ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

FACULTY OF SCIENCE

DEPARTMENT OF STATISTICS

SOCIO-DEMOGRAPHIC FACTORS AFFECTING BIPOLAR DISORDER PATIENTS: THE CASE OF BUTAJIRA, ETHIOPIA

By

Aklilu Zemicael

A thesis submitted to the Graduate Studies Program of Addis Ababa University in partial fulfillment of the requirements for the Degree of Masters of Science in BioStatistics

June, 2009

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

FACULTY OF SCIENCE

DEPARTMENT OF STATISTICS

SOCIO-DEMOGRAPHIC FACTORS AFFECTING BIPOLAR DISORDER PATIENTS: THE CASE OF BUTAJIRA, ETHIOPIA

By

Aklilu Zemicael

Approved by the Board of Examiners:

______________________	______________
Department Head	Signature
______________________	______________
Examiner	Signature
______________________	______________
Examiner	Signature

Tele: +251-11-239465 P.O. Box 1176 Addis Ababa, Ethiopia

ACKNOWLEGMENT

First and foremost I would like to express my deepest thanks and respect to my advisor Dr. Fentaw Abegaz for his patience, kind help, continual advice, useful comments and suggestions throughout the preparation of this paper. It is explicable that without his closer follow-up and continuous encouragement with valuable comments this thesis would not have been finalized in its present structure.

I would like to extend my thanks also to my Co-advisor Dr Tosheme Shibrie, Head Department of Psychiatry, Addis Ababa University and to other staff members.

My gratitude and sincere thanks is extended to all my lovely and cooperative families and friends for their continuous encouragement and plateful me in all aspects.

At last but not least I would like to thank all staff members of Statistics Department of Addis Ababa University.

Abstract

The purpose of this study is to identify socio-demographic factors affecting the bipolar disorder patients. The data included in this study are a cross sectional data collected by the Butajira mental health project. In this study, 828 psychiatric patients were included, of which 511(61.75%) belong to the non bipolar disorder symptom, 218(26.3%) belong to the moderate bipolar disorder symptom and 99(12%) belong to severe bipolar disorder symptom. In the analysis of data, polytomous logistic regression and multiple discriminant analysis are employed. It is found that place of residence, quality of life, marital status, number of children, and the interaction between residence and quality of life are significantly associated with the severity of bipolar disorder.

The polytomous logistic regression provides higher percentage of correctly classifying psychiatric patients into no, moderate, and severe bipolar disorder symptom groups (61.7%) as compared to multiple discriminant analysis (50.8%). So that polytomous logistic regression is more efficient than that of multiple discriminant analysis in classifying of the psychiatric patients into the no, moderate and severe bipolar disorder symptom groups. This study identifies that the potential socio-demographic factors such as place of residence, quality of life, number of children, marital status, and age of the patients and the interaction between residence and quality of life affects bipolar disorder patients.

Table of content

Page

CHAPTER IV STATISTICAL DATA ANALYSIS

CHAPTER FIVE CONCLUSION AND DISCUSSION

Acronyms

ADD	Attention Deficit Disorder
ADHD	Attention deficit hyperactivity disorder
BD	Bipolar Disorder
BD I	Bipolar Disorder type I
BD II	Bipolar Disorder type II
CIDI	Composite International diagnosis Interview
DSM-IV	Diagnostic and Statistical Manual of mental disorder 4th edition
ECA	Epidemiological Catchments Area
FLDA	Fisher Linear Discriminant Analysis
FMRI	Functional Magnetic Resonance Imaging
IQ	Intelligence Quotient
LDA	Linear Discriminant Analysis
MD	Major Depression
MDA	Multiple Discriminant Analysis
MRI	Magnetic Resonance Imaging
NCS	National Co morbidity Survey
OCD	Obsessive Compulsive Disorder
OLS	Ordinary Least Square
PET	Positron Emission Tomography
SES	Socio-economic status
WHO	World Health Organization
OPHCC	Office of Population and Housing Census Commission

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any University and all sources of material used for thesis have been duly acknowledged.

Name: Aklilu Zemicael

Signiture:__________

Place: Faculty of Science , Addis Ababa University

Date: June, 2009

This thesis has been submitted for examination with my approval as a university advisor.

Fentaw Abegaz (PHD)

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

Mental health problem is one of the serious problems in the developing and developed countries where many productive persons are suffering from such problems and they could not contribute to their society and country as much as what they could do. Mental illness is a medical condition that disrupts a person thinking, feeling, mood ability to relate to others, and daily functioning. Particularly in the developing, countries mental health is worsening since these countries are characterized by low income, high prevalence of communicable disease and malnutrition, low life expectancy and frequent natural disasters like famine, irregular raining and poorly staffed government and non-government services sectors (World Bank, 1998)

In developing countries there has not been a strengthening effort in the intervention of mental health. However, policy makers have been giving priority to the infectious diseases and malnutrition. Even though these are serious problems of developing countries such as Ethiopia, many people are subjected to morbidity and disabilities as a result of different mental health problems. In order to improve the mental health service the World Health Organization (WHO) adopted two resolutions for the first time in 1988 and 1990. Each state was expected to formulate mental health Policies, Program and Action Plans. Mental health is not the headache only for the developing countries, but it is also for the developed countries (Okasha and Karam, 1998).

The incident rate of the mental disorder in developed countries such as in America is about 26% and in the Asian countries such as China it is about 4%. This variation is explained by studies which show that a lack of a certain dietary nutrients contribute to the development of mental disorders such as essential vitamins, minerals, and omega-3 fatty acids are often deficient in the general population in America and other developed countries, and are exceptionally deficient in patients suffering from mental disorders. Studies have shown that daily supplements of vital nutrients often effectively reduce symptoms. Supplements that contain amino acids also reduce symptoms, because they are converted to neurotransmitters that alleviate depression and other mental disorders. Based on emerging scientific evidence, this form of nutritional supplement treatment may be appropriate for controlling common mental health disorder such as major depression, bipolar disorder, schizophrenia and anxiety disorders, eating disorders, attention deficit disorder/attention deficit hyperactivity disorder (ADD/ADHD), addiction, and autism. The four most common mental disorders currently affecting the most developed and developing countries are major depression, bipolar disorder, schizophrenia, and obsessive compulsive disorder (OCD) (Murray and Lopez, 1996)

About 450 million people worldwide are estimated to be suffering from neuropsychiatry conditions. These conditions include unipolar depressive disorder, bipolar affective disorder, epilepsy, alcohol , post traumatic stress disorder, obsessive and compulsive disorder (WHO, 2001) which also added that mental disorder are not the executive preserve of any special group, they are truly universal. Mental disorder are found people in all regions, all countries and all societies. The prevalence of the mental disorder in both developed and developing countries has been estimated to 25% (Regier et al., 1988).

Especially in the third world countries many people are suffering from the different mental illness such as schizophrenia, depression, mania, bipolar disorder and other disorders since there is no much awareness about the mental illness, due to dietary problem. In addition to these, there is also a shortage of a health centers to give service to the patients such as giving psychiatric service and giving treatment to the patients and others. As a result of these in the developing countries mental disorders were considered as non-lifethreating. Nowadays mental disorders are well recognized in public health problems both in developing and developed countries. Mental health problems contribute 8.1% of the global burden of diseases as estimated by the World Bank (1993). Particularly in the low income countries, mental disorders contribute 12% to the global burden of diseases. In line with these findings, a resent study by Abdulahi et al. (2001) in Ethiopia showed that mental illness contributes to over 12% of the burden of the diseases.

Most Ethiopians believe that mental illness is affliction by supper natural forces, and traditional methods are the most frequently used as a means to treat individuals who developed mental illness rather than the modern treatments. In Ethiopia there is a wide spread belief that severe mental illnesses are due to demon possessions, bewitchment by evil sprites, ancestors` sprite or the evil eye have existed for many years, but the attitude of the public to wards such illness has only recently been addressed (Alem et al., 1995).

In this study we will deal with the bipolar disorder which is one of the serious mental illnesses. Bipolar disorder is a brain disorder that causes unusual shifts in a person’s mood, energy, and ability to do something. The bipolar disorder patients can result in damaged relations, poor job or poor school performance, and even may commit suicide. People with this illness can have full and productive lives if they are not appropriately treated.

A clear understanding of bipolar disorder as a mental illness was recognized by early Chinese authors. The encyclopedist GaoLian (1583)[1] describes the malady in his Eight Treatises on the Nurturing of Life (Ts'un-sheng pa-chien). The earliest written descriptions of a relationship between mania and melancholia are attributed to Aretaeus of Cappadocia. Aretaeus was an eclectic medical philosopher who lived in Alexandria somewhere between 30 and 150 AD (Roccatagliata 1986; Akiskal 1996)[1]. Aretaeus is recognized as having authored most of the surviving texts referring to a unified concept of manic-depressive illness, viewing both melancholia and mania as having a common origin in ‘black bile’ (Marneros, 2001)[1].

The basis of the current conceptualization of manic-depressive illness can be traced back to the 1850s;Baillarger (1854)[1] described to the French Imperial Academy of Medicine a biphasic mental illness causing recurrent oscillations between mania and depression, which he termed as 'dual-form insanity'. Around the same time Falret (1854)[1] presented a description to the Academy on what was essentially the same disorder, and designated 'circular insanity' by him. The two bitterly disputed as to who had been the first to conceptualize the condition. These concepts were developed by the German psychiatrist Emil Kraepelin (1856–1926)[1], who, using Kahbaum concept of cyclothymia, categorized and studied the natural course of untreated bipolar patients.

Bipolar disorder typically develops in late adolescence or early adulthood. The late adolescence and early adulthood are the year of the onset of the bipolar disorder Chirstie et al., (1988). However, some people have their symptom during childhood, and some develop them in the late life. The symptom of bipolar disorder is by a dramatic mood swings from overly "high" and/ or irritable to sad and hopeless, and then back again, often with periods of normal mood in between severe change in energy and behavior go along with these changes in mood. Some decades ago the bipolar disorder was called as "manic-depression", which is the combination of the two pool episodes, thus are manic episode and depressive episode.

The manic episode is characterized by increased energy, activity, restlessness, excessively high, overly good, euphoric mood, extreme irritability, racing thoughts and talking very fast, jumping from one idea to another, distractibility, can't concentrate well, little sleep needed, unrealistic beliefs in one's abilities and power, poor judgment, speeding sprees, increased sexual derive, abuse of drugs, particularly cocaine, alcohol, provocative, intrusive, or aggressive behavior. On the other hand the symptom of the depressive episode are anxious, lasting sad, empty mood, feeling of hopelessness or pessimism, feeling of guilty, worthlessness, or helplessness, loss

[1] Cited at http://en.wikipedia.org/wiki/History_of_bipolar_disorder.

of interest or pleasure in activities once enjoyed, decreased energy, feeling of fatigue or of being “slow down”, restlessness or irritability, sleeping too much or cannot sleep, change in appetite and /or uninterested, weight loss or gain, thought of death or suicide, or suicide attempts (NIMH, 2007).

Some times, sever episodes of mania or depression includes symptoms of psychosis (or psychotic symptoms). The common psychotic symptoms are *hallucination* (hearing, seeing or otherwise sensing the presence of things not actually there) and *delusions* (false, strongly held belief not influenced by logical reasoning or explained by a person’s usual cultural concepts). Psychotic symptoms in bipolar disorder tend to reflect the extreme mood state at the time. There is a great confusion between the bipolar patients and schizophrenic patients’; people with bipolar disorder are sometimes incorrectly diagnosed as having schizophrenia or another severe mental illness (NIMH, 2007).

Bipolar disorder has two distinct groups. Theses are Bipolar I disorder and Bipolar II disorder. Bipolar disorder I is the classic form of the illness, which involves recurrent episodes of mania and depression. That means there is a swing of moods from depression to mania symptoms. Some people, however, never develop severe mania but instead experiences less episode of mania (called hypomania) that alternate with depression; this form of the illness is called bipolar II disorder (i.e. there is a swing of mood from depression to hypomania which is less than mania; in some people; however symptoms of mania and depression may occur together at the same time in what is called *a mixed bipolar state*. Symptom of the mixed state often includes agitation, trouble sleeping significant change in appetite, psychosis and suicidal thinking (NIMH, 2007)

Both children and adolescents can develop bipolar disorder. It is more likely to affect the children of parents who have the illness. Unlike many adults with bipolar disorder, whose episodes tend to be more clearly defined, children and young adolescents with the illness often experience very fast mood swings between depression and mania many times within a day. Children with mania are more likely to be irritable and prone to destructive tantrums than to be overly happy and elated. Mixed symptoms also are common in youths with bipolar disorder. Older adolescents who develop the illness may have more classic, adult-type episodes and symptoms. For any illness, however, effective treatment depends on appropriate diagnosis. Children or adolescents with emotional and behavioral symptoms should be carefully evaluated by a mental health professional. Any child or adolescent who has suicidal feelings, talks about suicide, or attempts suicide should be taken seriously and should receive immediate help from a mental health specialist (NHIM, 2007).

1.1.1 Causes of Bipolar Disorder

Scientists are learning about the possible causes of bipolar disorder through several kinds of studies. Most scientists now agree that there is no single cause for bipolar disorder, rather many factors act together to produce the illness. Because bipolar disorder tends to run in families, researchers have been searching for specific genes-the microscopic "building blocks" of DNA inside all cells that influence how the body and mind work and grow-passed down through generations that may increase a person's chance of developing the illness. But genes are not the whole story (NHIM, 2007).

Brain-imaging studies are helping scientists learn what goes wrong in the brain to produce bipolar disorder and other mental illnesses. They added that new brain-imaging techniques allow researchers to take pictures of the living brain at work, to examine its structure and activity, without the need for surgery or other invasive procedures. These techniques include magnetic resonance imaging (MRI), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI). There is evidence from imaging studies that the brains of people with bipolar disorder may differ from the brains of healthy individuals. As the differences are more clearly identified and defined through research, scientists will gain a better understanding of the underlying causes of the illness, and eventually may be able to predict which types of treatment willwork most effectively (Soares and Mann, 1997)

1.1.2 Treatment of Bipolar Disorder

Most people with bipolar disorder even those with the most severe forms can achieve substantial stabilization of their mood swings and related symptoms with proper treatment. Because bipolar disorder is a recurrent illness, long-term preventive treatment is strongly recommended and almost always indicated. A strategy that combines medication and psychosocial treatment is optimal for managing the disorder over time.

In most cases, bipolar disorder is much better controlled if treatment is continuous than if it is on and off. But even when there are no breaks in treatment, mood changes can occur and should be reported immediately to the doctor. The doctor may be able to prevent a full-blown episode by making adjustments to the treatment plan. Working closely with the doctor and communicating openly about treatment concerns and options can make a difference in treatment effectiveness (NIMH, 2007).

1.1.3. Medication

Medications for bipolar disorder are prescribed by psychiatrists medical doctors with expertise in the diagnosis and treatment of mental disorders (Sachs and Thase, 2000). While primary care physicians who do not specialize in psychiatry also may prescribe these medications, it is recommended that people with bipolar disorder see a psychiatrist for treatment. The medications known as “mood stabilizers” usually are prescribed to help control bipolar disorder. Several different types of mood stabilizers are available. In general, people with bipolar disorder continue treatment with mood stabilizers for extended periods of time (years). Other medications are added when necessary, typically for shorter periods, to treat episodes of mania or depression that break through despite the mood stabilizer.

1.1.4. Psychosocial Treatments

In addition to medication, psychosocial treatments including certain forms of psychotherapy (or "talk" therapy) are very helpful in providing support, education, and guidance to people with bipolar disorder and their families. Different studies have shown that psychosocial interventions can lead to increased mood stability, fewer hospitalizations, and improved functioning in several areas. A licensed psychologist, social worker, or counselor typically provides these therapies and often works together with the psychiatrist to monitor a patient's progress. The number, frequency, and type of sessions should be based on the treatment needs of each person. Psychosocial interventions commonly used for bipolar disorder are cognitive behavioral therapy, psycho education, family therapy, and a newer technique, interpersonal and social rhythm therapy (NIMH, 2007).

So far not that many studies have been undertaken about bipolar disorder in the country especially the rural part of the country as a result of which there is very limited knowledge regarding onset and course from countries where psychiatric services are very poor. This study will examine socio-demographic factors affecting bipolar disorder patients using polytomous logistic regression, or multiple discriminant analysis.

1.2. Objectives

1.2.1. **General objective**

- To identify the socio-demographic factors affecting the bipolar disorder patients and their interactions in the case of Butajira patients.
- To show the application of polytomous logistic regression and discriminant analysis to such medical data.

