COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF STATISTICS

Bayesian Joint Modeling on the Burden of Bipolar Symptoms and Time to Symptomatic Recovery of Bipolar Disordered Patients: The case of Jimma University Medical Center, Jimma, Ethiopia

---

By: Tefera Fufa

A Thesis Submitted to Jimma University, College of Natural Sciences, Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biostatistics.

September 10, 2020

Jimma, Ethiopia

Bayesian Joint Modeling on the Burden of Bipolar Symptoms and Time to Symptomatic Recovery of Bipolar Disordered Patients: the Case of Jimma University Medical Center, Jimma, Ethiopia

By: Tefera Fufa

Advisor: Geremew Muleta (Assistance Professor, PhD Scholar)
Co-advisor: Demeke Kifle (M.Sc.)

September 10, 2020
Jimma, Ethiopia

Approval Sheet

This is to certify that this thesis with title " Bayesian Joint Modeling on the Burden of Bipolar Symptoms and Time to Symptomatic Recovery of Bipolar Disordered Patients: the Case of Jimma University Medical Center, Jimma, Ethiopia" submitted in the partial fulfillment of the requirement for the degree of Master of Science in Biostatistics to the college of natural science Jimma University, and is record of original research carried out by Tefera Fufa Kulute,and under my supervision no part of the thesis has been submitted for another degree or diploma before. The assistance and the help received during the course of this investigation have been duly acknowledged. Therefore, I recommended that it would be accepted as fulfilling the thesis requirement.

Geremew Muleta (PhD Scholar)	.....	.....
Advisor	Signature	Date
Demeke Kifle(M.Sc.)	.....	.....
Co-advisor	Signature	Date

As the members of the board of examiners of M.Sc. thesis open defense examination of Tefera Fufa Kulute, we certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis has been accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

.....	.....	.....
Name of External Examiner	Signature	Date
.....	.....	.....
Name Of Internal Examiner	Signature	Date
.....	.....	.....
Name Of Chairman	Signature	Date

Statement of the Author

As author of this research study, I declare that the thesis is a result of my genuine work, support of my supervisors and help hands of other individuals. Thus, all those had who participated in the study and sources of materials used for writing this thesis have been duly acknowledged. I have submitted this thesis to Jimma University as a partial fulfillment for the requirements of Degree of Master of Science in Biostatistics. The library directorate of Jimma University can deposit the copy of the thesis in the university library so that students and researchers can refer it. Moreover, I declare that I have not so far submitted this thesis to any other institution anywhere for that award of any academic degree, diploma or certificate and or to get prove of society's problems.

Any brief quotations from this thesis are allowed without requiring special permission if an accurate acknowledgment and citation (after publication) of the source is made. In all other instances, however, permission must be obtained from the author.

Tefera Fufa ....

Name Signature Date

Acknowledgment

First and foremost,all the praises and thanks to be for the God, the Almighty, for his help and blessings throughout my work to complete the research successfully. I would like to express my deep and sincere gratitude to my thesis advisor, Geremew Muleta (Assistance professor,PhD Scholar), for providing me an invaluable guidance throughout work of this research and giving wonderful personality through the time without any reservation time. I would like to express my heartfelt thanks to my co-advisor, Demeke Kifle (M.Sc.), for his valuable comments and suggestion.

I wish to show my appreciation to Jimma University Medical Center for their permission to get the data for the study. My special appreciation and acknowledgment go to Assosa University for offering me the opportunity and financial support during my study at Jimma University.

Finally, I am extremely grateful to my parents for their love, prayers, caring and sacrifices for educating and preparing me for my future. I wish to extend my special thanks to, my father; Fufa Kulute , my mother; Turune Gudeta and to all my brothers, my sisters and all my best friends for their unconditional love and supports.

Contents

List of Tables

List of Figures

Abbreviations(Acronyms)

AFT. . . . .Accelerated Failure Time

BD. . . . .Bipolar Disorder

BD-I . . . .Bipolar Disorder Type One

BD-II . . . .Bipolar Disorder Type Two

CI . . . .Confidence Interval

DIC. . . . . .Deviance Information Criteria

DSM-5. . . . . .Statistical Manual of Mental Disorders, Fifth edition

EDA. . . . .Explanatory Data Analysis

EF. . . . . Exponential family

EM . . . .Expectation Maximization

GLM . . . .Generalized Linear Model

GLMM . . . . . Generalized Linear Mixed Effects Model

HIV. . . . .Human Immune Virus

HR. . . . .Hazard Ratio

JM . . . .Joint Model

JUM . . . .Jimma University Medical Center

LMM. . . . .Linear Mixed Effects Model

MCMC . . . .Markov Chain Monte Carlo

PH. . . . .Proportional Hazard

STEPBD. . . . . Systematic Treatment Enhancement Program for Bipolar Disorder

WHO. . . . .World Health Organization

Abstract

Background:- *Bipolar disorder also known as manic-depressive disorder is a mental health problem that primarily affects mood. Symptoms of bipolar disorder are extreme irritability or agitation, a period of feeling empty,loss of interest in normal activities, sleep problems and etc. Symptomatic recovery is a dimensional measure that refers to improvement in the magnitude of symptoms. Even though it is better if studying the association between burden of symptoms and time to symptomatic recovery of bipolar disorder, there is no study done that show it till now.*

Objectives:-*The objective of the study is to fit the joint model and examine the association between burden of symptoms and time to symptomatic recovery of bipolar disordered individuals.*

Methodology:-*The retrospective data from all the admitted follow up of bipolar disorder patients, who have followed at least three visits from first, September,2018 to first January,2020 in Jimma University Medical Center is used in this study. The study follows two stage,first fitting separate survival and longitudinal models and Joint Model is fitted next.*

Results:-*From the total of 257 bipolar disorders,about 116(45.1%) of them experienced event of recovery.The covariates time in month,age , the interaction between time in month and adolescent first onset of the disease,the interaction between time in month and event of relapse, the interaction between linear time in month and existence of other cofactors and the interaction of substance abuse and chewing khat are significantly affect the log-expected burden of bipolar symptoms.In survival sub-model the covariates;divorced ,event of relapse, mixed type of episodes are significantly affects the time to symptomatic recovery of individuals at 5% significance level*

Conclusions and Recommendations:- *From the Joint model,there is a negative relationship between event of recovery and burden of bipolar symptoms at the baseline time.This indicates that at the beginning time since burden of bipolar symptoms is high the chance of symptomatic recovery is low. Then at the base time the psychiatrists and the concerned body should give special service for the patients.*

***Key words:-B*ipolar disorder,random effects,Shared parameters**

1 Introduction

1.1 Background of the Study

Bipolar disorder is a persistent,serious psychiatric illness with an estimated prevalence of approximately 1% [1]. Its lifetime prevalence is about 3% worldwide [2, 3, 4]. Bipolar disorder, also known as manic-depressive disorder,is a mental illness problem that primarily affects mood. Symptoms of bipolar disorder are extreme irritability or agitation, a period of feeling empty,loss of interest in normal activities, sleep problems,etc. According to Findling RL, bipolar disorder affects over 5 million people in the United States [5].

Bipolar episodes are characterized by a drastic change in behavior and mood, and range from joyful and overexcited(manic episodes) to extremely sad and hopeless(depressive state). These disorders have different types of episodes such as manic episodes, hypomanic episodes, depressive episodes, and mixed episodes. Causes of bipolar disorders include childhood trauma, stressful life events, self-esteem problems, and genetic inheritance [6].

The WHO indicates that BD is the sixth leading cause of disability in the world. BD in youth is increasingly recognized as a significant public health problem often associated with impaired family and peer relationships, poor academic performance, high rates of chronic mood symptoms, psychosis, disruptive behavior disorders, anxiety disorders, substance use disorders, medical problems (e.g., obesity, thyroid problems, diabetes), hospitalizations, and suicide attempts and completions [7].

A 2000 study by the World Health Organization found that prevalence and incidence of bipolar disorder are very similar across the world. Age-standardized prevalence per 100,000 ranged from 421.0 in South Asia to 481.7 in Africa and Europe for men and from 450.3 in Africa and Europe to 491.6 in Oceania for women [8].

In 2017, the Global Burden of Disease Study estimated there were 4.5 million new cases and a total of 45.5 million cases globally [9].In most developing countries, particularly in sub-Saharan Africa, resources for mental health care are very scarce [10] and the delay in seeking treatment for bipolar disorder is long [11]. In sub-Saharan Africa, unipolar

depression was the third leading cause of disease burden, and by the year 2020 it is expected to become the second leading cause of disease burden worldwide [12]. The overall prevalence of mental illness in South Africa was 25% among adults [13].Moreover, mental health problems account for 12.45% of the burden of diseases in Ethiopia and 12% of the Ethiopian people are suffering from some form of mental health problems of which, 2% are severe cases [14].

Annual burden of bipolar disorder to Ethiopian society was estimated to be $ 331 million. Similarly, in the year 2005 the projected lost days of work to the Ethiopian society was estimated to be 112.8 million, 140 assuming a 2.9 percent life time prevalence rate of bipolar disorder in the general population and 93.52 cumulative lost days of work from each patient annually [15].A cross-sectional community based study in Jimma town had taken by Ermias $ Samuel in 2002; and based on their result they stated that mental distress is fairly common in Jimma town and the decentralization of mental health service and its integration with primary health care and use of community health agents in creating awareness among the community members is recommended [16].

Symptomatic recovery is a dimensional measure that refers to improvement in the magnitude of symptoms. Recovery should not be defined merely by symptomatic remission or even syndromal remission; rather, recovery should include symptomatic recovery, syndromal recovery, functional recovery, and a return to an acceptable quality of life for the patient [17]. Symptomatic recovery is the sustained resolution of the symptoms of the disorder. Functional recovery is the ability to return to an adequate level of functioning and includes an assessment of occupational status and living situation [18].

Longitudinal and survival data analysis are among the fastest expanding areas of statistics and biostatistics in the past three decades. Longitudinal data analysis generally refers to statistical techniques for analyzing repeated measurements data from a longitudinal study. Repeated measurements data include multiple observations of an outcome variable that are measured over time on the same study unit, during the course of follow up. The key issues for longitudinal data analysis are how to account for the within-subject correlations and how to handle missing observations. On the other hand, survival analysis deals with survival data or time-to-event data where outcome variable is time to event event.

Time-to-event data are usually incomplete, and thus cannot be handled by standard statistical tools as for complete data. A typical example is right censoring, which occurs when the survival time of interest is only known to be greater than some observed censoring time due to the end of follow up or the occurrence of early withdraw or competing events. Both types of data arise commonly in almost all scientific fields. Indeed, they extend naturally the survival model with time-dependent covariates and offer a flexible framework to explore the link between a longitudinal biomarker and a risk of event [19].

Joint modeling is a very rich and active research area examining the association between longitudinal and survival processes. It also enhances longitudinal modeling by allowing for the inclusion of non-ignorable dropout mechanisms through survival tools and survival modeling with the inclusion of internal time-dependent covariates [20].In longitudinal data analysis, joint models were primarily considered as a means to adjust for no ignorable missing data due to informative or outcome related dropouts which cannot be handled properly by the popularly used methods such as linear mixed and generalized linear mixed effects models [21].

Although, when the two outcome interested are correlated, joint modeling has been empirically demonstrated to lead to improved efficiency and reduced bias [22, 23, 24], there is not joint model done on bipolar disorder and time to recovery that may use to change the traditional view of peoples by justifying the association burden of bipolar disorder and time to recovery.Then the main objective of current study is to fit joint longitudinal and survival model and assess the association between burden of bipolar symptoms and time to symptomatic recovery in Jimma University Medical Center,Jimma, Ethiopia.

1.2 Statement of the Problem

Mental health problem is accepted as a public health problem in developed as well as developing countries [25]. Georgie [26] using samples from diverse populations, have suggested that the burden of psychiatric morbidity existing in Africa is very similar to that prevailing in Western countries. However, mental health problems account for 12.45% of the burden of diseases in Ethiopia and 12% of the Ethiopian people are suffering from some form of mental health problems of which, 2% are severe cases [27]. At this time,

there is no cure for bipolar disorder; however treatment can significantly decrease the associated morbidity and mortality.

When the two outcomes of interest are correlated, joint modeling has been empirically demonstrated to lead to improved efficiency and reduced bias [[22, 23, 24], improved prediction [29], and is applicable to outcome surrogacy [30]. But to the best of knowledge, there is virtually no advance literature using the joint models on bipolar disorder that leads to unknown association between burdens of bipolar and duration of recovery, except the studies about prevalence of bipolar in Ethiopia based on cross-sectional data. Even though some documents were there, they used multiple linear and logistic regression, or separate models survival and longitudinal to identify determinant factors and assess risk factors overtime separately without raising the association between two models. [78].

The lack of both longitudinal and survival data to monitor the pattern of causes and duration of recovery as well as extent load of symptoms make it a new avenue for investigation. The few studies on bipolar disorder do not track the longitudinal and survival history of the patients. Therefore, this study is expanded to fit and assess the joint model and examine the association between burden of bipolar symptoms and time to symptomatic recovery.

In general, this study answer the following major research questions:-

- What are the factors that significantly affect the time to symptomatic recovery of bipolar disorder ?
- What are the factors that significantly affect the burden of bipolar symptoms ?
- How strong is the association between the burden of bipolar symptoms and risk symptomatic recovery through bipolar disorder over time on the study area?
- Do the the bipolar disordered groups show similar relative risk of symptomatic recovery in the study area?

1.3 Objectives of the Study

1.3.1 General Objectives

The main objective of this study is to model the joint burden of bipolar symptoms and time to symptomatic recovery of the bipolar disordered patients in JUMC from September, 01, 2018 to January, 01, 2020.

1.3.2 Specific Objectives

The specific objectives of the study are:-

- To fit the cox-regression model for the time to symptomatic recovery and determine the significant factors in study area.
- To fit the longitudinal model for the burden of bipolar symptoms and determine the significant factors on the study area.
- To examine and assess the association between burden of bipolar symptoms and risk of symptomatic recovery through bipolar disorder on the study area.
- To examine and assess the hazard of the symptomatic recovery among the bipolar disordered groups using survival sub-model on the study area.

1.4 Significance of the Study

Studying the association between burden of bipolar symptoms and time to recovery is one way of overcoming the mental health problem in the community by addressing the effects in the burden of bipolar symptoms on the duration of recovery and significant factors of the burden of bipolar symptoms.

This study may helps physicians and researchers as a benchmark when investigating related studies. It also gives a chance to examine the progression and change trajectory of burden bipolar symptoms and its duration of recovery and awareness of this will provide all the concerning body to take appropriate measure to reducing the burden of bipolar symptoms and by reducing the risk factor or covariates.

1.5 The Organization of this Study

This study builds on five chapters. Chapter one is background of the study which provides a brief description on the bipolar disorder and Joint longitudinal and survival models,statements of the problem and the objectives of the study, which is used as our tool to validate the thesis statement. Chapter two elaborates an overview of previous related studies on bipolar disorder and related joint longitudinal and survival models. Chapter three provides a brief description on the study area,design and methodologies which are used as our tool to analysis of the data. Chapter four contains result of the data analysis using descriptive statistics and the joint model.To fit the joint model,first the separate longitudinal; generalized linear model is fitted and the variables are selected based on the backward elimination method then after, fit the generalized mixed model with different random effects and select the best model. For the separate survival model,fit the cox model with the selected variables.Then after the joint model is fitted using Bayesian approach. Chapter five contains discussion,conclusion and recommendation based on the results.

2 Literature Review

2.1 Description of Bipolar disorder

Bipolar disorder (BD) is a major mood disorder characterized by recurrent episodes of depression and hypomania [31]. According to the Diagnostic and Statistical Manual (DSM-5), the two main sub-types are BD-I (manic episodes, often combined with depression) and BD-II (hypomanic episodes, combined with depression) [32].Bipolar disorder is a chronic illness with substantial psychosocial and occupational morbidity. Several studies have shown that, after a manic episode, the majority of patients with bipolar disorder continue to exhibit significant impairment in role functioning, despite symptomatic recovery [15, 16, 17, 18, 19].

Symptomatic recovery is a dimensional measure that refers to improvement in the magnitude of symptoms. This differentiation permits the examination of psychopathology that persists despite symptomatic improvement to the point that patients no longer meet diagnostic criteria for an episode. Functional recovery refers to the return to previous levels of work and psychosocial function. These distinctions are important because separating these aspects of recovery may help clarify factors that differentially contribute to the recovery process [20].

2.2 Recovery of Bipolar Disorder

The study done on department of psychiatry, in University of Cincinnati, College of Medicine, by Keck Mcelroy SDtrakowski(1998) on studying the 12-month course of illness following hospitalization for a manic or mixed episode of bipolar disorder to identify potential outcome predictors by using multivariate analyses suggests that; during the 12-month follow-up period, there were no significant differences in outcome between patients with manic compared with mixed BD. Although syndromic recovery occurred in 48% of the overall group, symptomatic recovery occurred in only 26% and functional recovery in only 24% . Medication treatment compliance was inversely associated with the presence of comorbid substance use disorders. Symptomatic and functional recovery occurred more rapidly and in a greater percentage of patients from higher social classes.

A minority of patients with bipolar disorder achieved a favorable outcome in the year following hospitalization for a manic or mixed episode. Shorter duration of illness, higher social class, and treatment compliance were associated with higher rates of symptomatic recovery. The survival curve for symptomatic recovery was illustrated as ,of the 106 patients who completed the study, only 28 (26%) experienced symptomatic recovery at some time during the interval between hospital discharge and 12-month follow-up.

Logistic regression analysis revealed that only higher social class (χ 2= 6.2, df = 1, p = 0.01) was associated with symptomatic recovery. Using the Cox regression analysis, we found again that only higher social class (adjusted hazard ratio=1.17, 95% confidence interval=[1.02,1.34; Wald χ 2=5.6, df =1, p=0.02] was associated with less time to symptomatic recovery. The mean age at onset of manic phase of the illness was found to be 22.0 years (22.5 for men and 21.4 for women). The mean age at onset of depressive phase was 23.4 years (24.1 for men and 22.5 for women). There was no significant sex difference in the age of onset of manic or depressive phases. In 22.7% of the cases bipolar I illness started with a depressive episode and in 77.3% of the cases it started with a manic episode. Two or more episodes of the illness were reported by 64.1%.

Study conducted on Bipolar disorder in rural Ethiopia by Alemayehu Negash(2009) showed that relapse and remission assessed at follow-up average 2.5 years (range 1 - 4 years) ,65.9% relapsed at least once 47.8% manic, 44.3% depressive.Of those relapsed, 28.5% had psychotic features, 31% were continuously ill & 5% were mostly in remission during the follow up time.

