KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Analysis of Chronic Diseases in Ghana Using Logistic Regression

By

Azumah Karim

(Bsc Statistics)

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FUFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF M.PHIL MATHEMATICAL STATISTICS

May 29, 2014

DECLARATION

I hereby declare that this submission is my own work towards the award of the M. Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

Azumah Karim	....	....
Student	Signature	Date

Certified by:

Mr. F. K. Darkwah	....	....
Supervisor	Signature	Date

Certified by:

Prof. S. K. Amponsah	....	....
Head of Department	Signature	Date

Dedication

I dedicate this work to my family especially my mum ,for their love and care. Not forgetting Mr.F. K. Darkwah my tireless supervisor for giving me the needed support and advice, and to all the teaching staff of the Department of Mathematics, KNUST.

Abstract

The impact of chronic diseases on the development of a nation cannot be overemphasized. The influences of socio-demographic/socio-economic factors on chronic disease conditions in Ghana are not well comprehended. The main objective of this study is to examine the effect of socio-demographic/socio-economic factors of Ghanaians on their chronic disease conditions. A longitudinal study with nationally representative samples was undertaken. The data employed in this study were drawn from the World Health Organization Global Ageing and Adult Health (SAGE), Wave 1, 2008-2009. The study suggested some level of chronic diseases in the Ghanaians populations. Asthma recorded the highest prevalence conditions with 3.4%, while Lung Disease recorded the least, 0.6 %. However, Depression, Oral health, and Injuries fairly recorded 1.4%, 2.5% and 1.8% respectively. In addition, the strength of the age category groups tends heavily towards adults, and particularly older adults respectively. Most of the respondents belong to the Christian faith. There is fairly even distribution among males and females respectively. More than half of respondents were engaged in self-employment. Majority of the respondent belongs to the Akan ethnic group. Also, widowed respondents contracted depression and oral health disease conditions. This paper revealed that, in Ghana, the occurrences of chronic disease conditions are associated with some socio-demographic/socio-economic factors: age, sex, religion, ethnicity, marital status, occupation, level of education, and income levels.

Acknowledgements

My deep appreciation first goes to the Almighty Allah for his Love and Mercy in guiding me throughout the entire Programme as well as research work. My Sincere gratitude also goes to my supervisor Mr.F.K Adarkwah, his patience and commitment in reading through this work in spite of his tight schedules and Prof. S. K. Amponsah(H.O.D)Department of Mathematics,KNUST. My heartfelt gratitude goes to Dr. Bashiru I.I. Saeed(K-Poly). I am very grateful to the following individuals in my life : Mr. Bashiru Shehu, Opoku Nyantah Rahmat and to all my friends. Finally, I wish to express my deepest appreciation to the World Health Organization (WHO) study on Global AGEing and Adult Health for providing me with relevant information.

Contents

List of Tables

Chapter 1

Introduction

A chronic condition is a human health condition or disease that is persistent or otherwise long-lasting in its effects (WHO). It can also be seen as conditions that have a long duration, generally progress slowly and are incurable (Slama et al., 2012). Common chronic diseases include arthritis, asthma, cancer, diabetes and HIV/AIDS. In medicine, the opposite of chronic is acute. A chronic course is further distinguished from a recurrent course; recurrent diseases relapse repeatedly, with periods of remission in between. The non-communicable diseases are also usually lasting medical conditions but are separated by their non-infectious causes. In contrast, some chronic diseases, such as HIV/AIDS, are caused by transmission

1.1 Background of the Study

A chronic illness is a human health condition or disease that is persistent in its effects. The term chronic is usually applied when the course of the disease lasts for more than three months. Chronic diseases constitute a major cause of mortality and the World Health Organization (WHO) reports chronic non-communicable conditions to be by far the leading cause of mortality in the world, representing 35 million deaths in 2005 and over 60% of all deaths. Chronic illnesses cause about 70% of deaths in the US and in 2002 chronic conditions (heart disease, cancers, stroke, chronic respiratory diseases, and diabetes, Alzheimer's disease and kidney diseases) were 6 of the top ten causes of mortality in the general US population. 90% of seniors have at least one chronic disease, and 77% have two or more chronic conditions. For most people, medical conditions do not impair normal activities.

The experience of chronic illness is often emotionally challenging. People respond to these challenges in very different ways yet their responses are inevitably shaped by their resources and social position, the wider environment and clinical, personal and societal reactions to their symptoms and diagnosis (or lack of diagnosis). This special issue of Chronic Illness, which all consider how people experiencing chronic illness talk about, and live with, emotions in different cultural contexts. Hypertension and obesity are major risk factors for chronic diseases. Childhood obesity has increased 3.8 fold from 0.5% in 1988 to 1.9% in 1993/94. The Ghana Demographic and Health Surveys (DHS) demonstrate that prevalence of obesity or overweight among adult (non-pregnant) women across the country increased 2.5 fold in ten years from 10% in 1993 to 25.3% in 2003. Crucially, the 2003 DHS data shows that there are more obese women (25.3%) than malnourished women (9%). The DHS data is supported by a 2003 WHO sponsored national obesity survey which showed higher obesity rates in southern compared to northern regions, among women compared to men, among married individuals compared to unmarried and among older compared to young individuals.

In both national studies Greater Accra Region had the highest overweight and obesity rates and women constituted a high-risk group. The Accra Women's Health Study, a project conducted by the Schools of Public Health of Harvard University and University of Ghana in 2003 reported that 65.9% of a cohort of 3200 women, were overweight and 6.1% were morbidly obese. A relationship between obesity and pregnancy-related complications and deaths has been reported. Like global reports on obesity, increasing obesity rates in Ghana are linked to urbanization, modernization, affluence and changing lifestyles.

Inevitably, these conditions pose physical challenges: experiences range between minor physical ailments to severe physical disabilities. These physical challenges have psychological implications (coping with pain and its management, dealing with disrupted lives and identities) as well as concrete impact on mobility and productivity.

As in a 1981 assessment of the health impact of various diseases in Ghana, ranked in order of healthy days of particular sedentary occupations and consumption of a wider diversity of local and foreign foods. It was find out that,the urban wealthy are not the only high risk groups for chronic diseases in Ghana. Poverty do appear to be a risk factor for both communicable and non-communicable disease. There is a growing evidence that some infectious diseases precipitate chronic diseases and that some chronic conditions place sufferers at risk of infectious diseases.

In addition, the 1970s studies in poor communities in Accra have demonstrated stronger co-existence of communicable and non-communicable diseases compared to wealthier communities. These communities are also likely to suffer complications of, and die prematurely from, chronic diseases because they lack access to quality healthcare.

1.1.1 Ghana's Chronic Disease Burden

Chronic diseases have a longer history in Ghana than is usually thought. Cancer of the liver was recorded in 1817 Asante communities; sickle cell disease was described in 1866. Cases of stroke were presented and treated at Korle-Bu hospital when it opened in the 1920s. Between the 1920s and the 1960s data gathered from Korle-Bu hospital showed a steady increase of stroke and cardiovascular diseases.

Hospital-based and community-based studies conducted since the 1950s provide important information on prevalence and morbidity trends for hypertension, diabetes, cancers and sickle cell disease. Diabetes prevalence studies in southern Ghana have recorded a steady increase. The earliest studies in the 1960s recorded 0.2% prevalence in a population of men in Ho. Diabetes screening conducted by the Ghana Diabetes Association in the early 1990s suggested 2-3% prevalence in urban areas in southern Ghana; in the late 1990s a prevalence rate of 6.4% for diabetes and 10.7% for impaired glucose tolerance (IGT) was recorded in a community in Accra. At Korle-Bu hospital, the percentage of medical admissions due to diabetes increased almost two-fold from 3.5 in the mid-1970s to 6.4% in the mid-1980s.

In the 1970s, the World Health Organization (WHO) sponsored Mamprobi Cardiovascular Disease (CVD) study recorded hypertension prevalence of 13% in the community. A non-communicable disease survey conducted in 1998 recorded a national prevalence of 27.8% for hypertension. Studies conducted after the national survey show higher prevalence rates across different groups in different regions: 28.7% in Kumasi in the Ashanti Region; 32% prevalence in Bawku/Zebilla in the Upper East Region; 36.9% in Keta-Dzelukope in the Volta Region; and 47.8% among a cohort of women in Accra. Reported facility cases of hypertension increased by 67 per cent, from 58,677 in 1989 to 97,980 in 1998.

More so, in 2005, national out-patient hypertension cases totaled 250,000. During the same period (1950s to present) major causes of death have shifted from solely communicable diseases to a combination of communicable and chronic non-communicable diseases. In Accra, cardiovascular diseases rose from being the seventh and tenth cause of death in 1953 and 1966 respectively, to becoming the number one cause of death in 1991 and 2001. By 2003 at least four conditions - stroke, hypertension, diabetes and cancer - had become one of the top ten causes of death in at least each regional health facility.

1.1.2 Profile of Study Area

The Republic of Ghana is centrally located on the West African coast and has a total land area of 238,537 square kilometres. It is bordered by three French-speaking countries: Togo on the east, Burkina Faso on the north and northwest, and Côte d'Ivoire on the west. The Gulf of Guinea on the south forms a coastline extending 560 kilometres. Ghana is a lowland country except for a range of hills on the eastern border and Mt. Afadjato-the highest point above sea level (884 metres)-which is west of the Volta River. Ghana can be divided into three ecological zones: the sandy coastline backed by a coastal plain, which is crossed by several rivers and streams; the middle belt and western parts of the country, which are heavily forested and have many streams and rivers; and a northern savannah, which is drained by the Black and White Volta Rivers. The Volta Lake, created by the hydroelectric dam in the east, is one of the largest artificial lakes in the world. Ghana has a tropical climate with temperatures and rainfall varying according to distance from the coast and elevation. The average annual temperature is about 26°C (79°F). There are two distinct rainy seasons, April to June and September to November. In the north, however, the rainy season begins in March and lasts until September. Annual rainfall ranges from about 1,015 millimetres (40 inches) in the north to about 2,030 millimetres (80 inches) in the southwest. The harmattan, a dry desert wind, blows from the northeast between December and March, lowering the humidity and creating very warm days and cool nights in the north. In the south, the effects of the harmattan are felt mainly in January (Durairaj et al., 2010).

The final results of the 2010 Population and Housing Census (PHC) showed that the total population of Ghana as at 26th September, 2010 was 24,658,823. The results indicated that Ghana's population increased by 30.4 percent over the 2000 population figure of 18,912,079. The recorded annual intercensal growth rate in 2010 was 2.5 percent as against 2.7 percent recorded in 2000. The results revealed that there were 12,633,978 females and 12,024,845 males. This implied that females constituted 51.2 percent of the population and males 48.8 percent, resulting in sex ratio of 95 males to 100 females. It also showed increase in population density from 79 people per square km in 2000 to 103 per square km in 2010.

There were ten (10) administrative regions in Ghana during the 2010 Population and Housing Census as they were in 1984 and 2000. However, in 1988, Ghana changed from the local authority system of administration to the district assembly system. In that year, the then existing 140 local authorities were demarcated into 110 districts. twenty-eight (28) new districts were created in 2004; this increased the number of districts in the country to 138. Whilst in 2008, Thirty-two(32) additional districts were created bringing the total number of districts to 170. The 2010 Population and Housing Census was conducted in these 170 administrative districts (these are made-up of 164 districts/municipals and 6 metropolitan areas). The six metropolitan areas in all have 33 sub-metros which the Statistical Service considered as districts for the purpose of the exercise. Thus, the total number of statistical districts was 197.

From the 2010 Population and Housing Census final results, Greater Accra (16.3%) and Ashanti (19.4%) regions had the greater share of the population while upper East (4.2%) and Upper West (2.8%) regions had the smaller share of the population. Ghana has relatively better health system indicators in the region. In 2007, life expectancy was 57 years (52 years in the WHO African region and 57 years in low-income countries), maternal mortality ratio was 560 per 100,000 live births (900 in the region and 650 in low income countries), infant mortality rate was 73 per 1,000 live births (88 in the region and 80 in low income countries, antenatal care (at least 4 visits) coverage was 69% (45% in the region and 38% in low-income countries) and proportion of births attended by skilled health personnel was 50% (46% in the region and 41% in low-income countries). Similarly, number of outpatient care visits per capita is increasing at the rate of 3.9% per annum; there are a few concerns too. Ghana's human development index (HDI) worsened from 0.563 in 2001 to 0.520 in 2005.Similarly, infant mortality rate rose from 71 in 2000 to 73 in 2007. The level of health spending is low at 5.1% of GDP with household out-of-pocket spending accounting for 51.2% of it. The country has recently established a National Health Insurance Scheme (NHIS) to enhance the performance of its health system, particularly concerning the poor. This technical brief discusses the NHIS approach to social health protection.

1.1.3 Problem Statement

The prevalence of chronic diseases and their risk factors has increased over time and contributes significantly to Ghana's disease burden. Conditions like hypertension, asthma, depression, stroke and diabetes affect young and old, urban and rural, and the wealthy and poor communities. Chronic illness is a complex problem that demands mobilization in all areas of scientific research in order to facilitate appropriate decision-making. It is in this light that this study, seeks to use logistic regression to analyse the influential socio-demographic/socio-economic factors on chronic disease conditions in Ghana.

1.2 Objectives of the Study

This study seeks to examine the risk factors associated with chronic illness in Ghana. The study explores the feasibility for application of the binary logistic regression model to determine risk factors associated with chronic condition in Ghana. The consideration of future for the estimation of chronic illness would entail a variety of areas of investigation. This research is conducted as a way of undertaking this exploration

1.2.1 Specific Objectives of the Study

In particular, the goals of the research study are:

1. To establish a logistic regression model with the ultimate aim of determining the key risk factors associated with chronic diseases.
2. To use the model for prediction of chronic conditions.

1.2.2 The Research Questions

The research seeks to address the following questions:

1. Is the binary logistic regression analysis appropriate technique to determine the risk factors associated with chronic illness in Ghana?
2. What are the risk factors associated with chronic illness in Ghana?

1.3 Methodology

The data employed in this study were drawn from the World Health Organization Study on global AGEing and Adult Health (SAGE), Wave 1, 2008-2009. However, each chronic disease considered in the study (i.e. the dependent variable) were: lung disease, asthma, depression, oral health, and injuries. The predictor variables included gender, age, education, occupation, income level, ethnicity, marital status, and religion.

1.3.1 The Model

The generalized linear regression (GLR) models are being employed in carrying out of data analysis. The (GLR) models are a natural generalization of the familiar classical linear models. The class of GLR models includes, as special cases, linear regression, analysis-of-variance models, log-linear models for the analysis of contingency tables, logit models for binary data in the form of proportions and many others.

1.3.2 The Data

The data employed in this study were drawn from the World Health Organization Global Ageing and Adult Health (SAGE), Wave 1, 2008-2009. SAGE is a longitudinal study with nationally representative samples of persons aged 50+ years in Ghana with a smaller sample of adults aged 18-49 years. The independent variables considered in the study include gender, age, education, occupation, income level, ethnicity, marital status, religion and presence of obesity. For each chronic disease (i.e. the dependent variable) considered in this study were: lung disease, asthma, depression, oral health, injuries.

