_m and L_s are the maximized log likelihood of the simpler and saturated models respectively. This test statistic has a large sample chi-squared null distribution. The residual degree of freedom for the model is provided by subtracting the number of parameters in the model from the number of observations. A large value of test statistic and a small P-value provide evidence of model lack of fit.

3.7 Risk factors for tuberculosis which were investigated

Demographic and socio economic factors for predisposing tuberculosis were investigated. These factors include male and female gender. A gender which is highly susceptible to TB infection in the study area was investigated. The age group which is at a greater risk to TB infection was determined. In this study children of between 0-14 years and individuals from 15-80 years were included in the study.

Education level of the study sample was investigated to determine whether it is a predisposing to TB infection in the study area. This was stratified as those individuals with KCPE and below as well as those with KCSE and above.

Occupation as factor was investigated to determine whether TB was associated with formal employment or informal employment in the study area.

TB knowledge as to whether one had clear understanding of TB disease in terms of its causes, symptoms and how it spreads. Alongside TB history whether an individual had past record of TB treatment. These were investigated to whether they were predisposing factors to TB infection.

Housing conditions whether made of mud or cemented and ventilation by considering the number of airways were investigated as predisposing factors for TB infection. Congestion in a household was investigated to determine the number of individuals sharing a bedroom.

Alcoholism and smoking in the study sample were investigated as factors for predisposing TB infection. Nutrition as to whether an individual was taking number of

meals consumed per and to whether nutritionally balanced was also considered as a factor for investigation to TB infection in the study area.

Accessibility to a health facility as to whether one resided a distance of less than 3 KM or more than 3 KM. This would affect the time take one to seek medication. This factor was also investigated to determine as to whether it was a contributing factor to TB infection in the study area.

3.8 Ethical considerations

The authority to conduct the research was sought from the ministry of higher education through Kenyatta University ethics review committee and TB monitoring officer in Kacheliba Sub-county hospital. Before the questionnaires were administered, the objectives of the study and the planned use of the information were explained to the study participants.

The participants were informed of the purpose of the study. Since the study included participants of over and below 18 years, parents or guardians of those below 18 years consented for them. They were further informed that one would decide to participate in the study or decline. Those who declined were assured that they would get the same services of treatment within the hospital or elsewhere since there was no penalty for declining to participate in the study.

The participants were informed of the discomforts and risks of the nature of questions to be asked. They were told to make a decision of whether to answer the questions or not or they would decline the interview at any time.

The participants were informed of benefits of the study at individual level and at community level. They were further informed of the rewards on agreeing to participate in the study like free lunch and transport costs to and from home of which were provided.

The participants were assured of the confidentiality of the interview as it was conducted in a private setting within the clinic and the information collected in the questionnaires was kept private. The interviews included information on the socio-demographic information of the participants for instance age, sex, education level, occupation, ventilation, knowledge of TB transmission, signs and symptoms and appropriate data such as participant's case notification and definition entry, treatment phase and regimen was retrieved and recorded.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter comprises of results together with discussions. The results are presented in four sections. In section 4.1, the group demographics analyses the patient gender representation and age. Section 4.2 of this chapter presents chi-square and Fisher's exact test results comparing pulmonary tuberculosis prevalence between factor subgroups in $r \times c$ tables. A two – sample independent t-test result comparing mean ages of patients with and without PTB. A logistic regression model for prediction of PTB is also presented in this section. Section 4.3 and 4.4 presents chi-square and Fisher's exact test results comparing National and Pokot District PTB prevalence and Divisional prevalence rates respectively.

4.2 Group demographics

The total patient sample enrolled for the study was 300. The study sample consisted of 109 females and 191 males. Patients were aged between 6 months to 80 years with a median and mean age of 35 years and 35.85 years respectively. Both the male and female gender was considered in the study. The study sample included the formal employed and informal employed people. The study sample also consisted of individuals of different education level; those with KCPE and below and those with KCSE and above.

4.3 Factors for predicting pulmonary tuberculosis

There are demographic factors that were found out to contribute to pulmonary tuberculosis (PTB) prevalence in North Pokot Sub-county. The relationship between PTB prevalence and predisposing factors are investigated in this section.

4.3.1 Pulmonary tuberculosis prevalence and gender

The pulmonary tuberculosis prevalence rate was 69.63% and 65.14% among males and females respectively in North Pokot Sub-county (Table 4.1).

Table 4.1: Pulmonary tuberculosis prevalence by gender

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Gender	Males	133	58	191	69.63
	Females	71	38	109	65.14
Total		204	96	300	

To determine if the male and female PTB were significantly different, the hypothesis tested was:

H_0: $\pi_1 = \pi_2$ against H_1: $\pi_1 \neq \pi_2$

The result is presented in table 4.2

Table 4.2: Fisher's exact test result comparing male and female PTB prevalence

	Odds ratio	Proportion difference	P
	1.227	0.04496	0.249
95% CI	0.7441 - 2.024	-0.06482 - 0.1547	-

The null hypothesis was not rejected ($P = 0.249, OR = 1.227$) and it was concluded that male PTB prevalence was the same as that of the female PTB prevalence at $\alpha/2$=0.025 level of significance. However, odds ratio showed that PTB was 1.227 times more likely among males than females (Table 4.2). This is consistent with the study findings which indicated a higher prevalence of pulmonary TB in men than women in rural Western Kenya. (Crude OR, 1.5; 95% CI: 1.1-2.2) (Laserson *et al.*, 2011) and the WHO report that gender differences begin to appear between 10 and 16 years of age, with men reporting much higher rates of tuberculosis than women. The WHO recommends further research between men and women to understand these important differences. (WHO, 2002).

4.3.2 Pulmonary tuberculosis prevalence and smoking

Pulmonary tuberculosis prevalence rate was 78.10% and 62.56% among smokers and non-smokers respectively (Table 4.3).

Table 4.3: Pulmonary tuberculosis prevalence among smokers and non-smokers

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Smoker	Yes	82	23	105	78.10
	No	122	73	195	62.56
Total		204	96	300	

The hypothesis tested to determine whether PTB prevalence rate among smokers and non-smokers was significantly different is:

$H_0 : \pi_1 = \pi_2$ against $\pi_1 \neq \pi_2$

Where π_1 and π_2 are the PTB prevalence among smokers and non-smokers respectively.

The result is presented in table 4.4

Table 4.4: Fisher's exact test result comparing PTB prevalence among smokers and non-smokers

	Odds ratio	Proportion difference	P
	2.133	0.1553	0.004
95% CI	1.236 - 3.683	0.04461 - 0.2660	-

The null hypothesis was rejected $(\mathrm{P} = 0.004, \mathrm{OR} = 2.133)$ and it was concluded that smokers had a greater PTB prevalence than non-smokers at 5% level of significance. PTB was 2.133 times more likely among smokers than non-smokers (Table 4.4). This result was consistent with a study which was carried out in India which revealed a positive relationship between smoking of tobacco with pulmonary tuberculosis disease (OR 2.5) (Kolappan *et al.,* 2002) and another study done in Razi hospital, Ahuaz a city in South West Iran which indicated an association between pulmonary tuberculosis and cigarette smoking (OR 10.1; 95%, CI: 4.3-23.5), P<0.001 (Alavi *et al.,* 2009). It is biologically acceptable that tobacco smoking increases risks to pulmonary tuberculosis infection and disease. Mechanisms which have been suggested towards this increased risks include declined immune response and mechanical interruption of cilia performance in the pathways (Bates *et al.,* 2007).