1.2. 2 **Specific objectives**

- To examine /investigate the severity of the bipolar disorder across the different categories of each of socio demographic factors
- To asses the interaction among the socio-demographic factors affecting the bipolar disorder patients
- To show how to use the polytomous ordinal logistic regression and multiple discriminant analysis in the analysis of such medical data and to examine their performance up on classification of cases in to the different categories of the dependent variable.

1.3. Significance of the study

This study may help the Ministry of Health to make some adjustments in the intervention strategies for bipolar disorder patients since so far there has many studies have not been done about these areas especially in developing countries such as Ethiopia. There are only few studies conducted by the Department of Community Health and Psychiatry, Addis Ababa University, Ethiopia. The outcome of study could be taken as an addition to research these have so far been done. In the previous times there was not much attention given to mental illness in our country. For many years Ethiopia had only one mental health hospital namely the "St. Paul Hospital" which is found in Addis Ababa. The hospital gives inpatient and outpatient care for the mentally ill from allover the country.

1.4. **Limitation of the study**

1. The data are not collected by qualified statisticians but rather by the Psychiatrists (health professionals). So there could be a problem of handling and collecting of the data.
2. There are several potential variables which can have a relationship with the severity of bipolar disorder but could not be included here.

3. Patients of all groups are not included in this study; only the age group from 16-49 years old are included.
4. Since the questionnaire used in this study was designed by WHO, mainly focused on the developed countries, it may not be efficient for the developing countries such as Ethiopia.
6. There is a shortage of literature to be reviewed on the bipolar disorder.

CHAPTER II
LITERATURE REVIEW

2.1 Overall review Literature

Bipolar disorder is episodic, lifelong and a clinically severe mood disorder which poses a major public health concern due to high rate of mortality, morbidity, psychosocial impairment, psychosocial disability, health care costs, higher unemployment, decreased work productivity and secondary substance abuse (Altshuler et al., 2002; Angst et al., 2002).

The lifetime prevalence of bipolar disorder in different countries is estimated to range from 0.1% to 4.8%. There appear to be relatively few gender differences in bipolar disorder. The age of onset of bipolar disorder, most commonly around 20 years of age, is substantially (about 10 years) lower than that of unipolar depression (Angst et al., 2005).

Moreover, from the study conducted by Noaghiul and Hibbeln (2003) outside the US, prevalence rates of BD have varied somewhat. In a recent cross-national epidemiological study on adults, the lifetime prevalence of BD I or II in different countries was: Iceland 0.2%, Taiwan 0.4%, Korea 0.5%, Puerto Rico 0.9%, Canada 1.1%, New Zealand 2.4%, Israel 2.6%, USA 3.0%, Italy 3.4%, Switzerland 5.1% and Hungary 5.5%.

Sorvaniemi and Hintikka (2005) reported that the 1-month and 12-month prevalence of bipolar disorder were 0.9% and 2.1% respectively. And from other studies in a community mental health centres, the corresponding figures were 4.4% and 7.6% (Sorvaniemi, 2005). In addition to this Kieseppä et al. (2004) estimated that the annual incidence of bipolar disorder I was 5.8 in 100 000 (95% CI=5.4 to 6.3).

In contrast to Major Depression Disorder(MDD) the prevalence rates were remarkably consistent among countries (Canada, Finland, France, Germany, Hong Kong, Italy, Korea, New Zealand, Puerto Rico, Taiwan, and the United States), and there was little variation in rates by gender. In other studies in contrast to MDD, epidemiologic data show no significant sex differences in rates of bipolar disorder. No sex differences were reported in the ECA study by (Weissman et al., 1988).

Bipolar disorder is a recurrent and chronic illness. Even though the natural course of some patients includes full recovery after distinct phases, according to the latest clinical studies, patients spend half of their time in a symptomatic state (Judd et al., 2003), interposed symptoms are extremely common and practically all patients encounter affective recurrences. The disorder has a profound influence on functional outcome, psychosocial factors, and quality of life, and because of the burden of the illness on the individual, family, and society, it is among one of the most disabling disorders worldwide (Morselli and Elgie, 2003).

In the Epidomological Catchment Area (ECA) bipolar disorder was much less frequent among married people, as contrasted with divorced or never-married people. Bipolar disorder was more prevalent among those with multiple divorces (Weissman et al., 1982). In the NCS, bipolar disorder was more frequent among unmarried, poorly educated people. From unanimous study the persons who have married have less risk of bipolar disorder than those who have never married, divorced, and widowed (Weissman et al., 1991). In the ECA study in 1978 it was found that individual who never married, separated, divorced or widowed had high risk of having bipolar disorder as compared to married persons.

Dohrenwend and Dohrenwend (1967) identified 30 studies of SES and mental illness. However, the findings were inconsistent and did not always support links between bipolar disorder and social accomplishment. One methodological issue that could explain the inconsistency is that most studies were based on treatment samples recruited from either public or private hospitals. These hospitals were quite stratified economically, an effect that was further enhanced by the fact that the hospital system at that time typically had separate hospitals for African Americans, who often were paid less than Caucasian Americans (Malzberg, 1956).

On the other hand, mental illness of the biological parents was associated with creativity in the offspring; 27.7% of the nationally prominent individuals had mentally ill biological parents, compared to approximately 10% of the less creative individuals. These findings suggest that links between creativity and mental illness are more genetic than environmental. In contrast to this study offspring of persons with mania have been found to have IQ scores that are lower than or comparable to the general population and other psychiatric control groups (Decina et al., 1983; Waters, 1981).

The prevalence of Bipolar Disorder is approximately the same in both sexes; however, men are more likely to have a manic episode, and women are more likely to experience a depressive episode. In addition, the frequency of depressive episodes, illicit substance abuse and suicide attempts is higher among women. Leibenluft, (2000), Zarate et al., (1997) and other related studies in contrast to MMD, epidemiologic data show no significant sex differences in rates of bipolar disorder. No sex differences were reported in the ECA study by Weissman et al., (1996). In addition to these he got the mean age of onset for bipolar disorder was generally is ranging from 18 to 27 years of age in different countries.

Desjarlais et al.(1995) reviewed 15 studies focusing on psychiatric disorders as well as on psychological distress, carried out over the last decades in many parts of the world, including Africa, Asia, the Middle East and Latin America, and stated that "Comparative analysis of empirical studies of mental disorders reveals consistency across diverse societies and social contexts: Symptoms of depression and anxiety as well as unspecified psychiatric disorders and psychological distress are more prevalent among women, where as substance use are more prevalent among men." In other words "Men tend to externalize their suffering through substance abuse and aggressive behavior, resulting in an underreporting of psychological distress. Women, in turn, more often suffer distress in the form of depression, anxiety, “nerves” and the like.

The ECA (Weissman, et al., 1982) study has found that individuals who never married, separated, divorced or widowed had a higher risk of the bipolar disorder compared to the married. Other related studies have also reported similar findings. This may be due to fact that individuals with bipolar disorder may have a difficulty to initiate a relationship or sustain one. In addition to this based on based on the United States, the national health of ECA study, bipolar disorder was much less frequent among married people, as contrasted with divorced or never-married people. Bipolar disorder was more prevalent among those with multiple divorces.

Age of onset was defined by the first episode of depression, hypomania or mania. Early age of onset was defined as 30 years or below and late as above 30 years. Age of onset for early and late onset was 30.4 ± 10.1 years in the total bipolar group (I and II), 29.7 ± 10.5 years in bipolar I patients and 32.3 ± 8.7 years in bipolar II patients based on the study conducted on the personality and age of bipolar disorder patients by Engstrom et al. (2003). From similar study the risk of bipolar disorder has been shown that there is high risk of a bipolar disorder in younger adults (below 30)rather than those of the older adults (above 30) (Tohen et al.,1995).

A recent nationwide studies from Scandinavia (Ösby et al., 2001; Høyer et al., 2000) indicate that a standardized mortality ratio of about 20 for BD sufferers. It is commonly estimated that 25-50% of them attempt suicide at least once (Goodwin and Jamison, 1990; Jamison, 2000; Slama et al., 2004), and 30-75% have suicidal ideation (Suppes et al., 2001; MacKinnon et al., 2005). In addition to this study Chen and Dilsaver (1995) found that patients with bipolar disorder are at higher risk of suicidal behavior than those suffering from other disorders.

Guze and Robins (1970), who were the first to systematically review the suicide risk of bipolar disorder (or manic-depressive illness), found that approximately 15% of all deaths among manic-depressive patients were the result of suicide, and Goodwin and Jamison (1990) reported in their review that approximately 19% of manic-depressive patients committed suicide. Furthermore, both studies were based on hospitalized patients.

Judd (2003) compared patients with bipolar disorder with healthy subjects and reported that patients are significantly having more difficulties with work-related performance, leisure activities, as well as social and family interactions. However, treatment can improve many of these difficulties and concluded that individuals with bipolar disorder also demonstrate significant increases in lifetime health service utilization and the need for welfare and disability benefits, compared to populations with no mental disorder.

2.2. **Causes of Bipolar Disorder**

Based on the study conducted by Craddock and Jones (1999) in the eight family studies of bipolar I disorder that included a control group, a meta analysis that bipolar I disorder was seven times more likely among relatives of bipolar I proband than of controls. These studies also demonstrate an increased risk of Major depression Disorder (MD) in relatives of bipolar probands, although the relative risk is lower than for bipolar I.

In a pooled analysis of the six twin studies of bipolar I disorder conducted between 1962 and 1999, Craddock and Jones (1999) calculated an estimate of proband-wise concordance of 50%, although the authors believe this to be an underestimate. Integrating the results of family, twin, and adoption studies, Craddock and Jones conclude that there is a substantial genetic predisposition to bipolar disorder. Although biological and genetic factors have long been known to play a major role in the etiology of bipolar disorder, psychosocial factors are gaining attention. In particular, many studies have identified the association between stressful life events and social rhythm disruptions and onset of recurrence. It is unlikely that psychosocial factors play a major role in the risk of first onset of bipolar disorder, but they may have an important role in increasing the risk of recurrence.

A number of features of the bipolar disorder itself makes it hard to diagnose. First, a manic episode is usually preceded by one or more depressive episodes, and the diagnosis of bipolar disorder cannot be established until the first mania or hypomania has manifested itself. It is generally thought that 5% to 10% of the people with a depressive disorder end up developing bipolar disorder (Bebbington and Ramana, 1995). Second, hypomanic episodes are difficult to distinguish from normal feelings and behavior, whereas manic episodes accompanied by psychotic symptoms are difficult to distinguish from primary psychotic disorders (Jefferson and Greist, 1994). Thirdly, many of the problems associated with bipolar disorder, especially early ones such as work or school related problems, cannot always be associated straight away with a manic or depressive episode (see, for example, Lish et al., 1994). Fourthly, there is a failure on the part of the patients to seek help on time (Lish et al., 1994).

Quality of life is often compromised for individuals with bipolar disorder. Lower wages, higher unemployment, work absenteeism, reliance on workmen's compensation, higher rates of divorce, lower levels of educational attainment, higher arrest rates, and hospitalization are often the consequences (Gardner et al, 2006).

Bipolar Disorder (BD) is the sixth leading cause of disability in the world (Murray and Lopez 1996). Following initial optimistic estimations about the outcome of the bipolar disorder, now it is clear that substantial morbidity, like functional deficits and poor quality of life, remain in remitted BP patients (Fagioloni et al., 2005).

Based on the study conducted in Iran in order to asses the quality of life of the bipolar disorder patients have found 45% of the bipolar disorder patients present several recurrence of the disease, in spite of good quality of life. Incomplete remission is seen in 30% of the patients. It seems quality of life (QOL) in bipolar disorder patients is lower than the general population (Arnold, et al 2000). From other similar study bipolar disorder symptoms have been associated with significant functional impairments, often having a negative impact on the performance of work-related, leisure, and interpersonal activities (Calabres et al., 2003) and from several studies found that quality of life is compromised in the people with people with bipolar disorder, even during the periods of clinical remission (Robb et al., 1997).

CHAPTER III

DATA AND METHODLOGY

3.1. Data collection method_

The data were collected from the Butajira site mental health project using a few screening questions for bipolar and depressive disorder module of the Amharic version of the Composite International Interview (CIDI). Those individuals with a positive history of any sign of mania or depressive disorder were included for clinical confirmation. Trained and supervised key informants were used to identify individuals with mental illness. Those selected, as key informants were residents of the sub-distinct with intimate knowledge of their communities. Patients were selected and subjected to the second stage. There were 828 positive cases on the CIDI and those identified by the key informants who underwent evolutions by clinicians using the Amharic version of the Schedule for Clinical Assessment in Neuropsychiatry (SCAN). Of these 317 were identified as possible cases of bipolar disorder patients.

3.2. Description of the study area_

The study area is Butajira district 135 km south of the capital city of Ethiopia, Addis Ababa, Ethiopia. According to the 1994 census, the district had a total population of 227,135 (OPHCC, 1994). Most of the people (90%) were rural dwellers engaged in agriculture. The major ethnic groups are Gurage, and the population is predominantly Muslim.

3.3. Variables included in the Study

In this study a categorical dependent variable and continuous and categorical independent variables are included.

3.3.1. The Dependent Variable

The severity of bipolar disorder (no symptom, moderate, and sever symptom) is the dependent variable.

3.3.2 Independent variables/ factors

The possible socio demographic factors affecting the bipolar disorder are sex of the patient (male, female), age of the patient in year ($\leq 30, >30$), residence of the patient (urban, rural), marital status (married, single/ separated and widowed), education level (continuous), quality of life of subjects current dwelling (quality higher than average, and lower than average for catchments area). These variables are given with coding in Table 3.1

Table 3. 1. Independent variable included in the study

Independent Variables	Description of the variable	Category(catagorization of the variables)
X_1	Residence of the patient	1= Urban or Sub-Urban 2=Rural
X_2	Quality of life of subjects current	1=Quality higher than /on average for catchments area.

	dwelling	2=Quality lower than for catchment's area
X_3	Marital status	0=married 1= single/ widowed/separated
X_4	Number of the patient's children	0=has no children 1=has children
X_5	Sex of the patient	1=Male 2=Female
X_6	Age of the patient	1= <=30 2=>30
X_7	Education level	None(continuous)

The study includes 32 items these are supposed to be indicators of the bipolar disorder patient; they have been rated based on the degree of the following items from 0 to 4 that is 0= none, 1= mild , 2=moderate , 3= marked 4= severe.

main diagnosis(CIDI-10); Guilt; isuicide; insomnia initial; insomnia middle; insomnia delayed; work and; interests; retardation, agitation, restlessness associated with anxiety scaling; anxiety, psychotic scaling; Somatic symptoms, gastrointestinal, lose of appetite, heavy feeling in abdomen; Somatic symptoms, general (heaviness in limbs, back or head diffuce backache, loss of energy and fatiguability); genital symptoms, (loss of libido, menistral disterbances) scaling; hypochondriasis, self-absorption (bodily), preoccupation with health, quaerulous attitude, hypochondrial delusions; loss of weightscaling; insight scaling; diural variation, symptoms worse in morning or evening depersonalization and derealization; paranoid symptoms, (suspicous, idea of reference and persecution, and hallucinations persecutory (with a depressive quality)); obsessional symptoms, obsessive thoughts and compulsions; against which the patient struggles ; increased motor activity-energy; sexual interest;

sleeping; irritability; speech (rate and amount); language- thought disorder; content conditioning; disruptive aggressive behavior; appearance; insight; a*nxiety somatic.*

If the person is positive (rating ≥ 1) for any item of the above and are diagnostically checked the patients have a bipolar disorder. By averaging all the items result for each person and considering the possible value of the CIDI. Any person will be classified further into any of the three categories (no, moderate and severe symptom). If a patient is positive for the main diagnosis CIDI of bipolar disorder and his/her average (mean) value is different from zero, the patient will be grouped into either the moderate or severe bipolar disorder symptom categories. That is, if the patient is positive to the CIDI of bipolar disorder and the average of the patient is between 0 and mean (0.384) then the patient will be grouped to the moderate bipolar disorder symptom category. However, if the patient is positive for CIDI of bipolar disorder and the average of the patient is above the mean (0.384), then the patient will be grouped into the sever bipolar disorder symptom category. Otherwise the patient will be grouped to the non-bipolar disorder symptom category.

3.4. METHEDOLOGY

A parametric statistical procedure will be used for the purpose of determining the socio-demographic factors that affect bipolar disorder patients and classifying the patients into different groups of the bipolar disorder symptoms.