1.3.3 The Software

The Statistical Product for Service Solutions (SPSS) was employed in capturing the data. The Statistical Analysis (STATA) package will be used in analyzing and fitting the binary logistic regression models.

1.3.4 Source of Knowledge

The main source of knowledge for the successful completion of this study would be World Health Organization Global Ageing and Adult Health (SAGE) on chronic illness, the Kwame Nkrumah University of Science and Technology Library. However, the internet and other Health Journal largely on chronic illness publications would continue to help enrich the progress and outcome of the study

1.4 Significance of the Study

Several people have researched into chronic condition and contributed to knowledge immensely. These will enable the individual patient, the health professional as well as policy makers to manage, control and minimize the number of people dying from chronic diseases. The study is to bring to attention to the public with a view of, among other things, invigorating advance research

1.5 Limitations and Scope of the Study

Challenges inherent with any secondary source of data are some of the limitations such as definition of variables. This study suffered constraints of time and limited resources. The study would be confines to the structure of the Kwame Nkrumah University of Science and Technology thesis format.

1.6 Thesis Organization

This report is organised in five chapters. Chapter 1 is the introductory chapter to the entire study. It takes a critical look at the general background of chronic conditions in Ghana. The problem statement, research questions and objectives, research methodology, significance of the study as well as scope and limitations of the study are discussed in this chapter. Chapter 2 reviewed related literature based on the thesis objectives and preferred models to be used in achieving these objectives. The outcome of the study and other comparative results of similar studies are also discussed in this chapter. Chapter 3 described the theory of the model to be used, formulations and methods of solution. Chapter 4 is dedicated to data collection, analysis and results. Chapter 5 concludes the entire study by stating specific recommendations to stakeholders based on the major findings made in the study.

Chapter 2

Literature Review

2.1 General Review of Related Studies

This chapter confers the literature accessible on chronic illness in all-purpose. It also looks at digest of abstracts on various literatures with view to some of the related studies on the risk factors associated with chronic illness are reviewed below.

2.2 Chronic Illness Conditions

A report by the Ghana Health Service Report 2007, indicated that , evidences from Ghana points to the fact that Ghana is experiencing a "double burden of disease" with a high burden of both communicable and non-communicable diseases Hypertension now features among the top 10 causes of morbidity at the OPD level in all regions. This is a serious deterioration on the picture in 2006 when it featured among the top 10 causes of OPD attendance among adults in four regions. More females are affected than males in all regions and overall there are nearly two females for every male with the disease. Hypertension, heart failure, chronic liver disease and diabetes mellitus are among the top 10 causes of mortality with hypertension alone accounting for 4.7% of deaths.

2.3 Risk Factors

The causes (risk factors) of chronic diseases are well established and well known; a small set of common risk factors are responsible for most of the main chronic diseases. These risk factors are modifiable and the same in men and women: unhealthy diet; physical inactivity; tobacco use (WHO 2005). These causes are expressed through the intermediate risk factors of raised blood pressure, raised glucose levels, abnormal blood lipids, overweight and obesity. The major modifiable risk factors, in conjunction with the non-modifiable risk factors of age and heredity, explain the majority of new events of heart disease, stroke, chronic respiratory diseases and some important cancers. The relationship between the major modifiable risk factors and the main chronic diseases is similar in all regions of the world (WHO 2005).

2.3.1 Other Risk Factors

Many more risk factors for chronic diseases have been identified, but they account for a smaller proportion of disease (WHO 2005). Harmful alcohol use is an important contributor to the global burden of disease but its relationship to chronic disease is more complex (WHO 2005). Other risk factors for chronic disease include infectious agents that are responsible for cervical and liver cancers, and some environmental factors, such as cardiovascular diseases, mainly heart disease and

- stroke;
- cancer;
- chronic respiratory diseases;

- diabetes;

Others, such as mental disorders, vision and hearing impairment, oral diseases, bone and joint disorders, and genetic disorders.

2.3.2 Heart Disease

There are many forms of heart disease. Coronary heart disease, also known as coronary artery disease or ischaemic heart disease is the leading cause of death globally. It is caused by disease of the blood vessels (atherosclerosis) of the heart; air pollution, which contribute to a range of chronic diseases including asthma and other chronic respiratory diseases (WHO 2005).

2.4 Psychosocial and Genetic Factors

A report by World Health Organization,(WHO 2005), indicated that,there is now extensive evidence from many countries that conditions before birth and in early childhood influence health in adult life. For example, low birth weight is now known to be associated with increased rates of high blood pressure, heart disease, stroke and diabetes . Ageing is an important marker of the accumulation of modifiable risks for chronic disease: the impact of risk factors increases over the life course. The underlying determinants of chronic diseases are a reflection of the major forces driving social, economic and cultural change -globalization, urbanization, population ageing, and the general policy environment.

2.4.1 Poverty

Chronic diseases and poverty are interconnected in a vicious circle. At the same time, poverty and worsening of already existing poverty are caused by chronic diseases. The poor are more vulnerable for several reasons, including greater exposure to risks and decreased access to health services (WHO 2005). Psychosocial stress also plays a role. The total number of people dying from chronic diseases is double that of all infectious diseases (including HIV/AIDS, tuberculosis and malaria), maternal and perinatal conditions, and nutritional deficiencies combined (WHO 2005). It is indicated that 80% of chronic disease deaths occur in low and middle income countries and half are in women (WHO 2005). Without action to address the causes, deaths from chronic disease will increase by 17% between 2005 and 2015(WHO 2005). In addition, in the U.S. the impact of chronic condition were among the 15 leading causes of death in 2005. These were:heart disease,cancer, stroke, injuries,diabetes and Alzheimer's disease (Kung et al., 2008)

A study by de Graft Ama (2007), indicated that Hypertension and obesity are major risk factors for chronic diseases. Childhood obesity has increased 3.8 fold from 0.5% in 1988 to 1.9% in 1993/94. The Ghana Demographic and Health Surveys (DHS) demonstrate that prevalence of obesity or overweight among adult (non-pregnant) women across the country increased 2.5 fold in ten years from 10% in 1993 to 25.3% in 2003. Crucially, the 2003 DHS data shows that there are more obese women(25.3%) than malnourished women (9%). The DHS data is supported by a 2003 WHO sponsored national obesity survey which showed higher obesity rates in southern compared to northern regions, among women compared to men, among married individuals compared to unmarried and among older compared to young individuals. In both national studies Greater Accra Region had the highest overweight and obesity rates and women constituted a high-risk group . The Accra Women's Health Study, a project conducted by the Schools of Public Health of Harvard University and University of Ghana in 2003 reported that 65.9% of a cohort of 3200 women, were overweight and 6.1% were morbidly obese. A relationship between obesity and pregnancy-related complications and deaths has been reported. Like global reports on obesity, increasing obesity rates in Ghana are linked to urbanization, modernization, affluence and changing lifestyles (in with chronic disease in Ghana. First, these conditions pose physical challenges: experiences range between minor physical ailments to severe physical disabilities. These physical challenges have psychological implications (coping with pain and its management, dealing with disrupted lives and identities) as well as concrete impact on mobility and productivity. In a 1981 assessment of the health impact of various diseases in Ghana, ranked in order of healthy days of particular sedentary occupations and consumption of a wider diversity of local and foreign foods). The urban wealthy are not the only high risk groups for chronic diseases in Ghana. Poverty appears to be a risk factor for both communicable and non-communicable disease. There is growing evidence that some infectious diseases precipitate chronic diseases and that some chronic conditions place sufferers at risk of infectious diseases. Since the 1970s studies in poor communities in Accra have demonstrated stronger co-existence of communicable and non-communicable diseases compared to wealthier communities. These communities are also likely to suffer complications of, and die prematurely from, chronic diseases because they lack access to quality healthcare.

From Wikipedia, the free encyclopedia an article about medical conditions reports, while risk varies with age and gender, most of the common chronic diseases are caused by dietary, lifestyle and metabolic risk factors that are also responsible for the resulting mortality. Therefore these conditions might be prevented by behavioral changes, such as quitting smoking, adopting a healthy diet, and increasing physical activity. Social determinants are important risk factors for chronic diseases. Social factors such as socio-economic status, education level, and race/ethnicity, are also a major cause for the disparities observed in the care of chronic disease. Lack of access and delay in receiving care result in worse outcomes for patients from minorities and underserved populations. Those barriers to medical care complicate patients monitoring and continuity in treatment. Also, minorities and low-income populations are less likely to access and receive preventive services necessary to detect conditions at an early stage. In addition to this, Over use of antibiotic has recently been shown has been associated with chronic disease, for example Chronic Liver Diseases.

A study by Dabbs et al. (2013) argued that Chronic diseases, such as heart disease, stroke, cancer, respiratory diseases and diabetes, are by far the leading cause of morbidity and mortality in the world.The study also stated that, more than half of individuals with one chronic condition have multiple chronic conditions, increasing the complexity and burden of managing their health.

As the evidence is complex, the report should meet the needs of technical audiences for whom detailed knowledge is central as well as be accessible and useful to those for whom synthesised understandings are sufficient source,(Suhrcke et al., 2006)

A study by Mary (2003) Chronic Disease Prevention Organizing Project reported that chronic disease has become the leading killer and biggest single threat to quality of life in the United States. Diseases like cancer, heart disease and asthma disproportionately affect low income, ethnically diverse communities. After years of failing to engage these populations around chronic disease prevention, Contra Costa Health Services decided to try a new approach.

Heiner et al. (2012) argued that the associations between the intake of vegetables and fruit and the risk of several chronic diseases show that a high daily intake of these foods promotes health. Socio-economic status, oral symptoms and compliance with dietary guidelines were associated with general health status

(Singh and Brennan, 2012). Similarly, socio-economic status and oral health factors related to tooth loss and chewing ability have been related to compliance with dietary guidelines through food purchasing in older adults (Singh and Brennan, 2012). However, aged people show high prevalence of health problems and disability, good self-perception of health, and quality of life factors have been suggested as important (Ventegodt et al., 2003; Rodham et al., 2012; Lavie-Ajayi et al., 2012; Dow et al., 2012), food related self-efficacy and intentions (Gittelsohn et al., 2010), and identification with healthy eating (Strachan and Brawley, 2009). Also psycho-social factors have been suggested as important with sense of coherence (Lindmark et al., 2005)

Physical inactivity is a major factor in the development of many chronic illnesses, including hypertension by Idowu et al. (2012). Illness beliefs are important predictors of psychological outcome in people with chronic illness and evidence suggests these could also be significant in furthering our understanding of psychological functioning in people with Parkinson's disease, by (Simpson et al., 2013).

Research conducted in various countries shows that particularly many problems are associated with health-related quality of elderly people's lives (Van-Gool et al., 2007; Paúl et al., 2007; Dugan and Lee, 2013). Accordingly, the avoidance of illness as well as the preservation of physical and cognitive functions is some of the most important factors that significantly improve the quality of life in the elderly age. Different studies also conclude that health is one of the most prized values for elderly people (Bowling and Gabriel, 2007). Poor health is associated with the loss of control, autonomy and independence, and it makes people aware of the approaching death.

The debilitating and often fatal complications of cardiovascular disease (CVD) are usually seen in middle-aged or elderly men and women, (World Health

Organization 2007).

A study by Ziebland and Kokanovic (2012) , put it, the experience of chronic illness is often emotionally challenging. People respond to these challenges in very different ways yet their responses are inevitably shaped by their resources and social position, the wider environment and clinical, personal and societal reactions to their symptoms and diagnosis (or lack of diagnosis).

According to Pols et al. (2009) , chronic conditions will account for 80% of global disease burden by the year 2020. The increasing demand on health services from chronic conditions requires a shift from managing acute conditions to prevention of disease and maintaining a healthy life for people with chronic conditions. In Australia, healthcare systems are under substantial pressure from rising medical costs that are outstripping the capacity to afford them. In addition, in Australia, the proportion of the population aged over 65 years will rise from 12% in 2002 to 26% by 2051. Again a study in 2000 by Anderson and Horvath (2004) with the findings that , approximately 125 million Americans (45% of the population) had chronic conditions and 61 million (21% of the population) had multiple chronic conditions

Nieman (2009) in a research titled, 'Chronic condition self-management and two teaching models for chronic conditions', stated that, in the current era of globalization, medical educators worldwide face a common problem; namely, professional training in medical schools has not kept pace with the growth of chronic illnesses. In 2003, 33 million deaths were attributed to chronic diseases, that is, to diseases that do not resolve spontaneously but that often are resolved or prevented by lifestyle changes. However, Tommis et al., 2009, states that people or the carers with chronic conditions experienced a decline in their physical health over time whereas the physical health of carers without a chronic condition remained stable. There was also evidence of carers neglecting their own health because of their caring responsibilities. Despite their health problems, over time the carers with chronic conditions received no more support than other carers.

A study by Haslbeck and Schaeffer (2009) on 'Routines in medication management: the perspective of people with chronic conditions' reports that 'Chronic conditions are accompanied by numerous challenges for chronically ill people, including the management of medication regimens, which are often complex and difficult to handle. Empirically, medication management in chronic illness has received a lot of attention in the last three decades. Research on both compliance and adherence has focused on the causes and extent of the health, social, and economic consequences of non-compliance to medication regimens leading to a variety of factors'.

Munish et al. (2012), conducted a study "the use of practice guidelines by the American Society of Anesthesiologists for the identification of surgical patients at high risk of sleep apnea". Concluded that; Obstructive sleep apnea (OSA) is a clinical syndrome characterized by repeated occlusions of upper airway during sleep, resulting in sleep fragmentation and nocturnal hypoxemia. Symptomatic OSA affects approximately 2% and 4% of women and men, respectively, but the overall prevalence of the sleep disorder was estimated to be 9% for women and 24% for men between the ages of 30 and 60 years. In a survey study in elective surgical patients using Berlin questionnaire, 24% patients were found to be at risk of OSA. Due to the associated airway pathology and depressive effects of anesthetic agents on respiration and pharyngeal muscle tone, patients with OSA are at a high risk of respiratory complications. In addition, airway management of these patients has been shown to be challenging for the anesthesiologists in view of difficulty in securing the airway.

In another research by Gale et al. (2012) says patients with bronchiectasis had increased aortic stiffness compared with controls similar in age, BMI, smoking history, BP and cholesterol. Aortic PWV (aortic pulse wave velocity (PWV; (a measure of arterial stiffness and an independent predictor of CV (Cardiovascular) risk) was related to age and heart rate, and inversely related to lung function and oxygen saturation. The patients also had reduced exercise capacity and more patients had osteoporosis. These represent additional comorbidities that require further investigation to optimise outcomes for patients. Whilst a study by Benzo (2012) state “In reality, self-management has universal applications in the chronic disease process, to name a few are use of medications, physical activity, eating, emotional control, and so on”.