4.3.3 Pulmonary tuberculosis prevalence and alcoholism

Alcoholic and non-alcoholic patients had PTB prevalence of 74.73% and 57.63% respectively (Table 4.5).

Table 4.5: Pulmonary tuberculosis prevalence among alcoholic and non-alcoholic patients

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Alcoholic	Yes	136	46	182	74.73
	No	68	50	118	57.63
Total		204	96	300	

The following hypothesis was tested to determine if there was significant difference in PTB prevalence between alcoholic and non-alcoholic patients:

$$H_0: \pi_1 = \pi_2 \text{ against } H_1: \pi_1 \neq \pi_2$$

Where π_1 and π_2 are the PTB prevalence among alcoholic and non-alcoholic patients respectively.

The result is presented in table 4.6.

Table 4.6: Fisher's exact test result comparing PTB prevalence between alcoholic and non-alcoholic patients

	Odds ratio	Proportion difference	P
	2.174	0.1710	0.002
95% CI	1.325 - 3.567	0.06290 - 0.2791	

The result indicated that alcoholic patients had greater PTB prevalence than non-alcoholic patients $(P = 0.002, OR = 2.174)$ at 5% level of significance hence null hypothesis was rejected. PTB was 2.174 times more likely among alcoholic than non-alcoholic patients (Table 4.6). Studies have shown a strong association between heavy alcohol use and tuberculosis disease. A high level of analysis has been carried out to find out the risk of tuberculosis on alcohol consumption. .A relative risk of 2.94 (95%, CI: 1.89-4.59) was obtained. Several studies have shown a great effect of alcohol on an individual's immune system which cause vulnerability to tuberculosis in heavy alcohol drinkers. In the whole world the available data indicates that nearly 10% of tuberculosis cases are appraised to be associated to alcohol use (Christina *et al.,* 2009).

4.3.4 Pulmonary tuberculosis prevalence and daily nutritionally balanced meals

The patients were grouped into those who afforded only one, two and three nutritionally balanced meals a day. Of those who had one balanced meal, two and three balanced meals a day the PTB prevalence rate was 81.48%, 68.86% and 63.21% respectively (Table 4.7).

Table 4.7: Pulmonary tuberculosis prevalence and daily nutritionally balanced meals

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Balanced meals per day	1	22	5	27	81.48
	2	115	52	167	68.86
	3	67	39	106	63.21
Total		204	96	300	68.00

The following hypothesis was tested to determine if there was a significant difference in PTB prevalence between patients who had one balanced meal, two and three balanced meals daily:

$$H_0: \pi_1 = \pi_2 = \pi_3 \text{ against } H_1: \pi_1 \neq \pi_2 \neq \pi_3$$

Where π_1, π_2 and π_3 are the PTB prevalence among patients who had only one nutritionally balanced meal, two and three nutritionally balanced meals in a day respectively.

A Chi square test result demonstrated that there was no difference in PTB prevalence between patients who had one balanced meal, two and three balanced meals daily$\left(\chi^{2}_{0.05,2} = 3.431,\ \mathrm{P} = 0.1799\right)$ at 5% level of significance hence null hypothesis

was not rejected. It is recorded that lack of adequate nourishment increases the risk of tuberculosis (Krishna *et al.,* 2009).

4.3.5 Pulmonary tuberculosis prevalence and level of education

The patients were stratified into those with none formal education, primary, secondary and tertiary education. PTB prevalence for each education level is presented in table 4.8.

Table 4.8: Pulmonary tuberculosis prevalence and level of education

		None	**Primary**	**Secondary**	**Tertiary**
PTB infection	Yes	66	90	41	6
	No	22	43	24	7
Total		88	133	65	13
Positive %		75.00	67.67	63.08	46.15

The hypothesis tested to determine whether PTB prevalence was homogeneous across all levels of education is:

$$H_0: \pi_1 = \pi_2 = \cdots = \pi_4 \text{ against } H_1: \pi_1 \neq \pi_2 \neq \cdots \neq \pi_4$$

Where $\pi_1, \pi_2, \ldots, \pi_4$ are the PTB prevalence among patients with none formal education, primary, secondary, tertiary education respectively.

A Chi square test result indicated that there was no difference in PTB prevalence between patients with none formal education, primary, secondary and tertiary

education $\left(\chi^{2}_{0.05,\ 3} = 5.52,\ \mathrm{P} = 0.1356\right)$ at 5% level of significance hence we failed to reject null hypothesis. This showed that PTB prevalence is not determined by the level of education. This is not consistent with a study finding that education level is an important determinant of people's level of knowledge of TB, and persons with a higher level of education scored better on TB knowledge indices than those with less education or those who were illiterate (Mondal *et al*., 2014).

4.3.6 Pulmonary tuberculosis prevalence and infection history

The patients with and without history of PTB infection had a PTB prevalence of 62.16% and 73.68% respectively (Table 4.9).

Table 4.9: Pulmonary tuberculosis among patients with and without infection history

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Infection history	Yes	92	56	148	62.16
	No	112	40	152	73.68
Total		204	96	300	

PTB prevalence rate was 11.52% higher among patients without infection history and to determine if this was significant, the hypothesis tested is:

$$H_0: \pi_1 = \pi_2 \text{ against } H_1: \pi_1 \neq \pi_2$$

Where π_1 and π_2 are the PTB prevalence rate among patient with and without infection history respectively.

The result is presented in table 4.10.

Table 4.10: Fisher's exact test result comparing PTB prevalence between patients with and without infection history

	Odds ratio	Proportion difference	P
	0.5867	0.1152	0.022
95% CI	0.3592 - 0.9584	0.009613 - 0.2208	-

The results showed that patients with infection history had less PTB prevalence than those without infection history($P = 0.022, \mathrm{OR} = 0.5867$) at 5% level of significance hence null hypothesis was rejected. PTB was 0.5867 times less likely among patients with infection history than those without infection history (Table 4.10). The results are not consistent with a study carried out in Metropolitan City of South India which indicated that occurrence of pulmonary TB disease from both culture and sputum positive results showed a significant relationship with past history of tuberculosis treatment (Hussein *et al*., 2012).

4.3.8 Pulmonary tuberculosis prevalence and accessibility to health facility

The patients staying less than three kilometers from the health facility were grouped as having accessibility while those staying more than three kilometers from the health

facility as lacking accessibility. Of the patients with accessibility, PTB prevalence was 56.32% compared to 72.77% without accessibility (Table 4.11).

Table 4.11: Pulmonary tuberculosis prevalence by hospital accessibility

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Accessibility	Yes	49	38	87	56.32
	No	155	58	213	72.77
Total		204	96	300	68.00

The following hypothesis was tested to determine whether PTB prevalence among patients residing far (more than three kilometers) and near (less than three kilometers) from the health facility were significantly different:

$H_0: \pi_1 = \pi_2$ against $H_1: \pi_1 \neq \pi_2$ Where π_1 and π_2 are the PTB prevalence among patients with and without accessibility respectively.