3.4.1. INTRODUCTION

There are different statistical methods which can be used to predict a dependent variable from the corresponding sets of explanatory variables. These statistical methods could be the simple linear regression, multiple regression, logistic regression, discriminant analysis, exponential regression, and others. However, all of these regressions are not used for the same purpose. Each of them has their merits and demerits on predicting a dependent variable based on the independent variables/ factor.

When the dependent variable is a categorical variable, multiple linear regression and other such models are not be appropriate models. In such a case the logistic regression and discriminant analysis are appropriate for classification and predicting the membership of the cases. In this study the logistic regression and discriminant analysis will be used because of the nature of the data the dependent variable is a categorical.

3.4.2. Logistic regression model

A binary logistic regression model was developed primarily by Cox (1958). The relationship between the dependent variable and independent variables has a non-linear relation shape that is an "*S*" shaped in all logistic type regression. The model is either monotonic increasing or decreasing which depend upon the sign of the logistic regression coefficients.

Logistic regression is a type of regression which is used to build the relationship between a categorical dependent variable and any types of the explanatory (or independent) variables. According to the number of categories and the nature of categories of the dependent variable there are different types of logistic regressions. These are dichotomous (binary) logistic regression, polytomous logistic regression. Here in this study, the focus will be only on the polytomous logistic regression. In general logistic regression has two main uses; the first is the prediction of group membership since logistic regression calculates the probability of success over the probability of failure and the second is logistic regression is used to provide knowledge of the relationships type and strengths between the dependent variable and the explanatory variable variables.

3.4.2.1. Assumptions of the logistic Regression model

Logistic regression is popular in part because it enables the researcher to overcome many of the restrictive assumptions of OLS regression: The logistic regression does not require a linear relationship between the independent factors or covariates and the dependent, as does OLS regression. However, it does assume a linear relationship between the independent variables/ factors and the log odds (logit) of the dependent variable. The dependent variable need not be normally distributed but does assume its distribution is binomial or multinomial. In addition, it does not assume that there should be homogeneity of variance.

3.4.2.2. **Polytomous Ordinal logistic regression**

Some authors argue that, when the response variable is ordinal, inclusion of ordinality in the model to be estimated should improve model performance. As it has been emphasized by Anderson and Greenland (1985) proposed the same ordinal model that is in fact an ordinal logistic regression procedure and was obtained by the imposition of ordering constraints on the multinomial logistic model.

Polytomous ordinal logistic regression is a type of logistic regression which relates the polytomous ordinal dependent variable in which, there exists a natural ordering within the categories of the dependent variables with the independent (or explanatory) variable. The model is defined as follows.

Let the dependent variable Y have J ordered categories such that $Y_{1i}<Y_{2i},\ldots,<Y_{Ji}$ and P corresponding explanatory variables, say, $\underline{X_i} = (X_{i1},\ldots\ldots, X_{ip})^T$ for the ith respondent (or observation). In such case the usual logit will not be used instead the cumulative logistic will be used.

So that the cumulative logit model is given as

$$\log it(P(Y_i \le j)) = \ln(\frac{P(Y_i \le j)}{1-P(Y_i \le j)} = \alpha_j + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots. + \beta_K X_{ik} \quad , \quad j=1,2,\ldots J-1$$

and i=1,2,….n ……..(1)

This model was originally proposed by Walker and Duncan and it is later called the proportional odds model by McCullagh (1980). It follows that

$$\frac{P(Y_i \le j)}{1-P(Y_i \le j)} = \exp(\alpha_j + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots. + \beta_K X_{ik})$$

This equation is called the cumulative odds.

In the case of the polytomous ordinal logistic regression we will assume that the cumulative logits are parallel that means for each observation (or group) β^T = ($\beta_1, \beta_2, ...\beta_p$) is identical that means the coefficients of the explanatory variable for whatever observation is identical. However they have different intercepts that is $\alpha_j < \alpha_{j+1}$, j=1,2,….J-1.

So the above equation can be expressed in vector form as

$$\log it(P(Y_i \le j)) = \ln(\frac{P(Y_i \le j)}{1 - P(Y_i \le j)}) = \alpha_j + \beta^T \underline{X}_i \quad \text{…………………..(2)}$$

where

β^T is as defined above(not including intercept)

and $\underline{X_i} = (X_{i1}, ..., X_{ip})^T$

The cumulative logit model will be

$$\text{logit}(\gamma_{i1}) = \alpha_1 + \beta^T \underline{X}_i \quad \text{for j=1}$$

$$\text{logit}(\gamma_{i2}) = \alpha_2 + \beta^T \underline{X}_i \quad \text{for j=2}$$

$$\begin{array}{cc} . & . \\ . & . \\ . & . \end{array}$$

logit(γ_{ij})= $\alpha_j + X_i \beta^T \underline{X}_i$ for j=J-1 because $\gamma_{iJ} = p(Y_i \le J / \underline{X}) = 1$ each logit has the same slope except their intercept that is $\alpha_{j+1} > \alpha_j$. This model is composed of J – 1 parallel linear equations. And there are (J-1+P) parameters to be estimated by the maximum likelihood function. However, in the case of the polytomous nominal logistic regression we would estimate about (J-1+P(J-1)) parameters.

Let π_{ij} be the probability that the i^{th} observation will lie in the J group

That is , π_{ij} = p(Yi=j/ $\underline{X}$)) for i=1,2,…n,

j=1, 2, 3,…, J

Let γ_{ij} is the cumulative probability of the i^{th} observation and j^{th} category so then

γ_{ij} = p(Yi$\le$j/ $\underline{X}$)), j=1,2,3,…J-1

γ_{i1} = p(Yi$\le$1/ $\underline{X}$))= p(Yi=1/ $\underline{X}$))= π_{i1}

γ_{i2} = p(Yi$\le$2/ $\underline{X}$))= p(Yi=1/ $\underline{X}$))+ p(Yi=2/ $\underline{X}$))

$\gamma_{i2} = \pi_{i1} + \pi_{i2}$

Likewise the probability of the i^{th} observation to fall in group j is

γ_{ij} = p(Yi$\le$j/ $\underline{X}$))= $\pi_{i1} + \pi_{i2} + ... + \pi_{ij}$, j=1,2,…,J-1 So that equation (2) can be written as

$$p(\mathrm{Yi} \le \mathrm{j}/\ \underline{\mathrm{X}_\mathrm{i}})) = \frac{\exp(\alpha_\mathrm{j} + \beta^T \underline{X}_i)}{1 - \exp(\alpha_\mathrm{j} + \beta^T \underline{X}_i)} \quad \text{for} \quad \mathrm{j=1,2,\ldots J-1}$$

$$\gamma_{ij} = p(\mathrm{Yi} \le \mathrm{j}/\ \underline{\mathrm{X}})) = \frac{\exp(\alpha_\mathrm{j} + \beta^T \underline{X}_i)}{1 + \exp(\alpha_\mathrm{j} + \beta^T \underline{X}_i)} \quad \ldots\ldots\ldots(3)$$

This model is called the proportional odds model because the odds ratio of the event (Yi$\le$ J-1) is independent of the category events.

3.4.2.2.1 **Proportional odds model**

The proportional odds model is one of those models which are suggested to handle if the dependent variable is an ordered categories of the dependent variable. It has a special feature as discussed below which helps for dealing with the ordinal categorical response variable.

The first property of the proportional odds model is that it is invariant when the codes for the response Y are reversed (i.e. y_1 recoded as y_J , y_2 is recoded as y_{J-1} and so on), only the sign of the regression parameters is reversed; the second property is that it is invariant under collapsibility of the categories of the ordinal responses. This property implies that when the categories of Y are deleted or collapsed, the coefficients

β will remain unchanged, although the intercept parameters α_j will be affected. So, the collapsibility property of the proportional odds model enables one to model an ordinal outcome Y which may be continuous. The third property is that it is easy to interpret when compared to other models that are suggested for handling of the ordinal logistic regression. This model can be found in the different commercial statistical software packages (such as SAS, SPSS and others).

For a fixed j the response is a logistic curve for a binary response with outcomes $Y \leq y_j$ and $Y > y_j$ where y_j, j=1, 2, …, J-1, have the same shape. They share exactly the same rate of increase or decrease. However, are horizontally displaced from each other (Agresti, 2002).

Many researchers commonly interpret models with ordinal response variables probabilistically: regression parameters give information about how the probability (or odds) for some level of the categories increase or decrease with respect to the changes in X. This kind of motivation, then, leads to the consideration of Y*

$$Y_i^* = \alpha + \beta^T X_i + \varepsilon_i$$

However, Yi* is unobserved value (latent variable) and this latent variable we can connect with the categorized variable (Y) by using the cut points such as

$Y = 0 \equiv Yi^* \leq \alpha_0$

$Y = 1 \equiv \alpha_0 \leq Yi^* \leq \alpha_1$

. . .

$Y = J \equiv Yi^* \geq \alpha_{J-1}$

The cut points given by α_j , $j = 0,1,2,....,J-1$ partition the latent variable in terms of the J categories of Y. Hence, for a J-category response variable, J − 1 cut points fully partition Y*. In this study we do have three categories of the severity symptom of the bipolar disorder patients these are no, moderate, and severe symptoms of the bipolar disorder so that we will have

$$Y = 0 \equiv Y^* \leq \alpha_0$$
$$Y = 1 \equiv \alpha_0 \leq Y^* \leq \alpha_1$$
$$Y = 2 \equiv Y^* \geq \alpha_1$$

so that, in this case we can have two cut points since we do have three categories these are none, moderate, and sever bipolar disorder symptom.

Suppose Y is a dependent variable which has J-ordered categories and the probabilities of the J-categories are $\pi_1(\underline{X}), \pi_2(\underline{X}), ..., \pi_J(\underline{X})$ and there is an explanatory variable $\underline{X}_i = (X_{i1}, X_{i2}, ..., X_{ip})^T$. Assume that Qj=the cumulative odds, that is,

$$Q_j(X) = \frac{P(Y \leq j / \underline{X}_i)}{1 - P(Y \leq j / \underline{X}_i)} \text{ .... (4)}$$

It is known that the proportional odds model is defined in McCullagh (1980) as

$Q_j(\underline{X_i}) = Q_j e^{(\beta^T \underline{X}_i)}$ for j=1,2,…J-1.

And the ratio of the corresponding odds for the j category is

$$\frac{Q_j(\underline{X}_i)}{Q_j(\underline{X}_s)} = \frac{Q_j e(\beta^T \underline{X}_i)}{Q_j e(\beta^T \underline{X}_s)} = e^{(B^T(\underline{X}_i - \underline{X}_s)} \quad for\, i \neq s = 1,2,.....,n\, and\, j = 1,2,..J......(5)$$

This equation is independent of the j^th^ category of the dependent variable. However, it depends only on the difference between any two covariate values. That is why γ_{ij} is called the proportional odds model. This model provides a single odd ratio (OR) for all the categories compared, which can be obtained by exponentiation of the β coefficient. This estimate is quite convenient in terms of the ease of interpretation of the model and parsimony.

Simply by taking the difference between the corresponding cumulative logit models. It results an independent of the category of the dependent variable. Since for any two observations say i and s for a specified j category the difference between the corresponding cumulative logit is independent of the j category which is given by

$\text{Logit}(\gamma_j(\underline{X}_i) - \text{Logit}(\gamma_j(\underline{X}_s) = \alpha_j + \beta^T \underline{X}_i - (\alpha_j + \beta^T \underline{X}_s)$, because

$$\text{Logit}(\gamma_j(\underline{X_i}) = \alpha_j + \beta^T \underline{X}_i$$

$$= \beta^T(\underline{X}_i - \underline{X}_s) \text{ , for } i \neq s = 1, 2, \ldots, n \quad \ldots\ldots \ (6)$$

This is the same as the equation (5) result that means which is independent of the category event.

3.4.2.2.2 **Test of the Proportional odds model**

The relationship between the r^{th} explanatory variable is independent of the j^{th} category of the dependent variable can be given as

$$\text{logit}(Q_j(\underline{X_i})) = \alpha_j + \beta^T \underline{X_i} \quad \text{for } j=1,2,\ldots,J-1 \ldots\ldots(7)$$

The proportional assumption is a serious assumption as that of the Cox regression proportionality assumption in the case of survival data analysis. So that it needs a carefully checking of this assumption. The validity of this assumption can be checked based on X^2 score test by introducing a somewhat relaxed model as an alternative model that is

$$\text{logit}(Q_j(\underline{X_i})) = \alpha_j + \beta_j{}^T \underline{X_i} \quad \text{for } j=1,2,\ldots,J-1 \ldots\ldots(8)$$

where $Q_j(\underline{X_i})$ is the cumulative odds.

Here the relationship between the r^{th} explanatory variable and the dependent variable is depends on the j^{th} category of the dependent variable (H_1). This will be the alternative hypothesis when we are testing the null hypothesis, that is the coefficients of the r^{th} explanatory variable for each case is the same. Therefore, the hypothesis to be tested here will be

H_0: $\beta_{r1} = \beta_{r2} = ... = \beta_{rn} = \beta_r$, (The coefficient of the r^{th} explanatory variable for each observation is the same) for all r =1,2,…p.

V_s H_A: not H_0 (at least one of the r^{th} explanatory variable coefficients is different from the other).

3.4.2.2.3 Interpretation of the cumulative logit regression

i) When β_r is positive that means a one-unit increase in the r^{th} independent variable the effect of increasing the odds of being in a higher category for the dependent variable.

ii) When β_r is negative that means a one-unit increase in the r^{th} independent variable has the effect of decreasing the odds of being in a higher category for the dependent variable.

iii) The Parallel Slopes Assumption (also known as the proportional odds assumption) requires that the separate logit equations for each category differ only in their intercepts. That is, the slopes are assumed to be the same when going from one category to the next category.

3.4.2.2.4. Likelihood Function of polytomous ordinal logistic model

Let the response cell counts be n_{ij} with row total $n_1, n_2, \ldots, n_n$ and $n_{.j}$ is the column or category total. That is, a dependent variable Y which has J categories and have n independent observations which have the p-dimension explanatory variable. That is $\underline{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ip})^T$, i=1,2,….,n. This can be summarized by the following table

Table 3.2 Summarization of Allocation of Each observation to each category

Y-category

Y-category / Observation	1	2	…….	j	.	J	Total
1	n_{11}	n_{12}	.	n_{1j}	.	n_{1J}	n_1
2	n_{21}	n_{22}	.	n_{2j}	.	n_{2J}	n_2
.	.	.	.	.	.	.	.
I	n_{i1}	n_{i2}	.	n_{ij}	.	n_{iJ}	n_i
.	.	.	.	.	.	.	.
N	n_{n1}	n_{n2}	.	.	.	n_{nj}	n_n
Total	$n._1$	$n._2$	.	$n._j$	.	$n._j$	n

Let us assume that m_i is the cumulative rows sum, therefore $n_i=m_{iJ}$ is the i^{th} row total. Under the assumption of multiple nominal sampling in each row, the marginal distribution of m_{ij} conditional only on the row total n_i is binomial with n_i and probability γ_{ij}.

The contribution from a single multinomial observation $(n_1,n_2,….,n_J)$ to the likelihood function is $\pi_1^{n_1}\pi_2^{n_2}....\pi_J^{n_J}$ with probability π_j , j=1,2,…..,J-1.But here we are considering the cumulative probabilities.

$m_1 = n_1$

$m_2 = n_1+n_2$

$m_3 = n_1+n_2+n_3$

. .

. .

$$m_j=\sum_{j=1}^{J} n_j = n$$

In terms of the parameters of the cumulative transformation, the likelihood can be written as the product of K-1 quantities (Mc Cullagh (1980))

$$\left(\frac{\gamma_1}{\gamma_2}\right)^{m_1}\left(1-\frac{\gamma_1}{\gamma_2}\right)^{m_2-m_1}\left(\frac{\gamma_2}{\gamma_3}\right)^{m_2}\left(1-\frac{\gamma_2}{\gamma_3}\right)^{m_3-m_2}\cdots\left(\frac{\gamma_{J-1}}{\gamma_J}\right)^{m_{J-1}}\left(1-\frac{\gamma_{J-1}}{\gamma_J}\right)^{m_J-m_{J-1}}$$

Suppose Y is an ordered categorical dependent variable and a vector of explanatory variable $\underline{X}_i=(X_{i1},X_{i2},...,X_{ip})^T$, n_{ij} is the i[th] observation which come across with the j[th] category result, j=1,2,….,J and

Let{Y_i, i=1,2,..,n} are independent multinomial random variables Y_i ~multin($n_{i1},n_{i2,\ldots,}n_{iJ}$) with $E(Y_i)=n_{ij}\,\gamma_j(\underline{X}_i)$ where $n_{i1}+n_{i2}+\ldots+n_{iJ}=1$, for i=1,2,…,n. Since the number of counts that will fall into any of the J category for the i[th] observation is only one,

$m_{i1}=n_{i1}$

$m_{i2}=n_{i1}+n_{i2}$

$m_{i3}=n_{i1}+n_{i2}+n_{i3}$

. .