Gugiu et al. (2009) indicated that; Although the traditional healthcare system is designed to respond to episodic illnesses, chronic illnesses, such as diabetes mellitus, require ‘continuing medical care and patient self-management education to prevent acute complications and to reduce the risk of long-term complications’ (p. S13).Consequentially, investigators have begun to focus greater attention on the impact of treatment programmes based on models of continuous care and self-management, such as the Chronic Care Model (CCM), on health outcomes. In the past decade, the number of published studies has risen from a couple of articles per year to as many as 16 in 2008. Considering that the ’Baby Boomer’ generation is fast approaching the age of retirement, the need to improve chronic care services to meet the needs of an increasing elderly population will likely fuel further research on the Chronic Care Model (CCM).

Jerliu et al. (2013) indicated that factors associated with the presence of chronic conditions and/or multimorbidity were female sex, older age, self-perceived poverty and the inability to access medical care. In another study by Nepal et al. (2012) indicated the respondents who were 45 to 64 years had the highest proportion of single and multiple doctor-diagnosed chronic diseases (26.3%vs 43.8%). A comparatively high prevalence of doctor-diagnosed chronic disease ($P < 0.0001$) was, however, noted among participants who were 18 to 34 years, being 39.2% and 18.4% for single and multiple doctor-diagnosed chronic disease respectively. The doctor-diagnosed chronic disease was also significantly ($P < 0.05$) associated with race/ethnicity. Whites recorded the highest proportion of doctor-diagnosed chronic disease in both the single (37.4%) and multiple (36.7%) disease categories. Participants of Hispanic origin recorded 32.6% and were ranked second to whites for single doctor-diagnosed chronic disease whereas blacks were ranked second to whites for multiple doctor-diagnosed chronic diseases.

In addition, a research conducted by McKay et al. (2012) reported "That chronic obstructive pulmonary disease is largely environmentally-driven, and thus potentially preventable, is particularly important in the face of unfavourable epidemiological reports. Chronic obstructive pulmonary disease accounted for 5% of deaths globally in 2005, with most of these occurring in low- and middle-income countries, and it is predicted to become the third leading cause of death by 2030"

Another study that attempted to come out with how socio-economic affects chronic argued that effects of class membership and socio-economic status is one variable amongst a number of lifestyle, biomedical and environmental factors posited as vital in determining the aetiology of chronic illness in the community by Walker et al. (2001). .However, a study by Elizabeth Barrett-Connor (2013) argued that Women do have less heart disease than Men.

Chapter 3

Methodology

3.1 Review of Related Models

3.1.1 Introduction

This chapter describes the theory of models to be considered, formulations and methods of analyzing the available data to satisfy the objectives of the study. It focuses on the details, and comprehensive understanding of Generalised Linear Regression (GLR) models. The main focus is on the logistic regression model as the methodologies used in modelling, the software specifications, and the features that are incorporated in the model.

3.2 Logistic Regression Model

3.2.1 Introduction

The generalized linear regression (GLR) models together with the statistical package Stata have revolutionised graduate statistics and the practice of data analysis. The (GLR) models are a natural generalization of the familiar classical linear models. The class of GLR models includes, as special cases, linear regression, analysis-of-variance models, log-linear models for the analysis of contingency tables, logit models for binary data in the form of proportions and many others. The use of classical linear models in health data analysis is not new. Thus, such models have become an established part of the description of chronic indicators or incidence and odds ratios or relative risks - as evidenced by a number of papers in Study on AGEing population (SAGE) literature, including Ahern and Hendryx (2005). However, the use of generalised linear models in chronic health conditions is relatively new. Thus, Kyngäs (2007) used logistic regression to find the factors that predict good adherence to health regimens. One of these relates to Walker (2007) use of logistic regression techniques to study associations between: (1) co-morbidities and demographic, socio-economic and risk factor variables; and (2) quality of life (general and psychological distress) and demographic, socio-economic and health status indicators. Girón et al. (2009)to determine the prevalence of self-reported chronic mental health problems (MHPs) and mental health service use and their determinants, among the Spanish population over 14 years of age employed Generalised Linear model such as the Multivariate logistic regression analysis were used. In this thesis we use the generalised linear model, specifically, the binary logistic regression model to determine risk factors associated with chronic condition in Ghana. However, the dependent variable were: lung disease, asthma, depression, oral health, and injuries. The predictor variables included gender, age, education, occupation, income level, ethnicity, marital status, and religion.

3.3 Definitions of Logistic Regression Model

The logistic regression model is a generalized linear models with three components: The random component identifies the response variable Y and assumes a probability distribution for it.The random component is the response variable which is binary variable, $Y_i = 1$ or 0 (existence of chronic disease or non-existence of the chronic disease). The systematic component specifies the explanatory variables for the model,$X = (x_1, x_2, x_3, ..., x_q)$. The third component of a GLM, the link function connects the random and systematic components.

The systematic component is a linear predictor, $\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_qx_q$. In this case, we are interested in the probability that $Y_i = 1$, (that is, $\pi(x_i) = P(Y_i = 1)$) which has the Binomial distribution. The explanatory (predictor) variables may be quantitative (continuous), qualitative (discrete), or both (mixed). The outcome variable in this analysis is chronic condition. The fitted logistic regression models take the form:

$$logit(\pi(x_i)) = \log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + ... + \beta_qx_q \tag{3.1}$$

The logit function can take any real value, but the associated probability always lies in the required [0, 1] interval. In a logistic regression model, the parameter β_j associated with explanatory variable X_j is such that $\exp(\beta_j)$ is the odds that the response variable takes the value one when increases by one, conditional on the other explanatory variables remaining constant. The parameters of the logistic regression model are estimated by maximum likelihood method. The standard multiple linear regression model is inappropriate to model this data for the following reasons:

- The model's predicted probabilities could fall outside the range 0 to 1.
- The dependent variable is not normally distributed. It follows the Binomial probability model.
- If we consider the normal distribution as an approximation for the binomial model, the variance of the dependent variable is not constant across all risk variables.

The logistic regression model, (binary logistic regression model) is therefore developed to account for all these difficulties.

3.3.1 Odds and Odds Ratio

Odds are the probability of an event occurring divided by the probability of the event not occurring. An odds ratio is the odds of the event in one group, for example, those exposed to a drug, divided by the odds in another group not exposed. Odds are simply a different expression of the probability: the probability of an event divided by the probability of the event not happening. Thus, the odds of an event would be $(probability)/(1 - probability)$. Because this is a ratio, its values range from zero to infinity. Assume that, with a major storm blowing in, the weather forecast calls for an 80% probability of rain (0.80). The odds of rain would be $(0.80)/(1 - 0.80)$ or $0.80/0.2 = 4.0$. With the storm approaching, the odds of rain are 8/2 (eight chances "yes" to two chances "no"). The probability of rain is 0.80, while the odds of rain are 4.0, a twice fold difference.

For example Risk of developing Coccidioidmycosis associated with arthritis therapy? Groups: Patients receiving tumor necrosis factor a (TNFα) versus Patients not receiving TNFα (all patients arthritic)

Table 3.1: Coccidioidomycosis and TNFα-antagonists

	COC	No COC	Total
TNFα	7	240	247
Other	4	734	738
Total	11	974	985

Clearly, a patient on TNFα developing Coccidioidmycoisis odds = $(7)/(240) = 0.0293$ and patients on Other drugs developing Coccidioidmycoisis odds = $(4/734) = 0.0055$. The Odds Ratio of patients on Other drugs developing Coccidioidmycoisis and patient on TNFα developing Coccidioidmycoisis = $(0.0055)/(0.0293) = 0.188$. This means that patients on Other drugs are 81.2%

to developing Coccidioidmycoisis than those on the TNFα drug developing Coccidioidmycoisis.

3.3.2 Interpretation

Conclude that the probability that the outcome is present is higher (in the population) for group 1 if the entire interval is above 1. Conclude that the probability that the outcome is present is lower (in the population) for group 1 if the entire interval is below 1. Do not conclude that the probability of the outcome differs for the two groups if the interval contains 1

3.4 Statistical Model of Chronic Conditions

A closer look at statistical modeling of binary response variables for which the response measurement for each subject is the "existence" or "nonexistence" of chronic condition, or in other words "success" or "failure" of chronic conditions. Binary data are perhaps the most common form of categorical data and the methods are of fundamental importance. The most popular model for binary data is logistic regression. For a binary response Y and a quantitative explanatory variable X, let $\pi(x)$ denote the "existence of chronic condition" (success), probability when X takes value x. This probability is the parameter for the binomial distribution. The logistic regression model has linear form for the logit of this probability

$$\pi(x) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x \tag{3.2}$$

The formula implies that $\pi(x)$ increases or decreases as an S-shaped function of x. An alternative formula for logistic regression refers directly to the success probability. This formula uses the exponential function exp (x) $=e^x$ in the form

$$\pi(x) = \frac{exp(\alpha + \beta x)}{1 + exp(\alpha + \beta x)} \tag{3.3}$$

The parameter β determines the rate of increase or decrease of the S- shaped curve. The sign of β indicates whether the curve ascends or descends, and the rate of change increases as $|\beta|$ increases. When the model holds with $\beta = 0$ the right-hand side of equation (3.3) simplifies to a constant. Then $\pi(x)$ is identical at all x so the curve becomes a horizontal straight line. The binary response Y is then independent of X.

3.5 Confidence Intervals for Effects and Significance Testing

3.5.1 Confidence Interval

A large-sample confidence interval for the parameter β in the logistic regression model, $logit(\pi(x_i)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_q x_q$ is

$$CI = \hat{\beta} \pm Z_{\frac{\alpha}{2}} se\hat{\beta}$$

3.5.2 Test of Significance

We next discuss significance test for the effect of x on the binary response. For the logistic regression model, the null hypothesis $H_0 : \beta = 0$ states that the probability of success is independent of x. For large samples, the test statistic

$$Z = \frac{\hat{\beta}}{se(\hat{\beta})}$$

Consistently, for the two-sided alternative, the test statistic is given as

$$Z^2 = \left(\frac{\hat{\beta}}{se(\hat{\beta})} \right)^2$$

becomes the Wald statistic having a large-sample chi-squared distribution with degrees of freedom 1. Though the Wald test works well for very large samples, the likelihood-ratio test is more powerful and reliable for sample sizes used in practice. The test statistic compares the maximum L_0 null log-likelihood(that is the maximum likelihood for the intercept model only) to the maximum L_1 the final log-likelihood (that is the maximum likelihood for the full model).The log-lkelihood test statistic $-2(L_0 - L_1)$, also has a large-sample chi-squared distribution with degree of freedom (df)=1. Most software for logistic regression reports the maximized log- likelihood L_0 and L_1 and the likelihood-ratio statistic derived from those maxima.

3.5.3 Probability Distribution of Estimates

The estimated probability that Y at fixed setting x of X equals

$$\hat{\pi}(x) = \frac{exp(\hat{\alpha} + \hat{\beta}x)}{1 + exp(\hat{\alpha} + \hat{\beta}x)} \tag{3.4}$$

One can construct confidence intervals for the probabilities using the covariance matrix of the model parameter estimates. The term $\exp(\hat{\alpha}+\hat{\beta}x)$ in the exponents of the prediction equation (3.4) is the estimated linear predictor in the logit transform of equation (3.3) of $\pi(x)$. This estimated logit has large-sample standard error given by the estimated square root of

$$Var\left(\hat{\alpha}+\hat{\beta}x\right)=Var\left(\hat{\alpha}\right)+x^{2}Var\left(\hat{\beta}\right)+2xCov\left(\hat{\alpha},\hat{\beta}\right) \tag{3.5}$$

A 95% confidence interval for the true logit is $\left((\hat{\alpha}+\hat{\beta}x)\pm 1.96(se\hat{\beta})\right)$

3.5.4 Goodness of Fit and Likelihood-Ratio Model Comparison Test

The likelihood-ratio statistic $-2(L_0 - L_1)$ for testing whether certain parameters in a model equal zero. The test compares the maximized log likelihood (L_1) for the full model to the maximized log likelihood (L_0) for the intercept only model. Denote the fitted model by M_1 and the intercept only model for which those parameters equal zero by M_0 .The goodness-of- fit statistic G^2 for testing the fit of a logistic regression model$M$ is the special case of the likelihood-ratio statistic in which $M_0 = M$ and M_1 is the most complex model possible. That complex model has a separate parameter for each logit and provides a perfect fit to the sample logits. It is called the saturated model. In testing whether M fits, we test whether all parameters that are in the saturate model are not in M equal zero. Denote this statistic for testing the fit of M by $G^2(M)$. This is called the deviance of the model. LetL_s denote the maximized log likelihood for the saturated model. Then, the deviances for models M_0 and M_1 are:

$$G^{2}\left(M_{0}\right)=-2\left(L_{0}-L_{s}\right) \tag{3.6}$$

$$G^{2}\left(M_{1}\right)=-2\left(L_{1}-L_{s}\right) \tag{3.7}$$

Denote the likelihood-ratio statistic for testing M_0 given that M_1 holds, by $G^2(M_0|M_1)$. This statistic for comparing this model equals

$$G^2(M_0|M_1) = -2(L_0 - L_s) - (-2(L_1 - L_s)) = G^2(M_0) - G^2(M_1) \tag{3.8}$$

a difference in G^2 goodness-of ?fit statistics for the two models. That is, the likelihood-ratio statistic for comparing two models is simply the difference in the variances of those models. This statistic is large when M_0 fits poorly compared to M_1 . It is a large sample chi-squared statistic, with *df* equal to the difference between residual *df* values for two models. The likelihood-ratio method can be used to test hypotheses about parameters in multiple logistic regression models. More generally, one can compare maximized log- likelihood for any pair of models such that one is a special case of the other.

3.5.5 Residuals for Logit Models

Goodness-of fit statistics as G^2 and $\chi 2$ are summary indicators of the overall quality of fit. Additional diagnostic analyses are necessary to describe the nature of any lack of fit. Residuals comparing observed and fitted counts are useful for this purpose. Let y_i denote the number of "successes" for n_i trials at the i_{th} setting of the explanatory variables. Let $\hat{\pi}$ denote the predicted probability of success for the model fit. Then $n_i\hat{\pi}$ is the fitted number of success. For a GLM with binomial random component, the Pearson residual for the fit at setting is

$$e_i = \frac{y_i - n_i\hat{\pi}}{\sqrt{n_i\hat{\pi}(1-\hat{\pi})}} \tag{3.9}$$

Each residual divides the difference between an observed count and its fitted value by the estimated binomial standard deviation of the observed count. The Pearson statistic for testing the model fit satisfies $\chi 2 = \sum_{i=1} {e_i}^2$. Each squared Pearson residual is a component of $\chi 2$. When the binomial index n_i is large, the

Pearson residual e_i has an approximate normal distribution. When the model holds, it has an approximate expected value of zero but a smaller variance than a standard normal variants. If the number of model parameters is small compared to the number of sample logits, Pearson residuals are treated like standard normal deviates, with absolute values larger than 2 indicating possible lack of fit. Graphical displays are also useful for showing lack of fit one can compare observed and fitted proportions by plotting them against each other, or by plotting both of them against explanatory variables. We have noted that $\chi 2$ and G^2 are invalid when fitted values are very small. Similarly residuals have limited meaning in that case. When explanatory variables are continuous, often $n_i = 1$ at many settings. Then, y_i can equal only 0 or 1, and e_i can assume only two values. One must then be cautious about regarding either outcome as "extreme" and single residual is usually uninformative.