The result is presented in table 4.12.

Table 4.12: Fisher's exact test result comparing PTB prevalence between patients with and without accessibility

	Odds ratio	Proportion difference	P
	0.825	0.1645	0.005
95% CI	0.2868 - 0.8118	0.04812 - 0.2808	-

The results showed that PTB prevalence among patients with accessibility was less than that of those without accessibility to health facility($P = 0.005, \mathrm{OR} = 0.825$) at 5% level of significance hence null hypothesis was rejected. PTB was 0.825 times less likely among patients with accessibility to health facility than those without accessibility (Table 4.13). The results were consistent with those of a study conducted in Western Kenya which showed the highest PTB prevalence among patients residing a distance of more than two Kilometers from the health Centre offering treatment for tuberculosis, when comparing with individuals residing nearer to the TB health Centre (Anna *et al.*, 2011).

4.3.9 Pulmonary tuberculosis prevalence and occupation

The patients who were in formal employment, informal employment and the unemployed had PTB prevalence rate of 75.83%, 61.29% and 63.53%respectively (Table 4.13).

Table 4.13: Pulmonary tuberculosis prevalence and patient occupation

		Unemployment	**Formal employment**	**Informal employment**
PTB infection	Yes	54	91	57
	No	31	29	36
Total		85	120	93
Prevalence %		63.53	75.83	61.29

The following hypothesis was tested to determine if the PTB prevalence between the occupation categories was different:

$H_0: \pi_1 = \pi_2 = \pi_3$ against $H_1: \pi_1 \neq \pi_2 \neq \pi_3$ Where π_1, π_2 and π_3 are the PTB prevalence rates among patients in formal, informal employment and unemployment respectively.

A Chi square test result indicated that PTB prevalence was a significance difference among patients in formal employment, informal employment and unemployed $\left(\chi^2_{0.05,2} = 6.061,\ \mathrm{P} = 0.0483\right)$ at 5% level of significance hence null hypothesis was rejected.

4.3.10. Pulmonary tuberculosis prevalence and residential ventilation

Patients residing in good and poorly ventilated houses had 16.67% and 6.37% PTB prevalence rate respectively (Table 4.14).

Table 4.14: Pulmonary tuberculosis prevalence by residential ventilation

Tuberculosis Infection					
		Positive	**Negative**	**Total**	**Positive%**
Ventilation	Good	13	16	29	44.83
	Poor	191	80	271	70.48
Total		204	96	300	

The hypothesis tested to determine whether the PTB prevalence among patients residing in good and poorly ventilated houses was significantly different is:

$$H_0: \pi_1 = \pi_2 \text{ against } H_1: \pi_1 \neq \pi_2$$

Where π_1 and π_2 are the PTB prevalence rate among patients residing in good and poorly ventilated houses respectively.

The result is presented in table 4.15.

Table 4.15: Fisher's exact test result comparing PTB prevalence between patients residing in good and poorly ventilated houses

	Odds ratio	Proportion difference	P
	0.3403	0.2565	0.008
95% CI	0.1564 – 0.7403	0.0779 - 0.4352	-

The results showed that PTB was less prevalent among patients residing in houses with good ventilation than in those with poor ventilation($P = 0.008, \text{OR} = 0.3403$) at 5% level of significance hence null hypothesis was rejected (Table 4.16). PTB was 0.3403 times less likely among residence of houses with good ventilation than those with poor ventilation. Housing conditions clearly indicate the socio–economic status of people. Some conditions of house do contribute greatly poor respiratory health among the residents which as well to results to development and spread of tuberculosis disease. These conditions include inadequate ventilation, overcrowding, presence of smoke and poor quality of the house such made of mud which produces dust (Wanyeki *et al*., 2006). In all studies addressing housing conditions as contributing risk factors to tuberculosis infection none focused ventilation in private houses. However, sharing a home with an individual infected with tuberculosis who has started treatment has been believed that it exposes one to TB infection. It is advised that houses which have common ventilators should not be used by TB cases who are infectious (Health Canada, 2005).

4.3.11. Pulmonary tuberculosis prevalence and residential house

The prevalence of PTB among patients residing in cemented and mud walled houses was 55.10% and 70.52% respectively (Table 4.16).

Table 4.16: Pulmonary tuberculosis prevalence by type of residential house

Tuberculosis infection					
		TB	**No TB**	**Total**	**Positive%**
House	**Cemented**	27	22	49	55.10
	Mud	177	74	251	70.52
Total		204	96	300	

The following hypothesis was tested to determine whether the PTB prevalence rate among patients residing in cemented and mud walled houses was significantly different:

$H_0: \pi_1 = \pi_2$ against $H_1: \pi_1 \neq \pi_2$. Where π_1 and π_2 are the PTB prevalence among patients residing in cemented and mud walled houses respectively.

The result is presented in table 4.17.

Table 4.17: Fisher's exact test result comparing PTB prevalence between patients residing in cemented and mud walled houses

	Odds ratio	Proportion difference	P
	0.5131	0.1542	0.027
95% CI	0.2746 – 0.9586	0.0113 - 0.2970	-

The results showed that PTB was less prevalent among patients residing in cemented than in mud walled houses$(P = 0.027, \text{OR} = 0.5131)$ at 5% level of significance hence null hypothesis was rejected (Table 4.17). PTB was 0.5131 times less likely among residence of cemented than mud walled houses. The results are consistent with the study which indicated that houses built with mud increased the risk of PTB infection than that were built with cement (Lienhardt *et al*., 2005).

4.3.12. Pulmonary tuberculosis prevalence and infection knowledge

The PTB prevalence among patients with and without TB knowledge was 57.14% and 71.75% respectively (Table 4.18).

Table 4.18: Pulmonary tuberculosis prevalence and patient knowledge

Tuberculosis infection					
		Yes	**No**	**Total**	**Positive%**
Knowledge	**Yes**	44	33	77	57.14
	No	160	63	223	71.75
Total		204	96	300	

The following hypothesis was tested to determine whether PTB prevalence among patients with and without knowledge was significantly different:

$H_0: \pi_1 = \pi_2$ against $H_1: \pi_1 \neq \pi_2$. Where π_1 and π_2 are the PTB prevalence among patients with and without infection knowledge respectively. The result is presented in table 4.19.

Table 4.19: Fisher's exact test result comparing PTB prevalence between patients with and without infection knowledge

	Odds ratio	Proportion difference	P
	0.5250	0.1461	0.014
95% CI	0.3067 – 0.8986	0.0252 - 0.2669	-

The results showed that PTB prevalence among patients with knowledge of the infection was less than that of those without infection knowledge($P = 0.014, \mathrm{OR} = 0.5250$) at 5% level of significance hence null hypothesis was rejected. PTB was less 0.525 times less likely among with knowledge than those without infection knowledge (Table 4.19).

4.3.13. Pulmonary tuberculosis prevalence and house congestion

The PTB prevalence among patients residing in congested and non-congested houses was 73.37% and 57.43% respectively (Table 4.20).