. .

$n_{iJ}=n_{i1}+n_{i2}+\ldots.+n_{iJ}=1=n_i$

Scince we are dealing with the cumulative probabilities, the above likelihood can be written as the product of J-1 quantities. The joint probabilities mass function of (Y_1, $Y_{2,\ldots..,}$ Y_n) is proportional to the product of n multinomial functions.

The likelihood function will be

L($\underline{\alpha},\beta$)= $\prod_{i=1}^{n} f(\underline{X}_i)f(y_i/\underline{X}_i) \quad \propto\prod_{i=1}^{n} f(y_i/\underline{X}_i)$ since $f(\underline{X}_i)$ is independent of the parameter(α and β)

$$=\prod_{i=1}^{n}\left\{\left(\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i1}}\left(1-\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i2}-m_{i1}}\left(\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i2}}\left(1-\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i3}-m_{i2}}\cdots\cdots\left(\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{J-1}}\left(1-\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{iJ}-m_{i,J-1}}\right\}$$

The maximum likelihood for the ordinal logistic model will be

$$L(\alpha,\beta)=\prod_{i=1}^{n}\left\{\left(\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i1}}\left(1-\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i2}-m_{i1}}\left(\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i2}}\left(1-\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i3}-m_{i2}}\cdots\right.$$

$$\left.\cdots\left(\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{J-1}}\left(1-\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{iJ}-m_{i,J-1}}\right\}\ldots\ldots\ldots(9)$$

It is possible to obtain the regression coefficients which maximize the above likelihood. Since the maximum likelihood function and the log likelihood function attain their maximum value at the same point for simplicity use ln(L) rather than the likelihood (L) in order to estimate the logistic coefficients.

Ln(L)=

$$=\ln\left\{\left(\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i1}}\left(1-\frac{\gamma_{i1}}{\gamma_{i2}}\right)^{m_{i2}-m_{i1}}\left(\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i2}}\left(1-\frac{\gamma_{i2}}{\gamma_{i3}}\right)^{m_{i3}-m_{i2}}\ldots\ldots\left(\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{iJ}}\left(1-\frac{\gamma_{iJ-1}}{\gamma_{iJ}}\right)^{m_{iJ}-m_{i,J-1}}\right\}$$

$$=\sum_{i=1}^{n}\{m_{i1}\ln(\gamma_{i1})+(m_{i2}-m_{i1})\ln(\gamma_{i2}-\gamma_{i1})-m_{i2}\ln(\gamma_{i2})+m_{i2}\ln(\gamma_{i2})+(m_{i3}-m_{i2})\ln(\gamma_{i3}-\gamma_{i2})-$$
$$m_{i3}\ln(\gamma_{i3})\ldots,m_{i(J-1)}\ln(\gamma_{i(J-1)})+(m_{iJ}-m_{i(J-1)})\ln(\gamma_{iJ}-\gamma_{iJ-1})-m_{iJ}\ln(\gamma_{iJ}.\}$$

$$=\sum_{i=1}^{n}\{m_{i1}(\alpha_1+\beta^T\underline{X}_i-\ln(1+e^{\alpha_1+\beta^T\underline{X}_i})+(m_{i2}-m_{i1})[\beta^T\underline{X}_i+\ln(e^{\alpha_2}-e^{\alpha_1}))-\ln(1+e^{\alpha_2+\beta^T\underline{X}_i})$$
$$-\ln(1+e^{\alpha_1+\beta^T\underline{X}_i}]+(m_{i3}-m_{i2})[\beta^T\underline{X}_i+\ln(e^{\alpha_3}-e^{\alpha_2})-\ln(1+e^{\alpha_3+\beta^T\underline{X}_i})-\ln(1+e^{\alpha_2+\beta^T\underline{X}_i})]\ldots.$$
$$-(1-(m_{iJ}-m_{i(J-1)})\ln(1+e^{\alpha_{J-1}+\beta^T\underline{X}_i}).\}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

It is possible to obtain the estimated value of α_j as $\hat{\alpha}_j$ and β as $\hat{\beta}$ by partial differentiating the above eq(10) with respect the corresponding logistic regression coefficients and then equating to zero.

$\frac{\partial \ln(L)}{\partial \alpha_j} = 0$ That is each $\hat{\alpha}_j$ are estimated by partial differentiating eq(10) with respect to α_j , j=1,2,…J-1 and then equate to zero.

$$= \frac{\partial \sum_{i=1}^{n} \left\{ \begin{array}{l} m_{i1}(\alpha_1 + \beta^T \underline{X}_i - \ln(1+e^{\alpha_1+\beta^T \underline{X}_i}) + (m_{i2}-m_{i1})[\beta^T \underline{X}_i + \ln(e^{\alpha_2} - e^{\alpha_1})) - \ln(1+e^{\alpha_2+\beta^T \underline{X}_i}) \\ -\ln(1+e^{\alpha_1+\beta^T \underline{X}_i}] + (m_{i3}-m_{i2})[\beta^T \underline{X}_i + \ln(e^{\alpha_3} - e^{\alpha_2}) - \ln(1+e^{\alpha_3+\beta^T \underline{X}_i}) - \ln(1+e^{\alpha_2+\beta^T \underline{X}_i})].... \\ -(1-(m_{iJ} - m_{i(J-1)})\ln(1+e^{\alpha_{J-1}+\beta^T \underline{X}_i}). \end{array} \right\}}{\partial \alpha_j}$$

$=0$………. .(11)

For the logistic regression coefficients (i.e. β) are estimated by partial differentiating equ(16) with respect to the β_r and equating to zero(0).

$$\frac{\partial \ln(L)}{\partial \beta_r} =$$

$$\frac{\partial [\sum_{i=1}^{n} \left\{ \begin{array}{l} m_{i1}(\alpha_1 + \beta^T \underline{X}_i - \ln(1+e^{\alpha_1+\beta^T \underline{X}_i}) + (m_{i2}-m_{i1})[\beta^T \underline{X}_i + \ln(e^{\alpha_2} - e^{\alpha_1})) - \ln(1+e^{\alpha_2+\beta^T \underline{X}_i}) \\ -\ln(1+e^{\alpha_1+\beta^T \underline{X}_i}] + (m_{i3}-m_{i2})[\beta^T \underline{X}_i + \ln(e^{\alpha_3} - e^{\alpha_2}) - \ln(1+e^{\alpha_3+\beta^T \underline{X}_i}) - \ln(1+e^{\alpha_2+\beta^T \underline{X}_i})].... \\ -(1-(m_{iJ} - m_{i(J-1)})\ln(1+e^{\alpha_{J-1}+\beta^T \underline{X}_i}). \end{array} \right\}}{\partial \beta_r} = 0$$

Where r=1, 2 , …, p

It is possible to find the Fisher information matrix $I_0(\hat{\underline{\alpha}}, \hat{\underline{\beta}})$ from the above log-likelihood function equation (10) as follows

$I_0 = \lfloor I_{0jr} \rfloor$, j=1, 2,….,J-1 and r=1,2,….p

The off diagonal elements of I_0 are

$I_{0\,jr}$=-E$\left[\frac{\partial^2 \log L}{\partial\alpha_j \partial\beta_r}\right]$, j,r=1,2,…J-1 , $I_{0\,jk}$ =-E$\left[\frac{\partial^2 \log L}{\partial\alpha_j \partial\alpha_k}\right]$, j$\neq$k=1,2,..J-1 and

$I_{0\,rk}$ =-E$\left[\frac{\partial^2 \log L}{\partial\beta_r \partial\beta_k}\right]$, r$\neq$ k=1, 2,…,p

The diagonals of I_0 are

$I_{0\,jj}$ =-E$\left[\frac{\partial^2 \log L}{\partial\alpha^2{}_j}\right]$, j=1,2,…..J-1 and $I_{0\,rr}$=-E$\left[\frac{\partial^2 \log L}{\partial\beta^2{}_r}\right]$, r=1,2,…p

After having the information matrix from the log likelihood function then it is possible to obtain the estimated variance-covariance of the parameters from the observed information matrix. That is the asymptotic covariance matrix of the estimated parameters $(\underline{\hat{\alpha}}, \underline{\hat{\beta}})$ is $I_0^{-1}(\underline{\hat{\alpha}}, \underline{\hat{\beta}})$.

So that the maximum likelihood estimates $(\underline{\hat{\alpha}}, \underline{\hat{\beta}})$ has a multivariate normally distributed. That is $(\underline{\hat{\alpha}}, \underline{\hat{\beta}})$~N($\underline{\alpha}, \underline{\beta}, I_0^{-1}(\underline{\alpha}, \underline{\beta})$)

3.4.2.2.5. Model checking

As is done in any type of regression modeling, we must evaluate the quality of model fit. A careful assessment of model is important so as to have an adequate model which includes the most important explanatory variables and sometimes included interactions which have a significant relationship with the dependent variables. There are many types of procedures that have been mentioned in different books and journals to assess the goodness of fit of the estimated logistic regression model such as the Hosmer and Lemeshow, Chi-square test of goodness of fit, Omnibus tests of model coefficient, Pearson and deviance goodness of fit. The former two are restricted to the binary logistic and polytomous nominal logistic. However, the Pearson and deviance goodness are used for polytomous ordinal logistic regression. Therefore, we will use the Pearson and deviance goodness of fit.

i) Pearson and deviance goodness of fit.

The Pearson's cha-square and deviance test are used to test how much predicted values differ from observed frequencies. A well fitting model is one which has large value of the observed significance.

From the observed and expected frequencies, we can compute the usual Pearson and Deviance goodness-of-fit measures as follows.

The Pearson goodness-of-fit statistic is

$$X^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

The deviance measure is

$$G^2 = 2\sum f_0 \log\left(\frac{f_0}{f_e}\right)$$

where f_0 - the observed frequency and
f_e - the expected frequency.

3.4.3. Discriminant Analysis

Discriminant analysis is a technique used for classifying a set of observations into predefined classes. The model is built based on a set of observations for which the classes are known. This set of observations is sometimes referred to as the training set. Based on the training set, the technique constructs a set of linear functions of the predictors, known as discriminant functions, such that

$Z_j = \gamma_1 X_{1j} + \gamma_2 X_{2j} + ... + \gamma_p X_{pj}$ the γ_{ij} 's are discriminant coefficients, the x's are the input variables or predictors. The values resulting from the discriminant function are known as discriminant scores.

In discriminant analysis, the independent variables are the predictors and the dependent variable are the groups which are used to predict membership in naturally occurring groups. It answers the question: can a combination of variables be used to predict a membership of the observation? Usually, several variables are included in a study to see which ones contribute to the discrimination between groups. Ones first perform the multivariate test, and, if statistically significant, proceed to see which of the variables have significantly different means across the groups. Once group means are found to be statistically significant, classification of variables is undertaken. Discriminant analysis automatically determines some optimal combination of variable so that the first function provides the most overall discrimination between groups; the second provides the second most, and so on.

There are different types of discriminant analysis such as linear, quadratic, triple, and other discriminant analysis. The concern in this study is with the linear discriminant analysis. The simple and/or multiple discriminant analysis are used for several purposes. In addition to classifying the cases into different groups can be used to used for the following purposes. These are: to investigate differences between or among groups, to determine the most parsimonious way to distinguish among groups; to determine the percent of variance in the dependent variable explained by the independent variables over and above the variance accounted for by control variables using sequential discriminant analysis; to assess the relative importance of the independent variables in classifying the dependent variable.

3.4.3.1 Discriminant analysis Assumptions

Discriminant analysis has more restrictions as we compared to the logistic regression. It assumes that the sample data is drawn from a normally distributed population.

However, note that violations of the normality assumption are not "fatal" and the resultant significance test are still reliable as long as non-normality is caused by skewness and not outliers (Tabachnick and Fidell 1996). Discriminant analysis is sensitive to heterogeneity of variance–covariance matrices. There should not be multicolliniarity in the sample data set and the sample size of the smallest group needs to exceed the number of the predictor variables. As a rule of “thumb” the smallest sample size should be at least 20 for a few (4 or 5) predictors.

Based on the level (i.e. categories) of the dependent variable, we can classify the discriminant analysis into as simple discriminant analysis and multiple discriminant analysis.

3.4.3.2. Simple discriminant Analysis

Simple discriminant function analysis, (or DA), is a type of the discriminant analysis which is used to classify cases into the values of a categorical dependent variable, usually the dependent variable is dichotomous. If discriminant function analysis is effective for a set of data, the classification table (or called confusion matrix) of correct and incorrect estimates in simple discriminant analysis and multiple discriminant analysis will yield a high percentage correct classification.

If there are two group population π_1 *and* π_2, and each case has p characteristics that is

$$\underline{X}_i = (X_{i1}, X_{i2}, ..., X_{ip})^T$$

then we can have the following discriminant function.

$$Z_i = \gamma_1 X_{i1} + \gamma_2 X_{i2} + ... + \gamma_p X_{ip},$$

will be for the i^{th} observation

There are different methods are suggested for classifying in discriminant analysis such as Expected Cost Minimization (ECM), Bay’s classification and Fisher classification method.

The Fisher discriminant analysis is one of the different methods of discriminant analysis that will be used for discriminating the cases (observations) into different categories of the dependent variable. This is done based on maximizing the ratio between group variance of a given variables. Mathematically for the simple discriminant analysis we can see it as follow.

$$\frac{(Squared\ dis\tan ce\ between\ sample\ means\ of\ y^{/}s)}{(sample\ \mathrm{var}iance\ of\ y)} = \frac{(\bar{y}_1 - \bar{y}_2)^2}{S_y^{\ 2}} \quad , \quad y_i = 1,2.$$

$$= \frac{(\hat{a}\bar{X}_1 - \hat{a}\bar{X}_2)^2}{\hat{a}^{/} S^2_{\ pooled}\,\hat{a}}, \sin ce\ y_i = \hat{a}X_i$$

$$= \frac{(\hat{a}^{/} d)^2}{\hat{a}^{/} S^2_{\ pooled}\,\hat{a}}\ , d = \bar{X}_1 - \bar{X}_2$$

Where the maximum of the above ratio is $D^2 = (\bar{X}_1 - \bar{X}_2)^{/} S^{-1}_{\ pooled}(\bar{X}_1 - \bar{X}_2)$....(*) this is called the Mahalanobis distance(Johnson,2002)

The linear combination of the variables or characteristics which maximizes the $(\overline{X_1} - \overline{X_2})^{/} S^{-1}_{\ pooled} X$ will be maximizes the ratio above (*).

An allocation will be based on the Fisher`s Discriminant Function if the two population have the same variance covariance matrices (i.e. $\Sigma_1 = \Sigma_2$). Then discriminant function will be linear that separates the different groups.

Allocation X_0 to π_1 if

$$\hat{y}_0 = (\bar{X}_1 - \bar{X}_2)^{/} S^{-1}_{\ pooled} X_0$$

$$\geq \hat{m} = \frac{1}{2}(\bar{X}_1 - \bar{X}_2)^{/} S^{-1}_{\ pooled}(\bar{X}_1 + \bar{X}_2)\ or$$

$$\hat{y}_0 - \hat{m} \geq 0 \quad otherwise\ allocate\ to\ \pi_2$$

That means allocate X_0 to π_2 if

$$\hat{y}_0 < \hat{m} = \frac{1}{2}(\overline{X}_1 - \overline{X}_2)^{/} S^{-1}{}_{pooled} (\overline{X}_1 + \overline{X}_2) \text{ or } \hat{y}_0 - \hat{m} < 0$$

If the two population have different variance-covariance matrices (i.e. $\sum_1 \neq \sum_2$) then we will allocate by using the Quadratic Discriminant Function since a linear discriminant function will not discriminate the different groups. That is, allocate X_0 into π_1 if

$$-1/2x'_0\left(S_1^{-1} - S_2^{-1}\right)x_0 + \left(\overline{x}'_1 S_1^{-1} - \overline{x}'_2 S_2^{-1}\right)x_0 - k \geq 0$$

Otherwise allocate to π_2

where K is

$$k = 1/2\ln\left(\frac{|S_1|}{|S_2|}\right) + 1/2\left(\overline{x}'_1 S_1^{-1}\overline{x}_1 - \overline{x}'_2 S_2^{-1}\overline{x}_2\right)$$

$S_1$and S_2 are the estimated variance covariance matrices for π_1 and π_2 respectively. And $|S_i|$ for i=1, 2 is the determinant of the variance covariance matrix of the corresponding population of the population π_1 and π_2

3.4.3.3. Multiple-group Discriminant Analysis (MDA)

A Multiple Discriminant Analysis (MDA) is an extension of the simple discriminant analysis which is used to classify cases into more than two groups based on the corresponding explanatory variable or characteristics.