3.5.6 Logit Models for Qualitative Predictors

The Multiple Logistic regression models incorporating multiple explanatory variables. Moreover, some or all of those predictor variables can be qualitative, rather than quantitative. Suppose that a binary response Y has two binary predictors, X_1 and X_2. If the two levels for each variable are denoted by $(0, 1)$, then $logit(\pi(x_i)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ has separate main effects for the two predictors. It assumes an absence of interaction, the effect of one factor being the same at each level of the other factor. The variables X_1 and X_2 in this model are indicator variables that indicate categories for the predictors. At a fixed level x_2 of X_2, the effect on the logit of changing from $x_1 = 0$ to $x_1 = 1$ is $[\beta_0 + \beta_1(1) + \beta_2 x_2] - [\beta_0 + \beta_1(0) + \beta_2 x_2] = \beta_1$. This difference between two logits equals the difference of log odds, which equals the log of the odds ratio between X_1 and Y at a fixed level of X_2. Thus, $exp\,(\beta_i)$ describes the conditional odds ratio between X_1 and Y. Controlling for X_2, the odds of "success" at $x_1 = 1$

equal $exp\left(\beta_1\right)$ times the odds of success at $x_1 = 0$. This conditional odds ratio is the same at each level x_2 of X_2.

3.5.7 Model Selection and Test for Adequacy

Several model selection procedures exist, no one of which is "best". Cautions that apply to ordinary regression modeling of normal data hold for any generalized linear model. For instance, a model with several predictors has the potential for multicollinearity, strong correlations among predictors, making it seems that no one variable is important when all the others are in the model. A variable may seem to have little effect simply because it "overlaps" considerably with other predictors in the model. The following criteria are used to test for the model adequacy:

- Deviance Residuals: Deviance residual is the measure of deviance contributed from each observation and is given by$r_{D_i} = \text{sign}(r_i\sqrt{d_i})$ where di is the individual deviance contribution. The deviance residuals can be used to check the model fit at each observation for generalized linear models. The standardized and studentized deviance residuals are

$$r_{D_{si}} = \frac{r_{D_i}}{\sqrt{\hat{\phi}\left(1-h_i\right)}} \tag{3.10}$$

$$r_{D_{ti}} = \frac{r_{D_i}}{\sqrt{\hat{\phi}_{(i)}\left(1-h_i\right)}} \tag{3.11}$$

- Akaike's Information Criterion (AIC): This is used to select the best fitted model with the lowest AIC Value. In a categorical data this statistic is similar in selecting the model that minimizes the model: defined by the model deviance minus the twice the degree of freedom of the model given as $G_2 - 2v$

- Confidence Interval(CI): This indicate the acceptable limits for the estimated parameter defined as CI = $\hat{\beta} \pm Z_{\frac{\alpha}{2}} se\hat{\beta}$

3.6 Maximum Likelihood Estimation of Logistic Regression Models

3.6.1 Introduction

Logistic regression is widely used to model the outcomes of a categorical dependent variable. For categorical variables it is inappropriate to use linear regression because the response values are not measured on a ratio scale and the error terms are not normally distributed. In addition, the linear regression model can generate as predicted values any real number ranging from negative to positive infinity, whereas a categorical variable can only take on a limited number of discrete values within a specified range. The theory of generalized linear models of Nelder and Wedderburn identifies a number of key properties that are shared by a broad class of distributions. This has allowed for the development of modeling techniques that can be used for categorical variables in a way roughly analogous to that in which the linear regression model is used for continuous variables. Logistic regression has proven to be one of the most versatile techniques in the class of generalized linear models. Whereas linear regression models equate the expected value of the dependent variable to a linear combination of independent variables and their corresponding parameters, generalized linear models equate the linear component to some function of the probability of a given outcome on the dependent variable. In logistic regression, that function is the logit transform: the natural logarithm of the odds that some event will occur. In linear regression, parameters are estimated using the method of least squares by minimizing the sum of squared deviations of predicted values from observed values. This involves solving a system of N linear equations each having N unknown variables, which is usually an algebraically straightforward task. For logistic regression, least squares estimation is not capable of producing minimum variance unbiased estimators for the actual parameters. In its place, maximum likelihood estimation is used to solve for the parameters that best fit the data. In the next section, we will specify the logistic regression model for a binary dependent variable and show how the model is estimated using maximum likelihood.

3.6.2 Binomial Logistic Regression

Consider a random variable Y that can take on one of two possible values. Given a dataset with a total sample size of$N$, where each observation is independent, Y can be considered as a column vector of N binomial random variables Y_i. By convention, a value of 1 is used to indicate *success* and a value of either 0 is used to signify $failure$. To simplify computational details of estimation, it is convenient to aggregate the data such that each row represents one distinct combination of values of the independent variables. These rows are often referred to as *populations*. Let N represent the total number of populations and let n be a column vector with elements n_i representing the number of observations in population i for i = 1 to N where $\sum_{i=1}^{N} n_i = N$ the total sample size. Now, let$Y$ be a column vector of length N where each element Y_i is a random variable representing the number of successes of Y for population i. Let the column vector y contain elements y_i representing the observed counts of the number of successes for each population. Letπ be a column vector of the length N with elements $\pi = P(Y_i|i)$. That is the probability of success for any given observation in the i^{th} population. The linear component of the model contains the design matrix and the vector of parameters to be estimated. The design matrix of independent variables, X is composed of N rows and $K+1$ columns, where K is the number of independent variables specified in the model. For each row of the design matrix, the first element $x_{i0} = 1$. This is the intercept or the "β_0". The parameter vector, β, is a column vector of length $K + 1$. There is one parameter corresponding to each of the K columns of independent variable settings in X, plus one,β_0, for the intercept. The logistic regression model equates the logit transform, the log-odds of the probability of a success, to the linear component:

$$log\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{k=0}^{K} x_{ik}\beta i = 1, 2, ..., N \tag{3.12}$$

3.6.3 Parameter Estimation

The goal of logistic regression is to estimate the K +1 unknown parameters in equation(3.12). This is done with maximum likelihood estimation which entails finding the set of parameters for which the probability of the observed data is greatest. The maximum likelihood equation is derived from the probability distribution of the dependent variable. Since each y_i represents a binomial count in the $i_t h$ population, the joint probability density function of Y is:

$$f\,(y|\beta) = \prod_{i=1}^{N}\left(\frac{n_i!}{y_i!(n_i-y_i)!}{\pi_i}^{y_i}\,(1-\pi_i)^{n_i-y_i}\right) \tag{3.13}$$

For each population, there are $\begin{pmatrix} n_i \\ y_i \end{pmatrix}$ different ways to arrange y_i successes from among n_i trials. Since the probability of a success for any one of the n_i trials is π_i, the probability of y_i successes is ${\pi_i}^{y_i}$. Likewise, the probability of $n_i - y_i$ failures is $(1-\pi_i)^{n_i-y_i}$. The joint probability density function in $Eq.$3.13 expresses the values of y as a function of known, fixed values for β. (Note that β is related to π by Equation (3.12). The likelihood function has the same form as the probability density function, except that the parameters of the function are reversed: the likelihood function expresses the values of β in terms of known, fixed values for y. Thus,

$$L\left(\beta|\, y\right) = \prod_{i=1}^{N} \left(\frac{n_i!}{y_i!(n_i - y_i)!} \pi_i{}^{y_i} \left(1 - \pi_i\right)^{n_i - y_i} \right) \tag{3.14}$$

The maximum likelihood estimates are the values for

β that maximize the likelihood function in Equation (3.14). The critical points of a function (maxima and minima) occur when the first derivative equals 0. If the second derivative evaluated at that point is less than zero, then the critical point is a maximum. Thus, finding the maximum likelihood estimates requires computing the first and second derivatives of the likelihood function. Attempting to take the derivative of Equation (3.14) with respect to β

is a difficult task due to the complexity of multiplicative terms. Fortunately, the likelihood equation can be considerably simplified. First, note that the factorial terms do not contain any of the π_i. As a result, they are essentially constants that can be ignored: maximizing the equation without the factorial terms will come to the same result as if they were included. Second, note that since $a^{x-y} = a^x/a^y$, and after rearranging terms, the equation to be maximized can be written as:

$$\prod_{i=1}^{N} \left((\frac{\pi_i}{1 - \pi_i})^{y_i} (1 - \pi_i)^{n_i} \right) \tag{3.15}$$

Note that after taking e to both sides of Equation (3.12)

$$\left(\frac{\pi_i}{1 - \pi_i} \right) = e^{\sum_{k=0}^{K} x_{ik}\beta_k} \tag{3.16}$$

which, after solving for π_i becomes,

$$\pi_i = \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1 + e^{\sum_{k=0}^{K} x_{ik}\beta_k}} \right) \tag{3.17}$$

Substituting Equation (3.16) for the first term and Equation(3.17) for the second term, Equation (3.15) becomes:

$$\prod_{i=1}^{N} \left(e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)^{y_i} \left(1 - \frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1 + e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right)^{n_i} \tag{3.18}$$

$Eq.$3.18 can be simplify as:

$$\prod_{i=1}^{N} \left(e^{y_i} \sum_{k=0}^{K} x_{ik}\beta_k\right) \left(1 + e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right)^{-n_i} \tag{3.19}$$

The Equation (3.19) is the kernel of the likelihood function to maximize. However, it is still cumbersome to differentiate and can be simplified a great deal further by taking its log. Since the logarithm is a monotonic function, any maximum of the likelihood function will also be a maximum of the log likelihood function and vice versa. Thus, taking the natural log of Equation (3.19) yields the log likelihood function:

$$\ell(\beta) = \sum_{i=1}^{N} y_i \left(\sum_{k=0}^{K} x_{ik}\beta_k\right) - n_i \log\left(1 + e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right) \tag{3.20}$$

To find the critical points of the log likelihood function, set the first derivative with respect to each

equal to zero. In differentiating Equation(3.20) note that

$$\frac{\partial}{\partial \beta_k}\left(\sum_{k=0}^{K} x_{ik}\beta_k\right) = x_{ik} \tag{3.21}$$

since the other terms in the summation do not depend on β_k and can thus be treated as constants. In differentiating the second half of Equation (3.20), take note of the general rule that $\frac{\partial}{\partial x} logy = \frac{1}{y}\frac{\partial y}{\partial x}$. Thus, differentiating Equation

(3.20) with respect to each β_k,

$$\begin{aligned}
\frac{\partial \ell(\beta)}{\partial \beta_k} &= \sum_{i=1}^{N} y_i x_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} \frac{\partial}{\partial \beta_k}\left(1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}\right) \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} e^{\sum_{k=0}^{K} x_{ik}\beta_k} \frac{\partial}{\partial \beta_k}\left(\sum_{k=0}^{K} x_{ik}\beta_k\right) \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}} e^{\sum_{k=0}^{K} x_{ik}\beta_k} x_{ik} \\
&= \sum_{i=1}^{N} y_i x_{ik} - n_i \pi_i x_{ik} \qquad (3.22)
\end{aligned}$$

The maximum likelihood estimates for β

can be found by setting each of the K + 1 equations in Equation (3.22) equal to zero and solving for each β_k. Each such solution, if any exists, specifies a critical pointeither a maximum or a minimum. The critical point will be a maximum if the matrix of second partial derivatives is negative definite; that is, if every element on the diagonal of the matrix is less than zero. Another useful property of this matrix is that it forms the variance-covariance matrix of the parameter estimates. It is formed by differentiating each of the K + 1 equations in Equation (3.22) a second time with respect to each element of β, denoted by $\beta_{k'}$. The general form of the matrix of second partial derivatives is

$$\begin{aligned}
\frac{\partial^2 \ell(\beta)}{\partial \beta_k \partial \beta \acute{k}} &= \frac{\partial}{\beta \acute{k}} \sum_{i=1}^{N} y_i x_{ik} - n_i \pi_i x_{ik} \\
&= \frac{\partial}{\beta \acute{k}} \sum_{i=1}^{N} -n_i \pi_i x_{ik} \\
&= -\sum_{i=1}^{N} n_i x_{ik} \frac{\partial}{\beta \acute{k}} \left(\frac{e^{\sum_{k=0}^{K} x_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} x_{ik}\beta_k}}\right) \qquad (3.23)
\end{aligned}$$

To solve Equation(3.23) we will make use of two general rules for differentiation. First, a rule for differentiating exponential functions:

$$\frac{d}{dx} e^{u(x)} = e^{u(x)} \frac{d}{dx} u(x) \qquad (3.24)$$

In our case, let $u(x) = \sum_{k=0}^{K} x_{ik}\beta_k$ Second, the quotient rule for differentiating the quotient of two functions:

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{[g(a)]^2} \tag{3.25}$$

Applying these two rules together allows us to solve Equation(3.23).

$$\begin{aligned}
\frac{d}{dx}\frac{e^{u(x)}}{1+e^{u(x)}} &= \frac{\left(1+e^{u(x)}\right)e^{u(x)}\frac{d}{dx}u(x) - e^{u(x)}e^{u(x)}\frac{d}{dx}u(x)}{\left(1+e^{u(x)}\right)^2} \\
&= \frac{e^{u(x)}\frac{d}{dx}u(x)}{\left(1+e^{u(x)}\right)^2} \\
&= \frac{e^{u(x)}}{1+e^{u(x)}}\frac{1}{1+e^{u(x)}}\frac{d}{dx}u(x)
\end{aligned} \tag{3.26}$$

Thus, Equation (3.23) can now be written as:

$$-\sum_{i=1}^{N} n_i x_{ik}\pi_i\left(1-\pi_i\right)x_{ik'} \tag{3.27}$$

3.6.4 The Newton-Raphson Method

Setting the equations in Equation (3.22) equal to zero results in a system of K + 1 nonlinear equations each with K + 1 unknown variables. The solution to the system is a vector with elements, β_k. After verifying that the matrix of second partial derivatives is negative definite, and that the solution is the global maximum rather than a local maximum, then we can conclude that this vector contains the parameter estimates for which the observed data would have the highest probability of occurrence. However, solving a system of nonlinear equations is not easy: the solution cannot be derived algebraically as it can in the case of linear equations. The solution must be numerically estimated using an iterative process. Perhaps the most popular method for solving systems of nonlinear equations is Newton's method, also called the Newton-Raphson method. Newton's method begins with an initial guess for the solution then uses the first two terms of the Taylor polynomial evaluated at the initial guess to come up with another estimate that is closer to the solution. This process continues until it converges (hopefully) to the actual solution. Recall that the Taylor polynomial of degree n for f at the point $x = x_0$ is defined as the first n terms of the Taylor series for f:

$$\sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!} (x - x_0) \tag{3.28}$$

provided that the first n derivatives of f at x_0 all exist. The first degree Taylor polynomial is also the equation for the line tangent to f at the point$(x_0; f(x_0))$. The point at which the tangent line crosses the x-axis, $(x_1; 0)$, is used in the next approximation of the root to be found where $f(x) = 0$. The first step in Newton's method is to take the first degree Taylor polynomial as an approximation for f, which we want to set equal to zero:

$$f(x_0) + f'(x_0)(x - x_0) = 0 \tag{3.29}$$

Solving for x, we have:

$$x = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{3.30}$$

This new value of x is the next approximation for the root. We let $x_1 = x$ and continue in the same manner to generate x_2; x_3;..., until successive approximations converge. Generalizing Newton's method to a system of equations is not difficult. In our case, the equations whose roots we want to solve are those in Equation (3.22),the first derivative of the log-likelihood function. Since Equation (3.22) is actually a system of K +1 equations whose roots we want to find simultaneously, it is more convenient to use matrix notation to express each step of the Newton-Raphson method. We can write Equation (3.22) as $\ell'(\beta)$. Let $\beta^(0)$ represent the vector of initial approximations for each

β_k, then the first step of Newton-Raphson can be expressed as:

$$\beta^{(1)} = \beta^{(0)} + [-\ell'' \left(\beta^{(0)}\right)]^{-1} \ell' \beta^{(0)} \tag{3.31}$$

Let μ be a column vector of length N with elements $\mu_i = n_i \pi_i$. Note that each element of μ can also be written as $\mu_i = \mathrm{E}(y_i)$, the expected value of y_i. Using matrix multiplication, we can show that:

$$\ell'(\beta) = \mathbf{X}^T (y - \mu) \tag{3.32}$$

is a column vector of length K + 1 whose elements are $\dfrac{\partial^2 \ell(\beta)}{\partial \beta_k}$, as derived in Equation (3.22). Now, let $\mathbf{W}$ be a square matrix of order N, with elements $n_i \pi_i (1 - \pi_i)$ on the diagonal and zeros everywhere else. Again, using matrix multiplication, we can verify that

$$\ell''(\beta) = -\mathbf{X}^T \mathbf{W} \mathbf{X} \tag{3.33}$$

Continue applying Equation(3.33) until there is essentially no change between the elements of β from one iteration to the next. At that point, the maximum likelihood estimates are said to have converged, and Equation (3.32) will hold the variance-covariance matrix of the estimates.