Perlis, Ostacher(2006) to examined the prospective data from a cohort of patients with bipolar disorder participating in the multicenter Systematic Treatment Enhancement Program for Bipolar Disorder (STEPBD) study for up to 24 months. Accordingly, for those who were symptomatic at study entry but subsequently achieved recovery, time to recurrence of mania, hypomania, mixed state, or a depressive episode was examined with Cox regression. Of 1,469 participants symptomatic at study entry, 858 (58.4%) subsequently achieved recovery. During up to 2 years of follow-up, 416 (48.5%) of these individuals experienced recurrences, with more than twice as many developing depressive episodes (298, 34.7%) as those who developed manic, hypomanic, or mixed episodes (118, 13.8%

). The time until 25% of the individuals experienced a depressive episode was 21.4 weeks and until 25% experienced a manic/hypomanic/mixed episode was 85.0 weeks.

Study on Longitudinal course of Bipolar Disorder and duration of Mood Episodes by Solomon (2010) with the objective of describe the duration of bipolar I mood episodes and factors associated with recovery from these episodes while the probability of recovery over time from multiple successive mood episodes was examined with survival analytic techniques, showed that the median duration of bipolar I mood episodes was 13 weeks.

More than 75% of the subjects recovered from their mood episodes within 1 year of onset. The probability of recovery was significantly less for an episode with severe onset (psychosis or severe psychosocial impairment in week 1 of the episode) (hazard ratio [HR] =0.746; 95% confidence interval [CI], [0.578, 0.963]; P=.02) and for subjects with greater cumulative morbidity (total number of years spent ill with any mood episode was (HR=0.917; 95% CI, [0.886, 0.948]; P¡.001). Compared with the probability of recovery from a major depressive episode, there was a significantly greater probability of recovery from an episode of mania (HR=1.713; 95% CI, [1.373,2.137], P=.001), hypomania (HR=4.502; 95% CI, 3.466,5.849; P=.001), or minor depression (HR = 2.027; 95% CI, [1.622,2.534]; P=0.001) and, conversely, a significantly reduced probability of recovery from a cycling episode (switching from one pole to the other without an intervening period of recovery) (HR=0.438; 95% CI, [0.351,0.548]; P=.001).

Endalamaw Salalew(2019) studied the magnitude and associated factors of Perceived Stigma among adults with mental illness in jimma University Medical center, Ethiopia with objective of assessing the magnitude and associated factors of perceived stigma among adults with mental illness in an Ethiopian.He used facility-based, cross-sectional study design with a consecutive sampling technique employed from September 1 to 30, 2012. He suggested that; Among the total of 384 participants were invited and fully participated in the study making a response rate of 100% , the majority (271, 76%) of the participants were males. The mean age and standard deviation of the participants were 32.75 years ($\pm$ 10.24 years). The largest proportion (203, 52.9%) of the participants were from a rural area, and 180 (46.9%) were single. The majority (245, 63.8%) of the participants were Oromo by ethnicity, followed by Amhara (60, 15.6%).

Study hold on correlates of mental distress in Jimma town, with objective of determining the prevalence of mental distress and related socio-demographic and other risk factors by Ermias and Samuel in 2002 suggests that: - 689(68.5%) females and 317(31.5%) males were found. The mean and median age of the study population were found to be 33.89 and 29 years respectively. Sixty-seven individuals (6.7%) reported family history of mental illness out of which 38.8% showed mental distress compared to 21.0% among those individuals with no family history. Five hundred thirty-eight (53.7%) of the study population were married. The major ethnic groups identified in the study area were Oromo(34.6%), Amhara(26.1%).

2.3 Joint Modeling Approaches

Joint modeling of longitudinal and time-to-event data has witnessed an explosion in the literature of recent years by a many number of researches. One can use joint modeling approach when interested in assessing the association between an endogenous variable measured repeatedly over a time period and the interest of determining the risk of an event occurring. De and Tu proposed the longitudinal sub-model to model the progression of a CD4-lymphocyte count with some level of measurement errors and its association with other risk factors via the survival sub-model. They implemented their approach using the random effects model which enabled them describe the progression in the CD4 cell counts and the effects of the CD4 trajectory on survival of patients [33].

Tsiatis(1995) also proposed a two-stage approach in modeling CD4 counts, which according to the authors was a potential marker for human immune virus (HIV) trials, due to its observed correlation with clinical outcomes[32].The Cox proportional hazards regression model can be used to fit time-dependent covariates, but the CD4 counts are always measured periodically and with substantial errors in their measurements making it impossible for the Cox model to be used [34]. They therefore modeled the CD4 counts via a repeated random effects model while the other variables were considered under the relative risk (Cox proportional hazards) model.

Other studies that explored the joint modeling approach in different fields of study include Andrinopoulous looked at the valve function of cardiac-thoracic surgery which is moni-

tored over a period of time [34]. The joint modeling approach is the most appropriate in helping a physician scan for the trend in valve functions so that they are able to plan their next intervention, [35]. Their approach are implemented via P-splines using Bayesian methods of joint modeling that enable them to specify a time-varying coefficient to link the longitudinal and the survival processes.

Jean-Franc *in 2002*, addressed this problem in situations where the value of the covariate at dropout is unobserved. He suggested joint model which combines a first-order Markov model for the longitudinally measured covariate with a time-dependent Cox model for the dropout process by likelihood estimation of their model and show how estimation can be carried out via the EM algorithm.They state that joint model may have applications in the context of longitudinal data with non-ignorable dropout [36].

In the context of the application of joint modeling techniques to health insurance studies, previous work found by Piulachs, where the study focused on elderly policyholders and the counting process is approximated by a log-transformation of the longitudinal outcome. Given the discrete nature of emergency claims per year, the longitudinal response must account for non-Gaussian data. Previous approaches of this kind have been proposed [36]. For example, Rizopoulos and Ghosh defined a Bayesian JM to relate multiple longitudinal outcomes (discrete or continuous) to a time-to-event outcome [37].

Murawska, Rizopoulos and Lessaffre presented a two-stage JM where the longitudinal information was summarized by either a non-linear mixed-effects model or a generalized linear mixed model (GLMM) in the first stage, while in the second the EmXavier Piulachs, Ramon Alemany, Montserrat Guill´en and Dimitris Rizopoulos empirical Bayes estimates of the subject-specific parameters were included as predictors in the proportional hazards model [38]. Viviani, Alf´o and Rizopoulos implemented an expectation-maximization algorithm to incorporate non-Gaussian data in the longitudinal response, with particular attention to Poisson and binomial mixed models [39]. More recently, Ivanova, Molenberghs and Verbeke formulated a JM to handle different types of responses, i.e., continuous, discrete and ordinal. Parameters were estimated under a likelihood-based approach [40].

When the two outcome of interest are correlated JM is appropriate to see their association. But to the best of knowledge, there are virtually no studies those documented using joint longitudinal and time to event to assess the association between burdens of bipolar symptoms and time to symptom recovery the current days. Even though there are some studies on the prevalence of bipolar disorder in Ethiopia, they were based on cross-sectional data.They used multiple linear regression and logistic regression and only survival model to identify determinants factors that assess prevalence of bipolar episodes changes overtime separately without considering the correlations within the two outcomes and subject specific random effects or separate longitudinal without considering the joint association. But,it is better if studying how the of changes in the burden of bipolar symptoms associated with duration of symptomatic recovery of bipolar disorder. This study use the joint model approach to examine the association between changes in burden of bipolar symptom and time to symptomatic recovery of bipolar disorder in Jimma University Medical Center.

3 Data and Methodology

3.1 Description of the study area

The study is conducted in the Jimma University Medical Center which is which is located in Jimma Zone, Oromia Regional State and south west of Ethiopia. Jimma zone is located at a distance 325 Km from Addis Ababa which is the center of Ethiopian country. The hospital currently employs almost 1,000 people and each year provides tertiary care services for approximately 9,632 inpatients, 5,000 accident and emergency cases, and 80,000 outpatients from a catchment area population of 15 million. The psychiatry inpatient unit has 24 beds, which are mostly used for the management of acutely ill patients(www.hindawi.com/journals/psychiatry/2020/8739546/). Jimma University Medical Center serves as a teaching and referral center for the Jimma area community and, adjacent zones and regions of Southwest Ethiopia.

3.2 Data source and study design

In this study data from retrospective cohort admitted follow up of all bipolar disorder patients, who have followed at least three visits from first, September, 2018 to first, January, 2020 in Jimma University Medical Center is included. Both the longitudinal and survival data are extracted from the patient's card which contains epidemiological, laboratory and clinical information of all bipolar disorder patients after identification of patients who have admitted and follow-up.

3.2.1 Study population

All bipolar disordered patients in the JUMC who are admitted to follow up in time interval of the first September 2018 to first January 2020 are included in the study. Patients those have less than three follow-up were not included in the study.Hence the non-probability sampling is used to select the study population.

3.3 Variables of the study

3.3.1 Dependent variables

The response variables of this study are two ; “Burden of bipolar symptoms” which is obtained by counting the number of bipolar symptoms the patient shows that is repeatedly recorded by medical psychiatrist at every follow up time. At every follow up time bipolar disordered or caregivers are asked the symptoms the patient shown and that symptoms are recorded on patients card. And it is taken as a counting variable and the other response is “Time to symptomatic recovery of bipolar disorder ” which is the length of time in month to get recovery. It is recorded on patients’ card when all symptoms are improved or totally removed respectively.

3.3.2 Independent variables

The covariates used in this study are Age, sex, event of relapse, First onset, Bipolar type, Educational level, Other cofactors, Marital status, Types of episodes, Family history of disease, Substance abuse, Religion, Time in month and Time in months. These covariates are used both in separate models and Joint models and their description is as following table.

Table: 1. Description of the variables included in the study

Variables	Description
Id number:	Patient’s identifier, in total 257 patients are included using inclusion and exclusion criteria.
Sex :	Gender of patients, 0 = female, 1= male
Age :	Age of the patients (in years), 1=19 and below, 2=20-25, 3=26-49, 4= 50 and above
Event of relapse:	Indicator of whether patient face relapse history in previous, 1= yes, 0=No
Frist onset age:	Age of patient when first face the bipolar symptoms, 1=childhood, 2=indolence age, 3=adult age, 4=eldest
Bipolar type:	Type of bipolar the subject faced, 0=Bipolar I, 1=Bipolar II

Time in months: length of time in months between first observation date and the earlier occurrence of event

Time in month: Time interval in months for each follow up

Educational level: education level of the Patients: 1= tenth holder and below, 2=above tenth and diploma, 3=degree and above.

Other cofactors: - status of the patient had other diseases in between treatment, 0=No, 1=Yes.

Marital status: Patients marital status, 1=single, 2=married, 3=widowed, 4=divorced.

Types of episodes: patients type of bipolar episodes, 1=mania, 2=mixed, 3= depression.

Family history of mental illness: status of whether family of the patient had got mental illness before, 0=No, 1=Yes.

Substance Abuse: History of patient whether use alcoholic drinks, 0=No, 1=Yes.

Religion: Religion of the patients, 1=Muslims, 2=Orthodox, 3=Protestant, 4=others.

Chewing khat: History of the patient whether he/she use the khat, 0=No, 1=Yes

Employment: Employment status of the patients, 0=No, 1=Yes.

Ethnicity: Ethnic group of the patient, 1=Oromo, 2=Amhara, 3=others.

Status: Event indicator of the patients whether the subject got first symptomatic recovery or not.

0 = the subject is still not face symptomatic recovery the time in between the study was conducted (censored).

1= the subject had got event of symptomatic recovery in the time between the study was conducted (observed).

burden of bipolar symptoms: is the number of bipolar symptoms that the bipolar disorders shown during each follow up.

3.4 Statistical Methods of Data Analysis

The study consider both descriptive and inferential statistics.

3.4.1 Longitudinal Data Analysis

The collection of correlated data is very common in many fields of quantitative research. Following Verbeke and Molenberghs, the generic term correlated data encompasses several multivariate data structures, such as clustered data, repeated measurements, longitudinal data, and spatially correlated data [41].

Longitudinal data can be broadly defined as the data resulting from the observations of subjects (e.g., human beings, animals, or laboratory samples) that are measured repeatedly over time. Measurements on the same subject are expected to be positively correlated. This implies that standard statistical tools, such as the *t*-test and simple linear regression that assume independent observations, are not optimal for longitudinal data analysis. Then the direct approach to model correlated data were linear mixed model for continuous, normally distributed longitudinal response and generalized linear mixed models for categorical, discrete and non-normally distributed responses.

3.4.2 Exploratory Data Analysis

Exploratory analysis of longitudinal data seeks to discover patterns of systematic variation across groups of patients, as well as aspects of random variation that distinguish individual patients [42]. An exploratory data analysis is a good starting point for any analysis is to look at the data. This should include examining the univariate distribution of each variable to possibly identify anomalous observations, and unexpected aspects , specific procedures that are graphical, depend on the nature and role of the predictor variables. An added benefit of looking at data is that a researcher can further assess whether the final model fit to the data, yields fitted values that resemble the data and capture the main features of the data.

3.4.3 Generalized Linear Mixed-Effects Models

Generalized linear mixed effects models (GLMMs) are an extension of generalized linear models by incorporating random regression coefficients to characterize within-subject correlations in longitudinal or clustered data. GLMMs also extend linear mixed effects models to a rich class of distributions, which can be generally expressed in the form of

exponential families conditional on random coefficients [43, 44]. As such, they are of wide applicability and practical importance [45]. GLMM's enable the accommodation of non-normally distributed responses and specification of a possibly nonlinear link between the mean of the response and the predictors, and they can model over dispersion and correlation by incorporating random effect. GLMM is an alternative framework of LMM which is nowadays routinely used for the analysis of discrete repeated measures data [46].

To define GLMMs,let $y_{i(t)}$ denote the outcome measure at time t for subject i=1,2,..,n. We also denote by $y_i = y_{ij}, j = 1,2,3,..n$. The n_i-dimensional vector of observed longitudinal responses for the i^{th} subject. conditionally a q_b-dimensional of the random-effects vector over b_i, assumed to be independently drawn from MVN (0, D), where D is a variance and covariance matrix. The outcomes y_{ij} are independently distributed with densities from the exponential family of distributions.

$exp\{\frac{y_{ij}(b_i)-c(\psi_{ij})(b_i)}{\alpha(\varphi)-d(y_{ij})(y_{ij},\varphi)}\}$ ..... (1).

Where $\psi_{ij}(b_i)$ and φ denote the natural conical parameter and dispersion parameters in the exponential family, respectively, and $c(\cdot)$, d $(\cdot)$ and a $(\cdot)$ are known functions specifying the member of the exponential family. Different choices for these functions cover the Binomial, Poisson, Gamma, and normal distributions among others. We complete the specification of the model by defining the mean of $y_i(t)$ conditional on the random effects.
$Exp(y_i(t|b_i)=\frac{\partial c\psi_i(t|b_i)}{\partial \psi_i(t|b_i)}=g^{-1}\{x_i^T(t)\beta + z_i^T(t)b_i\}$ ..... (2)

Where $g(.)$ denotes a known monotonic link function,$x_i(t)$ and $z_i(t)$ denote the design vectors for the fixed effects β and $b_i$the random effects, respectively. Therefore, linear mixed effects models are special case of GLMM, in which β has a marginal interpretation because,$Exp(y_i j|\beta) = Exp(y_i j|b_i,\beta)$ =Exp($X_i j^T\beta + Z_i j^T b_i) = X_i j^T$ given that $Exp(b_i) = 0$. This indicates that the marginal mean of Y is a linear model with respect to. However, this relationship is not generally true when the link function g(.) is nonlinear. For other distributions in GLMMs,β is generally interpreted as the impact of covariates on the mean response of a specific subject conditional on the random effects.

3.4.4 Estimation of Generalized Linear Mixed Effects Model

In the estimation and inference GLMMs to get the marginal distribution of y, we integrate the joint density of (y, b_i) on b_i. The likelihood functions of
β,φ and D is evaluated by integrating the conditional probability distribution over b_i. Where $b_i \sim$ MVN(0,$\sigma^2 b$) and specifically we have;
$L(\beta, \phi, D)$=$\prod_{i=1}^{n} \int f(y_i|b_i) f(b_i) db_i$ ..... (3)

Maximum likelihood estimates can be obtained by maximizing the above equation (3). Because the integration is intractable, several numerical procedures have been proposed to approximate the integral. Broadly, inference for GLMMs can be conducted either through approximations of the likelihood or with Bayesian methods. Approximations of the likelihood either focus on approximating the integrand (the function being integrated) like penalized quasi-likelihood or try to approximate the integral itself as in Laplace approximation or Gaussian-Hermite quadrature. In practice, there are three likelihood approximation methods used: - Laplace approximation, adaptive Gaussian-Hermite Quadrature, and Penalized Quasi-Likelihood. These three likelihood-based methods are based on of Laplace's method for approximation, but alternatives such as the hierarchical likelihood introduced by Lee and Nelder [84] have been proposed.

Then in this study both Laplace approximation and Penalized Quasi-Likelihood Penalized quasi-likelihood approximation are used to estimate the GLMM . The glmer function in the R package {lme4} uses the Laplace approximation to do GLMM regression as the default method, and if setting the argument nAGQ to some positive integer, then it will evaluating the adaptive Gauss-Hermite approximation to the log-likelihood and nAGQ is the number of points per axis for the approximation. PQL is the fastest and most flexible approach for estimating GLMMs. The number of random effects and their structure are not as restricted as in AGHQ and should have asymptotic properties as the sample size increase Breslow and Clayton [85]. Though the fastest, it is the least accurate of the common likelihood-based methods, especially with small datasets [86]. The glmmPQL() function works by repeated calls to lme in the R package {lme4} uses the Laplace approximation to do GLMM regression as the default method so package nlme should be loaded at first use if necessary.

3.5 Survival Data Analysis

Survival analysis or time-to-event data analysis refers to statistical methods for time-to-event data. An event time, or survival time, is defined as the time from an initial event such as diagnosis of a disease to the occurrence of an event of interest such as death. The most important characteristic that distinguishes the analysis of survival times from other areas in statistics is censoring, which refers to a situation where the event time of interest is only partially known. Censoring occurred when we have some information about individual survival time, but we do not know the survival time exactly there are many types of censoring including right-censoring, left-censoring, interval censoring, and double censoring [50].Commonly used quantities to characterize an event time T include the survival function and hazard function defined by:-

Survival function:-;$S(t) = P(T > t)$ for any t, it is the probability that the event will occur after time t (survival beyond time t).

Hazard function:-

$$\lim_{\Delta t \to \infty} \frac{P(t \leq T + \Delta t / T \geq t)}{\Delta t}$$

;is the instantaneous failure rate at time t or the force of mortality Nelson & Plosser [51] indicating how likely a subject who has not experienced the event prior to time t will experience the event in the next instant. The survival also can be expressed in terms of the risk function as:- $S(t) = exp(-\lambda(t)) = exp[\int_0^t \lambda(t)dt]$;

Where $\lambda(.)$ is known as the cumulative risk (or cumulative hazard) function that describes the accumulated risk up until time t. Function H(t) also can be interpreted as the expected number of events to be observed by time t. When we are interested in estimating these two functions or any other characteristic of the event time distribution, from a random sample at hand, censoring must be taken into account. In particular, we let T_i denote the observed event time for subject i, defined as the minimum of the true event time and the censoring time C_i . We also introduce the event indicator δ_i = I ($T^* \leq C_i$) that takes the value 1 if the observed event time corresponds to a true event time and 0 otherwise, where I (.) denote the indicator function. In general, in survival analysis, we are interested in estimating characteristics of the distribution of using only the available information, i.e.,

using $\{T_i, \delta_i\}$. The censoring used in this paper is right censoring. Survival time is said to be right censored when it is recorded from its beginning to a defined time before its end time.