In this case we are interested in discriminating among more than two groups so that MDA is defined as the discriminating of more than two groups.

If there are J-discriminating groups and each case (element) has p characteristics, i.e. $\underline{X}_j = (X_{1j}, X_{2j}, ..., X_{pj})$ then we can have the following J-1 possible discriminant functions. These are

$$Z_1 = \gamma_{11}X_{1j} + \gamma_{12}X_{2j} + ... + \gamma_{1p}X_{pj}$$
$$Z_2 = \gamma_{21}X_{1j} + \gamma_{22}X_{2j} + ... + \gamma_{2p}X_{pj}$$
$$Z_3 = \gamma_{31}X_{1j} + \gamma_{32}X_{3j} + ... + \gamma_{3p}X_{pj}$$
$$.$$
$$.$$
$$.$$
$$Z_{J-1} = \gamma_{J-11}X_{1j} + \gamma_{J-12}X_{3j} + ... + \gamma_{J-1p}X_{pj}$$

The first discriminant function maximizes the differences between the values of the dependent variable. The second function is orthogonal to it (uncorrelated with it) and maximizes the differences between values of the dependent variable, controlling for the first factor and so on. Though mathematically different, each discriminant function has one dimension which differentiates a case into categories of the dependent variable based on its values on the independent variables. The first function will be the most powerful differentiating dimension, but later functions may also represent additional significant dimensions of differentiation.

There is one eigenvalue for each discriminant function. The ratio of the eigenvalues indicates the relative discriminating power of the discriminant functions.

The ratio

$$\phi = \frac{\lambda_i}{\lambda_l} \quad for \;\; i \neq l = 1,2,....J-1$$

is interpreted as the discriminant analysis which corresponds to the λ_i has a ($\phi - 1$)100% more (less) between-group variance in the dependent categories than does in the one corresponds to the λ_l eigenvalue.

The relative percentage of a discriminant function is obtained by deviding its eigen value by the sum of all eigenvalues of all discriminant functions in the model. Thus, it is the percent of discriminating power for the model associated with a given discriminant function. Mathematically it can defined as

$$R.P_i = \frac{\lambda_i}{\sum_{l=1}^{K-1} \lambda_l} x100\%, i=1,2,\ldots.J-1.$$

where λ_i is the i^{th} eigenvalue corresponds to the i^{th} discriminant analysis

$R.P_i$ is the relative percentage of the i^{th} discriminant function

This relative percentage is helpful in order to identify the number of the important discriminant functions.

Only a few of the discriminant functions usually are statistically significant. Significance tests that were provided by Bartlett (1947) are discussed by Rao (1952).

In order to investigate the relative importance of variables, the discriminant function coefficients often are standardized (Morrison, 1969). Also MDA includes the assumptions that the independent variables are multivariate normally distributed and that the groups have equal covariance matrices. The discriminant function scores are linear combinations of the original multivariate normal variables (Rao, 1973). However the Fisher discriminant analysis does not assume multivariate normality. It assumes the there should be a homogenous variance covariance in each group so as to use the pooled variance-covariance matrix.

3.4.3.4 Analytical Approach of MDA

The analysis of this MDA is similar to that of the two-group discriminant analysis. Here first a univariate analysis can be done to determine if each of the discriminating variables significantly discriminates among the other groups. This test can be achieved by an F-test. If the F-calculated is greater than that of the tabulated F-tabulated (that means at least one of the population mean of the group is different from any of the other population means), then we will proceed to use the discriminant analysis in order to discriminate the different population groups. Otherwise we will not proceed to the discriminant analysis since there is no significant difference between the means of the multiple-group populations.

Having identified the discriminating variables, the next step is estimating the discriminant function. Suppose the ith discriminant function is

$$Z_i = \gamma_{i1}X_1 + \gamma_{i2}X_2 + ...\gamma_{ij}X_j .... + \gamma_{ip}X_p, \quad i=1, 2, …., J-1$$

where γ_{ij} the weight of the jth variable and the ith discriminant function

X_j is the jth variable of any of the population group and

The weights of the ith discriminant function are estimated such that the

$$\lambda_i = \frac{Between - group\ SS\ of\ Z_i}{Within - gruop\ SS\ of\ Z_i}$$

is maximized(Sharma S.,1996)

where λ_i -the eigenvalue of the ith discriminant function

Usually in discriminant analysis there are two steps: the first is an F test (Wilks' lambda) which is used to test if the discriminant model as a whole is significant, and the second is, if the F test shows significance, then the individual independent variable is assessed to see which differ significantly in mean by group and these are used to classify the dependent variable.

Discriminant analysis shares all the usual assumptions of correlation, requiring linear and homoscedastic relationships. Like multiple regression, it also assumes proper model specification (inclusion of all important independents and exclusion of extraneous variables). Discriminant Analysis(DA) is an earlier alternative to logistic regression, which is now frequently used in place of DA as it usually involves fewer violations of assumptions (independent variables need not be normally distributed, linearly related, or have equal within-group variances), is robust, handles categorical as well as continuous variables, and has coefficients which many find easier to interpret. Logistic regression is preferred when data are not normal or group sizes are very unequal. However, discriminant analysis is preferred when the assumptions of linear regression are met. In such cases DA has more statistical power than logistic regression.

3.4.3.5. **Fisher Linear Multiple Discriminant Function (FLDA)**

The fisher discriminate analysis is one of the means of discriminating of the observations into different groups of the category of the dependent variable. The Fisher Discriminant Analysis (FDA) is employed based on maximizing the ratio of between-group variance to within-group variance of given variables. The maximization problem is frequently solved by reducing it to the symmetric generalized eigen problem defined by the between-group and the within-group covariance matrices. Unfortunately, these matrices are singular in classification tasks where the number of given variables exceeds the number of objects for training. The idea behind Fisher's Linear Discriminant Analysis is to reduce the dimensionality of the data to one or more dimensions. That is, to take p-dimensional, $X \in R^p$, and map it to J -1 dimension, where J is the number of population groups or categories. The discriminant function is defined

by $$Z = \underline{X}^T w = \sum_{j}^{p} w_j X_j$$

where $w = (w_1, w_2, ..., w_p)^T$ and $\underline{X}^T = (X_1, X_2, ..., X_p)$

The FLDA objective is to maximize the variation between groups in such away that minimizing the variation within each class based on the following objectives.

$$\text{Maximizing of } \quad \psi(w) = \frac{w^T S_B w}{w^T S_w w} \quad \text{..................................(12)}$$

where

$$S_B = \sum_i n_i (\overline{X}_i - \overline{X})(\overline{X}_i - \overline{X})^T$$

$$S_w = \sum_{i=1}^{K} (n_i - 1) S_i = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)(X_{ij} - \overline{X}_i)^T$$

n_i are the number of cases in the i^{th} group

$$\overline{X}_i = \frac{1}{n_i} \sum_{i \in j}^{n_i} X_i \quad and \ \overline{X} = \frac{1}{N} \sum_{i=1}^{K} \sum_{j}^{n_i} X_i = \frac{1}{N} \sum_i^{K} n_i \overline{X}_i, \ Where \quad N = \sum_{i=1}^{K} n_i$$

S_B is called the between class groups matrix and S_w is called the within groups scatter matrix. In addition to these the pooled variance-covariance matrix will be obtained by

$$S_{Pooled} = \frac{1}{\sum_{i=1}^{K} (n_i - 1)} \{S_w\} = \frac{1}{\sum_{i=1}^{K} (n_i - 1)} \left\{ \sum_{i=1}^{K} (n_i - 1) S_i \right\}$$

$$= \frac{1}{\sum_{i=1}^{K} (n_i - 1)} \left\{ \sum_{i=1}^{K} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)(X_{ij} - \overline{X}_i)^T \right\} \ this \ is \ an \, estimatore$$

$$of \ \Sigma(the population \ \mathrm{var} iance - \mathrm{cov} ariance \ matrix)$$

In this cases we are expecting to have a large gap between the groups in such away that minimizing the within group scatterness. The total scatter can be defined as

$$S_T = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})(X_{ij} - \overline{X})^T$$

This can also decomposed into the two components namely the between group scatter matrix (S_B) and the with in the group scatter matrix (S_w)

That is $S_T = S_w + S_B$

Thus the above eq(12) can written as

$$J(w)=\frac{w^T S_B w}{w^T S_w w}=\frac{w^T (S_T - S_w) w}{w^T S_w w}$$

$$=\frac{w^T S_T w - w^T S_w w}{w^T S_w w}$$

$$\Rightarrow J(w)=\frac{w^T S_T w}{w^T S_w w}-1 \text{ ....(13)}$$

And hence can be interpreted as maximizing the total scatter of the data while minimizing the within scatter of the groups

The vector of weights, w, can be obtained by differentiating eq(19) with respect to w, and equating to zero, that is,

$$\frac{\partial\lambda}{\partial w}=\frac{2(S_B w)(w^{/} S_w w)-2(w^{/} S_B w)(S_w w)}{(w^{/} S_w w)^2}=0$$ Or dividing through by $w^{/} S_w w$,

$$\frac{2(S_B w-\lambda S_w w)}{w^{/} S_w w}=0$$

$$(S_B-\lambda S_w)w=0$$

$$(S_w{}^{-1} S_B-\lambda I)w=0 \text{...(14)}$$

It is known that the w is a non-singular matrices so then eq(14) will be true if and only if the matrix $(S_w{}^{-1} S_B-\lambda I)$ is a singular matrix that means its determinant is equal to zero(0).

From eq(14) we can say that $\left|S_w{}^{-1} S_B-\lambda I\right|=0$, and the eigenvalues and eigenvectors will be found by solving these equations.

Classification with several variables

i) If there is an equal covariance matrices $\left(\Sigma_j=\Sigma\right)$ for j=1, 2,… J

allocate X_0 to π_l if the linear discriminant score

$\hat{d}_l{}^{(Q)}(x)=$ the largest of $\hat{d}^{(Q)}{}_1(x)$, $\hat{d}^{(Q)}{}_2(x)$,..... $\hat{d}_k{}^{(Q)}(x)$

where

$$\hat{d}_i(x) = \bar{\mathrm{x}}_{\mathrm{i}}^{'}\mathrm{S}_{\mathrm{pooled}}^{-1}\mathrm{x} - \frac{1}{2}\bar{\mathrm{x}}_{\mathrm{i}}^{'}\mathrm{S}_{\mathrm{pooled}}^{-1}\bar{\mathrm{x}}_i + \ln p_{\mathrm{i}} \quad i = 1,2,\ldots,k$$

And p_i is the prior probability of the i[th] group population (or category)

ii) In the case if the populations has no the same covariance matrices then

Allocate X_0 to π_l if the linear discriminant score

$\hat{d}_l(x) =$ the largest of $\hat{d}_1(x), \hat{d}_2(x), \ldots\ldots \hat{d}_k(x)$

Where

$$\hat{d}_i^{Q}(x) = -\frac{1}{2}\ln|S_i| - \frac{1}{2}(x-\bar{x}_i)^{'}S_i^{-1}(x-\bar{x}_i) + \ln p_{\mathrm{i}} \quad i = 1,2,\ldots,k$$

And $|S_i|$ is the determinant of the variance-covariance matrices of the i[th] population

3.4.3.6. Variable selection

There are several methods of variable selection, namely: stepwise selection, purposive selection, backward selection and forward selection. Out of these, the purposive selection will be used in this study. There are some explanatory variables to be included purposively because there is a strong belief that some of the variables have relationship with the dependent variable. Even though such variables are not significant in the univariate analysis these will be kept in the analysis till the final multivariate analysis

3.4.3.7 Standardized discriminant coefficients

The standardized coefficients are the coefficients of each explanatory variable in the discriminant analysis which has a similar meaning as of the Beta weights in a multiple regression in predicting the categories of the dependent variable(Z) based on the observation (cases) characteristics(X) that is, $Z_i = w_1X_1 + w_2X_2 + \ldots + w_pX_p$. Of course, one must realize that these coefficients reflect the contribution of one variant in the context of the other variants in the model. The explanatory variable which has a low standardized coefficient is the one which contributes small in grouping of the cases into the different cases. On the other hand, the one which has a large standardized coefficient is the one which has a great contribution in discriminating the cases into the different groups (or categories).

3.4.3.8. Test of Statistical significance

There are different tests that has been used to test the goodness of fit of the discriminant analysis model.

1. **Wilks' lambda** is used in an ANOVA (F) test of mean differences in Discriminant Analysis (DA), such that the smaller the lambda for an independent variable, the more that variable contributes to the discriminant function. Lambda varies from 0 to 1, with 0 meaning group means differ (thus the more the variable differentiates the groups), and 1 meaning all group means are the same. The F test of Wilks' lambda shows which variables' contributions are significant. This Wilks` lambda is defined by

$$\Lambda = \frac{S_{wi}}{S_T}$$

where $S_{wi} = S_i = \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)(X_{ij} - \overline{X}_i)^T \; the\, i^{th} \; with\, in\, class\, scatter\, matrix$

$$S_T = \sum_{i=1}^{J} \sum_{j=1}^{n_i} (X_{ij} - \overline{X})(X_{ij} - \overline{X})^T \quad \text{the total scaterdness matrix}$$

X_{ij} is the i^{th} explanatory variable in the j^{th} group(population).

$\overline{X}$ is the over all mean

Note all of the J discriminant functions may be statistically significant. That is, only r (where $r \leq J-1$) discriminant functions may be necessary to represent most of the differences among the groups. The following formula is used to compute the X^2 value for assessing the overall statistical significance of all the discriminant functions:

$$X^2 = [n - 1 - (P + G)/2] \sum_{k=1}^{J} \ln(1 + \lambda_k) \text{ ……………… } (15)$$

where λ_k -the eigenvalue of the K^{th} discriminant function.

G- the number of the group

p- the number of variable (or characteristics) in each group.

If $X^2_{calculated}$ value of eq(15) is greater than the tabulated X^2 that means at least the first discriminant function is significant; and other functions may or may not significant.

If the first discriminant function is significant then the statistical significance of the remaining discriminant functions determines whether they jointly explain a significant amount of difference among the all groups that has not been explained by the first discriminant function. Overall significant of the rest discriminant function will be tested by

$$X_l^{\ 2} = [n-1-(P+G)/2]\sum_{k=l}^{J}\ln(1+\lambda_k) \text{ for } l=2,3,\ldots J\text{-}1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (16)$$

If the above, $X^2_{calculated}$ eq(22) value is greater than the tabulated X^2 with (p-r+1)(J-r) degrees of freedom, at least the $(l+1)^{th}$ discriminant function is significant; the other discriminant functions may or may not be significant except the first discriminant function since we have stated that on the above the first discriminant function is significant. In the case of k discriminant functions the above procedure is repeated until the $X_l^{\ 2}$ value is not significant.

If all the discriminant functions are found significant that means all the groups are significantly different with respect to the means of the discriminant function. However; which pair of groups is different? This needs further analysis by examining the means of the discriminant functions.

2. Canonical correlation, R_c:

Squared canonical correlation, which is denoted by R_c^2, is the percent of variation in the dependent discriminated by the set of independent variables in DA or MDA. The canonical correlation of each discriminant function is also the correlation of that function with the discriminant scores. A canonical correlation close to 1 means that nearly all the variance in the discriminant scores can be attributed to group differences.

3. Box's M tests

This is one of the statistical tests which has been used in discriminant analysis. It helps to test the homogeneity of the variances of the different (groups) of the dependent variable. The null hypothesis to be tested is that “the covariance matrices do not differ between groups formed by the dependent.” The assumption of homogenous variance is one of the discriminant analysis assumptions. Box's M uses the F distribution. If the observed significance value is less than the level of significance (i.e. $P(M) < \alpha$), then the variances are significantly different. Here the null hypothesis will be rejected. The probability value of this F should be greater than α to demonstrate that the assumption of homo-scedasticity is held. This test is very sensitive to the assumption of multivariate normality.