3.6.5 Properties of Point Estimators

The most desirable properties of point estimators are; unbiasedness, efficiency, consistency, and sufficiency. This can be achieved by the methods of maximum likelihood estimation.

3.6.6 Unbiasedness

If $\hat{\theta}$ is an estimator for a parameter θ, and $E(\theta)$=θ then $\hat{\theta}$ is said to be unbiased estimator of θ.

3.6.7 Consistency

A second desirable property of an estimator is that as n, the sample size, gets larger, the estimator should be more accurate. That is we increase our chances of getting closer to θ as n increases. So the variability of the estimator $\hat{\theta}$ should reduce the as the sample size increases. estimators having this property are "consistent". That is $\hat{\theta}$ is a consistent estimator for θ if for any $\varepsilon > 0$

$$P\left(|\hat{\theta} - \theta| > \varepsilon\right) \to 0, \;\; as \;\; n \to \infty.$$

Although this property seems the same as the first property of unbiasedness it does not follow that one implies the other. An unbiased estimator $\hat{\theta}$ of θ is consistent if the variance of $\hat{\theta}$ tends to zero as n tends to infinity ,i.e. $E[\hat{\theta}] = \theta$ and $Var(\hat{\theta}) \to 0$ as $n \to \infty$

3.6.8 Efficiency

If two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased and consistent, then we can say that their first moment are identical (both have same expected value equal to θ). So, we go on and compare second moments to see if we can discriminate between $\hat{\theta}_1$ and $\hat{\theta}_2$. In fact,we consider the $Var(\hat{\theta}_1)$ and $Var(\hat{\theta}_2)$. Suppose, for a particular sample size n $Var(\hat{\theta}_1) > Var(\hat{\theta}_2)$ then $\hat{\theta}_2$ is likely to be closer to θ than $\hat{\theta}_1$ and so is preferable as an estimator. If $Var(\hat{\theta}_1) > Var(\hat{\theta}_2)$ for the a particular sample size n, then $\hat{\theta}_2$ is said to be more efficient than $\hat{\theta}_1$ for that sample size. if this is true for all values of n, then $\hat{\theta}_2$ is said to be more efficient than $\hat{\theta}_1$. The relative efficiency of $\hat{\theta}_2$ to $\hat{\theta}_1$ is given by

$$\frac{Var\hat{\theta}_1}{Var\hat{\theta}_2}$$

3.6.9 Sufficiency

An estimator is said to be sufficient for a parameter if it contains all the information which can be extracted from the sample concerning the parameter. The statistic$T(x_1, x_2, ..., x_n)$ is said to be sufficient for θ if and only if, for each value of t , the conditional distribution of the random sample $(x_1, x_2, ..., x_n)$ given $T = t$ does not depend on θ . that is: $f(x_1, x_2, ..., x_n|T = t) = \dfrac{f(x_1, x_2, ..., x_n, \theta}{h(t, \theta)}$ $= g(x_1, x_2, ..., x_n)$

3.6.10 General Likelihood Ratio Test

Likelihood ratio tests are useful to test a composite null hypothesis against a composite alternative hypothesis. Suppose that the null hypothesis specifies thatθ (may be a vector) lies in a particular set of possible values, say Θ_0, i.e.,$H_0 : \theta \in \Theta_0$; the alternative hypothesis specifies that θ lies in another set of possible values Θ_a , which does not overlap Θ_0, i.e., $H_a : \Theta_0 \in \Theta_a$. Let $\Theta = \Theta_0 \bigcup \Theta_a$. Either or both of the hypotheses H_0 and H_a can be compositional. Let $L(\hat{\Theta_0}$ be the maximum (actually the supremum) of the likelihood function for all $\theta \in \Theta_0$.That is, $L(\hat{\Theta_0} = max_{\theta\in\Theta_0} L(\theta)$. $L(\hat{\Theta_0}$ represents the best explanation for the observed data for all $\theta \in \Theta_0$. Similarly, $L(\hat{\Theta}) = max_{\theta\in\Theta_0} L(\theta)$ represents the best explanation for the observed data for all $\theta \in \Theta = \Theta_0 \bigcup \Theta_a$. If $L(\hat{\Theta}_0) = \hat{\Theta}$ then a best explanation for the observed data can be found inside Θ_0 and we should not reject the null hypothesis $H_0 : \theta \in \Theta_0$. If $L(\hat{\Theta_0}) < (\hat{\Theta})$ then the best explanation for the observed data could be found inside Θ_a, and we should consider rejecting H_0 in favour of H_a. A likelihood ratio test is based on the ratio $L(\hat{\Theta_0})/L(\hat{\Theta})$. Define the likelihood ratio statistic by

$$\Lambda = \frac{L(\hat{\Theta_0})}{L(\hat{\Theta})} = \frac{max_{\theta \in \Theta_0} L(\theta)}{max_{\theta \in \Theta} L(\theta)}$$

A likelihood ratio test of $H_0 : \theta \in \Theta_0$ vs. $H_a : \Theta_0 \in \Theta_a$ employs Λ as a test statistic, and the rejection region is determined by $\Lambda \leq k$. Clearly, $0 \leq \Lambda \leq 1$. A value of Λ close to zero indicates that the likelihood of the sample is much smaller under H_0 than it is under H_a, therefore the data suggest favoring H_a over H_0. The actually value of k is chosen so that α achieves the desired value.

Chapter 4

Data Collection, Analyses and Result

4.1 Introduction

This chapter presents data collection, analysis and discussion of the study. This was done by first considering the data, distribution of the various risk factors contributing to the chronic diseases. Applying the logistic regression model for the chronic disease conditions to obtain the estimates parameters in the model and the odds ratio analysis using the method of maximum likelihood. Statistical Analysis (Stata) software package (version 12.1) was used.

4.1.1 Data Collection

The data employed in this study were drawn from the World Health Organization Global Ageing and Adult Health (SAGE), Wave 1, 2008-2009. The dependent variables were: lung disease, asthma, depression, oral health, and injuries. The independent variables included gender, age, education, occupation, income level, ethnicity, marital status, and religion.

Table 4.1 shows the chronic diseases conditions coded as existence or non-existence of the disease in column one to column five. For the independent variables: Age captures in column six and coded as young adult, adult, and older adult. Religion in column seven coded as Christianity, Islam and Other religion.

Table 4.1: Sample of Data

Lung Disease	Asthma	Depression	Oral health	Injuries	Age	Religion	Sex	Occupation	Ethncity	Marital Status	Education	Income
NE	NE	NE	NE	NE	2	0	1	0	0	1	0	4
NE	NE	NE	NE	NE	1	2	1	0	0	0	1	3
NE	E	NE	NE	NE	2	0	1	0	0	4	2	4
NE	NE	NE	NE	NE	0	1	0	0	0	2	1	3
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
NE	NE	NE	NE	NE	1	0	0	1	1	0	1	3

Sex of the respondents coded as male and female in column eight, occupation coded as self-employed, public, private, and informal and others employments captured in column nine. Ethnicity were coded as Akan, Ewe, Ga-Adagbe, Mole-Dagbon, Mande-Busanga, and other ethnic groups also in column ten. While Marital status recorded in column eleven coded as currently married, never married, cohabiting, separated/divorced, and widowed couples. Educations were coded as primary, secondary, tertiary and others in column twelve. However Income quantile were coded from lowest to highest in column thirteen

4.2 Summary of Chronic Diseases

In Table 4.2 shows percentage of chronic conditions by respondents. Column one contains chronic diseases, column two contains the number of respondents, existence or non-existence of the chronic diseases and column three captures percentage (%) of the diseases.

Table 4.2: Percentage of Chronic Conditions in Ghana

Chronic Diseases	Frequency	Percentage(%)
lung diseases	**5090**	
existence of lung diseases	28	0.6
nonexistence of lung diseases	5062	99.4
Asthma	**5090**	
existence f Asthma	173	3.4
nonexistence of Asthma	4917	96.6
Depression	**5090**	
existence of Depression	69	1.4
nonexistence of Depression	5021	98.6
Oral Health	**5090**	
existence of Oral Health	126	2.5
nonxistence of Oral Health	4964	97.5
Inuries	**5090**	
existence of Injuries	90	1.8
nonexistence of Injuries	5000	98.2

The percentage of chronic disease conditions by respondents is represented in Table 4.2, where Asthma recorded the highest prevalence conditions with 3.4%, while Lung Disease recorded the least, 0.6%. However, Depression, Oral health, and Injuries fairly recoded 1.4%, 2.5% and 1.8% respectively.

4.3 Factors

4.3.1 The socio-demographic/socio-economic data

Table 4.3, represents the numerical strength as well as percentage for sample characteristics of the socio-demographic/socio-economic factors.

Table 4.3: Sample Characteristics of socio-demographic/socio-economic factors

Factors	Frequency	Percentage(%)
Age	**5,090**	**100.00**
Young Adult	799	15.70
Adult	1,684	33.08
Older Adult	2,607	51.22
Religion	**5,090**	**100.00**
Christianity	3,536	69.47
Islam	794	15.60
Others	760	14.93
Sex	**5,090**	**100.00**
male	2,682	52.69
female	2,408	47.31
Occupation	**5,090**	**100.00**
Self-employed	4,416	86.76
Public	60	1.18
Private	26	0.51
Others	241	4.73
Informal Employment	347	6.82
Ethnicity	**5,090**	**100.00**
Akan	2,553	50.16
Ewe	625	12.28
Ga-Adangbe	398	7.82
Mole-Dagbon	126	2.48
Mande-Busanga	1,012	19.88
Other	376	7.39
Marital Status	**5,090**	**100.00**
currently Married	2,999	58.92
Never Married	142	2.79
cohabitating	62	1.22
Separated/Divorced	666	13.08
widowed	1,221	23.99
Education	**5,090**	**100.00**
Primary	1,167	22.93
Secondary	1,143	22.46
Tertiary	183	3.60
Informal/no education	2,597	51.02
Income Quintile	**5,090**	**100.00**
Lowest	999	19.63
Second	1,000	19.65
Third	1,004	19.72
Fourth	1,047	20.57
Highest	1,040	20.43

From the Table 4.4, the strength of the age category groups tends heavily towards adults, and particularly older adults respectively. Most of the respondents belong to the Christian faith. There is fairly even distribution among males and females respectively. More than half of respondents were engaged in self-employment. Majority of the respondent belongs to the Akan ethnic group.

4.4 Cross tabulation Asthma Disease Condition

The Table 4.4 presents Cross tabulation of Asthma by Respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles. The first column contains the variables: Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles, as well as their categories, column two contains the number of existence of lungs disease and the percentage of the diseases, and last column captured the non-existence of lung disease and it percentages

Table 4.4: Cross tabulation of Asthma by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles

Variables	Existence of Asthma (%)	Non-existence of Asthma (%)
Age		
Young Adult	13(1.63)	786(98.37)
Adult	47(2.79)	1,637(97.21)
Older Adult	113(4.33)	2,494(95.67)
Total (%)	173(3.40)	4,917(96.60)
Religion		
Christianity	141(3.99)	3,395(96.01)
Islam	14(1.76)	780(98.24)
Others	18(2.37)	742(97.63)
Total (%)	173(3.40)	4,917(96.60)
Sex		
Male	82(3.06)	2,600(96.94)
Female	91(3.78)	2,317(96.22)
Total (%)	173(3.40)	4,917(96.60)
Occupation		
Self Employed	148(3.35)	4,268(96.65)
Public	3(5.00)	57(95.00)
Private	1(38.85)	25(96.15)
Informal	11(3.17)	336(96.83)
Others	10(4.15)	231(96.85)
Total (%)	173(3.40)	4,917(96.60)
Ethnicity		
Akan	108(4.23)	2,445(95.77)
Ewe	19(3.04)	606(96.96)
Ga-adangbe	18(4.52)	380(95.48)
Mole-Dagbon	4(3.12)	122(96.83)
Mande-Busanga	13(1.28)	999(98.72)
Others	11(2.93)	365(97.07)
Total (%)	173(3.40)	4,917(96.60)
Marital Status		
Currently Married	73(2.43)	2926(97.57)
Never Married	2(1.41)	140(98.59)
Cohabitating	4(6.45)	58(93.55)
Separated/Divorced	38(5.71)	628(94.29)
Widowed	56(4.59)	1,165(95.41)
Total (%)	173(3.40)	4,917(96.60)
Education		
Primary	43(3.68)	1,124(96.36)
Secondary	43(3.76)	1,100(96.24)
Tertiary	6(3.28)	177(96.72)
Others	81(3.12)	2,516(96.88)
Total (%)	173(3.40)	4,917(96.60)
Income Quintiles		
lowest	30(3.00)	969(97.00)
second	34(3.40)	966(96.60)
third	37(3.69)	967(96.31)
fourth	37(3.53)	1,010(96.47)
highest	35(3.37)	1005(96.63)
Total (%)	173(3.40)	4917(96.60)

The results from Table 4.4 shows that, there is a corresponding increase in Asthmatic condition as Ghanaians are ageing. Christians also reported Asthmatic conditions higher than the respondents with Islamic faith. Females and the self-employed do report higher levels of Asthmatic conditions. The Akan and the Ga-Adangbe tribes, cohabitating and separated/divorced couples have higher prevalence of Asthmatic conditions among the respondents. In addition, there was not much difference in prevalence rate in all levels of the respondents in terms of their level of education as well as the income quintiles