In reality right censoring can occur due to the following reasons:

- Death from unrelated causes
- Loss of follow-up
- Termination of study

  Due to censoring observation, various statistical methods for failure data are developed. These are:-

  1. Non-parametric Methods: (Kaplan-Meier method, log-rank test):- These methods are said to be non-parametric methods since they require no assumptions about the distribution of survival time.
  2. Semi-parametric method (Cox Proportional Hazard (PH) model, Accelerated failure model):- This model, also known as the Cox regression model, makes no assumptions about the form of $\lambda_0(t)$ (non-parametric part of model) but assumes parametric form for the effect of the Predictors on the hazard (parametric part of model). The model is therefore referred to as a semi-parametric model.
  3. Parametric model (PH model and AFT model). In this section, we will introduce parametric model, in which specific probability distribution is assumed for the survival times Cox PH model is the most common approach for modeling survival data. Parametric AFT model provides an alternative to PH model for statistical modeling of survival data Wei-bul when PH assumption fails [52].

3.5.1 Semi-Parametric (Cox Regression) Model

The non-parametric method does not control for covariates and it requires categorical predictors when we have several prognostic variables, we must use multivariate

approaches. But we cannot use multiple linear regression or logistic regression because they cannot deal with censored observations. We need another method to model survival data with the presence of censoring. One of the very popular models in survival data is the Cox proportional hazards model, which is proposed by Cox [49]. Due to the popularity of the Cox model (Cox, 1972) in modern survival analysis, proportional hazards models (also known as relative risk or relative hazard models) have prevailed. This model assumes that covariates have a multiplicative effect on the hazard for an event, and they are formulated as:-

$$(\lambda\,(t|x) = \lambda_0\,(t)\exp\,(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p) = \lambda_0\,(t)\,exp\left(\beta_x^{'}\right)\ldots\ldots \quad (4)$$

$$\text{Or } (\log\,(\lambda\,(t|x)) = (\,\log\,(\lambda_0\,(t)) + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p)$$

Where $(\lambda_0\,(t)$ is called the baseline hazard function, which is the hazard function for an individual for whom all the variables included in the model are zero. Where x = $(x_1,\ x_2,\ldots,x_p)^{'}$ is the values of the vector of explanatory variables for a particular individual, and $\beta^{'} = (\beta_1,\ \beta_2,\ldots,\beta_p)^{'}$ is a vector of regression coefficients.

The corresponding survival functions are related as follows: -
$S\,(t|x) = S_0\,(t)^{exp\left(\Sigma_i^p \beta_i x_i\right)}$. This model, also known as the Cox regression model, makes no assumptions about the form of $S_0\,(t)$ (non-parametric part of model) but assumes parametric form for the effect of the predictors on the hazard (parametric part of model). The model is therefore referred to as a semi-parametric model. The beauty of the Cox approach is that this vagueness creates no problems for estimation. Even though the baseline hazard is not specified, we can still get a good estimate for regression coefficients, hazard ratio, and adjusted hazard curves. The measure of effect is called hazard ratio. The hazard ratio of two individuals with different covariates **x** and x^* is

$\hat{HR} = \frac{\lambda_0(t)exp(\hat{\beta^{'}}x)}{\lambda_0(t)exp(\hat{\beta}'x^*)}) = \{\sum\hat{\beta}(x-x^*)\}$
This hazard ratio is time-independent, which is why this is called the proportional hazards model.

3.5.2 Proportional Hazard Assumption

The main assumption of the Cox proportional hazards model is proportional hazards. Proportional hazards means that the hazard function of one individual is proportional to the hazard function of the other individual, i.e., the hazard ratio is constant over time. There are several methods for verifying that a model satisfies the assumption of proportionality.The proportional hazards (PH) assumption can be checked using statistical tests and graphical diagnostics based on the scaled Schoenfeld residuals. In principle, the Schoenfeld residuals are independent of time. A plot that shows a non-random pattern against time is evidence of violation of the PH assumption. If this assumption is not satisfied by our data, we can use AFT model instead of PH model.But in this study cox-regression model is used.

3.5.3 Estimation of Cox Regression Model

In Cox proportional hazards model we can estimate the vector of parameters without having any assumptions about the baseline hazard $\lambda_0(t)$. The full likelihood for right censored data can be constructed as:- $L(\beta) = \prod_{i=1}^{n} \lambda(t_i, X, \beta)^{\delta_i} S(T_i, X, \beta)$ ... ... (5)

$L(\beta) = \prod_{i=1}^{n} \lambda_0(t_i) exp(\beta^T X_i)^{\delta_i} S_0(t_i)^{exp\beta^T x_i)}$ ....(6)

Where; δ_i = an indicator of censoring for the i^{th} individual given by 0 for censored and 1 for event, X = a vector of covariates for individual i ($x_1, x_2, \ldots, x_p$).

$\lambda(t_i, X, \beta) = \lambda_0(t_i)$ exp ($\beta' x_i$) is the hazard function for individual i.

$S(t_i, X, \beta) = S_0(t_i)^{exp\left(\beta' x_i\right)}$ is the survival function for individual i. It follows that under the relative risk (cox) model the distributional assumptions for T_i^* are hidden in the specification of the baseline hazard function. However, Cox (1972) showed that estimation of the primary parameters of interest, namely β , can be alternatively based on the partial log-likelihood function;

$pl(\beta) = \sum_{i=1}^{n} \delta_i [\beta^T X_i - log\{\sum_{T_j > T_i} exp(\beta^T X_i)\}]$ ..... (7)

Which that does not require specification of $\lambda_0(.)$, that is, without having to specify the

distribution of T_i^* . Even though this is not equivalent to a full log-likelihood, it can be treated as such. In particular, the maximum partial likelihood estimators are found by solving the partial log-likelihood score equations:-

$$\frac{\partial pl(\beta)}{\partial \beta^T} = \sum_{i=1}^{n} \delta_i \{X_i - \frac{\sum_{T_{ji}} X_i exp(\beta^T X_i)}{\sum_{T_{ji}} exp(\beta^T X_i)}\} = 0 \ldots\ldots\ldots\ldots .(8)$$

The function used to fit Cox models is coxph (), which has two main arguments; the formula argument that specifies the relationship between the observed failure times and covariates, and the data argument that specifies the data frame that contains these variables. In the left-hand side of the formula argument, function Surv() is used to specify the available information for the failure times, that is the observed failure times and the type of censoring (i.e., right, left, interval, and counting).The primary R package for the analysis of event time data is the **survival** package [56]. To load this package we use the command library ("survival"). A comprehensive list of other packages in CRAN related to survival analysis is available from the CRAN Task View; Survival Analysis [57].

3.6 The Joint Modeling Structure

An important area that fosters development of joint models is survival analysis with time-dependent covariates. Joint analysis is an elegant approach to model the association between time-dependent covariates and the event of interest when the covariate trajectory is not completely observed and subject to measurement error and/or biological variation.

When primary interest is in the association between such endogenous time-dependent covariates and survival, an alternative modeling framework has been introduced in the literature, known as the joint modeling framework for longitudinal and time-to-event data [58].The design followed by package JMbayes requires to first separately fitting a generalized linear mixed model for the longitudinal part and a Cox-PH model for the survival part.

3.6.1 The Longitudinal Sub Model

Let us assume a panel data context with repeated measurements over time, where $y_{ij}(t)$ denote the observed responses for the i^{th} subject, recorded in specific time interval points given the vector b_i of random effects for the i^{th} subjects, We assume that each subject is associated with a vector of random coefficients b_i. Then in this study we that assume that the observed measurements on this individual derived from a counting process generated by an exponential family EF distribution, $y_i(t)|b_i \sim EF(\Psi_i(t), \varphi)$, with probability mass function:-

$$P_y\{y_i(t)|b_i; \psi_i(t), \varphi\} = exp(\varphi^{-1}[y_i(t)\psi_i(t) - b\psi_i(t)] + c(y_i(t), \varphi)) \ldots\ldots\ldots (9)$$

Here, b(·) and c (·) are known functions, ψ_i and φ are termed the canonical and scale parameters, respectively. Based on the theory of exponential families, the conditional mean and variance of y_{ij} are $E(y_i(t)|b_i) = \mu_i(t) = b'(\psi_i(t))$ and $V(y_i(t)|b_i) = \sigma^2(t) = \varphi b"\psi_i(t)$ [60].In this study, the subject-specific count responses is observed within a specified time interval.Modeling the count data is more relevant than working with the raw counts, thus considering the expected longitudinal outcome μ_i in terms of counts per time unit. Difference from linear mixed effects models, GLMMs model the mean of y through a one-to-one continuous differentiable transformation $\eta_i j = g(\mu_i j)$ and assume that the transformed mean is characterized by a linear model.The most common choice for modeling panel count data is a logarithmic link, which ensures positive outcomes and provides a straightforward interpretation of the estimated regression parameters:-

$$\begin{cases} log\{\mu_i(t)\} = \eta_i(t) = X_i(t)\beta + Z_i(t)b_i \\ E\{y_i(t)|b_i\} = \mu_i(t) = exp\{\eta_i(t)\}, b_i \sim MVN(0, D_{q+1}) \end{cases} \ldots\ldots\ldots (10).$$

$Theterms\mathrm{x}_i(t)^T$ and $z_i(t)^T$ denote the row vectors of the fixed and random design matrices, respectively, while $\beta = (\beta_1, \beta_2, \ldots\ldots, \beta_p)^T$ and $b_i = (b_{i0}, b_{i1}, \ldots\ldots, b_{iq})^T$ are the corresponding fixed-effects and random-effects vectors. The random effects allow for the expression of individual deviations from the overall trend, and in most cases they can be assumed to follow a multivariate normal distribution with zero mean and unspecified variance-covariance matrixes. The basic option for modeling panel counts in equation (11) were considered a Poison mixed model, defined as:-

$$\begin{cases} y_i(t)|b_i \sim Poisson\{\mu_i(t)\}, \mu_i(t) \geq 0 \\ \mu_i(t) = exp\{\eta_i(t)\} = exp\{X_i^T(t) + Z_i^T(t)b_i\} \\ p_y\{y_i(t)|b_i;\mu_i(t)\} = \frac{exp\{-\mu_i(t)\}\mu_i(t)^{y_i(t)}}{y_i(t)!} \end{cases} \ldots \ldots \ldots \ldots (12)$$

The Poison mixed model allows for robust parameter estimates, even if the underlying distribution is not true, provided that the expectation is correctly specified [61].However, the observed response usually has a variance greater than the mean, so the longitudinal outcome will be affected by over dispersion.

Although there are several alternative models for dealing with the over dispersion related to correlated counts, the Negative Binomial mixed model appears in the literature as being the most natural choice [62].The Negative Binomial distribution for longitudinal data can be easily derived from the Poison distribution by placing a multiplicative gamma random noise ε_i in the conditional mean response. Such, latent variable is defined in terms of shape and rate parameters by $\varepsilon_i(t) \sim \gamma(k,k), k > 0$ with $E(\varepsilon_i(t)) = 1$ and $Var(\varepsilon_i(t)) = \frac{1}{k}$, so that the longitudinal counts are modeled by $y_i(t) \sim Po(\varepsilon_i(t)\mu_i(t))$. The Poisson-gamma mixture has a closed-form solution, leading to a Negative Binomial mixed model with dispersion parameter κ . The marginal mean response can related to the fixed and random effects using logarithmic link:-

$$\begin{cases} y_i(t)|b_i \sim NB\{\mu_i(t),k\}, \mu_i(t) \geq 0, k \geq 0 \\ \mu_i(t) = exp\{\eta_i(t)\} = exp\{X_i^T(t) + Z_i^T(t)b_i\} \\ p_y\{_i(t)|b_i;\mu_i(t),k\} = \frac{\gamma\{k+y_i(t)\}\mu_i(t)^{y_i(t)}K^K}{\gamma(k)y_i(t)!\{\mu_i(t)+k\}^{k+y_i(t)}} \\ E\{y_i(t)|b_i\} = \mu_i(t); V\{y_i(t)|b_i\} = \mu_i(t) + \frac{\mu_i(t)^2}{k} \end{cases} \ldots \ldots \ldots (13)$$

Where denotes the gamma function. The NB distribution has the general canonical form of the exponential family equations for any fixed k. Note that the NB distribution can actually be understood as an extension of the Poison distribution when over dispersion is accounted for by parameter κ , since it can be proven that NB converges to Poison as $k \longrightarrow \infty$ [63].

3.7 The Survival Sub Model

This measures the association between the longitudinal marker level and the risk for an event, while accounting for the special features of the former. To achieve this we introduce the term $m_i(t)$ that denotes the true and unobserved value of the longitudinal outcome at time t. Note that $m_i(t)$ is different from $m_i(t)$, with the latter being the contaminated with measurement error value of the longitudinal outcome at time t. To quantify the strength of the association between $m_i(t)$ and the risk for an event, a straightforward approach is to postulate a relative risk model where; $m_i(t) = m_i(s), 0 < S < t$.This indicates the history of the true unobserved longitudinal process up to time point t.

$$\begin{cases} \lambda_i(t|m_i x_i) = \lim_{dt\to 0} \frac{\{pr(t\leq T^* < t+dt | T^* \geq t, m_i(t), x_i)\}}{dt} \\ \lambda_0(t) exp\{\gamma^T X_i + \alpha m_i(t)\}, t > 0 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots (14).$$

where;$\lambda_0(t)$ denotes the baseline risk function,X_i is a vector of baseline covariates (such as a treatment indicator history of diseases, etc.),γ is a corresponding vector of regression coefficient similarly and α quantifies the effect of the underlying longitudinal outcome to the risk for an event. $exp(\alpha)$ denotes the ratio of hazards for one unit change in $x_i j$ at any time t,$exp(\gamma)$ denotes the relative increase in the risk for an event at time t that results from one unit increase in $m_i(t)$ at the same time point.

In particular, using the known relation between the survival function and the cumulative hazard function, we obtain that:-

$$\begin{cases} S_i(t|M_i(t), X_i) = pr(T_i^* > t | M_i(t), X_i) \\ \qquad\qquad = exp(\int_0^t \lambda_0(s) exp\{\gamma^T X_i + \alpha m_i(s)\} ds \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots (15)$$

This implies that the corresponding survival function depends on the whole covariate history $m(t_i)$. However, within thejoint modeling framework it turns out that following such a route may lead to an underestimation of the standard errors of the parameter estimate [59]. To avoid such problems we were explicitly define λ_0 (.). A standard option is to use a risk function corresponding to a known parametric distribution. In the survival analysis context, typically used distributions include the Weibull, the log-normal and the Gamma.

Alternatively, and even more preferably, we can option for a parametric but flexible specification of the baseline risk function. Several approaches had been proposed in the literature to flexibly model the baseline risk function. For instance, Whitmore and Killer (1986) used step-functions and linear splines to obtain a non-parametric estimate of the risk function, Rosenberg (1995) utilized a B-splines approximation, and Herndon and Harrell (1996) [80] used restricted cubic splines. Two simple options that often work quite satisfactorily in practice are the piecewise-constant and regression splines approaches. Under the piecewise-constant model, the baseline risk function takes the form:-

$\lambda_0(t) = \sum_{q=1}^{Q} \xi I(V_{q-1} < t \leq V_q)$

where 0 = $v_0 < v_1 < \cdots < v_Q$ denotes a split of the time scale, with V_Q being larger than the largest observed time, and ζ_q denotes the value of the hazard in the interval $|V_{q-1}, V_q|$ .As the number of knots increases the specification of the baseline hazard becomes more flexible. In the limiting case where each interval $|V_{q-1}, V_q|$ contains only a single true event time (assuming no ties), this model is equivalent to leaving $\lambda_0(.)$ completely unspecified and estimating it using nonparametric maximum likelihood. For the regression splines model the log baseline risk function $\log\lambda_0(t)$ is expanded into B-spline basis functions for cubic splines as follows:-

$log\lambda_0(t) = k_0 + \sum_{d=1}^{n} K_d B_d(t,q)$

where $k^T = (K_0, K_1, ...k_m)$ are the spline coefficients, q denotes the degree of the B-splines basis functions B(.), and $m = \ddot{m} + q - 1$, with $\ddot{m}$ denoting the number of interior knots. Similarly to the piece m̈ wise-constant model, increasing the number of knots increases the flexibility in approximating $\lambda_0(.)$. However in both approaches, we should keep a balance between bias and variance and avoid over fitting. After the number of knots has been decided, their location is typically based on percentiles of either the observed event times T_i^*, C_i or only the true event times $T_i : T_i^* \leq C_i, i = 1, 2, \ldots, n$, such that to allow for more flexibility in the region of greatest density.

3.7.1 Shared Parameters in the Joint Models

In selection of the models,the event time is associated with the longitudinal data through y_i or the random effects b_i. When b_i is used to link together y_i and T_i, the approach is referred to as shared parameter models, although b_i is a covariate, not a parameter, in the model for T_i. A more general extension of shared parameter models is proposed by Henderson [79] is used in this study. It is a latent zero-mean bivariate Gaussian process $U_i(t) = (U_{1i(t)}, U_{2i(t)})$ is used to characterize the association between Y and T. In particular, the model is given by;

•g(E(μ_y)) $= X_i^{(1)T}\beta_1 + U_{1i}(t_{ij})$ and

•$\lambda_i(t|U_2) = \lambda_0(t)exp(X_i^{(2)T}\beta_2 + U_{2i}(t_{ij}))$

Various functional forms have been proposed for the latent processes $U_{1i(t)}$ and $U_{2i(t)}$). for example $U_{1i(t)}$ can be b_{0i} or $b_{0i} + b_{1i}$ and etc. where b_{0i} and b_{1i} are random intercept and slope, respectively, and $U_{2i(t)}$) can be $\gamma U_{1i(t)}, \gamma_0 b_{0i} + \gamma_1 b_{1i} + \gamma_2 U_{1i(t)}$.

3.8 Bayesian Estimation of Joint Models

Before begin estimation the missing data is imputed using Multivariate Imputation by Chained Equations or mice() function [89] . Function jointModelBayes() in the JMbayes fits shared parameter joint models for longitudinal and survival outcomes under a Bayesian approach. For the longitudinal responses a linear mixed effects model represented by the lmeObject is assumed.