Table 4.20: Pulmonary tuberculosis prevalence and house congestion

Tuberculosis infection		**Yes**	**No**	**Total**	**Positive%**
Congestion	**Yes**	146	53	199	73.37
	No	58	43	101	57.43
Total		204	96	300	

The hypothesis tested to determine whether the PTB prevalence among patients residing in congested and non-congested houses were significantly different is given as:

$H_0: \pi_1 = \pi_2$ against $H_1: \pi_1 \neq \pi_2$. Where π_1 and π_2 are the PTB prevalence among patients residing in congested and non-congested houses respectively.

The result is presented in table 4.21.

Table 4.21: Fisher's exact test result comparing PTB prevalence between patients residing in congested and non-congested houses

	Odds ratio	Proportion difference	P
	2.042	0.1594	0.004
95% CI	1.233 – 3.382	0.0477 - 0.2711	-

The results showed that PTB prevalence among patients residing in congested houses was greater than that of those in non-congested houses$(P = 0.004, \mathrm{OR} = 2.042)$ at 5% level of significance hence null hypothesis was rejected. PTB was more 2.042 times more likely among patients residing in congested houses than in non-congested houses (Table 4.21). Crowding is one of the risk factor contributing to TB transmission. This is enhanced when there is poor movement of air in a crowded and enclosed area thus exposing people to *Mycobacterium tuberculosis*. The risk of exposure is also increased if there is limited air movement in an enclosed space. Bedroom is used to measure crowding in houses. This is according to the National Occupancy Standard. It states that a

house should have enough bedrooms for all households. There should be one bedroom for each couple, one for each adult over 18 years, one for a pair of children of the same sex and for children below five years a pair of opposite sexes should share a bedroom (Wanyeki *et al.*, 2006).

4.3.14 Logistic regression model of pulmonary tuberculosis

A logistic regression model of pulmonary tuberculosis prevalence was fitted with prediction factors as the covariates. The following hypothesis was tested to investigate the relationship between pulmonary tuberculosis prevalence and the prediction factors:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_{12} = 0$ against $H_1: \beta_j \neq 0$, for at least one j, j=1, 2,… 12

The result is presented in table 4.22.

Table 4.22: Result of logistic regression model of PTB prevalence - I

	$\widehat{\beta}_j$	Se($\widehat{\beta}_j$)	$e^{\widehat{\beta}_j}$	95% CI for $e^{\widehat{\beta}_j}$	Z	P
Intercept	0.1831	0.5142	1.2009	0.4383 - 3.2901	0.356	0.7217
Sex	-0.1382	0.3060	0.8709	0.4781 - 1.5865	-0.452	0.6515
Education	0.1167	0.2966	1.1237	0.6284 - 2.0098	0.393	0.6940
Occupation	-0.1396	0.3750	0.8697	0.4170 - 1.8137	-0.372	0.7098
Knowledge	-0.3248	0.3324	0.7227	0.3767 - 1.3864	-0.977	0.3285
History	-0.2007	0.3115	0.8182	0.4443 - 1.5066	-0.644	0.5194
House	0.1064	0.4621	1.1123	0.4496 - 2.7514	0.230	0.8180
Alcoholism	0.7484	0.3656	2.1136	1.0323 - 4.3274	2.047	0.0406
Meals	0.3208	0.3313	1.3782	0.7200 - 2.6383	0.968	0.3329
Accessibility	-0.5714	0.3080	0.5647	0.3088 - 1.0328	-1.855	0.0635
Smoking	0.5259	0.3186	1.6920	0.9062 - 3.1593	1.650	0.0989
Ventilations	-0.2115	0.3439	0.8094	0.4125 - 1.5881	-0.615	0.5384
Congestion	0.6490	0.3282	1.9136	1.0058 - 3.6410	1.978	0.0480
Null deviance:374.57 on 297 DF; Residual deviance: 345.02on 285 DF						

The result showed that at least one of the partial regression coefficients is not zero hence null hypothesis was rejected and it was concluded that the overall model is significant (deviance difference = 29.55, DF: 12; P = 0.00326). Pulmonary tuberculosis was more likely among alcoholic patients($P = 0.0406, OR = 2.1136$)and those residing in congested house($P = 0.048, OR = 1.9136$). Although marginally significant, PTB was more likely among smokers and less likely in patients with accessibility to the hospital. PTB was independent of sex of the patients, education, occupation, knowledge of PTB, infection history, type of residential house, balanced meals per day and ventilation of the house (Table 4.22).

Interaction terms were included and after backward elimination of insignificant covariates (P>0.05), a new model was fitted. The result is presented in table 4.23.

Table 4.23: Result of logistic regression model of PTB prevalence - II

	$\widehat{\beta}_j$	Se($\widehat{\beta}_j$)	$e^{\widehat{\beta}_j}$	95% CI for $e^{\widehat{\beta}_j}$	Z	P
Intercept	-0.5013	0.3634	-	-	-1.380	0.16774
Alcoholism	1.5131	0.4573	4.5408	1.8530 - 11.127	3.309	0.00093
Accessibility	-0.6527	0.2850	0.5206	0.2978 - 0.9102	-2.290	0.02200
Smoking	0.5620	0.3052	1.7542	0.9645 - 3.1905	1.841	0.06558
Congestion	1.4018	0.4257	4.0625	1.7638 - 9.3573	3.293	0.00099
Alcoholism: Congestion	-1.3789	0.5595	0.2519	0.0841 - 0.7540	2.464	0.01372
Null deviance:374.57 on 297 DF; Residual deviance:342.24 on 292 DF						

The result showed that the overall model is significant (deviance difference = 32.16, DF: 5; P<0.0001). It was now demonstrated in table 4.23 that pulmonary tuberculosis was less likely among patients having accessibility to the health facility and alcoholics given congestion in their houses of residence. Although PTB was more likely among smokers, it was marginally significant (Table 4.23).

The logistic regression model of pulmonary tuberculosis prevalence is given by:

$$\log\left[\frac{\hat{\pi}(x)}{1-\hat{\pi}(x)}\right] = -0.5013 + 1.5131x_1 - 0.6527x_2 + 0.5620x_3 + 1.4018x_4 - 1.3789x_1x_4$$

Where, $x_i(i = 1,2,\dots,4)$ are alcoholism, accessibility, smoking and congestion respectively.

4.4 North Pokot Sub-county pulmonary tuberculosis prevalence

The prevalence of PTB was compared to national (Kenya) prevalence and across age groups in this section.

4.4.1 North Pokot Sub-county pulmonary tuberculosis prevalence and age group

Patients were stratified by age and PTB prevalence of each age group is presented in table 4.24.

Table 4.24: Pulmonary tuberculosis prevalence by age group

Age (in years)									
		0 -14	15 – 24	25 - 34	35 – 44	45 – 54	55 - 64	65 – 80	**Total**
Tuberculosis Infection	Yes	14	32	44	46	41	11	16	204
	No	12	17	26	23	10	6	2	96
Total		26	49	70	69	51	17	18	300
Prevalence %		53.85	65.31	62.86	66.67	80.39	64.71	88.89	68.00

The hypothesis tested to determine whether PTB prevalence between age group were different is given as:

$H_0: \pi_1 = \pi_2 = \cdots = \pi_7$ against $H_1: \pi_1 \neq \pi_2 \neq \cdots \neq \pi_7$ Where $\pi_1, \pi_2, \ldots, \pi_7$ are the PTB prevalence among patients aged between 0 – 14, 15 – 24, 25 – 34, 35 – 44, 45 – 54, 55 – 64 and 65 – 80 respectively.