4.5 Cross tabulation of Depression Disease Condition

In the Table 4.5 below presents Cross tabulation of Depression by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles. The first column contains the variables: Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles, as well as their categories, column two contains the number of existence of lung disease and the percentage of the disease, and last column captured the non-existence of lung disease and it percentage

Table 4.5: Cross tabulation of Depression by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles

Variables	**Existence of Depression** (%)	**Non-existence of Depression** (%)
Age		
Young Adult	2(0.25)	797(99.75)
Adult	11(0.65)	1673(99.35)
Older Adult	56(2.15)	2551(97.85)
Total (%)	69(1.36)	5021(98.64)
Religion		
Christianity	53(1.50)	3483(98.50)
Islam	4(0.50)	790(99.50)
Others	12(1.58)	748(98.42)
Total (%)	69(1.36)	5021(98.64)
Sex		
Male	21(0.78)	2,661(99.22)
Female	48(1.99)	2360(98.01)
Total (%)	69(1.36)	5021(98.64)
Occupation		
Self Employed	59(1.34)	4357(98.66)
Public	0(0.00)	60(100.00)
Private	2(7.69)	24(92.31)
Informal	5(1.44)	342(98.56)
Others	3(1.24)	238(98.76)
Total (%)	63(1.36)	5021(98.64)
Ethnicity		
Akan	42(1.65)	2511(98.35)
Ewe	9(1.44)	616(98.56)
Ga-adangbe	10(2.51)	388(97.49)
Mole-Dagbon	1(0.79)	125(99.21)
Mande-Busanga	5(0.49)	1007(99.51)
Others	2(0.53)	374(99.47)
Total (%)	69(1.36)	5021(98.64)
Marital Status		
Currently Married	23(0.77)	2976(99.23)
Never Married	0(0.00)	142(100.00)
Cohabitating	2(3.23)	60(96.77)
Separated/Divorced	17(2.55)	649(97.45)
Widowed	27(2.21)	1,194(97.79)
Total (%)	69(1.36)	5021(98.64)
Education		
Primary	23(1.97)	1,144(98.03)
Secondary	8(0.70)	1,135(99.30)
Tertiary	3(1.64)	180(98.36)
Others	35(1.35)	2,562(98.65)
Total (%)	69(1.36)	5041(98.64)
Income Quintiles		
lowest	7(0.70)	992(99.30)
second	17(1.17)	983(98.30)
third	14(1.39)	990(98.61)
fourth	15(1.43)	1,032(98.57)
highest	16(1.54)	1024(98.46)
Total (%)	9(1.36)	5021(98.64)

In Table 4.5, Ghanaians do have a positive gradient in percentage of depression conditions as they are getting to their prime ages of life. Christians experienced a higher degree of suffering from depression as compared to their Muslims counterparts. Females do experience a little condition of the disease over their Ghanaians males. Also, Ghanaians who were in private employment surprisingly reported greater percentage of contracting depression. However, public employees did not experience the condition, while the other occupation reported the conditions fairly around the same region. Interestingly, respondents who were from the Ga-Adange ethnic tribe showed more depression condition as compared to the other groups. However, in terms of marital status, cohabiting couples reported more depression condition followed by separated/divorced couples; whiles respondents who never married do not report any depression condition. Income Quintiles exhibit fairly a linear gradient of the condition.

4.6 Logistic Regression Model of Chronic Disease Conditions

In this section we obtain the logistic regression model for the chronic disease condition of the respondents. We then estimate the parameters in the model using the method of maximum likelihood. Finally we performed the odds ratio analysis, predicting chronic illness for the various risk factors.

4.7 Model Formulation

The principal objectives of this study were to formulate a logistic regression model suitable for prediction of chronic disease conditions in Ghana. To achieve this, the collected data was analyzed using STATA (Version 12.1). The explanatory (socio-economic/socio-demographic) variables considered were all entered and the significance and the adequacy of the factors and model respectively determined. To determine the significant effect of socio-demographic and socio-economic factors on chronic illness development, a binary logistic regression model was specified. In the fitted model each chronic disease considered, as dependent variable was treated as existence or non-existence

4.7.1 Asthma Model

The Table 4.6 and Table4.7 showed the model fit statistics and the analyzed results when all the factors were entered into logistic model respectively. From Table4.7 the first column labelled parameters contains all the factors and their categories included in the model and the intercept. The column two labelled estimates contains the coefficients of the estimated parameters, whiles standard error estimates, the calculated test statistic and p-value are in third , fourth and fifth columns respectively.

Table 4.6: Model fit Statistics for Asthma Disease Condition

Number of Observations	Likelihood Ratio Test	Degrees of Freedom	P-Value
5090	63.42	26	<0.001

Table 4.7: Estimated Parameters for Asthmatic Chronic Illness Model

Parameters	Estimates	Std. error	Z	P-value
Intercept	-3.979	0.387	-10.27	**0.000**
Age				
Young Adult	ref			
Adult	0.473	0.326	1.45	0.146
Older Adult	0.931	0.320	2.91	**0.004**
Religion				
Christians	ref			
Muslims	-0.447	0.355	-1.26	0.208
Others	-0.261	0.274	-0.95	0.341
Sex				
Male	ref			
Female	-0.184	0.198	-0.93	0.353
Occupation				
Self-Employment	ref			
Public Employment	0.457	0.611	0.75	0.455
Private Employment	0.374	1.040	0.36	0.719
Others	0.342	0.343	1.00	0.319
Informal Employment	-0.087	0.321	-0.27	0.787
Ethnicity				
Akan	ref			
Ewe	-0.280	0.257	-1.09	0.275
Ga-Adagbe	0.051	0.265	0.19	0.849
Mole-Dagbon	0.185	0.605	0.31	0.760
Mande-Busanga	-0.864	0.337	-2.56	**0.010**
Others	-0.150	0.339	-0.44	0.658
Marital Status				
Currently Married	ref			
Never Married	-0.272	0.735	-0.37	0.712
Cohabitating	1.023	0.542	1.89	0.059
Separated/Divorced	0.713	0.230	3.10	**0.002**
Widowed	0.536	0.230	2.33	**0.020**
Educational Level				
Primary	ref			
Secondary	0.050	0.227	0.22	0.827
Tertiary	-0.091	0.462	-0.20	0.845
Others	-0.152	0.212	-0.72	0.472
Income Quintiles				
lowest	ref			
lower	0.044	0.258	0.17	0.865
Moderate	0.086	0.257	0.33	0.739
High	0.046	0.262	0.18	0.860
Highest	-0.021	0.278	-0.07	0.941

The likelihood ratio statistics measures the adequacy of the model. From the model fit statistic Table 4.6 the LRT = 63.42 with p-value less than 0.001

at 5% level of significance indicates the binary logistic model adequately fits the data. The final socio-demographic/socio-economic factors that contribute significantly to Asthmatic chronic conditions are given as: $logit(Y = 1) = -3.979 + 0.931age_3 - 0.864eth_5 + 0.713mar_4 + 0.536mar_5$ where age_3 = Older Adult Ghanaians,eth_5 = Mande-Busanga, mar_4 = Separated/Divorced couples, and mar_4 = Widowed Ghanaians

4.7.2 Depression Model

In the Table 4.8 and Table 4.9 showed the model fit statistics and the estimated parameters for depression chronic illness. The first column in the Table 4.9 labelled parameters contains all the factors and their categories included in the model and the intercept. The column two labelled estimates contains the coefficients of the estimated parameters, whiles standard error estimates, the calculated test statistic and p-value are in third , fourth and fifth columns respectively.

Table 4.8: Model fit Statistics for Depression Model

Number of Observations	Likelihood Ratio Test	Degrees of Freedom	P-Value
4894	78.69	24	<0.001

Table 4.9: Estimated Parameters for Depression Chronic Illness Model

Parameter	Estimates	Std. Err	Z	p-value
Intercept	-6.772	0.882	-7.68	0.000
Age				
Young Adult	ref			
Adult	0.939	0.790	1.19	0.235
Older Adult	2.242	0.764	2.94	0.003
Religion				
Christians	ref			
Muslims	-0.341	0.631	-0.54	0.589
Others	0.563	0.363	1.55	0.121
Sex				
Male	ref			
Female	0.729	0.334	2.18	0.029
Occupation				
Self-Employment	ref			
Public Employment	-	-		
Private Employment	2.138	0.825	2.59	0.010
Informal Employment	0.032	0.476	0.07	0.947
Others	-0.017	0.613	-0.03	0.978
Ethnicity				
Akan	ref			
Ewe	-0.240	0.384	-0.62	0.532
Ga-Adagbe	0.346	0.373	0.93	0.355
Mole-Dagbon	-0.240	1.153	-0.21	0.835
Mande-Busanga	-0.904	0.546	-1.66	0.098
Others	-0.863	0.748	-1.15	0.249
Marital Status				
Currently Married	ref			
Never Married	-	-		
Cohabitating	1.773	0.781	2.27	0.023
Separated/Divorced	0.553	0.368	1.50	0.133
Widowed	0.226	0.353	0.64	0.522
Educational Level				
Primary	ref			
Secondary	-0.896	0.428	-2.09	0.037
Tertiary	-0.109	0.654	-0.17	0.868
Others	-0.665	0.300	-2.22	0.027
Income Quintiles				
lowest	ref			
lower	0.808	0.458	1.76	0.078
Moderate	0.649	0.475	1.37	0.171
High	0.738	0.476	1.55	0.121
Highest	0.856	0.490	1.75	0.081

The likelihood ratio statistics measures the adequacy of the model. From the model fit statistic the Table 4.8 LRT = 78.69 with p-value less than 0.001 at 5%

level of significance indicated the binary logistic model adequately fitted the data. The final socio-demographic/socio-economic factors that contribute significantly to Depression chronic conditions are given as:
$logit(Y = 1) = -6.772 + 2.242age_3 + 0.729sex_2 + 2.138occ_2 + 1.773mar_3 - 0.665edu_4$. Where age_3 = Older Adult Ghanaians, sex_2 = Female respondent, occ_2 = public Worker, mar_3 = Ghanaians couples who are Cohabiting,with edu_4 = Informal Education

4.8 Odds Ratios

Odds are the probability of an event occurring divided by the probability of the event not occurring. An odds ratio is the odds of the event in one group, thus those exposed to Chronic disease conditions, divided by the odds in another group not exposed Chronic disease conditions. Odds are simply a different expression of the probability: the probability of an event (existence of chronic disease condition) divided by the probability of the event not happening(non-existence of chronic disease condition). Thus, the odds of chronic disease condition would be (probability of chronic disease)/(1-probability of chronic disease).

4.8.1 The Odds Ratios for Asthmatic and Depression conditions

The Table 4.10 below shows the factors in column one and the odds ratio for the Asthma and the Depression conditions in column two and three respectively

Table 4.10: Odds Ratios for Asthmatic and Depression conditions

Factors	Asthma	Depression
Age		
Young Adult	1.000	1.000
Adult	1.605	2.557
Older Adult	2.537	9.410
Religion		
Christianity	1.000	1.000
Islam	0.640	0.711
Others	0.770	1.756
Sex		
Male	1.000	1.000
Female	0.832	2.073
Occupation		
Self Employed	1.000	1.000
Public	1.579	1.000
Private	1.454	8.480
Informal	1.408	0.983
Others	0.917	1.032
Ethnicity		
Akan	1.000	1.000
Ewe	0.755	0.786
Ga-Adangbe	1.052	1.413
Mole-Dagbon	1.203	0.786
Mande-Busanga	0.421	0.405
Others	0.861	0.422
Marital Status		
Currently Married	1.000	1.000
Never Married	0.762	1.000
Cohabitating	2.780	5.888
Separated/Divorced	2.041	1.738
Widowed	1.7085	1.253
Education		
Primary	1.000	1.000
Secondary	1.051	0.408
Tertiary	0.914	0.897
Others	0.859	0.515
Income Quintiles		
lowest	1.000	1.000
second	1.045	2.243
third	1.089	1.914
fourth	1.048	2.092
highest	0.980	2.354
Model-Fit Statistics		
No. of obs.	5090	5090
LRT	63.42	78.69
Pseudo R^2	0.042	0.109
P-Value	<0.001	<0.001
AIC	1500.712	696.44

From Table 4.10 Ageing Ghanaians were 2.537 times as likely to suffer from asthmatic ailments as to their younger adults Ghanaians. Also, there is an indication that, older adults are 9.410 times as likely to suffer from depression conditions as to their young adults age category. Interestingly, female respondents are 2.073 times as likely to suffer depression to the male respondents. Again, respondents who are employed in the various sectors of the economy indicated that, those who are in the private sector of employment are 8.480 times as likely to be asthmatic patients. Ghanaians who are cohabiting are 2.780 and 5.888 times as likely to suffer from asthmatic and depression conditions as compared to their fellow Ghanaians who are currently married respectively. There is an indication that, Ghanaians who are separated/divorced are 2.041 times as likely to be asthmatic than those who are currently married. There is an interesting indication that, widowed in Ghana are 1.709 times as likely to have depression disease conditions than those Ghanaians who are currently married. For the Ghanaians who responded to stated level of education, respondents who attained the secondary level of education are 0.408 times as likely to have depression than those who attained the primary level of education. It is quite interesting that, Ghanaians who have had other levels of education than those stated are 0.515 times as likely to have depression conditions as to their fellow Ghanaians who attained the primary level of education. In another development, persons who are in the second income quintile are 2.243 times more likely to have depression than persons in the lowest income quintile. While Ghanaians who are in the highest income quintile are 2.354 times as likely to suffer from depression than those Ghanaians in the lowest income quintile Table.

From the Table 4.10 above, the Model-Fit Statistics showed that the binary logistic regression model adequately fitted the data, with the model likelihood ratio test (LRT)= 78.63 and the corresponding significance values or the p-values (<0.001) reported was less than 5% level of significance. However, Alkaike Information Criterion was also reported for each model, comparatively, model for

Depression was the best model that actually fitted the data.

From the results shown in the Appendix B (Table 5.1), Adult and Older Adult Ghanaians are 2.967 and 3.994 times as likely to suffer from lung disease as their younger adult counterparts respectively. Similarly, Ghanaians who are in other religious faith beside the stated ones are 0.132 times as likely to have lung disease as compared to the dominant Christian faith in Ghana. Also, there is an indication that, older adults are 3.285 times as likely to suffer from oral health conditions as to their young adults age category. Respondents who are employed in the various sectors of the economy indicated that, those who are in the private sector of employment are 9.971 times as likely to suffer from lung disease. There is an indication that, Ghanaians who are separated/divorced are 1.686 times as likely to report oral health diseases conditions to those who are currently married. Also, widowed in Ghana are 1.765 times as likely to have oral health disease conditions than those Ghanaians who are currently married. Surprisingly, Ghanaians who attained the tertiary level of education are much higher to have lung diseases than those at the primary level of education. Ghanaians who have had other levels of education than those stated are 1.598 times as likely to have oral health conditions as to their fellow Ghanaians who attained the primary level of education. While Ghanaians in highest income quintile are 2.130 times as likely to suffer from injuries to Ghanaians in the lowest income quintile.