Unless the user specifies is own probability density function using argument densLong. A function with arguments y,eta.y,scale,log, and data that calculates the density of the longitudinal outcome e.y denotes the longitudinal responses,eta.y the linear predictor that includes the fixed and random effects,scale a possible scale parameter (e.g., the measurement error standard deviation),log a logical argument that controls whether the density should be calculated in the log scale, and data a data frame which may be used to extract variables used in the definition of the density function (e.g., a censoring indicator for left censored longitudinal data).

lmeObject is an object of class 'lme' fitted by function lme() from package **nlme** or by

function glmmPQL() from package MASS. SurvObject is an object of class 'coxph' fitted by function coxph() from package survival. The survMod argument of specifies the type of survival submodel to be fitted; available options are a relative risk model with a Weibull baseline hazard (default) and a relative risk model with a B-spline approximation of the log baseline risk function. Bayesian estimation approach has received a lot of attention for analysing failure time data.

It makes use of ones prior knowledge about the parameters and takes into consideration the data available. In a situation where the researcher has a previous knowledge or can obtain an information from experts that is closely related to the current study, then a suitable prior to use is the informative prior if not a non-informative prior will be an alternative to use by assuming lack of previous knowledge.

And in this study all the standard prior distribution for all parameters which are aided by Rizopoulos [81] for the JMbayes package is considered. Forexample,the prior mean vector of the normal prior for the fixed effects of the mixed effects model,the prior mean vector of the normal prior for the regression coefficients of the survival model,, the prior precision matrix of the normal prior for the slope association parameter in the survival model,the prior shape parameter of the Gamma prior for the precision parameter of the penalty term when baseHaz = "P-splines",Wishart prior for the precision matrix of the random effects,the degrees of freedom of the Wishart prior for the precision matrix of the random effects.

In many regards the design of package JMbayes is similar to the one of package JM for fitting joint models under maximum likelihood. In particular, JMbayes has a basic model-fitting function called jointModelBayes(), which accepts as main arguments a linear mixed effects object fit as returned by function lme() of package nlme (Pinheiro, Bates, DebRoy, Sarkar, and R Core Team 2016) or by function glmmPQL() from package MASS [87], and a survival object fit as returned by function coxph() of package survival [88].The baseline hazard is by default approximated using penalized B-splines; regression splines can be instead invoked by appropriate setting argument baseHaz.

Compared to likelihood approaches, Bayesian methods in principle are relatively straight-

forward to implement. that the expectation is computed numerically through Monte Carlo simulations by repeatedly drawing random (θ, b_i) samples from its posterior distribution via Gibbs or Metropolis Hastings samplers.In addition, in Bayesian joint models, the variance-covariance matrix of model parameters can be estimated as a by-product of the sampling procedure, so there is no extra burden to compute standard errors. In general, Bayesian methods are computationally more efficient, when has a high dimension. However, there are some issues one should be aware of when fitting a Bayesian joint model [66].The posterior distribution may be improper when improper priors are used, especially in the no ignorable missing data setting [67].

The model could be weakly identifiable when a no ignorable missing data mechanism is assumed and thus the inference is quite sensitive to the choice of hyper parameters. Careful considerations are needed when specifying priors to avoid dominating the likelihood [66]. The computation is usually intensive and MCMC convergence may not be easily achieved. A detailed discussion of these issues is provided in Ibrahim and Molenberghs [68]. Let $\theta = (\theta_y, \theta_t, \theta_b)^T$ be the JM full parameter vector that collects the longitudinal parameters, the survival parameters, and the parameters for the random effects covariance matrix, respectively. In addition, let denote the information from our original dataset with n policyholders. Taking advantage of the conditional independence assumption, the overall joint likelihood conditioned on the random effects b_i can be properly formulated to tackle right censoring as:-

$$P(D_n|b_i,\theta) = \prod_{i=1}^{n_i}\prod_{j=1}^{n_i} P_y\{y_i(t_ij)/b_i,\theta\}\frac{p_t(T_i,\delta_i/b_i,\theta)}{pr(T_i>\iota_i/b_i,\theta)} \ldots \ldots \ldots (16).$$

Where $P_y(.)$ the conditional probability mass function to handle longitudinal rate is counts, and $P_t(.)$ is the conditional probability density function for the event times. The mean estimates of parameters and random effects are then derived by Markov chain Monte Carlo (MCMC) algorithms, which enable inferences to be made by efficiently drawing a sample from the posterior distribution of (θ, b_i) conditioned on the observed data:-

$$\pi(\theta,b_i|D_n) \propto P(D_n|b_i,\theta)\pi(\theta,b_i) = P(D_n|b_i,\theta)p_b(b|\theta)\pi(\theta) \ldots \ldots \ldots \ldots (17)$$

Where,$\mathrm{p}_b(.)$ is the conditional probability density function of the random effects, and $\pi(\theta)$ is the prior distribution of θ.

3.9 Asymptotic Inference and Model Selection for Separate and Joint Models

Having fitted the longitudinal sub-model, survival sub-model under maximum likelihood and partial maximum likelihood frame work and joint model under a Bayesian framework, the standard asymptotic likelihood inference tests are directly available. In general, if we are interested in testing the null hypothesis, we could use:- A Likelihood Ratio Test: with the test statistic defined as:-

$LRT = -2\{l(\hat{\theta_0}) - l(\theta)\}$ Where, $\hat{\theta_0}$ and θ denote the maximum likelihood estimates under the null and alternative hypothesis, respectively;

A Score Test, with the test statistic defined as:-

$U = S^T(\hat{\theta_0})\{I(\theta)\}^{-1}S(\hat{\theta_0})$ and,$I(.)$ $S(.)$ and $I(.)$ denotes the score function and the observed information matrix of the model under the alternative hypothesis; or A Wald Test, with the test statistic defined as:-

$W = (\hat{\theta} - \hat{\theta_0})^T I(\hat{\theta})(\hat{\theta} - \theta_0)$

Under the null hypothesis, the asymptotic distribution of each of these tests is a chisquared distribution on p degrees of freedom, with p denoting the number of parameters being tested. In this case, the likelihood ratio test is generally considered the most reliable and the Wald test the least reliable. The score and Wald test require fitting the model only under the null and alternative hypotheses, respectively, whereas the likelihood ratio test requires fitting the joint model under both hypothesis, and thus it is a bit more computationally expensive. If there are missing data in the variable we are interested to test for, then the score test will be more efficient since it requires fitting the model only under the null and therefore, avoids a case-wise deletion of missing values (i.e., excluding subjects who have a missing value in the variable of interest.

The three standard tests we have seen so far are only appropriate for the comparison of two nested models, in the sense that the model under the null hypothesis is a special case of the model under the alternative. When interest lies in comparing non-nested models, information criteria are typically used. The main idea behind these criteria is to compare two models based on their maximized log-likelihood value, but to penalize for the use of too many parameters. The two most commonly used information criteria are the

Akaike's Information Criterion (AIC) [69]and the Bayesian Information Criterion (BIC) [70]. These are included in the output of the anova() and summary() functions, and are defined as:-

$$\begin{cases} AIC = -2l(\hat{\theta} + npara) \\ BIC = -2l(\hat{\theta}) + nparalog(n) \end{cases} \quad \ldots\ldots\ldots (18)$$

Where npara denotes the number of parameters in the model. Under these definitions smaller is better. That is, if we are using either AIC or BIC to compare two models for the same data, we prefer the model with the lowest value for the corresponding criterion. To compare both the different longitudinal models and joint models, we focused on the analysis of the Bayesian deviance term, which in generic form can be expressed as:-

$D(\theta, b_i) = -2\sum_{i=1}^{n} log\{P(D_n|b_i, \theta)\} \ldots\ldots\ldots\ldots (19)$

In particular, we assessed the goodness-of fit of a specific model by using the deviance information criterion (DIC) suggested by Spiegel halter [71]. This criterion evaluates the fit of a model by balancing model adequacy with model complexity:-

$DIC(\theta, b_i) = D(\bar{\theta}, \bar{b_i}) + 2P_D \ldots\ldots\ldots\ldots (20)$

Thus reinforcing the idea that this criterion takes into account both the adequacy of the model, assessed through the posterior mean estimate of the deviance, and the number of parameters required, assessed through the penalty term. The score provided by DIC serves in general as the basis for ranking the fitted models, where lower scores correspond to a better model fit. To conclude this section, it is important to point out that the DIC score obtained for a specific model is not a fixed value, but it can be subject to a certain amount of random variability due to its dependency on the MCMC output of the model. Consequently, it will become a key point to get a DIC value derived from a relatively large number of iterations in the MCMC process before reaching convergence in each of the JM parameters.

3.10 Variable Selection Criteria

Variable selection has long been viewed as a necessary safeguard for model validity. Variable selection for joint models typically includes the selection of both fixed and random

effects. The two approaches for variable selection was used in this study to exhaustively compare all possible models based on a predefined criterion were typically an information-based criterion, such as the Akaike information criteria (AIC) or Bayesian information criterion (BIC) [68], [69].This approach has been used widely in the last several decades and a number of statistical tests, such as the likelihood ratio test, Wald-type test, or score test have been derived for variable selection. The second approach is the stepwise variable selection method. Although it is computationally more efficient than the first approach, it does not search the entire model space, thus leaving open the possibility that the true model could be missed. Most notably in this study backward variable selection method was used for fixed effects and random model and automatic variable selection using stepwise back ward using stepAIC() R function was used for survival variable selection.

3.11 Model Checking and Diagnosis

Even though in the above topics have focused on different formulations and several extensions of joint models when it comes to using these models in practice, a prerequisite step is to validate the model's assumptions. In JM we assess convergence using multiple chains.If parallel chains with varying starting values give the same solution that will increase our confidence for convergence. A simple (informal) method of assessing chain convergence is to look at the history of iterations using a time series plot. If the chains show a reasonable degree of randomness between iterations, it signifies that the Markov chain has found an area of high likelihood and is integrating over the target density and hence indicating that it has converged.

3.11.1 Residuals for the Longitudinal Models

In the standard linear mixed-effects model, two types of residuals are often used, namely the subject-specific (conditional) residuals, and the marginal (population averaged) residuals [75]. The subject-specific residuals aim to validate the assumptions of the hierarchical version of the model. These residuals predict the conditional errors and can be used for checking the homoscedasticity and normality assumptions. Both types of residuals

can be used to check the assumptions of the longitudinal part of a joint model as well. Some basic residuals diagnostic plots are directly available by calling the plot () method for joint Model objects; the Q-Q plot of the subject-specific residuals, and the marginal survival and cumulative risk functions for the event process.

3.11.2 Residuals for the Survival Models

A standard type of residuals for the relative risk sub model of the joint model is the martingale residuals. The theoretical framework behind the use of martingales to investigate the fit of relative risk models has been provided by Barlow and Prentice [76]. The main use of these residuals is for a direct identification of excess events (i.e., to reveal subjects that are poorly fit by the model) and for evaluating whether the appropriate functional form for a covariate interest has been used in the model. An alternative type of residuals for survival models, related to the martingale residuals, is the Cox-Snell residuals [77]. For each subject, these are calculated as the value of the estimated cumulative risk function evaluated at the observed event time

3.12 Ethical Consideration

The research ethics review board of Jimma University would provide an ethical clearance for the study. The data has been collected after written permeation was obtained from JMUC that department of statistics write an official co-operation letter to the Hospital for the permission. The study conducted without informed consent since retrospective study design has been applied. Confidentiality of any information related to the patients and their clinical history has been maintained by keeping both the hard-copy and soft-copy of every collected data in a locked cabinet and password secured computer. Only the researcher would access to the de-identified data that has been kept in a secure place. All data has been coded with numbers and hospital numbers and without personal identifiers. All analysis has been on de-identified and coded data. During the study, there is no contact between the patients and the researcher. The study is noninvasive and without any harm to the patients. Then, the data obtained from the hospital has been secured.

Statistical Software Used

The statistical software used in this study for the data analysis is R version 3.6.3 .

4 Result and Discussion

4.1 Descriptive Statistics

The data for this study consists of 257 patients who are bipolar disordered individuals under psychiatric follow up from first, September, 2018 to first, January, 2020 in JUMC. Among the total bipolar disordered individuals during the time period 116(45.1%) of them are faced symptomatic recovery whereas 141(54.9%) of them are censored .

As observed from the table 4.1 below, 100(38.9%) are females and 63(24.5%) of them have not recovered whereas 37(14.4%) of them are show symptomatic recovery and 157(61.1%) are males and 77(30.4%) of them have not recovered whereas 79(30.7%) of them were show symptomatic recovery. Majority of bipolar disorders 128(49.8%) are found between 26-49 age category and 67(26.1%) of them have not recovered whereas 61(23.7%) of them are show symptomatic recovery. Many of bipolar disorders 145(56.6%) are single in marital status and 90(35.2%) of them were not recovered whereas 55(21.5%) of them are show symptomatic recovery compared to other marital categories. Based on relapse status of bipolar disorders, patients who are not faced event of relapse before study are 66(25.7%) and more likely to face event of recovery relative to those who had faced event of relapse before the study 50(19.5%). Many of bipolar disorders have faced first the bipolar disorder at their adolescence age 119(46.5%) relative to other age categories and 62(24.2%) of them have not recovered whereas 57(22.3%) of them have recovered.

Table 4.1: Frequencies and percentages for the baseline covariates with the recovery status of bipolar disordered patients.

		Recovery status		
Variable Names	Category	Censored	Events	Total
Sex	Female	63(24.5%)	37(14.4%)	100(38.9%)
	male	77(30.4%)	79(30.7%)	157(61.1%)
Age	1-19	12(4.7%)	2(0.8%)	14(5.4%)
	20-25	61(23.7%)	50(19.5%)	111(43.2%)
	26-49	67(26.1%)	61(23.7%)	128(49.8%)
	50 and above	1(0.4%)	3(1.2%)	4(1.6%)
Marital status	Single	90(35.2%)	55(21.5%)	145(56.6%)
	Married	31(12.1%)	44(17.2%))	75(29.3%)
	Divorced	17(6.6%)	15(5.9%)	32(12.5%)
	Widowed	3(1.2%)	1(0.4%)	4(1.6%)
Event of relapse	No	69(26.8%)	66(25.7%)	135(52.5%)
	Yes	72(28%)	50(19.5%)	122(47.5%)
First onset age	Childhood	41(16%)	21(8.2%)	62(24.2%)
	Adolescent age	62(24.2%)	57(22.3%)	119(46.5%)
	Adult age and above	38(14.8%)	37(14.5%)	75(29.3%)
Types of episodes	Manic episodes	80(31.6%)	73(28.9%)	153(60.5%)
	Mixed episodes	41(16.2%)	20(7.9%)	61(24.4%)
	Depressed episodes	19(7.5%)	20(7.9%)	39(15.4%)
Family history of Mental illness	No	82(31.9%)	81(31.5%)	163(63.4%)
	Yes	52(21.1%)	35(14.2%)	87(35.2%)

Substance abuse	No	83(33.6%)	77(31.2%)	160(64.8%)
	Yes	59(23%)	35(13.6%)	94(36.6%)
Religion	Muslim	99(38.5%)	49(19.1%)	148(57.6%)
	Orthodox	22(8.6%)	44(16.3%)	64(24.9%)
	Protestant	17(6.6%)	18(7%)	35(13.6%)
	Other's	2(0.8%)	7(2.7%)	9(3.5%)
Chewing khat	No	73(29%)	67(26.6%)	140(55.6%)
	Yes	68(27%)	44(17.5%)	112(44.4%)
Educational level	Tenth and below	75(33.5%)	50(22.3%)	125(55.8%)
	Above tenth and diploma	30(13.4%)	34(15.2%)	64(28.6%)
	Degree and above	21(9.4%)	14(6.2%)	35(15.6%)
Bipolar type	Type-I	108(42%)	96(37.4%)	204(79.4%)
	Type-II	33(12.8%)	20(7.8%)	53(20.6%)
Ethnicity	Oromo	127(49.4%)	96(37.4%)	223(86.8%)
	Amhara	7(2.7%)	11(4.3%)	16(6.2%)
	Others	7(2.7%)	11(4.3%)	16(6.2%)
Employment	No	117(47.8%)	73(29.8%)	190(77.6%)
	Yes	23(9.4%)	32(13.1%)	55(22.4%)
Other cofactors	No	55(21.4%)	76(29.6%)	131(51%)
	Yes	86(33.5%)	40(15.6%)	126(49%)

Many of bipolar disorders with manic type of episodes are 153(60.5%) and 80(31.6%) of them have not recovered whereas 73(28.9%) of them have recovered relative other types of episodes. Of the total bipolar disorder, high number of patients 223(86.8 %) are Oromo by ethnicity and 127(49.4%) of them have not recovered whereas 96(37.45) of them have recovered and 16(6.2%) of them are Amhara of which 7(2.7%) have not

recovered whereas 11(4.3%) have recovered. The other ethnicities are 16(6.2%) of which 7(2.7%) of them have not recovered whereas 11(4.3%) of them have recovered.

Without considering the censoring status of the bipolar disordered individuals, the average number of bipolar symptoms the patient have with their standard deviation at each time point in months is reported as table 4.2 below. It shows that mean burden of bipolar symptoms the patient have faced at base line time is 9.76 and in general it is decreasing over time in month. From the standard deviation of burden of bipolar symptoms we observed that there is high variation in between study time point that small variation at baseline time and end time of study.

Table 4.2: Mean of Burden of bipolar symptoms with its standard deviation over time point in months.

Time	Mean	std.deviation	Time	Mean	std.deviation	Time	Mean	std.deviation
0	9.76	4.74	5	5.64	3.76	9	2.55	3.14
1	8.05	3.38	5.25	8	–	9.5	3.93	8.22
2	7.39	4.87	5.5	3.23	3.06	10	2.28	4.12
2.5	7.79	6.33	6	4.5	4.75	10.5	1.25	3.58
3	6.74	6.14	6.5	4.68	8.72	11	1.96	3.96
3.5	6.83	17.77	7	3.76	4.23	11.5	4	–
4	5.23	4.48	7.5	2.32	6.67	12	1.67	1.33
4.5	6.17	11.67	8	3.17	3.69	13	1.67	1.33
4.75	3.8	–	8.5	4.77	8.86	14	0.33	0.33

From the above table we can see that the mean burden of bipolar symptoms at the baseline and the first month of bipolar disorders decreases by (8.05-9.76) = 1.71 amount.

4.2 Results Using Separate Models

The data is analyzed separately first using longitudinal and survival models described in Section 3.4.1 and 3.5. This is important for the fully specification of the mean response of the model and determine the random effects and fixed effects to be included in the longitudinal sub model, and to identify the covariates that have a contribution on the hazard of an event in the survival sub model and also to provide initial values for the joint analysis.

4.2.1 Separate Analysis of the Longitudinal Data

Before directly going to the analyzing, first the data exploration is conducted for the longitudinally measured burden of bipolar symptoms that obtained by counting number of symptoms the patient shows during each observation time.

4.2.2 Exploring Individuals Profile and the Mean Structures

Exploratory analysis of longitudinal data seeks to discover patterns of systematic variation across groups of patients, as well as aspects of random variation that distinguish individual patients.