When the sample size is large even a small difference in covariance matrices may be found significant by Box's M, when in fact no substantial problem of violation of assumptions exists. Therefore, one needs to look at the *log determinants* of the group covariance matrices. If the group log determinants are similar, then a significant Box's M for a large sample is usually ignored. Dissimilar log determinants indicate violation of the assumption of equal variance covariance matrices, leading to greater classification errors. When violation occurs, quadratic discriminant a nalysis will be appropriate.

3.4.4. Logistic versus Discriminant analysis

Efron (1975) has shown that logistic regression estimators are between one-half and two-thirds as efficient as discriminant function estimators when the data are multivariate normal with equal covariance matrices. Thus, as long as the data are strictly normal with equal covariance matrices, linear disgiminant function estimators are more economical to calculate and are more efficient than logistic regression MLEs.

But, "Another important unanswered question -is the relative efficiency under some model other than [multivariate normality]"

Logistic regression answers the same questions as discriminant analysis. It is often preferred to discriminate analysis as it is more flexible in its assumptions and types of data that can be analyzed. Logistic regression can handle both categorical and continuous variables, and the predictors do not have to be normally distributed, linearly related, or of equal variance within each group (Tabachnick and Fidell 1996).

It is recommended that the sample size in discriminant analysis should provide at least 20 cases for each independent variable. A sample size smaller than this can result in discriminant coefficients that are not stable across samples and there it is not trust worthy (Hair et al.,1998). It is also recommended that the smallest group size in the dependent variable categories be at least 20, with an absolute minimum greater than the number of independent variables. If these conditions are held, unequal sample sizes are not problematic in themselves.

CHAPTER-IV

STATISTICAL DATA ANALYSISAND DISCUSSION

4.1 INTRODUCTION

Knowing the socio-demographic factors and their interaction terms which affect bipolar disorder patients is crucial. Even though there are several potential socio-demographic factors which affect bipolar disorder patients since due to the unavailability data on bipolar disorder patients, we have dealt with some factors on which data are available. So far there are few studies carried out on bipolar disorder patients in Ethiopia.

In this study, the response variable is the severity of bipolar disorder which has three possible values: These are 0, 1, and 2. As we have explained previously in the methodology section, 0 refers to no bipolar disorder symptom (i.e. subject has no bipolar disorder), 1 indicates a moderate symptom of bipolar disorder and 2 refers to severe bipolar disorder symptom. We use polytomous ordinal logistic regression since the severity of bipolar disorder have hierarchy even though it is difficult to get the actual difference between any two levels of severity of bipolar disorder. In addition to the polytomous ordinal logistic regression model we will use discriminant analysis as an alternative model.

4.2. Summary of descriptive statistics

As we have mentioned earlier, about 828 psychiatric patients were included in the study, of which 317 patients are identified as potential bipolar disorder patients. Table 4.1 displays the distribution of bipolar disorder patients by socio-demographic characteristics.

As can be seen from Table 4.1, a bout 76% (629)bipolar disorder patients were living in rural areas and the rest 199(24%) were living in Urban/sub-urban areas. About 65.7% (544) of the patients are males and the rest 284(36.3%) of the patients are females. A slight higher than half (57.2%) of the psychiatric patient were below 30 years of age (Youngster) and the rest 354(42.9%) were above 30 years of age (adults).

More than half of the psychiatric patients 434(52.4%) were married and the rest 394(47.6%) were not currently married. Of the psychiatric patients 479(57.9%) had one or more children and the rest about 349(42.1%) had no children. 531(64.2%) of the psychiatric patients are illiterate and the rest 297(35.8%) of the patients were literate. About 428(51.7%) of the patients were living in a situation where quality of life is lower than average and the rest 400(48.3%) had an average or better quality of life.

Table 4.1 Distribution of psychiatric patients by demographic characteristics.

Patients characteristics	Catagories	Frequencies	Percentages (%)
Residence	Urban	199	24%
	Rural	629	76%
Quality of life	Quality higher than /on average	400	48.3%
	Quality lower than average	428	51.7%
Marital status	Married	434	52.4%
	Separated/widowed/divorced	394	47.6%
Number of children	No children(0)	349	42.1%
	Have one or more children	479	57.9%
Education	Illiterate	531	64.2%
	Literate	297	35.8%
Sex	Male	544	65.7%

	Female	284	34.5%
Age	<=30	474	57.2%
	>30	354	42.8%

4.3 Bivariate Analysis of the ordinal logistic regression

In this analysis, the dependent variable is “severity of the bipolar disorder”. It has three ordinal categories namely no bipolar disorder, moderate bipolar disorder and severe bipolar disorder. There is some order among the categories as no bipolar disorder symptom is less than moderate bipolar disorder symptom and less than sever bipolar disorder symptom. In SPSS ordinal logistic regression the higher order is taken as a reference category by default. All categories found before that higher ordered categories will be compared with respect to that category.

To determine the socio-demographic factors which have significant relationship with the severity of bipolar disorder patient, first we need to test the association between each socio-demographic factor with the dependent variable (severity of bipolar disorder) using ordinal logistic regression. This is a preliminary analysis to which variable are significant related to severity of bipolar disorder patient. The results are presented in Table 4.2.

Table 4.2 Results from bivariate ordinal logistic regression analysis

		Estimate	Std. Error	Wald	Df	Sig.	95%CI	
							LCB	UCB
Intercept	NoneBPD	.504	.082	37.937	1	.000	.344	.665
	ModerBPD	2.025	.115	311.495	1	.000	1.800	2.249
Variables	Urban]	.119	.162	.543	1	.461	-.198	.436
	Rural	0(a)	.	.	0	.	.	.
Intercept	None BPD	.633	.104	37.116	1	.000	.430	.837
	Mod BPD	2.159	.133	263.101	1	.000	1.898	2.419
Variable	Lower Qual	.293	.140	4.366	1	.037*	.018	.568
	Higher Qual	0(a)	.	.	0	.	.	.
intercept	None BPD	.238	.117	4.131	1	.042	.008	.467
	ModerBPD	1.766	.139	161.987	1	.000	1.494	2.038
Variables	male]	-.370	.145	6.507	1	.011*	-.655	-.086
	Female	0(a)	.	.	0	.	.	.
intercept	None BPD]	.638	.110	33.314	1	.000	.421	.854
	ModerBPD	2.162	.138	244.872	1	.000	1.891	2.433
Variables	<=30 years	.280	.142	3.880	1	.049*	.001	.559
	>30 years	0(a)	.	.	0	.	.	.
intercept	None BPD	.824	.109	57.705	1	.000	.612	1.037
	ModerBPD	2.369	.139	289.655	1	.000	2.096	2.642
Variables	Married	.629	.143	19.410	1	.000*	.349	.909
	curr not married	0(a)	.	.	0	.	.	.
intercept	None BPD	.243	.091	7.138	1	.008	.065	.422
	ModerBPD	1.783	.118	227.107	1	.000	1.551	2.015
Variables	No-children	-.582	.145	15.996	1	.000*	-.867	-.297
	One or more children	0(a)	.	.	0	.	.	.
intercept	None BPD	.466	.083	31.632	1	.000	.304	.628
	ModerBPD	1.986	.115	298.907	1	.000	1.761	2.211
Variables	Eudication	-.005	.021	.052	1	.819	-.046	.037

* Significant at 0.05 level of significant

As can be seen from Table 4.2 the effect of socio-demographic characteristics like quality of life, sex, age, marital status, and number of children are found to be statistically significant ($p<0.05$) whereas place of residence and education are not statistically significant at 0.05 level of significance. Since we are using here is a purposive variable selection even though a variable is not significant at 0.05 level of significant it is possible to include clinically important variables/factors in the multivariate analysis variable(s) is clinically important. In this study place of residence is one of the clinically important variables even though it is not significant at 0.05 level of significant in the bivariate analysis. This will be included in the multivariate analysis.

4.4. **Results based on Polytomous ordinal multiple logistic regression analysis**

As is known, the main problem in the bivariate data analysis is to measure the relationship (or association) between the dependent variable and an explanatory variable. In practice, however, those individual variables weakly associated with the dependent variable (the severity of the bipolar disorder) may be highly associated with the dependent variable when that variable is taken together with other variable(s) (Hosmer and Lemshow, 1989).

4.4.1. Parameter estimation of the polytomous ordinal logistic regression

The results from the pliminary polytomous ordinal multivariate analysis are presented in Table 4.3.

Table 4.3. Preliminary model analysis results without including interaction term

		Estimate	Std. Error	Wald	df	Sig.	95% CI	
							LCB	UCB
intercept	None-BPDS	.869	.229	14.372	1	.000	.420	1.318
	Moderate BPDS	2.443	.246	98.529	1	.000	1.961	2.926
variables	Urban	.155	.165	.880	1	.348	-.169	.479
	Rural	0(a)	.	.	0	.	.	.
	Lower A Qual	.173	.144	1.434	1	.231	-.110	.455
	Higher A Qual	0(a)	.	.	0	.	.	.
	Male	-.203	.152	1.791	1	.181	-.502	.095
	Female	0(a)	.	.	0	.	.	.
	< 30 years	.640	.162	15.549	1	.000	.322	.958
	>= 30 years old	0(a)	.	.	0	.	.	.
	Married	.421	.193	4.755	1	.029	.043	.800
	Currently not married	0(a)	.	.	0	.	.	.
	Has no children	-.527	.212	6.194	1	.013	-.941	-.112
	Has one or more children	0(a)	.	.	0	.	.	.

* significant at 5%

From this preliminary analysis one can see that age, marital status and number of children of the patient have significant relationship with the severity of bipolar disorder of patients and others factors has not ($p<0.05$).

In addition to the investigation of main effects, we considered some meaningful two-way interaction terms. The way of selecting the interaction is first by regressing the dependent variable with all main effects and an interaction of two independent factors. Those significant interactions at $\alpha = 0.05$ and which are found meaningful are included to the preliminary model. It was found that the only meaningful interaction term which is significant is quality of life and residence of the psychiatric patient.

After including the significant two-way meaningfull interaction term to the model, we have got the final preliminary model results which are displayed in Table 4.4. It is referred that the results in Table 4.4 as estimates of the preliminary final model until we assess and ascertain model assumptions.

Table 4.4 Results of the final polytomous ordinal regression Analysis

		Estimate	Std. Error	Wald	df	Sig.	95% CI	
							LCB	UCB
intercepts	None-BPDS	.981	.235	17.378	1	.000	.520	1.442
	Moder BPDS	2.566	.253	103.130	1	.000	2.071	3.061
Variables	Urban	.583	.231	6.367	1	.012	.130	1.036
	Rural	0(a)	.	.	0	.	.	.
	Lower than Aver Qual	.388	.168	5.336	1	.021	.059	.717
	Higher than Aver Qual	0(a)	.	.	0	.	.	.
	Male	-.196	.152	1.653	1	.199	-.494	.103
	Female	0(a)	.	.	0	.	.	.
	Less than 30 years	.633	.163	15.114	1	.000	.314	.952
	Higher than 30 years	0(a)	.	.	0	.	.	.
	Married	.408	.194	4.424	1	.035	.028	.789
	Currently not married	0(a)	.	.	0	.	.	.
	Has no children	-.534	.213	6.320	1	.012	-.951	-.118
	Has one or more children	0(a)	.	.	0	.	.	.
	Urban*lower than aver qual	-.847	.331	6.539	1	.011	-1.497	-.198
	Urban*higher than aver qual	0(a)	.	.	0	.	.	.
	Rural *lower than aver quality of life	0(a)	.	.	0	.	.	.
	Rural* higher than aver quality of life	0(a)	.	.	0	.	.	.

* Significant at 0.05 level of significant

From this preliminary final ordered logistic regression analysis estimates one can see significant relationship between severity of bipolar disorder and place of residence, quality of life, age, number of children and marital status. In addition to these factors there is a significant relationship between the severity of bipolar disorder and with the two-way interaction term between residence and quality of life.

In addition to parameter estimates, their standard errors, Wald test statistical values and the corresponding p-values, Table 4.4 provides 95% confidence interval estimates of these parameters. One can also make a test on the significance of parameters based on the confidence interval estimates. That is if the 95% confidence interval contains zero, then the parameter is not significantly different from zero indicating that the effect of the corresponding factors is not significant on the severity of bipolar disorder. On the other hand, if the confidence interval does not contain zero; then the parameter is significantly different from zero and the corresponding factor is significantly related to severity of bipolar disorder.

Interms of the logit representation of the polytomous ordinal multiple logistic regression model, we have the following model representation.

Logit($\hat{p}(Y \leq j)$)= $\hat{\alpha}_j$ +0.583X_{resid}+0.388$X_{quality}$ +0.408X_{mar} + 0.633X_{age}-0.534X_{NCh} - 0.847$X_{resid}X_{qual}$, j=0,1.

Further can be written as

Logit($\hat{p}(Y \leq j)$ =0.981+0.583X_{resid}+0.388$X_{quality}$ +0.408X_{mar} + 0.633X_{age}-0.534X_{NCh} - 0.847$X_{resid}X_{qual}$

Logit($\hat{p}(Y \leq j)$ =2.566+0.583X_{resid}+0.388$X_{quality}$ +0.408X_{mar} + 0.633X_{age}-0.534X_{NCh} - 0.847$X_{resid}X_{qual}$

4.4.2 Testing of the proportional odds model

If the assumption of polytomous odds model fails, it is difficult to use the cumulative logit link function. Rather it is better to use any of the other suggested models by which the ordinal logistic regression can be analyzed such as the continuation-ratio model, partial proportional odds model, and adjacent-category logistic model. Test of the assumption of the proportional odds or the use of cumulative logit link function can be done based on test of parallelism.

Note that when we fit an ordinal logistic regression, it is assumed that the relationship between the independent variables and the logits are the same for all the logits. Here the null hypothesis to be tested is that all possible logits are parallel. This assumption is the core assumption to use and not to use the cumulative logit. If this assumption is attainable, we will use the cumulative logit representation to analyze the ordinal logistic regression.

This assumption will be tested by using the Chi-square test which is the difference between the full model and the constant model. If the lines are parallel, the observed significance will be large since the general model does not improve the fit very much. The null hypothesis "the logits are parallel" is rejected when the observed significance is smaller than the level of significance (α is less than the p-value). That means the general model improves very much by including the explanatory variable. The results on test of parallel lines is displayed in Table 4.5

Table 4.5 Test of Parallel Lines

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	314.580			
General	301.682	12.898	7	.075

In other words the null hypothesis can be tested as the location parameters or the slope coefficients are the same across the response category. That means the logit which corresponds for the threshold or intercepts are parallel. And the alternative hypothesis is that the set of logits which correspond to the threshold will not be parallel. From Table 4.5 one can see that the observed significance value is 0.75 which is greater than the level of significance 0.05. Hence, the assumption of parallelism is fulfilled and the logit link function is appropriate.

4.4.3 Goodness of fit tests for polytomous ordinal multiple logistic regression models

Before we proceed to examine the individual coefficients, we have to test the null hypothesis "the regression coefficients are all zero". However, it did not assure that all the factors included to the model were significantly related to the severity of the bipolar disorder. It only tests the overall significance of the model. This was obtained based on the change in – 2log-likelihood when the variables were added to the model to the model containing only the intercept. The change in the likelihood has a Chi-square distribution even when there are cells with small observed and predicted counts. In order to have an overall good of fit model the observed significance should be less than the level of significance fixed in this study. For this analysis the model–fitting information is given in the Table 4.6

Table 4.6 Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	364.340			
Final	314.580	49.760	7	.000

Link function: Logit.

As it is shown in Table 4.6 the observed significance is 0.00 which is much less than the level of significance 0.05 what we have considered in this study. So that we could conclude that the overall goodness of fit has been met.

In addition, the Pearson and deviance Chi-square tests are used to test how much predicted values differ from observed frequencies. A well fitted model is one which has large value of the observed significance. The results of the goodness of fit for this study are shown in Table 4.7.

Table 4.7 Goodness-of-fit

	Chi-Square	Df	Sig.
Pearson	106.486	115	.702
Deviance	114.921	115	.485

Link function: Logit.

As can be seen in Table 4.7, the Pearson's Chi-square value result is large, observed significance value is much larger than level of significance (0.05) that is 0.702. We can conclude that the predicted values and the observed value in each cell were not much different and the model provides a good fit.

Table 4.8 Classification table of the polytomous ordinal logistic regression

	Predicted			Correctly classification
	None BPD	Moderate BPD	Sever BPD	
None BPD	342	119	50	66.93%
Moderate BPD	50	126	42	56.89%
Sever BPD	34	22	43	43.54%

The result presented in Table 4.8 can be taken as estimates of the final model to be fitted for the data in this study. The interpretations and discussion of result presented in Section 4 are based on the final model estimates shown in Table 4.4.