From Table 5.2, (Appendix B) Lung Disease conditions increases as respondents' were ageing. Also Christians are reporting higher prevalence of lung diseases as compared to their Muslim counterparts. Male respondents do experience higher degree of lung disease to their female respondents. However, in all, 55% of the respondents by their contributing factors namely; age, religion, sex and occupation categories, have lung disease. Also, Lung disease conditions were higher in the Mande-Busanga ethnic group, and lower in Ewe, followed by the Ga-Adangbes. However, the Akan, Mole-Dagbani, and the other ethnic groups reported fairly stable gradient. Interestingly, Ghanaians who were never married and cohabitating showed no lung disease condition. Currently married respondents reported the least existence of lung disease. While separated/divorced, widowed couples were around the same gradient. An increasing condition in disease was reported as level education increases from primary to tertiary levels of education, except other educational level with the minimum. Income Quintiles showed a gradient.

A Cross tabulation of Oral Health by respondents Age, Religion, Sex, and Occupation are presented in Table 5.3,(Appendix B). From the table it showed that older adult have higher association with Oral health conditions, while respondents of the Islamic faith do have higher marginal existence of oral health conditions. Also male respondents are associated with the conditions and as well as privately employed respondents as compared to their peers in the same categories. It also shows that, Akan ethnic group do have a high prevalence oral health likewise widowed respondents and other educational levels. However there was not much difference in terms of respondents income levels.

In Table 5.4,(Appendix B) there was a marginal difference in percentage of Injury conditions among the age categories with adult respondents reporting higher rate. Surprisingly, public workers reports higher prevalence of the conditions as well as Muslims. It is also observed that the Mole-Dagbon and others tribes, separated/divorced couples, tertiary level of education as well as the fifth income quintiles do experience higher prevalence of Injuries conditions among Ghanaians.

4.9 Prediction of Chronic Illness Conditions

4.9.1 Prediction for Asthmatic conditions

The Model for Asthmatic chronic conditions is given as: $logit(Y = 1) = -3.979 + 0.931age_3 - 0.864eth_5 + 0.713mar_4 + 0.536mar_5$ where age_3=Older Adult Ghanaians,eth_5=Mande-Busanga, mar_4=Separated/Divorced couples, and mar_4=Widowed Ghanaians For a Ghanaian Older Adult Mande-Busanga, Separated/Divorced, then the odds of the conditions is

$$Odds = e^{-3.979+0.931(1)-0.864(1)+0.713(1)+0.536(0)} = 0.041$$

Given that $\pi(x)$ = probability of existence of asthmatic condition, the case of interest, $\pi(x)=\dfrac{Odds}{1+Odds}$
$\pi(x) = 0.039$
1-$\pi(x)$= 0.961,
This simply means that 3.9% of Older Adult Ghanaians, separated/divorced, from the Mande-Busanga ethnic extraction, would suffer Asthmatic condition and 96.1% of Ghanaians with these socio-demographic/socio-economic factors will not suffer the Asthmatic condition.

4.9.2 Prediction for Depression conditions

The Model for Depression chronic conditions is given as:
$logit(Y = 1) = -6.772 + 2.242age_3 + 0.729sex_2 + 2.138occ_2 + 1.773mar_3 - 0.665edu_4$. Where age_3 = Older Adult Ghanaians,sex_2 = Female respondent,occ_2 = public Worker, mar_3 = Ghanaians couples who are Cohabiting,with edu_4 = Informal Education. For an ageg Ghanaian, who works at the public sector,

separated /divorced, with informal education, then odds of the condition is:

$$Odds = e^{-6.772+2.242(1)+0.729(1)+2.138(1)+1.773(1)-0.665(1)} = 0.574$$

Given that $\pi(x)$ = probability of existence of depression condition, the case of interest, $\pi(x)=\dfrac{Odds}{1+Odds}$ $\pi(x) = 0.37$ 1-$\pi(x)$= 0.63, This means that 37% of aged female Ghanaians, employed at the public sector, who is separated/divorced with informal education would suffer from Depression condition and 63% of Ghanaians with the same socio-demographic/socio-economic factors will not suffer from the Depression condition

4.10 Discussion

This study established the effect of influential socio-demographic/socio-economic factors on chronic illness conditions of Ghanaians on social and medical science literature. Research conducted in various countries shows that particularly many problems are associated with health-related quality of elderly people's lives (Van-Gool et al., 2007; Paúl et al., 2007; Dugan and Lee, 2013). Accordingly, the avoidance of illness as well as the preservation of physical and cognitive functions is some of the most important factors that significantly improve the quality of life in the elderly age. Different studies also conclude that health is one of the most prized values for elderly people (Bowling and Gabriel, 2007). Poor health is associated with the loss of control, autonomy and independence, and it makes people aware of the approaching death. Asthma recorded the highest prevalence conditions with 3.4%, while Lung Disease recorded the least, 0.6% However, Depression, Oral health, and Injuries fairly recoded 1.4%, 2.5% and 1.8% respectively. Also, Ghanaians considered for the research tends heavily towards adults, and particularly older adults respectively. Most of the respondents belong to the Christian faith. There is fairly even distribution among males and females respectively. More than half of respondents were engaged in self-employment. Again, majority of the respondent belongs to the Akan ethnic group. In addition, Adult and Older Adult Ghanaians are likely to suffer lung disease as to their younger adult counterparts. Ageing Ghanaians were likely to suffer asthmatic ailments as to their younger adults Ghanaians. Also, there is an indication that, older adults are likely to suffer depression and oral health conditions as to their young adults' age category. Similarly, Ghanaians who are in other religious faith beside the stated ones do have lung disease as compared to the dominant Christian faith in Ghana. Interestingly, female respondents were reporting depression than the male respondents. Again, respondents who are employed in the various sectors of the economy indicated that, those who are in the private sector of employment were likely to suffer from lung disease as well as asthma.

A study by Mary (2003) indicated that diseases like cancer, heart disease and asthma disproportionately affect low income, ethnically diverse communities as it is partly collaborated this study.

However, there is an indication that, cohabiting, separated/divorced and widowed couples in Ghana do have a greater chance to be asthmatic and report oral health and well as depression diseases conditions than those who are currently married, these was also reported in a study by (Ziebland and Kokanovic, 2012)

The study however, did not support, Chronic diseases and poverty are interconnected in a vicious circle, (WHO 2005). As higher level of education and income quintiles are associated with lung disease, depression and injuries. Our study partly support the 1981 assessment of the health impact of various diseases in Ghana, ranked in order of healthy days of particular sedentary occupations and consumption of a wider diversity of local and foreign foods, the urban wealthy are not the only high risk groups for chronic diseases in Ghana.

The Model-Fit Statistics showed that the binary logistic regression model adequately fitted the data, with the model likelihood ratio test (LRT)= 78.63 and the corresponding significance values or the p-values (<0.001) reported was less than 5% level of significance. However, Alkaike Information Criterion was also reported for each model, comparatively, model for Depression was the best model that actually fitted the data

Chapter 5

Conclusion and Recommendation

5.1 Conclusion

This chapter presents the main outcomes of the study and conclusion and recommendation on the chronic diseases condition in Ghana. The thesis has sought to determined the impact of socio-demographic/socio-economic factors on chronic disease conditions in Ghana. The influences of socio-demographic/socio-economic factors on chronic disease conditions are effectively treated and the contribution of the factors determined. In addition, the significance of each of the contributing factors and as well as the adequacy of the model determined.

The thesis has therefore treated adequately the thesis topic: Analysis of Chronic Disease Condition in Ghana Using Logistic Regression. The study has largely formulated many models and the adequacy of each models that best fit the data determined.

Generally the objectives of the study was largely met and revealed some level of chronic diseases in the Ghanaian populations. Asthma recorded the highest prevalence conditions with 3.4%, while Lung Disease recorded the least, 0.6 %. However, Depression, Oral health, and Injuries fairly recoded 1.4%, 2.5% and 1.8% respectively. In addition, most of the respondents belong to the Christian faith. There is fairly even distribution among males and females respectively. More than half of respondents were engaged in self-employment and the majority of the respondent belongs to the Akan ethnic group. Also, widowed respondents contracted depression and oral health disease conditions. This thesis revealed that, in Ghana, the occurrences of chronic disease conditions are associated with some socio-demographic/socio-economic factors: age, sex, religion, ethnicity, marital status, occupation, level of education, and income levels. Model-Fit Statistics shows that the binary logistic regression model adequately fits the, since in almost all cases the model likelihood ratio test (LRT) and the corresponding significance values or the p-values reported for each model was less than 5% level of significance.

- The risk factors that significantly contributed to Asthmatic Chronic conditions are: $logit(Y = 1) = -3.979 + 0.931 age_3 - 0.864 eth_5 + 0.713 mar_4 + 0.536 mar_5$.

Where the risk factors are Older Adult Ghanaian (age_3), Mande-Busanga (eth_5), Separated/Divorced (mar_4)

- The risk factors that significantly contributed to Depression Chronic conditions are: $logit(Y = 1) = -6.772 + 2.242 age_3 + 0.729 sex_2 + 2.138 occ_2 + 1.773 mar_3 - 0.665 edu_4$.

Where age_3 = Older Adult Ghanaians, sex_2 = Female respondent, occ_2 = public Worker, mar_3 = Ghanaians couples who are Cohabiting, with edu_4 = Informal Education

- Asthma recorded the highest prevalence conditions followed by Oral health; Injuries, Depression and Lung disease.

5.2 Recommendation

- The stake holders such as Government, MOH and WHO, should consider developing a Chronic disease policy.
- The National Health Insurance Scheme (NHIS) should consider coverage for chronic conditions in Ghana to reduce death as a result of chronic illness.
- This could benefit both chronic disease patients and health practitioners concerned in dealing with the various health conditions.
- Further studies are needed to determine psycho-social and physical risk factors at work, associated with the chronic disease conditions.

REFERENCES

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Willey-Interscience: A John Willey and Sons,Inc.

Ahern, M. M. and Hendryx, M. S. (2005). Social capital and risk for chronic illnesses. *Chronic Illness*, 1:183–190.

Anderson, G. and Horvath, J. (2004). The growing burden of chronic disease in america. *Practice Articles*, 119.

Benzo, R. (2012). Collaborative self-management in chronic obstructive pulmonary disease: Learning ways to promote patient motivation and behavioral change. *Chronic Respiratory Disease*, 9(4):257–258.

Bowling, A. and Gabriel, Z. (2007). Lay theories of quality of life in older age. *Ageing and Society*, 27:827–848.

Dabbs, A. D., Song, M.-K., Geest, S. D., and Davidson, P. (2013). Promoting patient and caregiver engagement in self-management of chronic illness. *Nursing Research and Practice*, 2013, Article ID 180757, 2 pages.

de Graft Ama, A. (2007). Ghana's neglected chronic disease epidemic: a developmental challenge. *Ghana Medical Journal*, 41(4).

Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC.

Dow, C. M., Roche, P. A., and Ziebland, S. (2012). Talk of frustration in the narratives of people with chronic pain. *Chronic Illness*, 8(3):176–191.

Dugan, E. and Lee, C. M. (2013). Biopsychosocial risk factors for driving cessation: Findings from the health and retirement study. *Journal of Aging and Health*, 25(8):1313–1328.

Durairaj, V., D'Almeida, S., and Kirigia, J. (2010). Ghana's approach to social health protection. Technical report, World Health Report (2010). Background Paper, 2.

Gale, N. S., Bolton, C. E., Duckers, J. M., Enright, S., Cockcroft, J. R., and Shale, D. J. (2012). Systemic comorbidities in bronchiectasis. *Chronic Respiratory Disease*, 9(4):231–238.

Ghana, H. S. (2007). Annual report. Technical report, Ghana Health Service.

Ghana, S. S. and Ghana, H. S. (2009). Ghana demographic and health survey 2008. Technical report, Ghana Statistical Service, Ghana Health Service and ICF Macro ,Accra, Ghana.

Girón, M., Nolasco, A., Pereyra-Zamora, P., Munarriz, M., Salazar, J., Tabarés, R., Moncho-Vasallo, J., and Gómez-Beneyto, M. (2009). Self-reported chronic mental health problems and mental health service use in spain. *Chronic Illness*, 5:197–208.

Gittelsohn, J., Song, H.-J., Suratkar, S., Kumar, M. B., Henry, E. G., Sharma, S., Mattingly, M., and Anliker, J. A. (2010). An urban food store intervention positively affects food-related psychosocial variables and food behaviors. *Health Education & Behavior*, 37(3):390–402.

G.S.S. (2012). 2010 Population and Housing Census. Technical report, Ghana Statistical Service.

Gugiu, P. C., Coryn, C., Clark, R., and Kuehn, A. (2009). Development and evaluation of the short version of the patient assessment of chronic illness care instrument. *Chronic Illness*, 5:268–276.

Haslbeck, J. W. and Schaeffer, D. (2009). Routines in medication management: the perspective of people with chronic conditions. *Chronic Illness*, 5:184–196.

Heiner, B., Angela, B., Achim, B., Sabine, E., Dirk, H., Anja, K., Eva, L.-B., Manfred, J. M., Helmut, O., Matthias, S., Peter, S., and Bernhard, W. (2012). Critical review: vegetables and fruit in the prevention of chronic diseases. *Eur J Nutr*, 51:637–663.

Idowu, O., Adeniyi, A., Atijosan, O., and Ogwumike, O. (2012). Physical inactivity is associated with low self efficacy and social support among patients with hypertension in nigeria. *Chronic Illness*, 9(2):156–164.

Jerliu, N., Toçi, E., Burazeri, G., Ramadani, N., and Brand, H. (2013). Prevalence and socioeconomic correlates of chronic morbidity among elderly people in kosovo: a population-based survey. *BMC Geriatrics*, 13:22.

Kung, H.-C., Hoyert, D. L., Xu, J., and Murphy., S. L. (2008). U.s. department of health & human services.national vital statistics system results. *Centers for Disease Control and Prevention National Center for Health Statistics*, (56)10.

Kyngäs, H. A. (2007). Predictors of good adherence of adolescents with diabetes(insulin-dependent diabetes mellitus). *Chronic Illness*, 3:20–28.

Lavie-Ajayi, M., Almog, N., and Krumer-Nevo, M. (2012). Chronic pain as a narratological distress: a phenomenological study. *Chronic Illness*, 8(3):192–200.

Lindmark, U., Stegmayr, B., Nilsson, B., Lindahl, B., and Johansson, I. (2005). Food selection associated with sense of coherence in adults. *Nutrtion Journal*, (4)9.

Mary, A. M. (2003). Health departments and communities mobilize to prevent chronic disease. *Contra Costa Health Services*.

McKay, A. J., PA, M., Patel, R. K. K., and Majeed, A. (2012). Management of chronic obstructive pulmonary disease in india: a systematic review. *Journal of the Royal Society of Medicine Short Reports*, 3:85.