Figure 4.1 on the left hand side below indicates the individual profile plot of 16 randomly selected individuals with burden of bipolar symptoms. It indicates within and between subjects variability on burden of bipolar symptoms over time. Close inspection of the shapes of the burden of bipolar symptom profiles indicates that for some individuals, these seem to be nonlinear.Therefore, during fitting later in JMbayes,regression splines is required to estimate the baseline relative risks. Whereas on the right hand side, is the overall individual curves with loess smoothing technique suggest the linear growth effect in the mean structure of burden of bipolar symptoms over time. Therefore the linear time effects should include as fixed-effects and random effects in the model.And also indicates the mean burden of bipolar symptoms is decreasing over time.

Exploring the Mean Structure, to understand the possible relationships among the burden

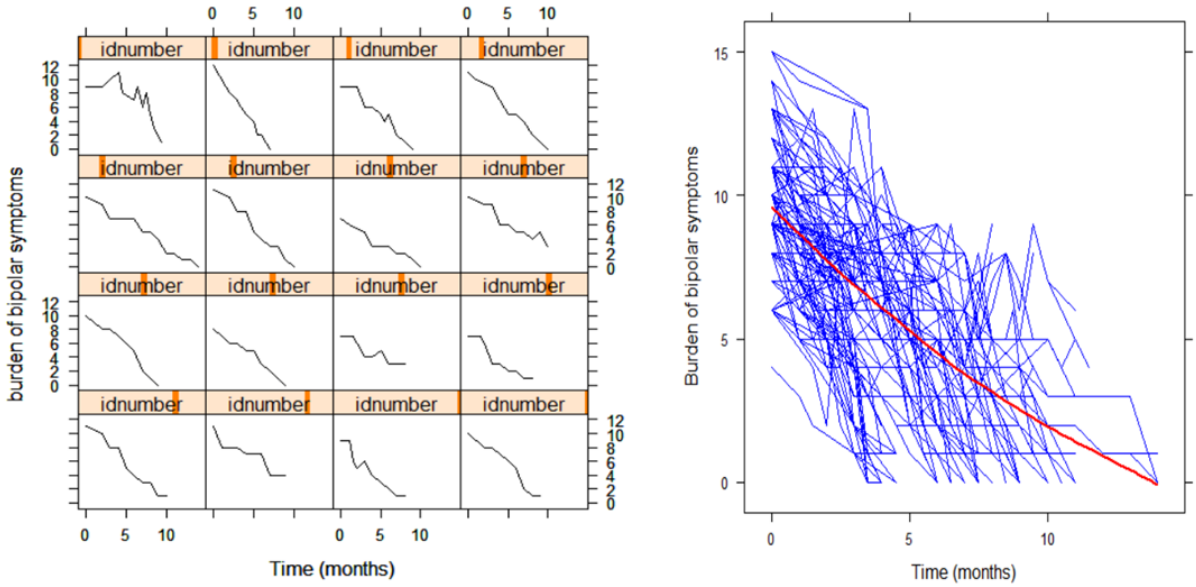

Figure 4.1: individual profile plot of 16 random sampled individuals and overall individual profile plot with loess curve over time.

of bipolar symptoms means over time, a plot of a line connecting the average values computed at each time point likes the Figure 4.2 below. Then mean structure plot suggests that the mean burden of bipolar symptoms profiles have a nonlinear growth over time.

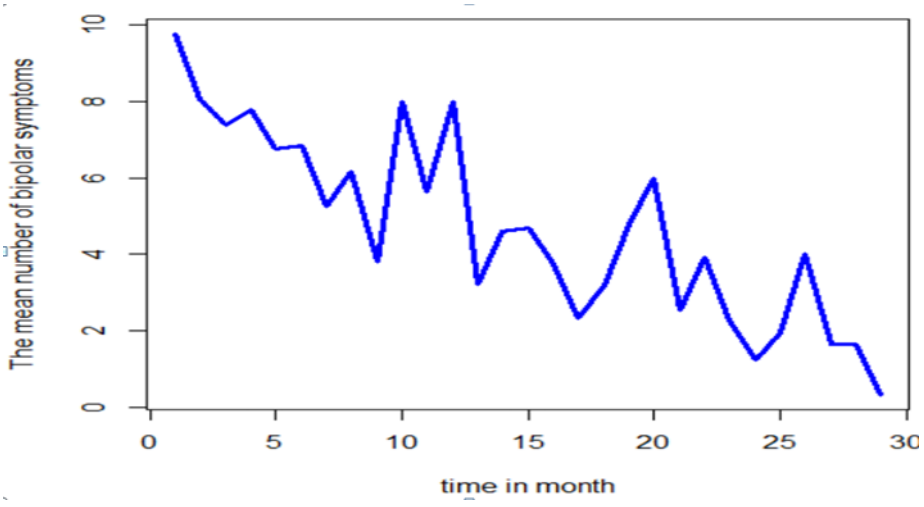

Figure 4.2: Mean of Burden of bipolar symptoms over time in months

.

4.2.3 Generalized Linear Mixed Effect Model Result

From the individual profile plots and mean structure exploratory analysis, linear and quadratic time effects seem to be useful in modeling the random effects. Then the generalized linear mixed model is built within two stages in which the first stage involves fitting generalized linear regression models which only considers within variability of the subjects and an appropriate fixed effect for the outcome variable based on the AIC values of the fitted candidate GLM. Therefore let y_{ij} is the burden of bipolar symptoms the patients have faced at each time t_{ij} , i=1, 2... N, j=1, 2,...,n_i. where , $Y_{ij} = (y_{i1}, y_{i2}....., y_{ij})$ and $\varepsilon \sim N(0.D_{ij})$. Then generalized linear model is first fitted by Poisson model and if over dispersion is observed Negative Binomial distribution is appropriate. The fitted model in the table 5.1 explain the burden of bipolar symptoms by accounting only the between variability of the bipolar disordered patients.The variables are selected based on Akaike information criteria. we see that dispersion parameter for poison distribution taken to be 1 and We also see that the residual deviance is equal to the degrees of freedom, so that we have not over-dispersion. that indicates there is not over dispersion. This means variance of the poison distribution is equal to its mean. Then the poison distribution is appropriate for this study to fit both GLMM .

To have an appropriate generalized linear mixed model the selected fixed effects are fitted with different random effects starting with only random intercept up to random intercept; linear and quadratic slopes. Indeed; the final appropriate generalized linear mixed model is selected with smallest AIC and BIC of the fitted models. The summaries of fitted different GLMM using glmer() function in nlme4 package by considering different random effects are reported on the table 4.3 below.

As reported on table 4.3 below seven generalized linear mixed models are fitted with different time random effects starting from random intercept to random intercept; linear and quadratic time slopes using Laplace approximation. Of the seven fitted GLMM, random linear slope time effect is selected as best time random effect with small 8494.9 AIC and 8663.5 BIC values and based significance value of chi-square test.

Then b_{i1} is the selected random linear time slope for count burden of bipolar symptoms,

Table 4.3: : AIC and BIC value of the fitted generalized linear mixed effects model by considering different random effects.

Random affects	AIC	BIC	$Pr(>Chisq)$
Random intercept only	8615.3	8772.7	–
Random linear slope	8494.9	8663.5	0.000
Random intercept and linear slope	8598.6	8750.3	1.000
Random intercept and quadratic slope	8494.9	8663.5	1.000
Random linear and quadratic slope	8494.9	8663.5	1.000
Random intercept, linear and quadratic slope	8494.9	8663.5	1.000

that the model assumes the subject specific linear time trend in the underlying trajectory of burden of bipolar symptoms, affect the hazard of symptomatic recovery. After the appropriate random effect is selected for GLMM an appropriate covariates are selected among the selected fixed effects of the fitted generalized linear model for the optimality of the generalized linear mixed model. Then final appropriate GLMM for the longitudinal separate model is fitted using both Laplace approximation which use (In R: glmer() in (package lme4)) and then Penalized quasi-likelihood approximation which use (In R: glmmPQL() in (package MASS)) as reported in the table 4.4 below.

The independence and normality assumption diagnosis of the random effects was checked by using residual versus fitted value and quantile-quantile plots shown in Appendix part of figure 5.12 and the plots show no problems of the normality assumptions. From the selected generalized linear mixed model in table below, we can see that the estimated parameter for both estimation methods is nearly the same. But we discuss the about the significant parameters using penalized quasi-likelihood estimation since it is applicable in Joint model estimation later using JMbayes() package in R.

The parameters intercept term, religion, other cofactors, event of relapse, interaction between linear time and event of relapse, interaction between linear time and other cofactors

and the interaction between substance abuse and chewing khat have a positive significant effect on the log expected of burden of bipolar symptoms at 95% confidence level. Where the parameters linear time in month, age and the interaction between linear time in month and first onset have a negative significant effect on the log expected burden of bipolar symptoms at 95% confidence level.

β_0=2.313 estimated parameter for intercept, indicates that an expected log burden of bipolar symptoms during the first follow up time.β_1 = - 0.245 time in month, indicates that for one unit increase in time in month, the log expected burden of bipolar symptoms decreases by 0.245 taking other factors constant.

β_{23}= 0.256 protestant religion, it is the log expected difference in burden of bipolar symptoms between protestant religion and Muslims (reference). It indicates that the log expected burden of bipolar symptoms for Protestant religion was 0.2256 greater than Muslim religion when other covariates remain constant.β_{24} = 0.243 other religions, it was the log expected difference in burden of bipolar symptoms between other religions and Muslims (reference). It indicates that the log expected burden of bipolar symptoms for other religions is 0.225 greater than Muslims when other covariates remain constant.β_{34} = - 0.586, 50 and above age groups, it indicates that the log expected difference in burden of bipolar symptoms between 50 and above age group and 19 and below age groups. It indicates that the log expected burden of bipolar symptoms for 50 and above age group is 0.586 less than 19 and below age group when other covariates remain constant.

β_{71}=0.151 event of relapse, it is the log expected difference in burden of bipolar symptoms between those who face event of relapse and those patients who do not face event of relapse ever. It indicates that the log expected burden of bipolar symptoms for those who faced event of relapse is 0.151 greater than those do not faced event of relapse when other covariates remain constant. β_{11}= -0.029, interaction between time in month and adult age first onset of bipolar disorder, It indicates that the log expected burden of bipolar symptoms for the interaction between time in month and adult age first onset of bipolar disease is 0.029 less than the childhood first onset of bipolar diorder when other variables remain constant.

β_{12} = 0.049 the interaction between linear time in month and event of relapse, it indictes the log expected difference in burden of bipolar symptoms between the interaction of linear time in month and event of relapse realative to those patients who do not face event of relapse ever and time interraction. It indicates that the log expected burden of bipolar symptoms for the interaction of time in month and event of relapse is 0.049 greater than those did not face event of relapse and time interaction when other covariates remain constant.β_{14} = 0.049 the interaction between linear time in month and existence of other cofactors, it was the log expected difference in burden of bipolar symptoms between the interaction effect of linear time in month and other cofactors and no other cofactors. It indicates that the log expected burden of bipolar symptoms for the interaction between time in month and other cofactors is 0.049 greater than no other cofactors when other covariates remain constant.

β_{15} = 0.150 the interaction between substance abuse and chewing khat , it is the log expected difference in burden of bipolar symptoms between the interaction of substance abuse and chewing khat and not chewing. It indicates that the log expected burden of bipolar symptoms for the interaction of substance abuse with chewing khat is 0.150 greater than those who are not chewing khat when other covariates remain constant.

The σ_{b0} and σ_{b1} are 0.148 and 0.068 respectively and indicates the heterogeneity in the longitudinal measurement of the burden of bipolar symptoms that must be accounted for.

Table 4.4: : The final selected GLMM and estimated parameters with their 95% confidence interval

		Laplace approximation		Penalized quasi-likelihood
Fixed effects	coeff(β)	95% CIs	coeff(β)	95% CIs
β_0	2.343	[2.207, 2.479]$*$	2.313	[2.172,2.454]$*$
β_1	-0.232	[-0.260,-0.203]$*$	-0.245	[-0.272,-0.218]$*$
β_{22}	0.031	[-0.040, 0.102]	0.044	[-0.031, 0.119]
β_{23}	0.218	[0.129,0.308]$*$	0.256	[0.161, 0.350]$*$
β_{24}	0.190	[0.052, 0.328]$*$	0.243	[0.099, 0.387]$*$
β_{32}	-0.151	[-0.273, -0.028]$*$	-0.103	[-0.236, 0.028]
β_{33}	-0.157	[-0.295, -0.018]$*$	-0.124	[-0.273, 0.023]
β_{34}	-0.525	[-0.773, -0.278]$*$	-0.586	[-0.823,-0.350]$*$
β_{41}	-0.008	[-0.080, 0.063]	-0.008	[-0.075, 0.059]
β_{51}	0.031	[-0.060, 0.123]	0.045	[-0.046, 0.059]
β_{62}	0.021	[-0.066, 0.109]	0.003	[-0.081, 0.087]
β_{63}	0.037	[-0.069, 0.145]	0.083	[-0.021, 0.187]
β_{71}	0.181	[0.102, 0.260]$*$	0.151	[0.075, 0.228]$*$
β_{81}	0.044	[-0.041,0.130]	0.015	[-0.075,0.106]
β_{91}	-0.049	[-0.121, 0.022]	-0.045	[-0.121, 0.031]
β_{10}	0.027	[0.001, 0.054]	0.026	[0.0002, 0.053]
β_{11}	-0.014	[-0.043, 0.014]	-0.029	[-0.058,-0.0003]$*$
β_{12}	0.041	[0.017, 0.064]$*$	0.049	[0.026, 0.071]$*$
β_{13}	-0.128	[-0.235, -0.02]$*$	-0.046	[-0.159, 0.066]
β_{14}	0.042	[0.017, 0.066]$*$	0.049	[0.025, 0.073]$*$
β_{15}	0.192	[0.070, 0.314]$*$	0.150	[0.025, 0.276]$*$
			Random effects	Random effects
σ_{b0}	0.000	–	0.148	[0.121, 0.182]
σ_{b1}	0.053	–	0.068	[0.058, 0.081]
		AIC= 8611.9		

Where;$\beta_0 = intercept, \beta_1 = timeinmonth, \beta_{22} = religion(orthodox)$,

$\beta_{23} = religion(protestants), \beta_{24} religion(others), \beta_{32} = age(20-25), \beta_{33} = age(26-49), \beta_{34} =$
$age(50 andabove), \beta_{41} = othercofactors(yes), \beta_{51} = substanceabuse(yes), \beta_{62} = firstonset(adolescenceag$
$firstonset(adultageandabove), \beta_{71} = eventofrelapse(yes),$
$beta_{81} = familyhistoryofmentalillness(yes), \beta_{91} = chewingchat(yes),$
$\beta_{10} = \beta_1 * \beta_{62}, \beta_{11} = \beta_1 * \beta_{63}, \beta_{12} = \beta_1 * \beta_{71}, \beta_{13} = \beta_{71} * \beta_{81}, \beta_{14} = \beta_1 * \beta_{14}, \beta_{15} = \beta_{51} * \beta_{91}$

4.2.4 Separate Survival Data Analysis

4.2.5 Kaplan-Meier Survival Function Estimates

The most common-parametric technique for modeling the survival function is the Kaplan-Meier estimate. The Kaplan-Meir estimated median value that the half of the bipolar disorders experiences the event was 9 months. It was applied to estimate the survival curves for categorical covariates and the estimated survival probability curve of some selected categorical covariates; which is displayed in the following figures:-

From the figure 4.3, we can see that those single individuals has highest probability curve of time to recovery. The divorced individual has the lowest probability curve of time to recovery relative to marital status of other groups. The patients who faces event of relapse has high probability curve of time to recovery relative to those who do not faced event of relapse. The patients who are substance abuse have higher time to recovery relative to those who do not use substance abuse. The patients who had faced event of relapse before had higher time to recovery relative to those who had not faced any event relapse.To test the significance difference of the plotted curves by different covariates the log rang tested were employed and the result of log rank test was reported as table below. The hypothesis in Log-Rank Survival Estimates is as follows: H_0: Survival curve for all groups are the same vs H_1 Survival curve for all groups are different.

As indicated on log rank test of each covariate on table above there is a significance difference in survival probability curve by Family history of mental illness; religion; age; event of relapse; other factors and substance abuse of patients since the computed P-value of log rank test statistics for these covariates are less than 5% significance level where as there is no significance difference in survival probability curve by sex; freestones; chewing; ethnicity ;marital status ;employment and bipolar type covariates since the p-value of the computed log rank test statistics for these covariate groups are greater than 5% significance level.

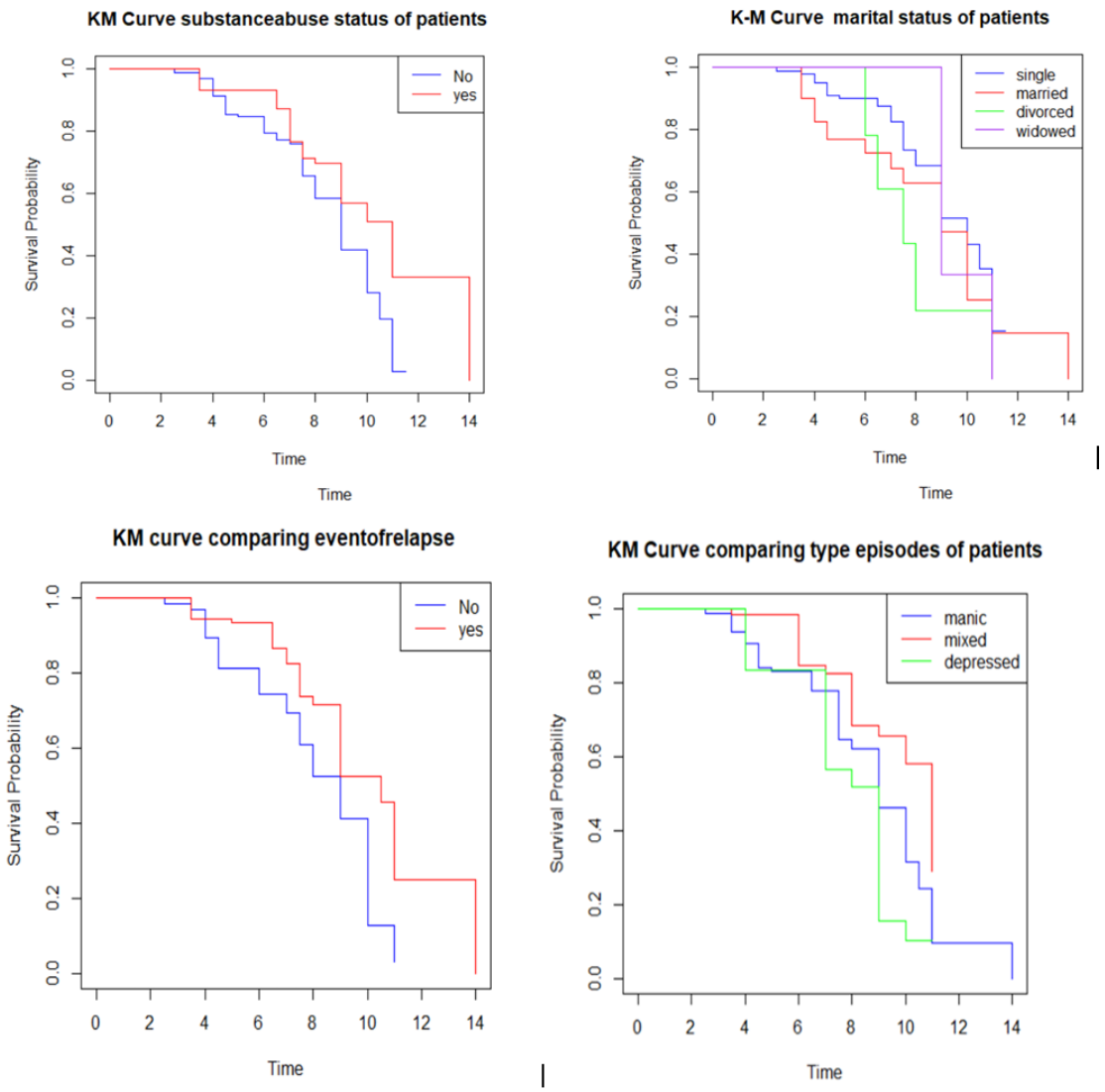

Figure 4.3: Estimated Kaplan-Meier survival curve substance abuse, type of episodes, marital status and substance abuse the of patients.