4.5. Resulted based on Multiple Discriminant analysis

In this section, we discuss the construction of discriminant functions, assessment of model assumptions and other related topics.

4.5.1. Testing of the equality of the group means.

We begin with the analysis by testing whether the means of each explanatory variable is equal in the three groups of severity of bipolar disorder. As it is discussed in the methodology section, the variable which has a small Wilks` lambda is the one which has a high power of discriminating among the groups of severity of bipolar disorders. On the other hand, the explanatory variable which has a large Wilks` lambda is the one which contributes least for discriminating between the different groups. Table4.9. shows the results of testing the equality of group means based on Wilk`s lambda.

Table 4.9. Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Residence	.999	.366	2	825	.694
Quality of life	.986	5.723	2	825	.003*
Sex	.991	3.898	2	825	.021*
Age	.995	1.991	2	825	.137
Marital status	.969	13.336	2	825	.000*
Number of children	.979	8.778	2	825	.000*

* significant 5%

From the results in Table 4.9., we can conclude that quality of life, marital status, number of children and sex of the psychiatric patient are important variables for discriminating the psychiatric patient into the different severity categories of the bipolar disorder. Of these factors, marital status has the highest contribution for the discriminating functions since its value of Wilks` lambda is the smallest as we compare with the others Wilk`s lambda factors. On the other hand the sex of a patient has less discriminating power in the discriminant function as we compare with other variables since its Wilks` lambda is statistically significant.

4.5.2. **Test for equality of group covariance matrices**

The Box`s M test is used for testing the assumption of equal group covariance matrix. The assumption of equal group covariance matrix will be attained if the observed significance is larger than that of the level of significance α . The null hypothesis to be tested here is that there is an equal group covariance matrix in each group and the alternative hypothesis to be tested is equality of group covariance matrices is not attained. Result of Box`s M test are presented in Table 4.10. As can be seen from

Table 4.10, since the observed significance (0.236) is larger than that of the level of significance (0.05), we do not reject the equality of group covariance matrices. Hence the linear discriminant analysis will be used so as to classify the cases into different classes. If this assumption had not been fulfilled, we would have used the quadratic discriminant analysis for the classification of cases in to the different groups.

Table 4.10. Box`s Test of equality of covariance matrices

.

Box's M		49.045
F	Approx.	1.148
	df1	42
	df2	285801.746
	Sig.	.236

4.3.2.3. Construction of discriminant functions

As is known that the eigenvalue, also called the characteristic root of each discriminant function, reflects the ratio of importance of the dimensions which classify cases of the dependent variable. There is one eigenvalue for each discriminant function. The results are displayed in Table 4.11

4:11: Eigenvalues

Function	Eigenvalue	% of Variance	Cumulativ e %	Canonical Correlation
1	.062(a)	87.7	87.7	.242
2	.009(a)	12.3	100.0	.093

The first function in Table 4.11 accounts for about 87.7 % of the total among-groups variability and the second accounts for about 12.3% of the total among-group variability. Next we test the null hypothesis that overall discriminant functions are significant as a whole or not discriminating the three groups of severity of bipolar disorder. As one could see from Table 4.12 the discriminant functions jointly have a significant chi-square. However the second discriminant function alone is not significant. That is, the two discriminant function jointly discriminate well the cases (patients) in the three groups.

Table4.12 Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.933	56.940	12	.000
2	.991	7.168	5	.208

Next we compute the standardized discriminant function coefficients corresponding to each variable in the two discriminant functions. The standardized discriminant function coefficients are presented in Table 4.13.

Table 4. 13. Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
Residence	.087	-.182
Quality of life	.292	.629
Sex of patient	-.234	.012
Age of patient	.613	-.625
Marital status	.560	.590
No. of children	-.419	.886

From the above Table 4.13 we can conclude that place of residence in the first function contributes least for the discriminating cases into the different groups. In another words, place of residence has the least effect for predicting group membership than other variables. However, age in the first function contributes the greatest for discriminating cases into the different groups. That is, age has the greatest contribution for predicting group membership. Age is followed by marital status, number of children and quality of life in the first function.

Sex in the second function contributes the least for discriminating cases into the groups. This is to mean sex has the least effect for predicting group membership than other variables in the second function. However, number of children of a patient in the second function contributes the greatest for discriminating cases into the different groups. That is, number of children of a patient has the greatest contribution for predicting group membership in the second function. Number of children of a patient is followed by quality of life, age, marital status, place of residence.

Another way to look at the discriminating power of the discriminant function is using group centroids. Group centroids are the mean discriminant scores for each of the dependent variable categories for each of the discriminant functions in the multiple discriminant analysis. If there are J groups there will be J number of mean centroids. If the group means are so apart, then the corresponding groups can be well discriminated. The closer the means, the more errors of classification likely to occur. Based on group centroids presented in Table 4.14 it is possible to say something about the discrimination of the groups. That is the first function tends to discriminate none bipolar disorder group from the moderate and severe bipolar disorder group. Similarly, the second function tends to discriminate the severe bipolar disorder group from none and moderate disorder groups.

Table 4.14 Functions at Group Centroids

Severity of BPD	Function	
	1	2
NoneBPD	.194	-.012
ModerateBPD	-.357	-.081
Sever BPD	-.214	.240

The discriminating power of the different discriminant functions can also be seen and say something simply by studying /looking at plots like separate-group plots, a combined–group plot, a territorial map. Separate and combined group plots show where cases are located in the property space formed by two functions. The territorial map shows inter-group distance on the discriminant functions. The combined and territorial maps are shown in Appendix I.

Moreover the structure matrix also called the structure correlation or discriminant loadings tells how closely a variable is related to each discriminant function. The discriminant loadings are the correlations between a given independent variable and the discriminant scores associated with a given discriminant function. The canonical structure matrix for this study is depicted in Table 4.15

Table 4.15 Structure Matrix/ discriminant loading

Variables	Function	
	1	2
Marital status	.712(*)	.282
No of children	-.581(*)	.146
Sex of the patient	-.389(*)	.013
Quality of life of patient	.408	.631(*)
Age of patient	.230	-.418(*)
Residence of patient	.074	-.249(*)

*Largest absolute correlation between each variable and any discriminant function

From the discriminant loading in Table 4.15, we observe that marital status, number of children, sex of the patient have a strong relationship or correlation with the first discriminant function. That means these factors have high contribution to the discriminating power in the first function as compared with the second discriminant function. On the other hand quality of life, age, place of residence of the patient are strongly associated with the second function but weakly associated with the first discriminant function. Hence these factors have high discriminant power in the second discriminant function as compared to the first discriminant function.

4.5.4. Classification using the Fisher linear discriminant functions

After the discriminant functions are determined and groups are differentiated, the utility of these functions can be examined in terms of their ability to correctly classify each case (observation) into their prior group. There are different methods of classifying observation into their prior groups. Of these, the Fisher`s classification function uses the unstandardized coefficients generalized distance functions which are based on the Mahalanobis distance. The Fisher classification function coefficients are presented in Table 4.16.

Table 4.16 Fisher Classification Function Coefficients

Variable	Bpdmean		
	.00	1.00	2.00
Residence	9.645	9.562	9.454
Quality of life	6.404	5.994	6.485
Sex of patient	6.331	6.602	6.539
Age of patient	5.408	4.812	4.584
Marital status	6.346	5.637	6.185
No of children	2.474	2.820	3.279
(Constant)	-24.914	-23.583	-24.222

Based on the classification coefficients associated with each explanatory variable, we can write the Fisher linear discriminant function for the three groups: No bipolar, moderate bipolar disorder and sever bipolar disorder as

$$Z_{i1}= -24.914 +9.645X_{i,res}+ 6.404X_{i,qual}+6.331X_{isex} +6.346X_{i,mar}+2.474X_{i\ Nch}+ 5.408X_{iage}$$

$$Z_{i2}= -23.583 + 9.562X_{i,res}+ 5.994X_{i,qual}+6.602X_{isex} + 5.637X_{i,mar}+2.82X_{i\ Nch}+ 4.812X_{iage}$$

$$Z_{i3}= -24.222 + 9.454X_{i,res}+6.485X_{i,qual}+6.539X_{isex} +6.185X_{i,mar}+ 3.279X_{i\ Nch} + 4.584X_{iage}$$

Note that the first, second, and third functions are corresponding to no, moderate, and severe bipolar disorder, respectively. The classification rule based on these three functions is as follows. First one has to evaluate each of the above functions for each case (or patient) as a function of his/her characteristics or factors. Then a patient will be allocated to the group where his or her score function is large. That means if a patient has high score on the first function corresponding to the first group(no bipolar disorder) then the patient will be allocated to no bipolar disorder group and in a similar way we will continue the allocation of other patients in each of the three groups.

The result of classification is presented in Table 4.17. The classification table is a table in which the rows are the observed group of membership and the columns are the predicted group membership. When prediction is perfect all cases will be on the diagonal. The percentage of cases on the diagonal refers to the percentage of correct classification. This percentage is called hit ratio.

Table 4.17.Classification Results

Bpdmean			Predicted Group Membership Total			
			NoneBPD	Moderate BPD	SeverBPD	total
Original	Count	.00	289	137	85	511
		1.00	73	101	44	218
		2.00	42	31	26	99
	%	.00	56.6	26.8	16.6	100.0
		1.00	33.5	46.3	20.2	100.0
		2.00	42.4	31.3	26.3	100.0

*50.8% of the original grouped cases correctly classified

As can be seen from Table 4.17 the discriminant analysis correctly classified about 50.8% the psychiatric patients into the different severity of bipolar disorder groups. About 56.6% of the cases (or psychiatric patients) are correctly classified into non-bipolar disorder group, about 46.3% of the cases in to the moderate bipolar disorder group patients are correctly classified, about 26.3% of the cases into the severe bipolar disorder group.

4.6. Discussion and interpretation of the results

The result of the bivariate analysis based on ordinal logistic regression as given in Table 4.2 shows that quality of life, marital status, number of children, sex and age of the patients have significant effect on severity of bipolar disorder. On the other hand education level of the patient and the place of residence of the patient have no significant relationship with severity of bipolar disorder. Similarly, from the final polytomous ordinal logistic regression analysis Table 4.4 we obtained that place of residence, quality of life, age, number of children and marital status of the patient have significant relationship with the severity of bipolar disorder. In addition there is a meaningful two-way interaction terms that is residence and quality of life of the patient which is significantly related to the severity of bipolar disorder.

Referring to the final model result presented in Table 4.4 the result is discussed as follows.
Sex is not significantly related with the severity of bipolar disorder which means that the probability of having bipolar disorder is the same in both sexes. The prevalence of bipolar disorder is the same among males and females. This is consistent with the previous studies by Weissman et al., (1988,1996).

There is a relationship ($\hat{\beta}$ =0.633, p<0.05) between the severity of bipolar disorder and the age of the psychiatric patients. That means, as age increases the probability of being in one of the higher category of the bipolar disorder symptom will decrease by $e^{0.633}$ = 1.88. This indicates that the prevalence of bipolar disorder is higher in the age of youngsters (age<=30 years) as we compare to adults (age >30) years. This is consistent with some previous studies (Weissman et al., 1996).

There is a relationship ($\hat{\beta}$ =0.408, p < 0.05) between the severity of bipolar disorder and marital status. Here the odds ratio of patients who are currently married to those patients who are not currently married is $e^{0.408}$ = 1.503. This is to mean that the odds of being in the higher category of the severity of bipolar disorder for the currently married is 1.503 increase in odds of being in the higher category of the bipolar disorder for the psychiatric patients who are not married. This suggests that currently married patients are more prone than those of not currently married to bipolar disorder. This result is not consistent with the previous studies by Weissman et al (1982).

There is a relationship ($\hat{\beta}$ =-0.534, p<0.05) between the severity of bipolar disorder and the number of children. From this estimated value, the odds of the psychiatric patients who had no children to that of psychiatric patients who had one or more children is $e^{-0.534}$ = 0.586. This is to mean that the odds of being in the higher category of the severity of bipolar disorder of psychiatric patients who had no children increases by 0.586, the odds of being in the higher category of the bipolar disorder symptom the psychiatric patients who had one or more children. That is psychiatric patients who had no children are less prone to those patients who had one or more children.

From the significant interaction effect we observe that the effect of quality of life is modified by the effect of residence area where the psychiatric patients are living. For psychiatric patients living in rural areas , the odds of being in low quality of life to that of higher quality of life is $e^{0.124}$ =1.32. That is the odds of being in the higher category of the severity of bipolar disorder for psychiatric patients who have low quality of life and live in rural is 1.32 increase in the odds of being in the higher category of the severity of bipolar disorder for psychiatric patients who have high quality of life and living in rural area. Similarly for psychiatric patients living in urban area, the odds of being in low quality of life to that of higher quality of life is $e^{0.388}$ =1.474 which is a constant ratio for all the three severity groups of bipolar disorder. That is the odds of being in the higher category of the severity of bipolar disorder for psychiatric patients who have low quality of life and live in urban areas increases by 1.47, the odds of being in the higher category of the severity of bipolar disorder for psychiatric patients who have high quality of life and living in urban area. The result suggests that psychiatric patients with low quality of life are slightly more exposed to bipolar disorder in urban areas than rural areas.

From Table 4.9 one can see that quality of life, sex, marital status, and number of children of the patient have a high contribution for the classification of the patients (cases) into the different categories of the severity of bipolar disorder. These are, no, moderate, and sever bipolar disorder. All of these factors were significant in polytomous ordinal binary analysis of the ordinal logistic regression; nevertheless, sex is not significant in the case of the final polytomous ordinal multivariate analysis.

Based on discriminant analysis the structure matrix (Table (4.15)) we conclude that marital status, number of children and quality of the patients have high correlation with the first function which discriminates the no bipolar disorder from the other two bipolar disorder symptoms. On the other hand, quality of life, age, residence and the education status of a psychiatric patients have high correlation with the second function which discriminates the sever bipolar disorder from the other categories of the psychiatric patients.

From Table 4.13 standardized discriminant analysis coefficients we can conclude that marital status and age of the patients have a higher contribution in discriminating the no bipolar disorder from the other two categories, but residence has the least effect in discriminating these categories. On the other hand, number of children and quality of life have higher contribution in the discriminating sever bipolar disorder from the other categories.

CHAPTER V

CONCULUSION AND RECOMMENDATION

5.1 Conclusion

The main purpose of this study is in order to identify some factors and their interactions affecting bipolar disorder patients. In addition, the study attempted in choosing the best fitted model for the data under consideration from the polytomous ordinal logistic regression and the discriminant analysis. Three groups of bipolar disorder patients are considered according to the severity of symptom- no, moderate and sever bipolar disorder. The distribution of patients in these three groups are: 61.7% of the patients fall to the no bipolar disorder symptom and 46.3% belong to moderate bipolar disorder symptom and 26.3% of the patients into sever bipolar disorder symptom group.

From the polytomous ordinal logistic regression we observed that quality of life, marital status, place of residence, number of children age , and the two-way interaction term residence*quality of life are statistically significant. However, there is no significant relationship between sex of the patient and severity of bipolar disorder. That is there is no difference in the prevalence of severity of bipolar disorder in both sexes.

There is an indirect relationship between the severity of bipolar disorder and age. That is the severity of bipolar disorder increases for the case of youngsters than those of the adults. On the other hand, there is a direct relationship between the severity of bipolar disorder and the number of children the patient. Patients who have one or more children are more vulnerable to the severity of bipolar disorder than those have no children.

Moreover, based on multiple discriminant analysis for the three groups of severity of bipolar disorder, we constructed two discriminant functions. The first discriminant function is highly influenced by marital status, number of children, and sex of the patient can be referred to as “demographic factors” that discriminates no bipolar disorder patients from the other two groups of psychiatric patients than the other one. The second function which can be referred to as “economic factor” which is highly determined by quality of life, age, place of residence of a psychiatric patients that discriminates the severe bipolar disorder from the other two groups of psychiatric patients.

To compare the productive ability of the polytomous ordinal logistic regression and multiple discriminant analysis we used the percentage of correctly classifying cases. Using polytomous ordinal logistic regression we found 61.7% of the were patients correctly classified in the severity of bipolar disorder groups. Whereas using multiple discriminant analysis we found 50.8% correctly classified cases. Thus for the data under consideration, the performance of the polytomous logistic regression is better than the multiple discriminant analysis.