A Chi square test result indicated that there was no difference in PTB prevalence between age groups of the patients$\left(\chi^2_{0.05,6} = 10.76,\ \mathrm{P} = 0.0962\right)$ at 5% level of significance hence we failed to reject null hypothesis. Several studies investigating the relevance of age and sex to the prevalence of TB observed that the prevalence rises with age in both sexes (Kadri *et al*., 2003).Another study indicated that majority of PTB patients were found among age groups of 15-24. In addition smear positive pulmonary tuberculosis prevalence distribution is significantly different among classes of age groups (Hussein *et*

al., 2012).A study which was carried out in rural Western Kenya indicated well that PTB prevalence was not the same in different age as well as in the sexes. PTB prevalence in women was highest in the age groups of 25-34 years, while in men, it was highest in the age group of 35-54 year. The obtained results in this contrasts the study conducted in Western province and Nakuru provincial general hospital where a total of 872 clients participated in the study. The mean age was 35 years while the median age was 32 years. Majority were in the 25-34 age group, then those in the 35-44 age group and 15-24 brackets respectively. The results agree with the age bracket 0-14 years which in both had the lowest PTB prevalence rate (Anna *et al.,* 2011).

4.4.2 North Pokot Sub-county and National pulmonary tuberculosis prevalence

In 2014 the total national PTB cases notified was 72385 in a population of 38.9 million (W.H.O 2014) Kacheliba Sub-county Hospital with a catchment patient population of 23679 in the same period had 204 confirmed PTB cases. The national and North Pokot Sub-county prevalence was 2 (per 1000) and 9 (per 1000) respectively (Table 4.25).

Table 4.25: District and national pulmonary tuberculosis prevalence

Tuberculosis infection					
		Yes	**No**	**Total**	**Prevalence (per 1000)**
Region	**District**	204	23475	23679	9
	National	72385	38827615	38900000	2

The hypothesis tested to determine whether North Pokot Sub-county and national PTB prevalence were significantly different is given as:

$$H_0: \pi_1 = \pi_2 \text{ against } H_1: \pi_1 \neq \pi_2$$

Where π_1 and π_2 are the North Pokot Sub-county and national PTB prevalence respectively.

The result is presented in table 4.26.

Table 4.26: A Chi square test result comparing Pokot District and national PTB prevalence

	Odds ratio	Proportion difference	$\chi^2_{0.05,1}$	P
	4.6614	0.0068	576.3803	< 2.2e-16
95% CI	4.6163 - 4.7069	0.0056 - 0.0080	-	-

The results showed that there was a difference between North Pokot Sub-county and national PTB prevalence $(P < 0.0001, \text{OR} = 4.6614)$ at 5% level of significance hence null hypothesis was rejected. PTB was 4.6614 times more likely in North Pokot Sub-county than nationally (Table 4.26). This results shows a higher PTB prevalence in the study area comparing to other regions and in the country. For instance in a study estimating the prevalence of tuberculosis in rural Western Kenya in May 2011 indicates that out of 20566 participants, 123 had pulmonary tuberculosis. Tuberculosis prevalence was 6.0 per 1000 people for all PTB and 2.5 per 1000 people for smear positive

pulmonary tuberculosis (Laserson *et al*, 2011)Notified tuberculosis in three high burden provinces of Kenya indicates an even decrease for the past six years. The Great rift region consistently led in the absolute numbers that were reported to the national program followed by Nairobi and Nyanza regions. This is based on the fact that the Great rift region is a vast region as this order is reversed when case notification rates are calculated and Nairobi leads followed by Nyanza and Great rift region respectively. This is expected because Nairobi is basically an urban setting with a huge population living below poverty line. (Sitienei *et al*., 2013)

4.5 Divisional pulmonary tuberculosis prevalence.

The pulmonary tuberculosis prevalence in 5 Divisions in North Pokot Sub-county is presented in table 4.27.

Table 4.27: Divisional pulmonary tuberculosis prevalence

Division						
		Alale	Kacheliba	Kasei	Kongelai	Konyao
Tuberculosis Infection	Yes	26	98	26	32	22
	No	16	41	15	12	12
Total		42	139	41	44	34
Prevalence %		61.90	70.50	63.41	72.73	64.71

The hypothesis tested to determine whether PTB prevalence between the Divisions were significantly different is given as:

$$H_0: \pi_1 = \pi_2 = \cdots = \pi_5 \text{ against } H_1: \pi_1 \neq \pi_2 \neq \cdots \neq \pi_5$$

Where $\pi_1, \pi_2, \ldots, \pi_5$ are the PTB prevalence among patients from Alale, Kacheliba, Kasei, Kongelai and Konyao Divisions respectively.

A Chi square test result indicated that there was no difference between Divisional PTB prevalence $\left(\chi^2_{0.05,4} = 2.135,\ \mathrm{P} = 0.7109\right)$ at 5% level of significance hence we failed to reject null hypothesis. This is a clear indication that the predisposing risk factors to *Pulmonary tuberculosis* are similar in all the divisions and each division contribute equally to *PTB* prevalence in North Pokot sub County.

CHAPTER FIVE

CONLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter contains three areas. Area 5.1 consists of the summary of the study, area 5.2 is comprised of conclusions of study objectives and section 5.3 consists of the recommendations.

5.2 Summary

In this thesis the significant predisposing risk factors for predicting pulmonary tuberculosis in North Pokot Sub County were investigated using appropriate statistical tools such PTB prevalence in North Pokot Sub-county was also investigated in this thesis work and compared with the national PTB prevalence. Lastly the comparison of PTB prevalence of the divisions within North Pokot Sub-county was investigated.

5.3 Conclusions

The research has demonstrated that there is a higher PTB prevalence in a resource poor community; North Pokot Sub County compared to that of the National PTB prevalence. This has contributed greatly to TB burden in a developing country Kenya.

The statistically significant factors which predisposes individuals to TB infection in the area under study were found to both host related and environmental as recorded by (Rieder, 1999). According to this study smoking, alcoholism, congestion, ventilation, accessibility, occupation and TB knowledge were the significant risk factors predisposing individuals to PTB infection. When all the factors were investigated together using a logistic regression model to determine their relationship with PTB infection, congestion,

alcoholism and smoking were found out to be the predicting factors for TB in North Pokot Sub County, West Pokot County. Surprisingly the male gender was highly susceptible to TB infection than the female counterpart. This would be attributed to the fact that the male gender is out in the field looking after the cattle, therefore may miss meals and do not observe proper medication. The male also engage more in alcohol consumption and smoking which predispose them to YB infection. 25-34 and 35-44 age brackets were found to be at a higher risk to TB infection.

The study has clearly indicated that PTB prevalence among the divisions in the region under study was not significantly different. Thus they contribute equally to the PTB prevalence. In this case same measures should be put in place to prevent PTB infection and development of the disease.

This research is timely significant has it is a way to attain vision 2030 of eliminating TB. The information that has been obtained will be of great help to the country in planning attainable strategies of controlling TB in a resource poor community.