.

4.2.6 Result of Cox-regression Model

To determine the variables to be included in the survival model, an automatic variable selection method stepAIC in R using backward elimination method is applied. Regardless of the survival time distributions,Covariates selected are first onset; religion; bipolar type; substance abuse; age and event of relapse are extracted to be included in the model. In the joint modeling cox regression is used as default in in the JMbayes package using jointModelBayes() function when regression splines are used to estimate the baseline

hazards.

Table 4.5: :The estimated parametes for cox regression model models with their 95% CI

	Cox PH		
Covariates	Estimated values	P-value	95 % CIs
Intercept	—	—	
(maritalstatus)2	-0.025	0.567	[-0.656 ,0.605]
(maritalstatus)3	0.905	0.021	[0.208 ,1.601]
(maritalstatus)4	0.203	0.530	[-1.196,1.604]
(religion)2	0.287	0.337	[-0.226,0.802]
(religion)3	-0.065	0.959	[-0.746,0.615]
(religion)4	1.488	0.011	[0.438,2.539]
(age)2	2.046	0.006	[0.675,3.416]
(age)3	1.504	0.107	[-0.092,2.915]
(age)4	3.662	0.001	[1.619,5.704]
(eventofrelapse)1	-1.132	0.000	[-1.651,-0.613]
(bipolartype)1	-0.836	0.013	[-1.504 ,-0.167]
(typeofepisodes)2	-0.542	0.104	[-1.111, 0.026].
(typeofepisodes)3	0.925	0.024	[0.264, 1.586]
(substanceabuse)1	-0.892	0.003	[-1.423,-0.361]
AIC	Null model =1187.26	Full model =726.49	

From the table 4.5 , the covariates marital status,religion,age,event of relapse, bipolar type ,types of episodes and substance abuse are significantly affect the time to symptomatic recovery of bipolar disorder.

4.3 Result Using Joint Model

In this study, the ”param” specifies the association structure between the longitudinal and survival processes. The ”shared-RE” association structure is used in which only the random effects of the generalized linear mixed model are included in the linear predictor of the survival submodel. And the GLMM is obtained by using argument densLong, since the response variable for this study is count and follows Poisson distribution follows. glmmPQL() function from MASS package is used to generate GLMM using penalized quasi-likelihood. glmmPQL () works by repeated calls to *lme*, so package nlme should be loaded at first use if necessary. Based on the GLMM that incorporate subject specific variance under longitudinal sub-model and semi-parametric model under survival sub-model, we explore several joint models with a variety of latent processes.

In this study the Bayesian method results are based on four parallel MCMC sampling chains of 20,000 iterations each, following a 3000 discarded as burn-in to achieve convergence. Several joint models using different shared parameter association structure with different combinations of the random effect processes are fitted. Then in this study shared parameter association structure is selected based on the smallest DIC value of the joint models which are reported on table 4.6 below.

For checking convergence of the MCMC chains, we have used time series plot of the history of iterations of the final joint model , which shows a reasonable degree of randomness between iterations and the overlaps of the three chains indicates that the same solutions are obtained for each initial values.Therefore, the Gibbs sampler has converged to the target density.

Table 4.6: JM Selection using different shared parameter association structure(random effects).

Models	$U_1 i(t)$	$U_2 i(t)$	PD	DIC
Random intercept only				
I	b_{0i}	$m_i(t)$	277.586	9216.139
II	b_{0i}	$\alpha_0 b_{0i}$	281.82	9207.546
III	b_{0i}	$\alpha_0 b_{0i}$+$\alpha_1 b_{1i}$	278.11	9217.67
Random intercept and linear slope				
IV	b_{0i} +$b_1 itij$	$m_i(t)$	439.486	10021.36
V	b_{0i} +$b_1 itij$	$\alpha_0 b_{0i}$	439.486	10021.36
VI	b_{0i} +$b_1 itij$	$\alpha_1 b_{1i}$	380.268	9970.937
VII	b_{0i} +$b_1 itij$	$\alpha_0 b_{0i}$+ $\alpha_1 b_{1i}$	380.268	9970.937
Random linear slope				
VIII	$b_1 itij$	$m_{i(t)}$	439.486	10021.36
IX	$b_1 itij$	$\alpha_1 b_{1i}$	380.268	9970.937
Random quadratic slope				
X	$b_2 itij^2$	$m_i(t)$	439.486	10021.36
XI	$b_2 itij^2$	$\alpha_2 b_2 itij^2$	380.268	9970.937

Random linear and quadratic slope				
XII	$b_1itij+b_2itij^2$	$m_i(t)$	439.486	10021.36
XIII	$b_1itij+b_2itij^2$	$\alpha_1 b_1itij$	380.268	9970.937
XIV	$b_1itij+b_2itij^2$	$\alpha_2 b_2itij^2$	380.268	9970.937
XV	$b_1itij+b_2itij^2$	$\alpha_1 b_1itij+\alpha_2 b_2itij^2$	380.268	9970.937
Random intercept; linear and quadratic slope				
XVI	$b_{0i}+\alpha_1 b_1itij+\alpha_2 b_2itij^2$	$m_i(t)$	439.486	10021.36
XVII	$b_{0i}+\alpha_1 b_1itij+\alpha_2 b_2itij^2$	$b_{0i}+\alpha_1 b_1itij+\alpha_2 b_2itij^2$	380.286	9970.937

As mentioned above, the precise nature of the two sub models has chosen; the longitudinal to be GLMM with subject-specific variances and the survival model is the Cox-PH model. The table 4.6 above reports PD and DIC scores for 17 joint models with different random effects that are used as the latent shared processes $U_1(it)$ and $U_2(it)$. The GLMM incorporates patient-specific burden of bipolar symptoms for the longitudinal sub-model and time to symptomatic recovery of Cox model is used for survival sub-model.The simple joint models I, II and III with random intercepts for longitudinal sub-model is fitted first and the incorporation of random intercepts in the longitudinal sub-model improves the total DIC. In model III we add random intercept and linear time in month slope which do not improves the decrement in DIC. Models IV, V, VI, and VII are fitted with both linear and intercept random slopes.

In models VI, VII there is no improvement in DIC by incorporating linear time in month effects and intercept to $U_2(it)$ part. Model VIII and IX are fitted with linear random slope and there is no improvement in DIC. Models X, XI are fitted with quadratic random slope and there is no improvement in DIC. Models XII, XIII, XIV and XV are fitted with linear and quadratic random slopes and there is no improvement in DIC. Also, Models XVI and XVII are fitted with intercept, linear and quadratic random slopes and there no improvement in DIC in general. Because Model II emerges with the smallest total DIC 9207.546

among all other models, it is selected as the finaljoint model for the burden of bipolar symptoms and time to symptomatic recovery of bipolar disorder obtained from JUMC in this study.

Table 4.7: :The estimated joint model of the longitudinal burden of bipolar symptoms and Time-to- symptomatic recovery of bipolar disordered patients

Longitudinal sub-model			Survival sub-model	
Fixed effects	coef(β)	95% CIs	Covariates	Coef(γ) 95% CIs
β_0	2.316	[2.244,2.385]*	γ_{12}	0.425 [-0.475, 1.339]
β_1	-0.011	-0.152, 0.123]	γ_{13}	2.358 [1.216,3.449]*
β_{22}	0.175	[0.002, 0.379]*	γ_{14}	1.239 [-0.625, 3.078]
β_{23}	0.196	[-0.041, 0.457]	γ_{22}	-0.734 [-2.267, 0.456]
β_{24}	-0.106	[-0.257, 0.046]	γ_{23}	-0.747 [-2.527, 0.757]
β_{32}	-0.165	[-0.340, -0.007]*	γ_{24}	0.104 [-2.765, 2.156]
β_{33}	-0.473·	[-0.811, -0.149]*	γ_{32}	1.961 [0.241, 3.743]*
β_{34}	-0.005	[-0.090, 0.080]	γ_{33}	0.846 [-0.696, 2.599]
β_{41}	0.059	[-0.059, 0.190]	γ_{34}	3.800 [0.501, 8.902]*
β_{51}	0.007	[-0.086, 0.107]	γ_{41}	-1.857 [-2.794,-0.918]*
β_{62}	0.005	[-0.134, 0.134]	γ_{51}	-1.195 [-2.222, -0.285]*
β_{63}	-0.216	[-0.234, -0.195]*	γ_{61}	-1.584 [-2.539,-0.839]*
β_{71}	0.122	[-0.004, 0.246]	γ_{62}	1.241 [0.552, 2.015]*
β_{81}	-0.025	[-0.122, 0.097]	γ_{71}	-1.137 [-2.061,-0.334]*
β_{91}	-0.024	[-0.117, 0.064]	Random effects	
β_{10}	0.029	[0.014, 0.046]*	$\sigma_{\sigma 0} = 0.410$	
β_{11}	-0.006	[-0.022, 0.013]	Association	
β_{12}	0.040	[0.023, 0.056]*		
			α_0= -8.403	[-11.157,-6.576]*
β_{13}	0.010	[-0.129, 0.133]		
β_{14}	0.031	[0.016, 0.047]*		
β_{15}	0.213	[0.052, 0.49]*		
				DIC = 9207.546

* Indicates significance of the covariates at 5% level of significance

CIs is the 95% credibility intervals indicates an interval within which an unobserved parameter value falls with a particular probability.

Where;β_0 =intercept, $\beta_1 = timeinmonth$,β_{22}=religion (orthodox),
β_{23} =religion (protestants),$\beta_{24} religion(others)$,$\beta_{32}$= age
(20-25),β_{33}=age(26-49),β_{34}=age(50andabove),β_{41}=other cofactors(yes),β_{51}= substance
abuse(yes),β_{62}=first onset(adolescence age),β_{63}=first onset(adult age and
above),β_{71}=event of relapse(yes),
$beta_{81}$=family history of mental illness(yes), β_{91}=chewing chat(yes),
β_{10}=$\beta_1 * \beta_{62}, \beta_{11}$ =$\beta_1 * \beta_{63}$,$\beta_{12} = \beta_1 * \beta_{71}$,$\beta_{13} = \beta_{71} * \beta_{81}$,$\beta_{14} = \beta_1 * \beta_{14}$,$\beta_{15} = \beta_{51} * \beta_{91}$
and for survival coefficients were, γ_{12}=marital status(married),γ_{13}=marital
status(divorced),γ_{14}=marital status(widowed), γ_{22}=
religion(orthodox),γ_{23}=religion(protestant), γ_{24}=religion(others),γ_{32} =age(20-25),γ_{33}
=age(26-49),γ_{34} =age(50 and above),γ_{41}=event of relapse(yes),γ_{51} =bipolar type,γ_{61}
=types of episodes(mixed), γ_{62}=types of episodes(depression), γ_{71}=substance abuse(yes).

As observed from the appropriate joint model with a minimum DIC score values than the remaining joint models is selected. The longitudinal sub-model specification is the same to that of the selected generalized linear mixed model whereas the survival sub-model specification incorporates the association parameters to the selected cox-regression model.

The posterior estimates of the regression coefficients of longitudinal sub-model and survival sub-model with their 95% credible interval are summarized in the Table 4.7 above. In the longitudinal sub-model the covariates; The intercept, time in month, 50 and above age group, the interaction between time in month and adolecentage first onset of the bipolar disorder, the interaction between time in month and event of relapse, the interaction between linear time in month and existence of other cofactors and the interaction of substance abuse and chewing khat are significantly affect the log expected burden bipolar symptoms. Where as in the survival sub-model the covariates; divorced marital status, event of relapse, mixed type of episodes are significantly affects the symptomatic recovery of bipolar symptoms .

The association between the longitudinal outcomes and the time-to-event outcome is expliained by α_0=-8.403 with [-11.157,-6.576] credible interval. It associates longitudinal count, burden of bipolar symptoms and time-to-symptomatic recovery of bipolar disorder

using shared random effect parameters . We observe significant strong negative association between the subject-specific random intercept (base line) of the burden of bipolar symptoms and the risk ratio of symptomatic recovery since their 95% credible intervals excludes zero. The σ_{b0}=0.410, is an indication of heterogeneity in the longitudinal measurement of the burden of bipolar symptoms that must be accounted for.The mean burden of bipolar symptom scores at baseline for a reference individual are estimated at 0.410.

Table 4.8: Posterior estimated risk ratio and their 95% credible intervals of the survival sub-model fitted in the joint model.

Covariates	Hazard ratio	95% CIs
γ_{12}	1.529	[0.621, 3.817]
γ_{13}	10.572	[3.376, 31.496]$*$
γ_{14}	3.452	[0.534, 21.752]
γ_{22}	0.479	[0.103, 1.577]
γ_{23}	0.473	[0.079, 2.133]
γ_{24}	1.109	[0.062, 8.643]
γ_{32}	7.110	[1.273, 42.235]$*$
γ_{33}	2.332	[0.498, 13.454]
γ_{34}	0.4338	[1.650, 734.928]$*$
γ_{41}	0.156	[0.006, 0.399]$*$
γ_{51}	0.302	[0.108, 0.075]$*$
γ_{61}	0.205	[0.078, 0.432]$*$
γ_{62}	3.460	[1.737, 7.500]$*$
γ_{71}	0.320	[0.127, 0.715]$*$
α_0	0.00022	[0.000014, 0.0013]$*$

were,γ_{12}=marital status(married),γ_{13}=marital status(divorced),γ_{14}=marital status(widowed), γ_{22}= religion(orthodox),γ_{23}=religion(protestant), γ_{24}=religion(others),γ_{32} =age(20-25),γ_{33} =age(26-49),γ_{34} =age(50 and above),γ_{41}=event of relapse(yes),γ_{51} =bipolar type,γ_{61} =types of episodes(mixed), γ_{62}=types of episodes(depression), γ_{71}=substance abuse(yes)

The above table 4.8 shows that the Coefficient and hazard ratios of the standard relative risk model to determine the hazard of symptomatic recovery using baseline covariates and with the assumption of a time-independent covariate. The risk ratio value of 1 means no difference in relative risk between groups. When the reiative risk is less than one, the corresponding covariate has negative effect on the event of interest and when it is greater than one , the corresponding covariate has a positive effect on the event of interest.

γ_{13}= 10.572, shows that the risk of symptomatic recovery time for divorced bipolar disorders is greater than the single bipolar disorders given burden of burden bipolar of symptoms.

.γ_{32} =7.110, shows that the risk of time to symptomatic recovery for between twenty and twenty five age group is 7.110 greater than those nineteen and below nineteen age groups given burden of bipolar symptoms. γ_{34} = 44.78, shows the risk ratio of recovery time for fifty and above age group of bipolar disorders is greater than those nineteen and below age groups given burden bipolar symptoms.

γ_{41}=0.156, shows that the risk of recovery time for bipolar disorders those faced event of relapse before is less than for those who did not faced event of relapse before given the burden of bipolar symptoms. γ_{51} = 0.302, shows that the relative risk of recovery time for bipolar disorders with bipolar II is 30.2 % less than those with bipolar I given the burden of bipolar symptoms.

γ_{61}=0.205, shows that the relative risk of recovery for bipolar disorders with mixed type of episodes is less than those with manic type of episodes given the burden of bipolar symptoms. γ_{62}=3.460, shows that the hazard of recovery for bipolar disorders with depressed type of episodes is greater than those with manic type of episodes given the burden of bipolar symptoms. γ_{71}=0.320, shows that the risk of recovery for substance abuse bipolar disorders is less than those not abuse substances given the burden of bipolar symptoms.

$exp(\alpha_0)$= 0.00022, shows that the relative risk of symptomatic recovery time for bipolar disorders at base line time(random intercept). It indicates that the one unit increase in the burden of bipolar symptoms at starting point in time results in a 0.00022(0.022%) decreased risk of time to symptomatic recovery with 95% credible interval of [0.000014,

0.0013]. This results are statistically significant since its credible interval excludes zero, indicating that burden of bipolar symptoms is a good predictor of time to symptomatic recovery.

4.4 Discussion

In this study, the advantages and applications of joint models for longitudinal and survival data on bipolar disorder are discussed. When the objective of the study is to investigate the effect of the longitudinal outcome on the time to the event data, the joint modeling approach leads to unbiased and more efficient estimates of the longitudinal effect considering the correct model for the time-to-event data. Then main aim of this study is to identify the association between burden of bipolar symptoms and time to symptomatic recovery from first, September, 2018 to first, January, 2020 in Jimma University Medical Center. This study illustrates that how joint modeling can be used as an alternative when dealing with repeated measurements and time-to-event data in the predicting an individual's symptoms of bipolar disorders. The burden of bipolar symptoms and time to symptomatic recovery may be improved by the use of these types of models which take into consideration all the individual bipolar symptoms that the patient shows and have recorded over a time period.

In this study of 257 bipolar disorders individuals under psychiatric follow up during the time period 116(45.1%) of them are faced symptomatic recovery whereas 141(54.9%) of them are censored. Similarly the study done on department of psychiatry, in University of Cincinnati, College of Medicine, by Keck Mcelroy SDtrakowski(1998) shows symptomatic recovery occurred in only 26% disorders among 134 total bipolar disorders. Also study by Perlis, Ostacher(2006) shows that of 1,469 participants symptomatic at study entry, 858 (58.4%) subsequently achieved recovery supports the idea which aslo support this study.