5.2. **Recommendation**

As there are a limited number of mental health centers in Ethiopia the society have little perception about mental illness. So as to have more awareness, government and non-government organizations must expand mental health centers and increase the number of qualified professionals in that field of psychiatry. Budget must be allocated to conduct researches on mental health. In Ethiopia the perception of mental illness is related with the sin of the parents or relatives of the patient or the patient it self. In order to cast this idea the society needs a great intervention from educated persons, NGOs, and government.

The concerned bodies should give due attention to factors such as place of residence, quality of life, marital status, age, number of children of the patient which have a significant relationship with the severity of bipolar disorder patients.

The Ministry of Health and other concerned bodies should give due attention to psychiatric patients living in urban areas because the likelihood of severe bipolar disorder is higher in urban than in rural areas. Health professionals also need to carry out research in order to know the reasons why urban psychiatric patients are more vulnerable to severe bipolar disorder than of the rural dwelling psychiatric patients.

The government and other concerned bodies need to adjust the condition that will enhance the well being of the psychiatric patients wherever a patient lives. This is because, in many cases, psychiatric patient who have low quality of life are prone to develop bipolar disorder regardless of one's living area compared with those who have better quality of life. In addition, there should be an intervention program like family planning to limit the number children of psychiatric patients should have since the study indicates that psychiatric patients who have children are more susceptible to have severe bipolar disorder.

It is advisable that health professionals and other members of the society need to be aware that the prevalence rate of bipolar disorder differ among persons in different age categories and marital status. As to the study, the severity of bipolar disorder is high among youngsters and married psychiatric patients. In this regard, due attention should be given to youngsters. Similar attention be given to married psychiatric patients as they have family burden and responsibility. However, the findings indicate that there is no gender difference with respect to the severity of bipolar disorder.

References.

- Abdulahi H, Hailemariam D, and Kebede D. (2001). Burden of disease analysis in rural Ethiopia. *Journal Ethiop. Med.* 39 (4), 271-81.
- Agresti , A(2002):An Introduction to Categorical Data Dnalysis. John Wiley and Sons, Inc., New york.

- Alem,A.,Menilik,D.,Mesifin,A.(1995).Mental health in in Ethiopia EPHAexpert group report. *Ethiop.j.Health Dev*.,9 (1), 47-62.

- Altshuler LL, Gitlin MJ, Mintz J, Leight KL, Frye MA (2000). Subsyndromal depression is American Psychiatric Association: Practice guideline for the treatment of patients withBipolar disorder. *Am J Psychiatry,* 159 (April suppl).
- Anderson.J and Greenland, S. (1985). An application of Logistic Models to the Analysis of Ordinal Regression. *Biometrika Journal* , 27,189-197
- Angst F, Stassen HH, Clayton PJ, Angst J.(2002) Mortality of patients with mood disorders: follow-up over 34-38 years. *J Affect Disord* , 68, 167-181.
- Angst J, Sellaro H, Stassen A, Gamma A.(2005). Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions. *J Affect Disord*, 84,149-157.
- Arnold Lm, Witzeman KA, Swank ML,(2000). Health related quality of life using the SF-36 in patients with bipolar disorder compaired with patients with cronic back pain and the general population. *J Affect Disorder*, 57, 235-239.
- Bartlett, M. S (1947) Multivariate Analysis. Journal of the Royal Statistical Society Supplement, 9, 176-190.
- Bebbington, P., Ramana, R., (1995). The epidemiology of bipolar affective disorder. *Social Psychiatry and Psychiatric Epidemiology* 30, 279-292.
- Calabrese JC, Hirschfeld RM, Read M (2003).Impact of bipolar disorder on a US community sample. *J clin psychiatry*: 425-432.
- Chen YW, Dilsaver SC (1995) Comorbidity for obsessive-compulsive disorder in bipolar and unipolar disorders. *Psychiatric Res*, 59, 57-64.
- Christie KA,Burke JD Jr, Regier DA, Rae DS, Boyd JH, Locke BZ (1988). "(abstract) Epidemiologic evidence for early onset of mental disorders and higher risk of drug abuse in young adults". *Am J Psychiatry* 145: 971–975.

- Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society* B, 20, 215-242.
- Craddock N, Jones (1999) Genetics of bipolar disorder. *J MedGenet*, 36(8): 585–594.

- Decina, P., Kestenbaum, C. J., Farber, S., Kron, L., Gargan, M., Sackeim, H. A., et al. (1983). Clinical and psychological assessment of children of bipolar probands. *American Journal of Psychiatry*, 140, 548–553.
- Desjarais, EisenbergL, Good B, Kleinman A. (1995) World Mental Health.Problems and Priorities in Low-Income Countries. New york, Oxford University press.
- Desjarlais R.(1995). World mental health: problems and priorities in low-income countries. New York: Oxford University Pres
- Dohrenwend, B. S., & Dohrenwend, B. P. (1967). Field studies of social factors in relation to three types of psychological disorder. *Journal of Abnormal Psychology*, 72, 369–378
- Efron.B (1975). The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis, *Journal of the American Statistical Association*, 70, 892-898.
- Engstro¨m C, Bra¨ ndstro¨m S, Sigvardsson S, Cloninger R, Nylander P (2003) ; personality and age of on set bipolar disorder. *J Bipolar Disord* , 5: 340–348.
- Fagiolini A, Kupfer DJ, Masalehdan A (2005). Functional impairment in remission phase of bipolar disorder. *Bipolar Disord* , 7: 281-285
- Gardner, H. H., Kleinman, N. L., Brook, R. A., Rajagopalan, K., Brizee, T. J., & Smeeding, J. E. (2006). The economic impact of bipolar disorder in an employed population from an employer perspective. *Journal of Clinical Psychiatry*, 67(8), 1209–1218.
- Goodwin FK, Jamison KR.(1990) Manic-Depressive illness. New York, Oxford University Press.
- Guze SB, Robins E.(1970): Suicide and primary affective disorders. *Br J Psychiatry* ;117: 437-438.
- Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C. (1998). Multivariate data Analysis. Upper saddle River, New Jersey: Prentice-Hall.
- Hosmer, D.W.,and Lemshow, S. (1989), Applied Logistic regression, New York: John Wiley and Sons, Inc.

- Høyer EH, Mortensen PB, Olesen AV.(2000) Mortality and causes of death in a total national sample of patients with affective disorders admitted for the first time between 1973 and 1993. *Br J Psychiatry* 176:76-82.

- http://en.wikipedia.org/wiki/History_of_bipolar_disorder.
- Jamison KR(2000). Suicide and bipolar disorder. *J Clin Psychiatry* 61(Suppl.9):47-51.

- Jefferson, J.W., Greist, J.H., (1994). Mood Disorders. In: The American Psychiatric Press textbook of psychiatry (eds R.E. Hales & S.C. Yudofsky & J.A. Talbott), p.p. 465-494. American Psychiatric Press, Washington.
- Johnson A. and Wighern W.(2002) : Applied Multivariate Statistical Analysis 5th ed. Printice-Hall, inc.
- Judd L, Akiskal H.(2003). The prevalence and disability of bipolar spectrum disorders in the US population: re-analysis of the ECA database taking into account subthreshold cases. *J Affect Disord*er 73: 123–131.
- Kieseppä T, Partonen T, Haukka J, Kaprio J, Lönnqvist J. (2004) High concordance of bipolar I disorder in a nationwide sample of twins. *Am. J Psychiatry*,161: 1814-1821.
- Leibenluft E.(2000) Women and bipolar disorder: an update. *Bull Menninger Clin* ; 64 (1):5–17.[Medline]
- Lish, J.D., Dime-Meenan, S., Whybrow, P.C., Arlen Price, R., Hirschfeld, R.M.A., 1994. The National Depressive and Manic-Depressive Association (DMDA) survey of bipolar members. *J. Affect. Disorder*, 31,281-294.
- MacKinnon DF, Potash JB, McMahon FJ, Simpson SG, Depaulo JR Jr, Zandi PP(2005). Rapid mood switching and suicidality in familial bipolar disorder. *Bipolar Disord ,* 7:441-448.
- Malzberg, B. (1956). Mental disease in relation to economic status. *Journal of Nervous and Mental Disease*, 123, 257–261.
- McCullagh, P.(1980). Regression models for ordinal data. *Journal of Royal Statistical Society*, Series B 42, 109-142.
- Morrison, D. G (1969)On the interpretation ofdiscriminant analysis. *Journal of Marketing Research,* 6: 156-163.

- Morselli, P. L., and Elgie, R. (2003). GAMIAN-Europe/BEAM survey I-global analysis of a patient questionnaire circulated to 3450 members of 12 European advocacy groups operating in the field of mood disorders. *J. Bipolar Disord* 5, 265-78.
- Murray CJL, Lopez AD (1996). The Global Burden of Disease and Health Statistics. Boston: Harvard University Press.
- NIMH (2007): Bipolar disorder (National Institute of Mental Health publication).
- Noaghiul S, Hibbeln JR (2003). Cross-national comparisons of seafood consumption and rates of bipolar disorders. *Am J Psychiatry* 160: 2222–2227.
- Okasha A, Karam E.(1998) Mental health services and research in the Arab World. *Acta Psychiatr Scand*; 98: 406-13.
- OPHCC, 1994),(1994).The 1994 population and housing census of Ethiopia .Results for Southern Nations ,Nationalities and peoples religion. Abridged report. Addis Ababa, Vol I partV
- Ösby U, Brandt L, Correia N, Ekbom A, Sparen P(2001) Excess mortality in bipolar and unipolar disorder in Sweden. *Arch Gen Psychiatry* 58:844-850.
- Rao, C. R(1973). Linear statistical inference and its applications. New York: Wiley.
- Rao, C. R.Advanced statistical methods in biometric research. New York:Wiley, 1952.
- Reiger, D.A, Boyd, J.H, Bucke, (1988) one month prevalence of mental disorder in the united States. *Gen psychiatry* 45, 977-86.
- Robb Jc, Cooke RG, Devins GM(1997) quality of life and life style distribution in euthymic bipolar disorder. *J psychiatry Res* 31,509-517.
- Sachs GS, Thase ME(2000) Bipolar disorder therapeutics: maintenance treatment. *Biological Psychiatry* ,48(6): 573-81.
- Sharma S. (1996):Applied Multivariate Techniques Jhon Wiely and Sons,Inc. Canada.
- Slama et al (2004): Bipolar patients with suicidal behavior: toward the identification of a clinical subgroup. *J Clin Psychiatry,* 65, 1035-1039.
- Soares JC, Mann JJ(1997) The anatomy of mood disorders—review of structural neuroimaging studies. *Biological Psychiatry*,41(1): 86-106.

- Sorvaniemi M, Hintikka J.(2005) Recorded psychiatric comorbidity with bipolar disorder-a Finnish hospital discharge register study. *Nord J Psychiatry* , 5: 531-533
- Suppes T et al., (2001):The Stanley Foundation Bipolar Treatment Outcome Network II. Demographics and illness characteristics of the first 261 patients. *J Affect Disord* 67: 45-59.
- Tabachnick, B.G. and L.S. Fidell. (1996). Using Multivariate Statistics. Harper Collins College Publishers: New York.
- Tohen, M.and Good win, F.K. (1995) Eipodomology of bipolar disorder. In:Tsuang, M.T, Tohen. M., Zahner GEP, eds. Textbook in Psychiatric Epidemiology. New York: Wiley-Sons, Inc..
- Waters, B. G. H., Marchenko-Bouer, I., & Smiley, D. (1981). Educational achievement, IQ and affective disorder in the adult offspring of bipolar manic-depressives. *British Journal of Psychiatry*, 139, 457–462.
- Weissman M.M., Bland, R.C., Canino, G.J., (1996) Cross national epidomology of major depression and bipolar disorder. *J. Am. Med. Assoc*. 276, 293-299.
- Weissman M.M., Kidd KK, Prusoff BA.(1982). Variability in rates of affective disorders in relatives of depressed and normal probands. *Arch Gen Psychiatry*,39 (12):1397–1403.
- Weissman MM, Leaf PJ, Tischler GL, (1988): Affective disorders in five United States communities. *Psychol Med*, 18(1): 141–153.
- Weissman, M. M., Bruce, M. L., Leaf, P. J., Florio, L. P. & Holzer III, C. (1991). Affective disorders. In *Psychiatric Disorders in America* (ed. L. N. Robins and D. A. Regier), pp. 53–80.
- WHO (2001). The world health report 2001. Mental Health: New Understanding New hope. World health organization,Geneva.
- World Bank (1993) World Development Report1993: investing in health. New York, Oxford University press.
- World Bank (1998). African Development Indicators 1998/99. Washington: World Bank.

- Zarate CA Jr, Tohen M, Baraibar G, Zarate SB, Baldessarini RJ (1997) Shifts in hospital diagnostic frequencies: bipolar disorder subtypes, 1981–1993. *J Affect Disord* ,43 (1):79–84.[Medline]

Annex I

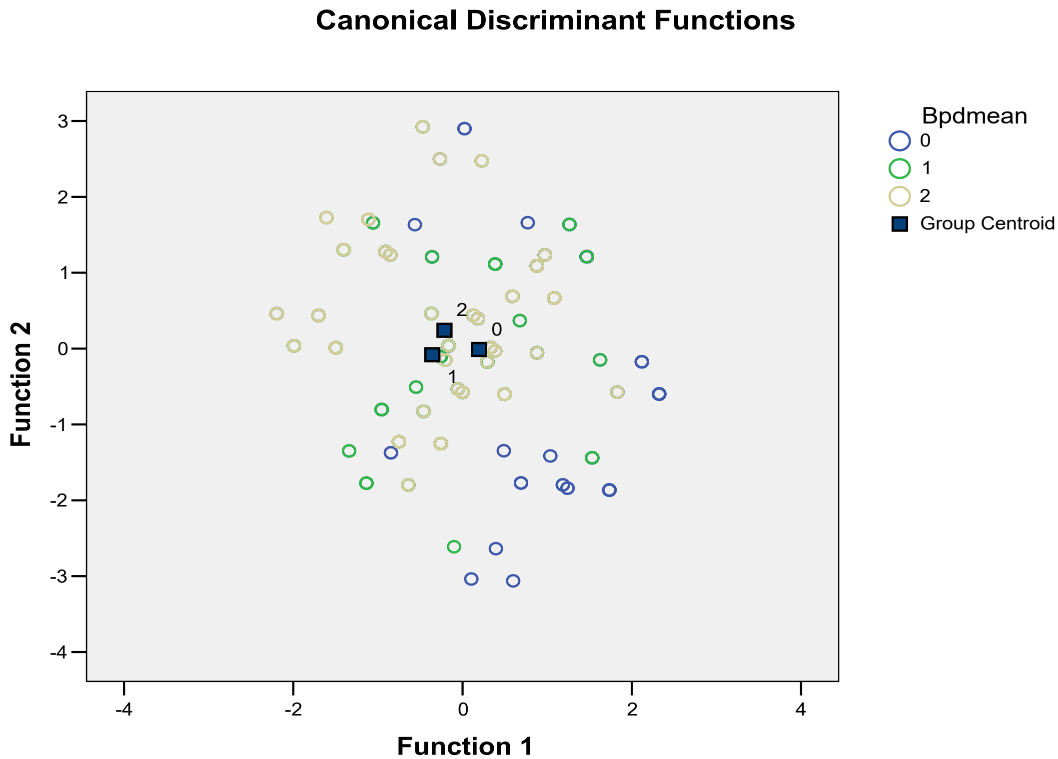
Figure -1 Territorial Map

```
                                  Territorial Map
Canonical Discriminant
Function 2
       -3.0      -2.0      -1.0        .0       1.0       2.0       3.0
```

```
           223333                               31
     1.0     2222333                             31
                 2223333                      31
                    22223333                 31
                        22223333            31
                            2222333        31
                                2223333*  31
      .0                            22*23331*
                                       221
                                        21
                                         21
                                         21
                                         21
    -1.0                                  21
                                         21


       -3.0      -2.0      -1.0        .0       1.0       2.0       3.0
                         Canonical Discriminant Function 1
```

```
Symbols used in territorial map

Symbol   Group  Label
------  -----  --------------------

   1        0   No BPD
   2        1   Moderate BPD
   3        2   Severe BPD
   *           Indicates a group centroid
```

Figure 2 Combined Group plot



Label 0=No BPD
1=Moderate BPD
2=Severe BPD

The percentage distribution of the three groups

Table A: Percentage distribution of the three severe of bipolar disorder

	Frequency	Percent
No BPD	511	61.7
Moderate BPD	218	26.3
Severe BPD	99	12.0
Total	828	100.0