5.4.1 Recommendations from the study

From the research findings there is higher PTB prevalence in North Pokot sub County, West Pokot County thus contributing greatly to the country's TB burden. Proper measures should be put in place to control the determined predisposing risk factors predicting tuberculosis in this study so as to prevent tuberculosis infection and disease. The study therefore recommends the following:

1. The risk age group of 15-44 years to TB infection in this study need to be informed on tuberculosis knowledge symptoms, how it spreads, diagnosis and treatment by the education programs provided by the government.
2. Accessibility as a contributing factor to TB infection should be addressed in that the communities that are remote from the main hospitals would benefit if the government would construct peripheral health infrastructures like health centers and dispensaries and be equipped with microscopy services for tuberculosis and enough trained personnel who can supervise and motivate the peripheral staff, particularly the community health workers, this would help solve the problem of accessibility as a contributing factor to TB infection.
3. Alcoholism and smoking being the predisposing factors to TB infection in this study should be addressed by the government to put measures that will reduce smoking and alcohol drinking.
4. Housing conditions such congestion, ventilation should be addressed by the government funding the community to aid in the construction of enough houses which are well ventilated to prevent the spread of TB.

5.4.2 Recommendations for further study

Although extensive research has been studied in this thesis, a number of gaps are noted in the literature review, therefore the study recommends further research on these areas:

1. Extent and persistence of contact with an infected person in the household as a factor for TB transmission.
2. Contribution of HIV to TB infection.
3. Prevalence of PTB in children with and without BCG vaccination.

4. Prevalence of MDR-TB and Extensive Drug Resistance TB.
5. The prevalence of extra pulmonary tuberculosis.

REFERENCES

Abdikadir, M.M. (2004). Prevalence of tuberculosis in Central division Wajir district, North Eastern Province, Kenya; Unpublished Master's Thesis in Public Health Kenyatta University.

Agresti, A. (2007). *An Introduction to Categorical Data Analysis,* Second Edition : A John Wiley & Sons, Inc.

Arch, G. and Mainous (2010). Management of Antimicrobials in infectious Diseases: Impact of Antibiotic Resistance Humeral.

Ahmed, N. and Hasman, S. (2011). "Molecular epidemiology of tuberculosis in India: moving forward with a system biology approach." *Tuberculosis Journal;* **91**(5):407-413.

Alavi, S.M. and Ershadian, S. (2009). "Association between cigarette smoking and pulmonary tuberculosis". *Pakistan Journal of Medical Sciences*; **25,** 6, 912-915.

Anna, H.V., Kayla F.L., Willie, A.G., Helen, K., M., Janet, A.A., Lazurus, O.O., Benson, G.M., Barbara, J.M., Kevin, M.D. and Martien, W.B. (2011). "High Prevalence of Tuberculosis and Inadequate Case Finding in Rural Western Kenya". *American Journal of Respiratory and Critical Care Medicine*; **183,** 9, 1245-1253.

Aver, C. Sarrol, J.J. Tanner, M. and Weiss, M. (2000). Health seeking and perceived causes of tuberculosis among patients in Manila; *Journal Medicine and Health*; **5**: 648-656.

Barner, P.F. (2001). Diagnosing latent tuberculosis infection: the 100-year upgrade. *American Journal of Respiratory and Critical Care Medicine*; **163**:807-808.

Bates, M.N., Asheena, K., Madhukar, P., Lisa, C., Ferneda, L. and Kirk, R.S. (2007). Risk of Tuberculosis from Exposure to Tobacco smoke: *Archives of International Medical Journal*; **167**(4); 335- 342.Doi: 10.1001/archninte.167.4. 335 .

Chap, L. (2003). *Introductory Biostatistics:* A John Wiley & Sons Inc.

Centre of Disease Control (2000). "Targeted tuberculin testing and treatment of latent tuberculosis infection American Thoracic Society". *Recommendations and Reports* **49** (RR-6):1-51.

Centre of Disease Control (2011). "Core Curriculum on Tuberculosis: What the Clinician should know".*Division of Tuberculosis Elimination;* GA 30329-4027 USA.

Christina, T.F., Carol D.H. and Jason, E.S. (2009). Alcohol use and clinical manifestation of tuberculosis. *Journal of Tuberculosis Infection*; **58**(5): 395-401.

Dolin, Gerald L. Mandell, John E. Bnnett, Raphael (2010) "Principles and practice of infectious diseases" ISBN 978-0-443-06839-3.

Dye, C., Sheele, S., Dolin, P. Pathania, V. and Raviglione, M.K. (1999). Global burden of tuberculosis estimated incidence, prevalence and mortality by country. *Journal of American Medical Association;* **282**:677-683.

Ekwaro, A. O., Clea, M., Jemimmah, K., Simm, B., Christine, A., Everlyn, N., Morris, L., Nikki, J., David, N. (2012). "Socio demographic determinants and prevalence of Tuberculosis Knowledge in three slum populations of Uganda". *BioMed Central Public Health Journal; **12***:536 doi: 10.11861477-2458-12-536.

Fennelly, K.P. Martyny, J.W. and Fulton, K.F. (2006). Cough generated aerosols of Mycobacterium tuberculosis in confined spaces: A new method to study infectiousness:*American Journal of Respiratory and Critical Care Medicine*: **169**:60-49.

Health Canada (2005). "Guidelines for preventing the transmission of tuberculosis in Canadian health care facilities and other institutional settings" URL: http://ww.phacqspc.gc.ca/publicat/ccdr-rmtc/96 vol **22/2251.**

Hussein, A., Ahmed, Z., Abiyu, M., Solomon, A. and Solomon, A. (2012). "Smear positive pulmonary tuberculosis amongst patients at Agaro Teaching Health center, South West Ethiopia" *Ethiopia Journal of Health Sciences*. 2012 mar; **22**(1): 71-76.

Halezeroglu, S. and Okur, E. (2013). "Thoracic surgery for haemoptysis in the context of tuberculosis: What is the best management approach?" *Journal of Thoracic Disease* **6**(3): 182-185.

Jindal (2011). Text book of pulmonary and critical care medicine. New Delhi: Jaypee Brothers Medical Publishers; P.549.

Kolappan, C. and Gobi, P.G. (2002). "Tobacco smoking and pulmonary tuberculosis". *Thorax*;**57**:964-966.

Konstantinos, A. (2010). "Testing for tuberculosis". *Australian Prescriber;* **33**(1): 12-18

Kumar, v., Abbas A.K., Fausto, N., Mitchell, R.N. (2007). Robbins Basic Pathology. Saunders Elsevier. P.561-522.

Laserson, K.F., Vant, H.A.H., Githui, W.A., Meme, H.K., Aqaya, J.A., Odeny, L.O., Muchiri,B.G.,Marston, B.J., Decock, K.M. and Borqdorff, M.W. (2011). "High prevalence of pulmonary tuberculosis and inadequate case finding in rural Western Kenya". *Indian Journal of Respiratory and Critical Care Medicine;* **183**(9):1245- 1253.doi:10.1164/vccm.2010008-1269UC.Epub?

Lienhardt, C. Fielding, K. Sillah, J.S. Bah, B. Gustafston, P. Warndorff, D. Palayew, M., Lassies, I., Diallo, S., Mianneh, K., Adegbola, R., Aaby, P., Bah, O., Bennett, S. and McAdam, K. (2005). Investigation of risk factors for tuberculosis: A case-control study in three countries in West Africa; *International Journal of Epidemiology*; **34**:914-923.