The study done on the Longitudinal Course of Bipolar-I Disorder and Duration of Mood Episodes by Solomon et al,(2010) with the objective of describe the duration of bipolar I mood episodes and factors associated with recovery from these episodes while the probability of recovery over time from multiple successive mood episodes was examined with survival analytic techniques, showed that the median duration of bipolar I mood episodes was 13 weeks which support the result of this study.

This study shows that the hazard of recovery for bipolar disorders with depressed type

of episodes is greater than those with manic type of episodes given the burden of bipolar symptoms with a 95% crediable interval of [1.737, 7.500], which is agree with study by Solomon (2010) showed that the probability of recovery from a major depressive episode was a signicantly greater probability of recovery from an episode of mania (HR=1.713; 95% and Conficence interval of [1.373,2.137].

This study shows that the relative risk of recovery for bipolar disorders with mixed type of episodes is less than those with manic type of episodes given the burden of bipolar symptoms which disagree with the study done by by Keck Mcelroy SDtrakowski(1998) by using multivariate analyses suggests that;during the 12-month follow-up period, there were no significant differences in outcome between patients with manic compared with mixed BD.This may due ineffient of the model he use which means using Joint model may improve the efficiecy and unbiasedness of the estimator.

Bayesian methods of joint modelling enabled to specify a time-varying coefficient to link the longitudinal and the survival processes [80]. where, Andrinopoulou [81], looked at the valve function of cardiac-thoracic surgery which is monitored over a period of time. The approach was implemented via P-splines using Bayesian methods of joint modelling that enabled to specify a time-varying coefficient to link the longitudinal and the survival processes which support .Similarly this study used Bayesian approach estimation with regression splines to estimamate the base line hazard function.Two candidate association structures; shared parameter structure associates by Henderson [81] are used to associates the longitudinally counted burden of bipolar symptoms at the same time point to the survival sub-model .

The joint models are estimated under Bayesian framework using a four chain of 20,000 MCMC iterations from which we discarded the first 3,000 samples as burn-in; finally trace time series plots are plotted for the convergence diagnosis of MCMC samples and in the plots there is not any problem of convergence failure.Our results support the hypothesis of an association between time to symptomatic recovery and burden of bipolar symptoms. We found that the hazard of symptomatic recovery increases with by decreasing the burden of bipolar symptoms indicates the negative relationship between them.

5 Conclusion and Recommandations

5.1 Conclusion

The data for this study consists of 257 patients who are bipolar disordered individuals. Among the total bipolar disordered individuals during the time period 116(45.1%) of them are faced symptomatic recovery whereas 141(54.9%) of them are censored .

Then the main objective of current study is to fit Bayesian joint longitudinal and survival model and assess the association between burden of bipolar symptoms and time to symptomatic recovery. The intercept term, religion, other cofactors, event of relapse, interaction between linear time and event of relapse, interaction between linear time and other cofactors and the interaction between substance abuse and chewing khat have a positive significant effect on the log expected of burden of bipolar symptoms. Where the variables linear time in month, age and the interaction between linear time in month and first onset have a negative significant effect on the log expected burden of bipolar symptoms.

The covariates marital status,religion,age,event of relapse,bipolar type ,types of episodes and substance abuse are significantly affect the time to symptomatic recovery of bipolar disorder.

Findings from this work has revealed that the risk increment of symptomatic recovery with bipolar disorder can best be predicted using burden of bipolar symptom scores and analysed via joint longitudinal and survival models. It has also been demonstrated that a decrease in burden of bipolar symptoms results has a significant decrease in the time of symptomatic recovery. The use of standard or extended relative risk , logistic regression and multivariate analysis models may not be the most appropriate statistical tools to consider with a variable measured repeatedly over a time period.

5.2 Recommendations

The study revealed that at base line burden of bipolar symptoms is negatively associated with the hazard of symptomatic recovery. This indicates that at beginning time since burden of bipolar symptoms is high the chance of symptomatic recovery is low. Then the

psychiatrists and the concerned body should give special service(such as take care for the drug side effect) for the patients at beginning time till some symptomatic recovery achieved.

The hazard of symptomatic recovery for between twenty and twenty five age group and fifty and above age groups is greater than those nineteen and below age groups. Then for lower age bipolar disorders special treatment should be given in order to improve their symptomatic recovery.

The relative risk of symptomatic recovery for substance abuse bipolar disorders is less than those not abuse substances. Then it is better if reducing over using substances such as alcoholic drinks that affects our mental health

The hazard of symptomatic recovery for bipolar disorders those faced event of relapse before is less than those who do not faced event of relapse before study holding. So when the patients had faced their first symptomatic recovery, the family of the patients and caregivers should responsible to take away the patients from events that may cause mental problem again.

To characterize the more information about bipolar disorder this study is not sufficient. Then it is better if the further statistical models would be applied there.

The additional information of the patients history such as hypomanic type of episodes,drug side effects and severity of bipolar disorder using charting the bipolar illness should be recorded in patients card and included in the further studies[90].

5.3 Limitations of the Study

Some of the limitations of the study are:-

- The study is conducted based on secondary data which might have incomplete and biased information that the data is not well organized during collection.
- Lack of prior research studies on the topic.

References

[1] Merikangas KR, Pato M. Recent developments in the epidemiology of bipolar disorder in adults and children: Magnitude, correlates, and future directions. Clinical Psychology: Science and Practice. 2009 Jun;16(2):121-33.

[2] Ketter TA. Diagnostic features, prevalence, and impact of bipolar disorder. J Clin Psychiatry. 2010 Jun 15;71(6):e14.

[3] Merikangas KR, Pato M. Recent developments in the epidemiology of bipolar disorder in adults and children: Magnitude, correlates, and future directions. Clinical Psychology: Science and Practice. 2009 Jun;16(2):121-33.

[4] Ketter TA. Diagnostic features, prevalence, and impact of bipolar disorder. J Clin Psychiatry. 2010 Jun 15;71(6):e14.

[5] Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM, Petukhova M, Kessler RC. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. Archives of general psychiatry. 2007 May 1;64(5):543-52.

[6] Bauer M, Pfennig A. Epidemiology of bipolar disorders. Epilepsia. 2005 Jun;46:8-13.

[7] Findling RL, Çavuº I, Pappadopulos E, Vanderburg DG, Schwartz JH, Gundapaneni BK, DelBello MP. Efficacy, long-term safety, and tolerability of ziprasidone in children and adolescents with bipolar disorder. Journal of child and adolescent psychopharmacology. 2013 Oct 1;23(8):545-57.

[8] Obo CS, Sori LM, Abegaz TM, Molla BT. Risky sexual behavior and associated factors among patients with bipolar disorders in Ethiopia. BMC psychiatry. 2019 Dec 1;19(1):313.

[9] Hafeman DM, Merranko J, Goldstein TR, Axelson D, Goldstein BI, Monk K, Hickey MB, Sakolsky D, Diler R, Iyengar S, Brent DA. Assessment of a person-level risk calculator to predict new-onset bipolar spectrum disorder in youth at familial risk. JAMA psychiatry. 2017 Aug 1; 74(8):841-7.

[10] Stanley KS. Asian American mental health: Assessment theories and methods. Springer Science & Business Media; 2002 Aug 31.

[11] Ferrari AJ, Baxter AJ, Whiteford HA. A systematic review of the global distribution and availability of prevalence data for bipolar disorder. Journal of affective disorders. 2011 Nov 1;134(1-3):1-3.

[12] Ayuso-Mateos JL. Global Burden of panic disorder in the year 2000: Version 1 Estimates. GBD 2000 Working Paper. GBD 2000 Working Paper. 2002. WHO (http://www. who. int/evidence/bod), Geneva; 2012.

[13] Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abebo TA, Abera SF, Aboyans V. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990?2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017 Sep 16;390(10100):1211-59.

[14] Hadera E, Salelew E, Girma E, Dehning S, Adorjan K, Tesfaye M. Magnitude and Associated Factors of Perceived Stigma among Adults with Mental Illness in Ethiopia. Psychiatry journal. 2019;2019.

[15] Kleintjes S, Flisher AJ, Fick M, Railoun A, Lund C, Molteno C, Robertson BA. The prevalence of mental disorders among children, adolescents and adults in the Western Cape, South Africa. African Journal of Psychiatry. 2006;9(3):157-60.

[16] Hanlon C, Medhin G, Alem A, Araya M, Abdulahi A, Hughes M, Tesfaye M, Wondimagegn D, Patel V, Prince M. Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. Journal of affective disorders. 2008 Jun 1;108(3):251-62.

[17] Zergaw A. *Economic burden of schizophrenia and bipolar disorders in Ethiopia* (Doctoral dissertation, Addis Ababa University).

[18] Mekonnen E, Esayas S. Correlates of mental distress in Jimma town, Ethiopia. Ethiopian journal of health sciences. 2003;13(1).

[19] Harvey PD. Defining and achieving recovery from bipolar disorder. The Journal of clinical psychiatry. 2006;67:14-8.

[20] de Barros Pellegrinelli K, de O. Costa LF, Silval KI, Dias VV, Roso MC, Bandeira M, Colom F, Moreno RA. Efficacy of psychoeducation on symptomatic and

functional recovery in bipolar disorder. Acta psychiatrica scandinavica. 2013 Feb; 127(2):153-8.

[21] Arbeev KG, Akushevich I, Kulminski AM, Land KC, Yashin AI. Approaches to Statistical Analysis of Longitudinal Data on Aging, Health, and Longevity: Biodemographic Perspectives. InBiodemography of Aging 2016 (pp. 241-261). Springer, Dordrecht.

[22] Kalbfleisch JD, Prentice RL. The statistical analysis of failure data. Wiley; 2002.

[23] Rizopoulos D. Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule. Computational Statistics & Data Analysis. 2012 Mar 1;56(3):491-501.r

[24] Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. Journal of Clinical Oncology. 2010 Jun 1;28(16):2796.

[25] Chen LM, Ibrahim JG, Chu H. Sample size and power determination in joint modeling of longitudinal and survival data. Statistics in medicine. 2011 Aug 15;30(18):2295-309.

[26] Hogan JW, Laird NM. Increasing efficiency from censored survival data by using random effects to model longitudinal covariates. Statistical Methods in Medical Research. 1998 Feb;7(1):28-48.

[27] Wilson MS, Busick DN, inventors. Composite bipolar plate for electrochemical cells. United States patent US 6,248,467. 2001 Jun 19.

[28] Thomas DK. West African Immingrants' Attitude Toward Seeking Psychological Help.

[29] Fond G, Boyer L, Gaman A, Laouamri H, Attiba D, Richard JR, Delavest M, Houenou J, Le Corvoisier P, Charron D, Krishnamoorthy R. Treatment with anti-toxoplasmic activity (TATA) for toxoplasma positive patients with bipolar disorders or schizophrenia: a cross-sectional study. Journal of psychiatric research. 2015 Apr 1; 63:58-64.

[30] Hirschfeld RM, Dunner DL, Keitner G, Klein DN, Koran LM, Kornstein SG, Markowitz JC, Miller I, Nemeroff CB, Ninan PT, Rush AJ. Does psychosocial functioning improve independent of depressive symptoms A comparison of nefazodone,

psychotherapy, and their combination. Biological psychiatry. 2002 Jan 15;51(2):123-33.

[31] Rizopoulos D, Hatfield LA, Carlin BP, Takkenberg JJ. Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. Journal of the American Statistical Association. 2014 Oct 2;109(508):1385-97.

[32] Tsiatis AA, Degruttola V, Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. Journal of the American statistical association. 1995 Mar 1; 90(429):27-37.

[33] Benazzi F. Bipolar disorder focuses on bipolar II disorder and mixed depression. The Lancet. 2007 Mar 17; 369(9565):935-45

[34] £ojko D, Suwalska A, Rybakowski J. Bipolar and related disorders and depressive disorders in DSM-5. Psychiatr. Pol. 2014; 48(2):245-60.

[35] DeGruttola V, Tu XM. Modeling the relationship between disease progression and survival time. Biometrics. 1994; 50:1003-14.

[36] Andrinopoulou ER, Eilers PH, Takkenberg JJ, Rizopoulos D. Improved dynamic predictions from joint models of longitudinal and survival data with time varying effects using P?splines. Biometrics. 2018 Jun;74(2):685-93.

[37] Berchiche N, Franc JP, Michel JM. A cavitation erosion model for ductile materials. J. Fluids Eng.. 2002 Sep 1;124(3):601-6

[38] Piulachs, X., Alemany, R., Guillén, M. and Serrat, C., 2015. Joint modeling of health care usage and longevity uncertainty for an insurance portfolio. In *Scientific Methods for the Treatment of Uncertainty in Social Sciences* (pp. 289-297). Springer, Cham.

[39] Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time?to?event. Statistics in medicine. 2011 May 30;30(12):1366-80.

[40] Murawska M, Rizopoulos D, Lesaffre E. A two-stage joint model for nonlinear longitudinal response and a time-to-event with application in transplantation studies. Journal of Probability and Statistics. 2012;2012.

[41] Viviani S, Alfó M, Rizopoulos D. Generalized linear mixed joint model for longitudinal and survival outcomes. Statistics and Computing. 2014 May 1;24(3):417-27.

[42] Ivanova A, Molenberghs G, Verbeke G. Mixed models approaches for joint modeling of different types of responses. Journal of biopharmaceutical statistics. 2016 Jul 3;26(4):601-18.

[43] Molenberghs G, Verbeke G. Models for discrete longitudinal data. Springer Science & Business Media; 2006 Jan 28.

[44] Molassiotis A, Wilson B, Brunton L, Chaudhary H, Gattamaneni R, McBain C. Symptom experience in patients with primary brain tumours: a longitudinal exploratory study. European Journal of Oncology Nursing. 2010 Dec 1;14(5):410-6.

[45] Coran AY, Patel RP, Williams D. Rubber-thermoplastic compositions. Part V. Selecting polymers for thermoplastic vulcanizates. Rubber Chemistry and Technology. 1982 Mar;55(1):116-36.

[46] Mason KO, Parmar AN, White NE. Simultaneous X-ray and optical observations of the X-ray dip source X1755-338. Monthly Notices of the Royal Astronomical Society. 1985 Oct;216:1033-41.

[47] Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. Journal of the American statistical Association. 1993 Mar 1;88(421):9-25.

[48] Stiratelli R, Laird N, Ware JH. Random-effects models for serial observations with binary response. Biometrics. 1984 Dec 1:961-71.

[49] m Lin X, Breslow NE. Bias correction in generalized linear mixed models with multiple components of dispersion. Journal of the American Statistical Association. 1996 Sep 1;91(435):1007-16.

[50] Breslow NE, Lin X. Bias correction in generalised linear mixed models with a single component of dispersion. Biometrika. 1995 Mar 1;82(1):81-91.

[51] Bates MJ. Birger Hj?rland's manichean misconstruction of Marcia Bates' work. Journal of the American Society for Information Science and Technology. 2011 Oct;62(10):2038-44.

[52] Klein JP, Moeschberger ML. Semiparametric proportional hazards regression with fixed covariates. Survival analysis: techniques for censored and truncated data. 2003:243-93.

[53] Nelson CR, Plosser CR. Trends and random walks in macroeconmic time series: some evidence and implications. Journal of monetary economics. 1982 Jan 1;10(2):139-62.

[54] Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics. 1992 Jun 1:361-72.

[55] Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. John Wiley & Sons; 2011 Jan 25.

[56] Gill R. Understanding Cox's regression model. Experientia. Supplementum. 1982;41:187-99.

[57] Grambsch PM, Therneau TM, Fleming TR. Diagnostic plots to reveal functional form for covariates in multiplicative intensity models. Biometrics. 1995 Dec 1:1469-82.

[58] Therneau T, Lumley T. Package survival: Survival analysis, including penalised likelihood. R package version 2.3612.

[59] Thiele JC, Kurth W, Grimm V. RNetLogo: An R package for running and exploring individual?based models implemented in NetLogo. Methods in Ecology and Evolution. 2012 Jun;3(3):480-3.

[60] Rizopoulos D. Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive Gaussian quadrature rule. Computational Statistics & Data Analysis. 2012 Mar 1;56(3):491-501.

[61] Hsieh H, Boehm J, Sato C, Iwatsubo T, Tomita T, Sisodia S, Malinow R. AMPAR removal underlies Aβ -induced synaptic depression and dendritic spine loss. Neuron. 2006 Dec 7;52(5):831-43.

[62] Hens N, Aerts M, Molenberghs G, Thijs H, Verbeke G. Kernel weighted influence measures. Computational statistics & data analysis. 2005 Mar 1;48(3):467-87.

[63] Gourieroux C, Monfort A, Trognon A. Pseudo maximum likelihood methods: Theory. Econometrica: Journal of the Econometric Society. 1984 May 1:681-700.

[64] Ismail, N. and Zamani, H., 2013. Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models. In *Casualty Actuarial Society E-Forum* (Vol. 41, No. 20, pp. 1-28).

[65] Hinde J, Demétrio CG. Overdispersion: models and estimation. Computational Statistics and Data Analysis. 1998 Apr 1;27(2):151-70.

[66] Daniels MJ, Hogan JW. Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis. CRC Press; 2008 Mar 11.

[67] Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. Test. 2009 May 1;18(1):1-43.

[68] Aalen O. Nonparametric estimation of partial transition probabilities in multiple decrement models. The Annals of Statistics. 1978 May 1:534-45.

[69] Aalen O. Nonparametric inference for a family of counting processes. The Annals of Statistics. 1978 Jul 1:701-26.

[70] Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. Test. 2009 May 1;18(1):1-43.

[71] Cavanaugh JE. Unifying the derivations for the Akaike and corrected Akaike information criteria. Statistics & Probability Letters. 1997 Apr 30;33(2):201-8.

[72] Acquah HD, Carlo M. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. Journal of Development and Agricultural Economics. 2010 Jan;2(1):1-6.

[73] Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. Journal of the royal statistical society: Series b (statistical methodology). 2002 Oct;64(4):583-639.

[74] Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996 Jan;58(1):267-88.

[75] Bondell HD, Reich BJ, Wang H. Noncrossing quantile regression curve estimation. Biometrika. 2010 Dec 1;97(4):825-38.

[76] Ibrahim S. Poverty, aspirations and well-being: Afraid to aspire and unable to reach a better life?voices from Egypt. Brooks World Poverty Institute Working Paper. 2011(141).

[77] Santos Nobre J, da Motta Singer J. Residual analysis for linear mixed models. Biometrical Journal: Journal of Mathematical Methods in Biosciences. 2007 Jun;49(6):863-75.

[78] Lin DY, Wei LJ, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika. 1993 Sep 1;80(3):557-72.

[79] Cox DR, Snell EJ. A general definition of residuals. Journal of the Royal Statistical Society: Series B (Methodological). 1968 Jul;30(2):248-65.

[80] Solomon DA, Leon AC, Coryell WH, Endicott J, Li C, Fiedorowicz JG, Boyken L, Keller MB. Longitudinal course of bipolar I disorder: duration of mood episodes. Archives of general psychiatry. 2010 Apr 1;67(4):339-47.