Munish, M., Sharma, V., Yarussi, K. M., Sifain, A., Porhomayon, J., and Nader, N. (2012). The use of practice guidelines by the american society of anesthesiologists for the identification of surgical patients at high risk of sleep apnea. *Chronic Respiratory Disease*, 9(4):221–230.

Nepal, V. P., Mgbere, O., Banerjee, D., and Arafat, R. (2012). Determinants of fruits and vegetables consumption among persons with doctor-diagnosed chronic diseases. *Journal of Primary Care & Community Health*, 3(2):132–141.

Nieman, L. Z. (2009). Chronic condition self-management and two teaching models for chronic conditions. *Chronic Illness*, 5:15–17.

Paúl, C., Ayis, S., and Ebrahim, S. (2007). Disability and psychosocial outcomes in old age. *Journal of Aging and Health*, 19 (5):723–741.

Pols, R. G., Battersby, M. W., Regan-Smith, M., Markwick, M. J., Lawrence, J., Auret, K., Carter, J., Cole, A., Disler, P., Hassed, C., McGuiness, C., and Hung (2009). Chronic condition self-management support: proposed competencies for medical students. *Chronic Illness*, 5:7–14.

Rodham, K., McCabe, C., Pilkington, M., and Regan, L. (2012). Coping with chronic complex regional pain syndrome: advice from patients for patients. *Chronic Illness*, 9(1):29–42.

Simpson, J., Lekwuwa, G., and Crawford, T. (2013). Illness beliefs and psychological outcome in people with parkinson's disease. *Chronic Illness*, 9(2):165–176.

Singh, K. A. and Brennan, D. S. (2012). Chewing disability in older adults attributable to tooth loss and other oral conditions. *The Gerodontology Society and John Wiley & Sons A/S, Gerodontology*, 29:106–110.

Slama, S., Loutan, L., Wernli, D., Verma, S., and Beran, D. (2012). Chronic conditions: lessons from the frontlines. *Chronic Illness*, 9(2):83–86.

Strachan, S. M. and Brawley, L. R. (2009). Healthy-eater identity and self-efficacy predict healthy eating behavior: A prospective view. *Journal of Health Psychology*, 14(5):684–695.

Suhrcke, M., Nugent, R. A., Stuckler, D., and Rocco, L. (2006). Chronic disease: An economic perspective. *Oxford Health Alliance*.

Tommis, Y., Robinson, C. A., Seddon, D., Woods, B., Perry, J., and Russell, I. T. (2009). Carers with chronic conditions: changes over time in their physical health. *Chronic Illness*, 5:155–164.

Van-Gool, C. H., Kempen, G. I. J. M., Penninx, B. W. J. H., Deeg, D. J., and van Eijk, J. T. M. (2007). Chronic disease and lifestyle transitions results from the longitudinal aging study amsterdam. *Journal of Aging and Health*, 19 (3):416–438.

Ventegodt, S., Merrick, J., and Andersen, N. J. (2003). Quality of life as medicine: A pilot study of patients with chronic illness and pain. *TheScientificWorldJOURNAL*, 3:520–532.

Walker, A. E. (2007). Multiple chronic diseases and quality of life: patterns emerging from a large national sample, australia. *Chronic Illness*, 3:202–218.

Walker, C., Peterson, C., and Millen, N. (2001). Investigating socio-economic status and chronic illness: Directions and dilemmas. In *TASA 2001 Conference, The University of Sydney*.

W.H.O. (2005). Preventing chronic diseases : a vital investment : WHO Global Report. Technical report.

W.H.O. (2007). Prevention of cardiovascular disease: guidelines for assessment and management of total cardiovascular risk. Technical report, World Health Organization.

WHO (2008). 2008-2013 action plan for the global strategy for the prevention and control of noncommunicable diseases. : prevent and control cardiovascular diseases, cancers, chronic respiratory diseases and diabetes. Technical report, World Health Organization.

W.H.O (2013). Fact sheet 317. cardiovascular diseases (cvds). who action plan for the strategy for prevention and control of noncommunicable. Technical report, World Health Organization (Updated March 2013).

Ziebland, S. and Kokanovic, R. (2012). Emotions and chronic illness. *Chronic Illness*, 8:159–162.

Appendix A

5.3 Definition of Variables

The chronic diseases were defined as

$$Y = \begin{cases} 1 = & \text{existence of chronic disease} \\ 0 = & \text{non-existence of chronic disease} \end{cases}$$

Transformation of the factors

Age Categories

$age_1 = 0$ for age = younger adult (reference category)

$age_2 = 1$ for age = adult, 0 otherwise

$age_3 = 1$ for age = older adult, 0 otherwise

Religion Categories

rel_1=0 for religion = Christianity (reference category)

rel_2=1 for religion = Islam, 0, otherwise

rel_3=1 for religion = Others, 0, otherwise

Sex Categories

sex_1=0 for sex = male (reference category)

sex_2=1 for sex = female, 0, otherwise

Occupation Categories

occ_1=0 for occupation = self employment (reference category)

occ_2=1 for occupation = public employment,0,otherwise

occ_3=1 for occupation = private employment,0,otherwise

occ_4=1 for occupation = informal employment,0,otherwise

occ_5=1 for occupation = other employment,0,otherwise

Ethnicity Categories

eth_1=0 for ethnicity = Akan (reference category)

eth_2=1 for ethnicity = Ewe, 0,otherwise

eth_3=1 for ethnicity = Ga-adangbe, 0,otherwise

eth_4=1 for ethnicity = Mole-dagbon, 0,otherwise

eth_5=1 for ethnicity = Mande-Busanga, 0,otherwise

eth_6=1 for ethnicity = Other tribes, 0,otherwise

Marital Status Categories

mar_1=0 for marital status = currently married (reference category)

mar_2=1 for marital status = never married, 0, otherwise

mar_3=1 for marital status = cohabiting' 0, otherwise

mar_4=1 for marital status = separated/divorced, 0, otherwise

mar_5=1 for marital status = widowed, 0, otherwise

Educational Level Categories

edu_1=0 for educational level = primary (reference category)

edu_2=1 for educational level = secondary, 0, otherwise

edu_3=1 for educational level = tertiary, 0, otherwise

edu_4=1 for educational level = no education, 0, otherwise

Income Quantiles Categories

iq_1=0 for income quintile = lowest (reference category)

iq_3=1 for income quintile = second,0,otherwise

iq_4=1 for income quintile = third,0,otherwise

iq_5=1 for income quintile = fourth,0,otherwise

iq_6=1 for income quintile = highest,0,otherwise

Appendix B

5.4 Estimated Odds rations of the selected chronic diseases and the socio-demographic and socio-economic factors

Table 5.1: Odds Rations for Lung Diseases, Oral health and Injuries

Factors	Lungs Diseases	Oral Health	Injuries
Age			
Young Adult	1.000	1.000	1.000
Adult	2.967	1.752	1.773
Older Adult	3.994	3.283*	1.486
Religion			
Christianity	1.000	1.000	1.000
Islam	0.399	1.355	0.975
Others	0.132***	0.846	0.675
Sex			
Male	1.000	1.000	1.000
Female	0.559	0.879	1.480
Occupation			
Self Employed	1.000	1.000	1.000
Public	1.000	0.809	1.027
Private	9.971**	1.503	2.203
Informal	1.601	0.344	0.440
Others	1.650	0.606	1.287
Ethnicity			
Akan	1.000	1.000	1.000
Ewe	0.312	0.675	0.775
Ga-Adangbe	0.395	0.758	0.524
Mole-Dagbon	2.786	0.916	1.614
Mande-Busanga	1.840	0.587*	1.509
Others	1.673	0.509	1.786
Marital Status			
Never Married	1.000	1.000	1.000
Cohabitating	1.000	0.576	0.432
Separated/Divorced	2.287	1.686***	1.088
Widowed	1.558	1.765**	0.698
Education			
Primary	1.000	1.000	1.000
Secondary	1.880	1.231	1.158
Tertiary	4.822***	1.387	1.126
Others	2.148	1.598***	1.026
Income Quintiles			
lowest	1.000	1.000	1.000
second	0.675	0.864	1.272
third	1.324	0.942	1.442
fourth	0.323	0.910	2.160**
highest	1.149	1.249	2.130***
Model-Fit Statistics			
No. of obs.	5090	5090	5090
LRT	29.45	55.64	25.32
P-Value	>0.01	<0.001	<0.001
Pseudo R^2	0.086	0.048	0.028
AIC	395.405	1165.5	933.420

NOTE::***,** and * Indicates 1%, 5% and 10% significance level.

5.5 Cross tabulation for Lung Diseases, Oral Health and Injuries

Table 5.2: Cross tabulation of Lung Disease by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles

Variables	Existence of Lung Disease (%)	Non-existence of Lung Disease (%)
Age		
Young Adult	1(0.13)	798(99.87)
Adult	8(0.48)	1676(99.52)
Older Adult	19(0.73)	2588(99.27)
Total (%)	28(0.55)	5062(99.45)
Religion		
Christianity	23(0.63)	3513(99.35)
Islam	4(0.50)	790(99.50)
Others	1(0.13)	759(99.45)
Total (%)	28(0.55)	5062(99.45)
Sex		
Male	16(0.60)	2666(99.40)
Female	12(0.50)	2396(99.50)
Total (%)	28(0.55)	5062(99.45)
Occupation		
Self Employed	22(.50)	4394(99.50)
Public	0(0.00)	60(100)
Private	1(3.85)	25(96.15)
Informal	3(0.86)	344(99.14)
Others	2(0.83)	239(99.17)
Total (%)	28(0.55)	5062(99.45)
Ethnicity		
Akan	16(0.63)	2537(99.37)
Ewe	1(0.16)	624(99.840
Ga-Adangbe	1(0.25)	397(99.75)
Mole-Dagbon	1(0.79)	125(99.21)
Mande-Busanga	3(0.80)	373(99.20)
Others	6(0.59)	1006(99.41)
Total (%)	28(0.55)	5062(99.45)
Marital Status		
Currently Married	14(0.14)	2985(99.53)
Never Married	0(0.00)	142(100)
Cohabitating	0(0.00)	62(100)
Widowed	8(0.66)	1213(99.34)
Total (%)	28(0.55)	5062(99.45)
Education		
Primary	3(0.26)	1164(99.74)
Secondary	6(0.52)	1137(99.48)
Tertiary	3(1.64)	180(98.36)
Others	16(0.62)	2581(99.38)
Total (%)	28(0.55)	5062(99.45)
lowest	6(0.60)	993(99.40)
second	4(0.40)	996(99.60)
third	8(0.80)	996(99.20)
fourth	2(0.19)	1045(99.81)
highest	8(0.77)	1032(99.23)
Total (%)	28(0.55)	5062(99.45)

Table 5.3: Cross tabulation of Oral Health by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles

Variables	**Existence of Oral Health** (%)	**Nonexistence of Oral Health** (%)
Age		
Young Adult	6(0.75)	793(99.25)
Adult	27(1.60)	1657(98.40)
Older Adult	93(3.57)	2514(96.43)
Total (%)	126(2.48)	4964(97.52)
Religion		
Christianity	89(2.52)	3447(97.48)
Islam	22(2.77)	772(99.23)
Others	15(1.97)	745(98.03)
Total (%)	126(2.48)	4964(97.52)
Sex		
Male	56(2.09)	2,626(97.91)
Female	70(1.91)	2338(97.08)
Total (%)	126(2.48)	4964(97.52)
Occupation		
Self Employed	11(2.63)	4300(97.37)
Public	1(1.67)	59(98.33)
Private	1(3.85)	25(96.15)
Informal	2(0.83)	239(99.17)
Others	6(1.73)	341(98.27)
Total (%)	126(2.48)	4964(97.52)
Ethnicity		
Akan	75(2.94)	2478(97.06)
Ewe	12(1.92)	613(98.08)
Ga-adangbe	8(2.01)	390(97.99)
Mole-Dagbon	5(3.97)	121(96.03)
Mande-Busanga	20(1.98)	992(98.02)
Others	6(1.60)	370(98.40)
Total (%)	126(2.48)	4964(97.52)
Marital Status		
Currently Married	54(1.80)	2,945(98.20)
Never Married	1(0.70)	141(99.30)
Cohabitating	0(0.00)	62(100.00)
Separated/Divorced	22(3.30)	644(96.70)
Widowed	49(4.01)	1,172(95.99)
Total (%)	126(2.48)	4,964(97.52)
Education		
Primary	19(1.63)	1,148(98.37)
Secondary	22(1.92)	1,121(98.08)
Tertiary	4(2.12)	179(97.81)
Others	81(3.12)	2,516(96.88)
Total (%)	126(2.48)	4964(97.52)
Income Quintiles		
lowest	27(2.70)	972(97.30)
second	23(2.30)	977(97.70)
third	25(2.49)	979(97.80)
fourth	23(2.20)	1,024(97.80)
highest	28(2.69)	1012(97.31)
Total (%)	126(2.48)	4964(97.52)

Table 5.4: Cross tabulation of Injuries by respondents Age, Religion, Sex, Occupation, Ethnicity, Marital Status, Education and Income Quintiles

Variables	**Existence of Injuries** (%)	**Nonexistence of Injuries** (%)
Age		
Young Adult	10(1.25)	789(98.75)
Adult	37(2.20)	1647(97.80)
Older Adult	43(1.65)	2564(98.23)
Total (%)	90(1.77)	5000(98.23)
Religion		
Christianity	64(1.81)	3472(98.19)
Islam	18(2.77)	776(97.73)
Others	8(1.05)	752(98.95)
Total (%)	90(1.77)	5000(98.23)
Sex		
Male	43(1.60)	2,639(97.91)
Female	47(1.95)	2361(97.08)
Total (%)	90(1.77)	5000(97.52)
Occupation		
Self Employed	78(1.77)	4338(98.23)
Public	1(1.67)	59(98.33)
Private	1(3.85)	25(96.15)
Informal	8(2.31)	339(97.69)
Others	2(0.83)	239(99.17)
Total (%)	90(1.77)	5000(98.23)
Ethnicity		
Akan	46(1.80)	2507(98.20)
Ewe	8(1.28)	617(98.72)
Ga-adangbe	4(1.01)	394(98.99)
Mole-Dagbon	3(2.38)	123(97.62)
Mande-Busanga	19(1.88)	993(98.12)
Others	10(2.66)	366(97.34)
Total (%)	90(1.77)	5000(98.23)
Marital Status		
Currently Married	56(1.87)	2,943(98.13)
Never Married	1(0.70)	141(99.30)
Cohabitating	1(1.61)	61(98.39)
separated/Divorced	14(2.10)	652(97.90)
Widowed	18(1.47)	1,203(98.53)
Total (%)	90(1.77)	5000(97.52)
Education		
Primary	19(1.63)	1,148(98.37)
Secondary	23(2.01)	1,120(97.99)
Tertiary	4(2.19)	179(97.81)
Others	44(1.69)	2,553(98.31)
Total (%)	90(1.77)	5000(98.23)
Income Quintiles		
lowest	11(1.10)	988(98.90)
second	14(1.40)	986(98.41)
third	16(1.59)	988(98.61)
fourth	25(2.39)	1,022(97.80)
highest	24(2.31)	1016(97.69)
Total (%)	90(1.77)	5000(98.23)