Lienhardt, C. (2001). From exposure to disease, the role of environmental factors in susceptibility and development of tuberculosis. *Epidemiological Review Journal*. **22**:288-301

Lawn, S.D. and Zumla (2011). "Tuberculosis" Lancet **378** (9785): 57-72

Mondal, M. N., Hoque, M. N., MRK Chowdhory and J Howard (2014) "Socio-demographic factors affecting Knowledge level of Tuberculosis patients". *African Health Sciences* **14.** DOI**:** http://dx.doi.org/10.4314/ahs.v144.13

Mayo Clinic (2006) ."Causes of Tuberculosis" DOI:https:// www.mayoclinic. org/ diseases/tuberculosis/symptoms-causes/syc-20351250.

Menzies, D., Fanning, A. and Yuan, L. (2000). Hospital ventilation and risk factors for tuberculosis infection in Canadian healthcare workers. *An International Medical Journal*;**133**: 779-789.

Menzies, D., Al Jahdal, H. and Otaibi, B. (2011). "Recent development in treatment of latent tuberculosis infection". *The Indian Journal of Medical Research* **133**(3): 257-266 PMC 3103149.

Mcshane, H. (2011). "Tuberculosis vaccines: beyond bacilli Calmette Guerin". *Philosophical transactions of the Royal society of London. Series B, Biological sciences* **366** (1579): 2782-9 doi:10-1098/rstb.2011.0097 PMC 3146779. PMID 21893541.

Pai, M., Zwerling, A., Menzies, D. (2008). "Systematic Review: T-cell-based Assays for the diagnosis of latent tuberculosis infection": *Annals of Internal Medicine*; **149**(3):1-9.

Gibson. P. and Michael, A. (2005). "Evidence based respiratory medicine". Oxford: Blackwell.p.321 ISBN 978-0-7279-1605-1.

Kiesitra, P. and Zoe Tabar (2014). "A time for greater action against tuberculosis" Economists Insights.

Rieder, H.L. (1999). Epidemiological basis for TB control; IVATLD Paris, France.

Schiffman (2012). Tuberculosiscauses.Availableat, *http/www.medicinehealth.com /script /main/art asp? Article key*=60180.

Sheskin, D. (2000). *Handbook of Parametric and Nonparametric Statistical Procedures*, SecondEdition: Chapman & Hall/CRC.

Singh, M. Mynak, M.K. and Kumar, L. (2005). Prevalence and risk factors for tuberculosis transmission among children in household contact with adults having pulmonary tuberculosis: *Archives of Disease in Childhood*; **90**:624-628.

Shitandi, A. and Nyakeriga, E. (2010). Spatial Prevalence Rate of Tuberculosis as presenting to Homa-bay District, Kenya. *Journal of Applied Sciences Research*; **6**: 1008-1011.

Sitienei, J., Nyambati, V. and Borus, P. (2013). "The epidemiology of smear-positive tuberculosis in three TB/HIV High Burden Provinces in Kenya". *Epidemiology Research International* **2013** http://dx.doi.org/10.1155/2013/417038.

Tsara, V., Serasli, E. and Christaki, P. (2009). "Problems in diagnosis and treatment of tuberculosis infection"; *American Journal for Critical Care*; **174**:736-742.

USAID and TB CARE 1-2013 http://www.w.org/TR/xhtm11/DTD/xhtm11-transitional.dtd.

Van Zyl Smit, R.N., Pai, M.Y. and Laung, W.W. (2010). "Global lung health: the colliding epidemics of tuberculosis, tobacco smoking and HIV".*European Respiratory Journal* **35**(1): 27-33.

Wanyaki, I., Olson, S. and Brassard, P. (2006) "Dwellings, crowding and tuberculosis in Montreal".*Socio Sciences Medical Journal;* **63**:501-511.

WHO (2002), "Gender and Tuberculosis". Factsheet.

WHO (2002). "Global Tuberculosis Control Report".

WHO (2010) "Tuberculosis Fact sheet 104".

WHO (2009) "Epidemiology". Global tuberculosis control: epidemiology, strategy, financing pp.6-33 ISBN 978-92-4-156380-2.

WHO (2009). "The Stop TB strategy, Case reports, treatment outcomes and estimates of TB burden" Global tuberculosis control: epidemiology, strategy, financing pp.187-300.

WHO (2011). "The sixteenth global report on tuberculosis".

WHO (2011). "Global Tuberculosis control Report".

APPENDICES

Appendix 1: Informed Consent.

I am Lydia Nyabate Mboga pursuing masters in Biostatistics at the University of Kenyatta. I am carrying out a research on "A statistical investigation on the prevalence of active tuberculosis and the statistically significant factors for predicting tuberculosis in North Pokot District, West Pokot County". All information which will be obtained will be useful to the community as well as the Ministry of Health in provision of the required public health services to control tuberculosis infection and spread in the area of study.

Steps to be followed

You will be asked few questions then be observed carefully so as obtain information from regarding tuberculosis. A sputum will be obtained from you for further examination. All information acquired from you will be recorded in a questionnaire.

One is not forced to participate in the study. Whoever who declines to participate will not be discriminated from medical attention. There is no penalty if one refuses to be a participant in the study.

Participation in the study is at will and questions concerning the study can be asked at any time.

One may wish to respond to the asked questions or object or may stop proceeding any time from the interview without penalty.

Discomforts and risks

Some questions may be personal or shameful, here one can choose to or refuse answer them without any penalty. It will take to up to half an hour to receive your results after an interview is completed.

Benefits

If one wills to participate in the study the information obtained will be useful in the provision of proper public health services to control TB infection and spread in the region. In case one has TB infection will be accorded treatment promptly.

Reward

One who agrees to take part in the study will be provided with lunch and the transport to and from home.

Confidentiality

The procedure will be carried out in a private room where there will no disturbance. The name of the participant will not be recorded. The information in the questionnaire will be locked in a private box where nobody can access it apart from the interviewer.

Contact information

For any information concerning the study you can contact Prof. Leo Odongo on 0722387299 or Dr. John Maingi on o722880280 or the Kenyatta University Ethical Review Committee Secretariat on chairman.kuerc@ku.ac.ke or secretary.kuerc@ku.ac.ke

Partipant's Statement

I have clearly followed the information concerning my participation in the study. There were available chances to ask questions for clarification which were well answered to my satisfaction. It is my choice to participate in the study. I have been assured that all the information I shall give will be kept private. If I decide to decline from participating in the study I shall still be provided with the same medical services in the same clinic or elsewhere without prejudice. I have been not been restricted to participate in the study throughout since I may stop at any time without penalty.

Name of the participant…….

_____ _____

Signature or Thumbprint Date

Investigator's statement

I as the investigator have elaborated all the information concerning the procedures, risks and benefits involved in participating in the study in a language she or he understands best.

Name of interviewer……..

_____ _____

Signature Date

Appendix 2 Questionnaire

INTERVIEWER TO READ TO THE RESPONDENT

Request for participation

We would like you to participate in this survey. You have been chosen together with others and everything you tell us will be kept confidential .The information that you will provide us will be combined with information from other people living in the TB manyatta with TB, and will not be identifiable as coming from you. One will take part in the study at will .If you do you answer a particular question, or even stop the interview at any point .Your family will not get into trouble if you decide not to participate

Would you like any more information before making your decision to participate or not?

Social- Economic and demographic on the respondents

1. How old are you?

2. Tick your sex?

□ Male □ female

3. What is your level of education?

□ Primary incomplete

□ Primary complete

□ Secondary incomplete

□ Secondary complete

□ Tertiary

□ None

4. What do you do for your living?

□ Formal employment

□ Non formal employment (pastoralism, business, etc.)

5. Has any member of your household ever suffered from TB?

□ Yes

□ No

5.1. If yes how close do you interact?

□ Share a bedroom

□ Sleep separately but stay together during the day.

6. How many members of your household share a bedroom?

□ More than two

□ .Less than two

7. Does every room in your house have a window?

□ Yes

□ No

8. Have you ever taken alcohol?

□ Yes

□ No

8.1 If yes how often?

□ Daily

□Not daily

9. Do you smoke?

□ Yes

□ No

If yes how often?

□ Daily □ Not daily

10. How many balanced diet meals do you take in a day?

□ One

□ Two

□ Three

□None in some days

11. Have you ever had signs and symptoms of TB?

□Yes □ No

11.1 If yes how long did you take to seek medical advice?

□Two weeks

□More than two weeks

12. How accessible is the hospital from your home?

□ Near, (less than 3 KM)

□ Far (greater than 3 KM)

13. Which means of transport do you use to access the hospital?

□ Walking

□ Bicycle

□ Motorcycle

□ Taxi

□ Matatu

13. Is there any TB history in your family?

□ Yes □No

14. How is your house structured?

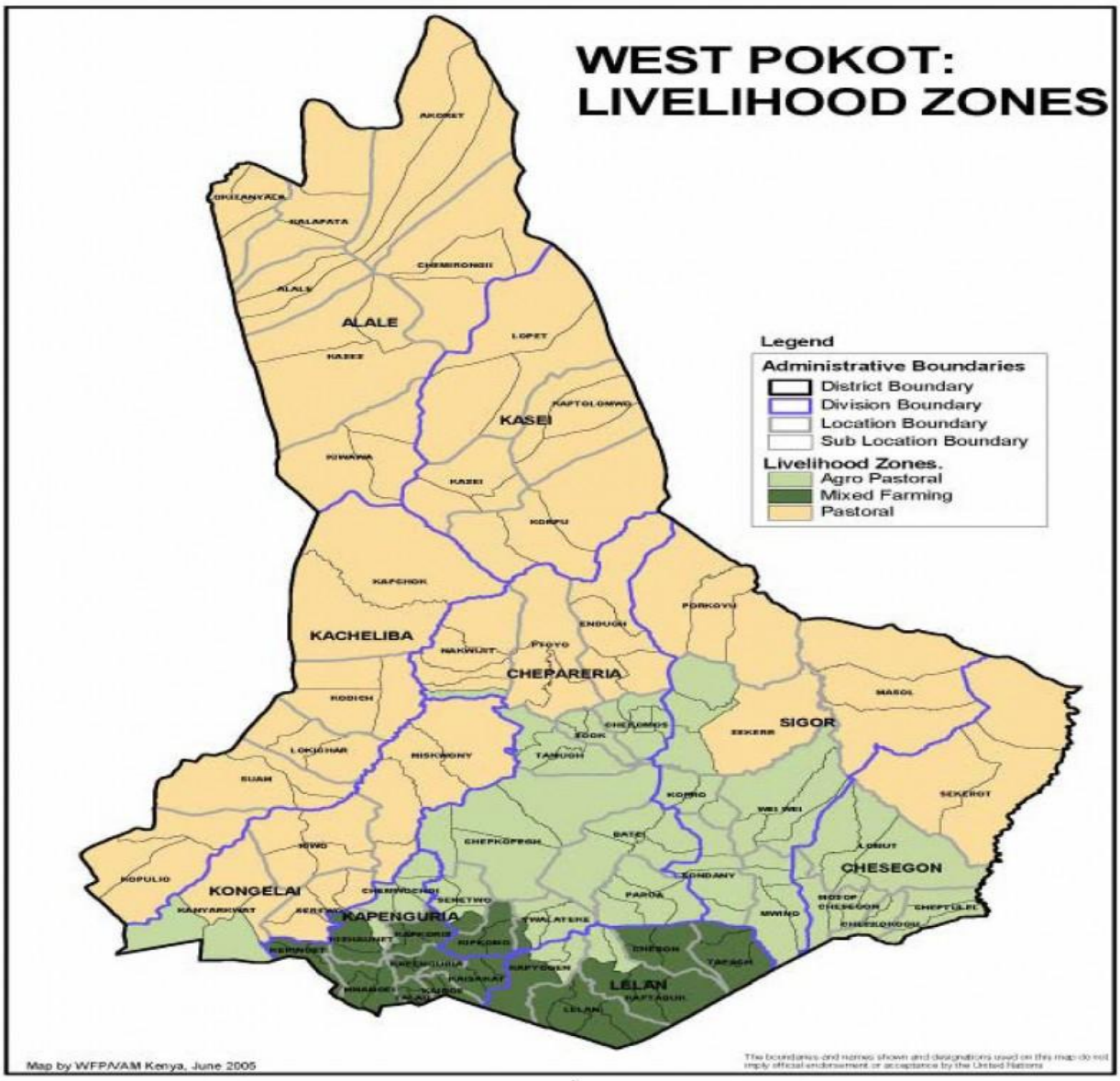
□Cemented □Mud

Appendix 3: Data coding

Variable	Observation	Code
Gender	Male	1
	Female	0
Education level	KCPE and below	0
	KCSE and above	1
Occupation	None	0
	Employed	1
TB knowledge	Adequacy	1
	None	0
History	Yes	1
	No	0
House structure	Cemented	1
	Mud	0
Alcoholism	Yes	1
	No	0
Balanced meals	1	1

	2	2
	3	3
Accessibility	Less 3 KM and a walking distance (near)	0
	More than 3 KM (far)	1
TB-status	Yes	1
	No	0
Smoking	Yes	1
	No	0
Ventilation	Presence of a window	1
	Absence of a window	0
Congestion	Yes	1
	No	0

Appendix 4: West Pokot County Map



Appendix 5: R program

```
TB<-read.table("C:\\Users\\Lydia\\Desktop\\TBDATA.txt",header=TRUE)

attach(TB)

names(TB)

model<-
glm(TB.Status~Age+Sex+Education+Occupation+Knowledge+History+House+Alcoholi
c+Meals+Accessibility+Smoker+Ventilations+Congestion,family="binomial")

summary (model)

model1<-
glm(TB.Status~Alcoholic+Accessibility+Smoker+Congestion+Alcoholic:Congestion,fa
mily="binomial")

summary(model1)

b<--

OR<-exp(b)

OR

alpha<-0.05

zstar <- qnorm(1 - alpha/2)

zstar
```

```
st.error<-

Conf.Int<-exp(b+c(-zstar*st.error,zstar*st.error))

Conf.Int

pchisq(374.57-342.24,5)

anova(model, model1,test="Chisq")

cor.matrix<-model.matrix(model)

cor.matrix

cor(cor.matrix)
```