[81] Rizopoulos D, Verbeke G, Molenberghs G. Shared parameter models under random effects misspecification. Biometrika. 2008 Mar 1;95(1):63-74.

[82] Younes N, Lachin J. Link-based models for survival data with interval and continuous time censoring. Biometrics. 1997 Dec 1:1199-211.

[83] Rizopoulos D. The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. arXiv preprint arXiv:1404.7625. 2014 Apr 30.*studies, 5:145-149*.

[84] Noh M, Wu L, Lee Y. Hierarchical likelihood methods for nonlinear and generalized linear mixed models with missing data and measurement errors in covariates. Journal of Multivariate Analysis. 2012 Aug 1;109:42-51

[85] Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. Journal of the American statistical Association. 1993 Mar 1;88(421):9-25.

[86] Skaug H, Fournier D, Nielsen A, Magnusson A, Bolker B. Generalized linear mixed models using AD model builder. R package version 0.7. 2013 Feb 20;7.

[87] Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D, Ripley MB. Package 'mass'. Cran R. 2013 Jan 8;538.

[88] Therneau TM, Lumley T. survival: Survival analysis. 2015. URL https://CRAN. R-project. org/package= survival. R package version. 2016:2-41.

[89] Buuren SV, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of statistical software. 2010:1-68.

[90] Leverich GS, Post RM. Charting the course of bipolar illness and its response to treatment. Medscape Psychiatry Mental Health. 1998;3(3).

Appendices

Appendix-I:-Some selective relevant information and graphs

Appendix-1:-Information Sheet

Introduction:- this information sheet was prepared for Jimma University Medical center,Jimma, Ethiopia. The aim of the letter was to make clear about the purpose of thesis, data collection procedures and to get permission for data collection.

Objective:- The aim of the study is to investigate the Joint association between burden of bipolar sy in Jimma Medical center using Joint longitudinal and survival models.

Data Collection Procedure:- In order to achieve the above objective, information, which is necessary for the study, will be taken from the registration log book and patients‘ registration card; if any inadequate information is countered it is checked from the file and excluded from analysis if proven to be inadequate. In order to come up with the above mentioned findings, total document of program clients enrolled during first September 2018 to first January 2020 be seen and a review of the required information from the records are made by using the checklist.

Risk:- Since the study will be conducted by taking appropriate information from medical chart, it will not inflict any harm on the patients. The name or any other identifying information will not be recorded and all information taken from the chart will be kept strictly confidential and in a safe place. The information extracted will be kept secured and the information retrieved will only be used for the study purpose.

Benefits:- the thesis has no direct benefit for those whose document/ record is included in this thesis. However, indirectly the result of this study might be used to improve awareness on the factors that triggers the recurrence of breast cancer patients. It also enables to provide scientific information about the finding to Ministry of health in Ethiopia that helps policy makers to enhance the awareness of the society about factors that increase the probability of recurrence due to breast cancer which is protected and curable if it is screened and treated in its earlier stage with appropriate treatment.

Confidentiality:- The information collected from this thesis will be kept confidential and will be stored in a file. In addition, it will not be revealed to anyone except the investigator and it will be kept in key and locked system with computer password.

Person to contact:- This thesis will be reviewed and approved by the institutional review board of college of Natural sciences, Jimma University.

Permission:- Lastly but not least, you are kindly requested to permit and forward your permission to concerned body in your organization so that I can get cooperation from the data clerks and other responsible bodies in place.

Data Extraction Form

Data extraction form, for the Joint modelling on the burden of bipolar symptoms and time to symptomatic recovery(Starting from2018 to first January 2020 .

Appendix-1:-Some Important Tables and Figures

$*$ indicates the significance of covariates at 5% significance level.

Table 5.1: Generalized linear model with linear and quadratic time effects selected for the fixed effects with 95% confidence interval of the estimated coefficients.

Coefficients	Estimated coff(β)	95% CIs	pvalue
(Intercept)	2.430	[2.141,2.702]	0.000
timeinmonth	-0.197	[-0.216,-0.178]	0.000
I($timeinmonth^2$)	-0.005	[-0.007,-0.002]	0.000
(religion)2	-0.085	[-0.141,-0.029]	0.003
(religion)3	0.152	[0.08,0.225]	0.000
(religion)4	-0.072	[-0.185,0.039]	0.578
(age)2	-0.061	[-0.173,0.052]	0.218
(age)3	-0.079	[-0.203,0.046]	0.255
(age)4	-0.768	[-0.978,-0.562]	0.000
(eventofrelapse)1	0.195	[0.119,0.271]	0.000
(othercofactors)1	-0.020	[-0.089,0.048]	0.520
(substanceabuse)1	0.445	[0.293,0.598]	0.000
(bipolartype)1	0.061	[0.006, 0.115]	0.094
(chewingchat)1	-0.343	[-0.827,0.109]	0.034
(typeofepisodes)2	-0.109	[-0.159,-0.06]	0.000
(typeofepisodes)3	0.012	[-0.044 ,0.068]	0.438
(fristonset)2	0.024	[-0.058,0.108]	0.782
(familyhistoryofmentalillness)1	0.110	[0.039,0.181]	0.005
(sex)female	-0.169	[-0.413,0.095]	0.119
(sex)male	-0.131	[-0.375,0.131]	0.000
timeinmonth: (fristonset)2	0.028	[0.011,0.046]	0.147
timeinmonth: (fristonset)3	0.007	[-0.011,0.026]	0.15
timeinmonth:(eventofrelapse)1	0.038	[0.022,0.054]	0.147
$eventofrelapse1 : familyhistoryofmentalillness_1$	-0.205	[-0.288,-0.121]	0.000
timeinmonth:othercofactors1	0.027	[0.011,0.043]	0.000
(chewingkhat)1:as.factor(sex)male	0.353	[-0.102,0.839]	0.000
(age)2:(substanceabuse)1	-0.477	[-0.647,-0.307]	0.000
(age)3:(substanceabuse)1	-0.272	[-0.430,-0.113]	0.000
(age)4:(substance abuse)1	–	–	
	AIC= 8724.647	$\varphi = 1$	

Table 5.2: : Log rank tests of survival curve differences for categorical covariates

Covariates	chi-square	DF	$Pr > chi - square$
Sex	2	1	0.2
First onset	1.4	2	0.5
Educational level	1.1	2	0.6
Marital status	7	3	0.07
Family history of mental illness	4.9	1	0.03
Religion	17.1	3	7e-04
Employment	0.5	1	0.5
Age	21.2	3	1e-04
Event of relapse	13.9	1	0.036
Bipolar type	1.3	1	0.3
Other cofactors	9.9	1	0.002
Type of episodes	7.8	2	0.02
Substance abuse	10.6	1	0.001
Chewing chat	2.5	1	0.1
Ethnicity	4.1	2	0.1

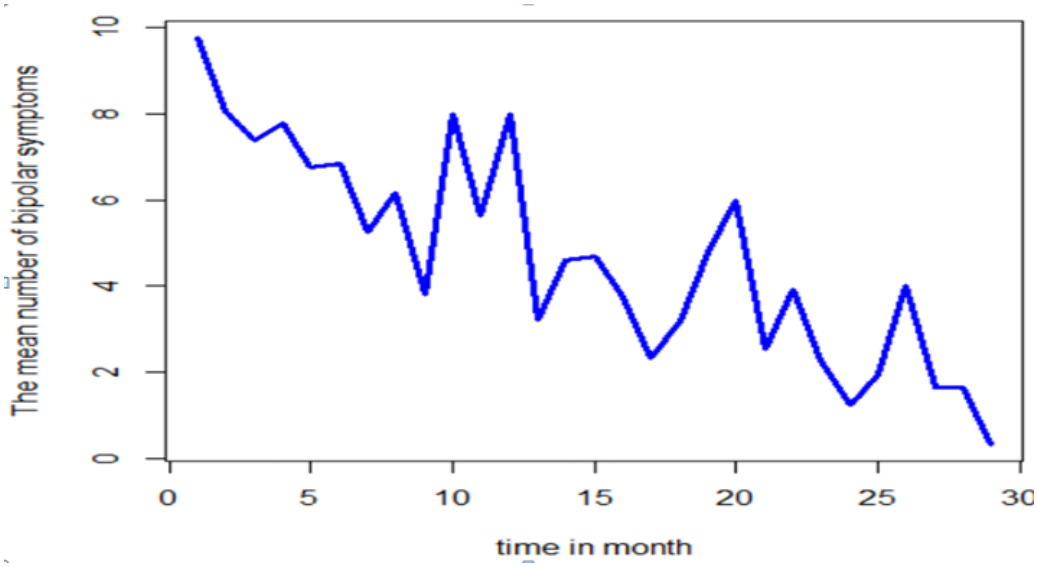

Figure 5.1: Mean burden of bipolar symptoms over time

.

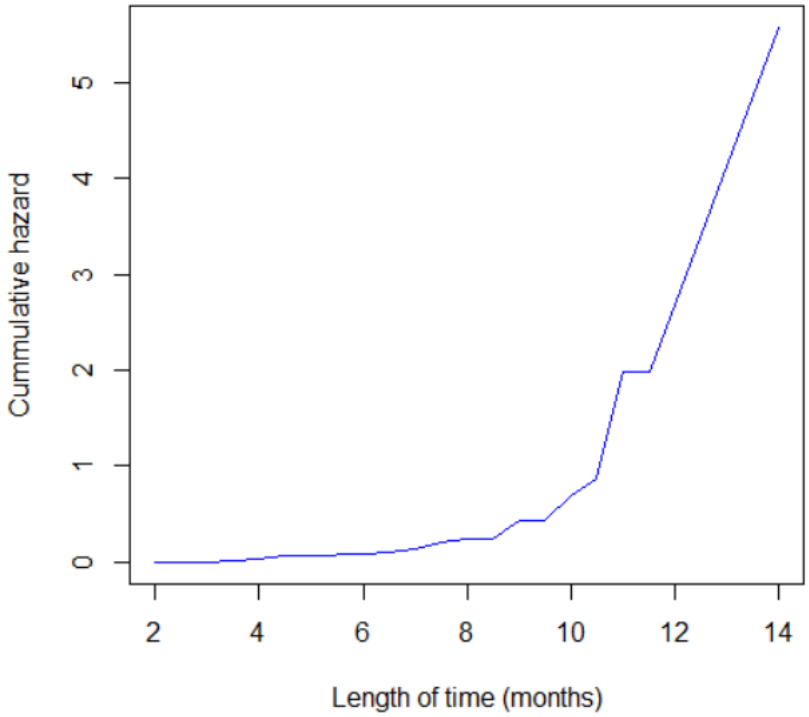

Figure 5.2: cumulative hazard ratio plot recovery of individuals

.

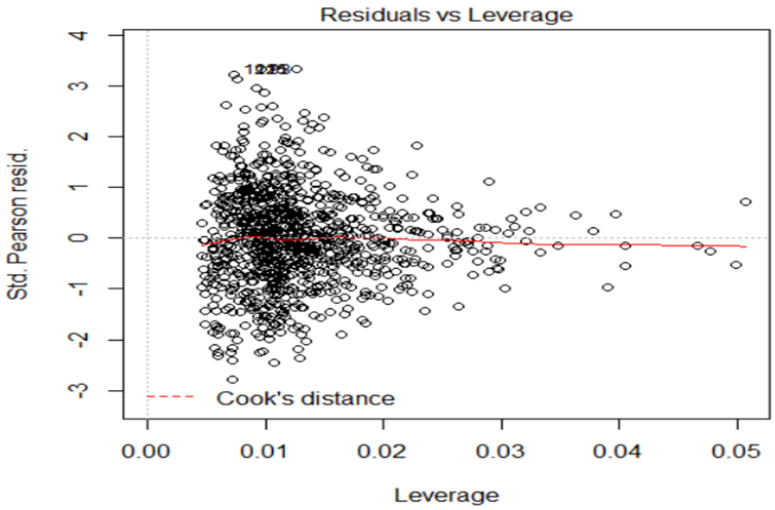

Figure 5.3: plot of leverage for GLM

.

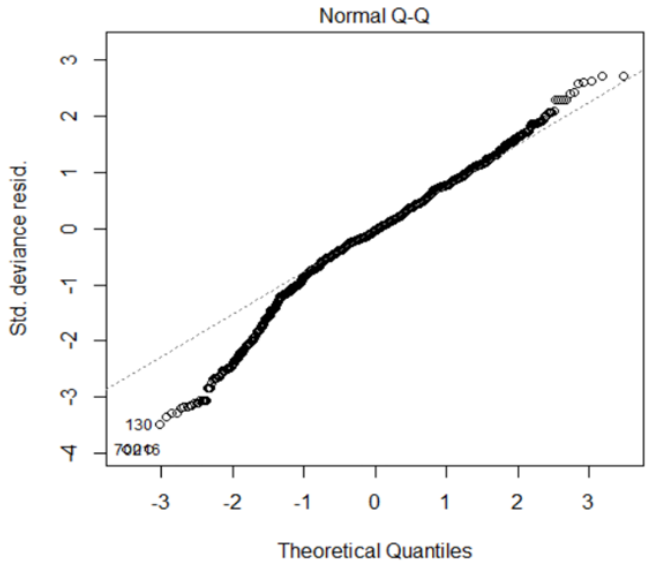

Figure 5.4: Q-Q plot for residuals of GLM model

.

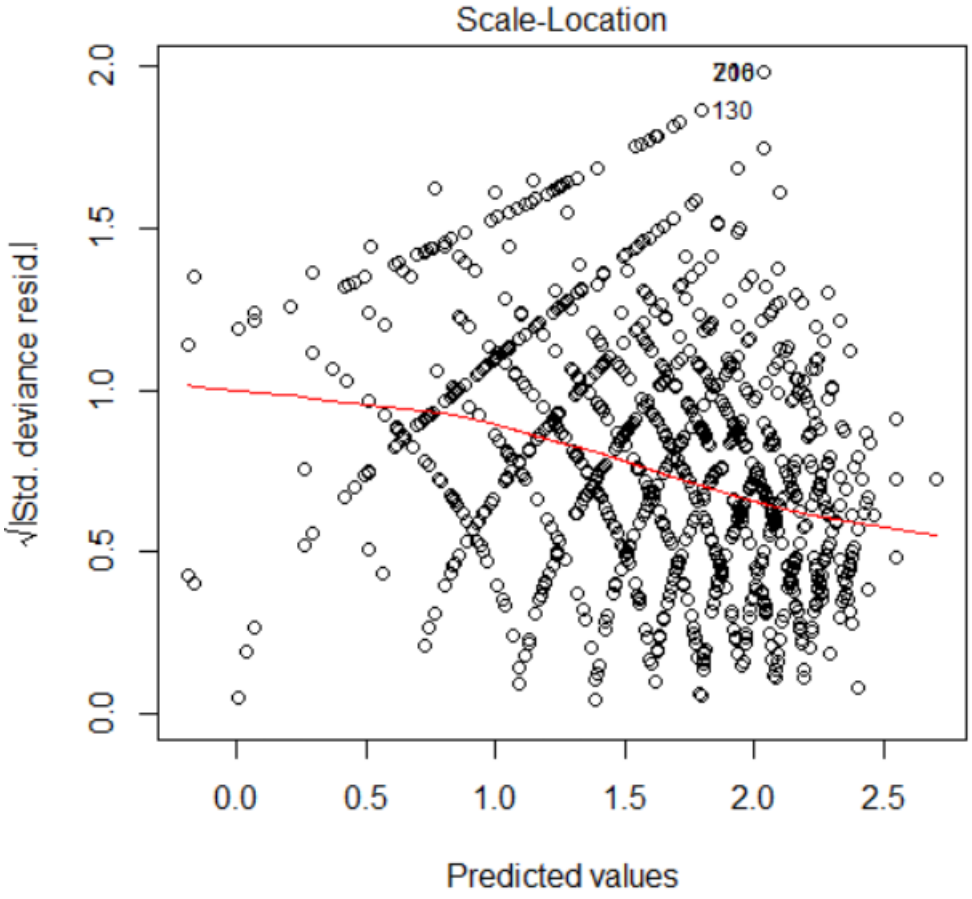

Figure 5.5: plot of deviance with fitted value to check over dispersion

.

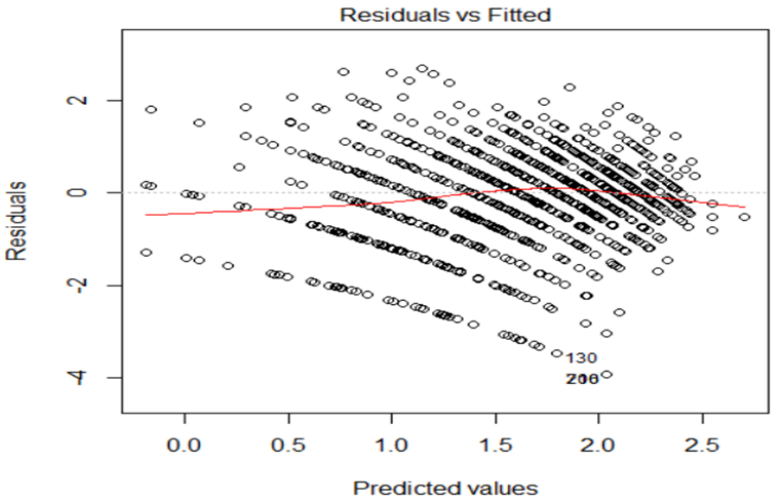

Figure 5.6: plot of Martingale residuals to check linearity

.

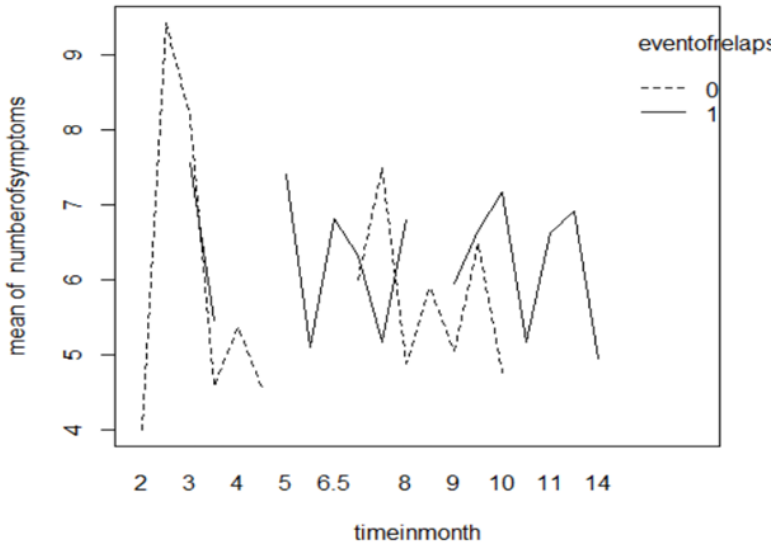

Figure 5.7: plot of interraction of mean burden of symptoms and time

.

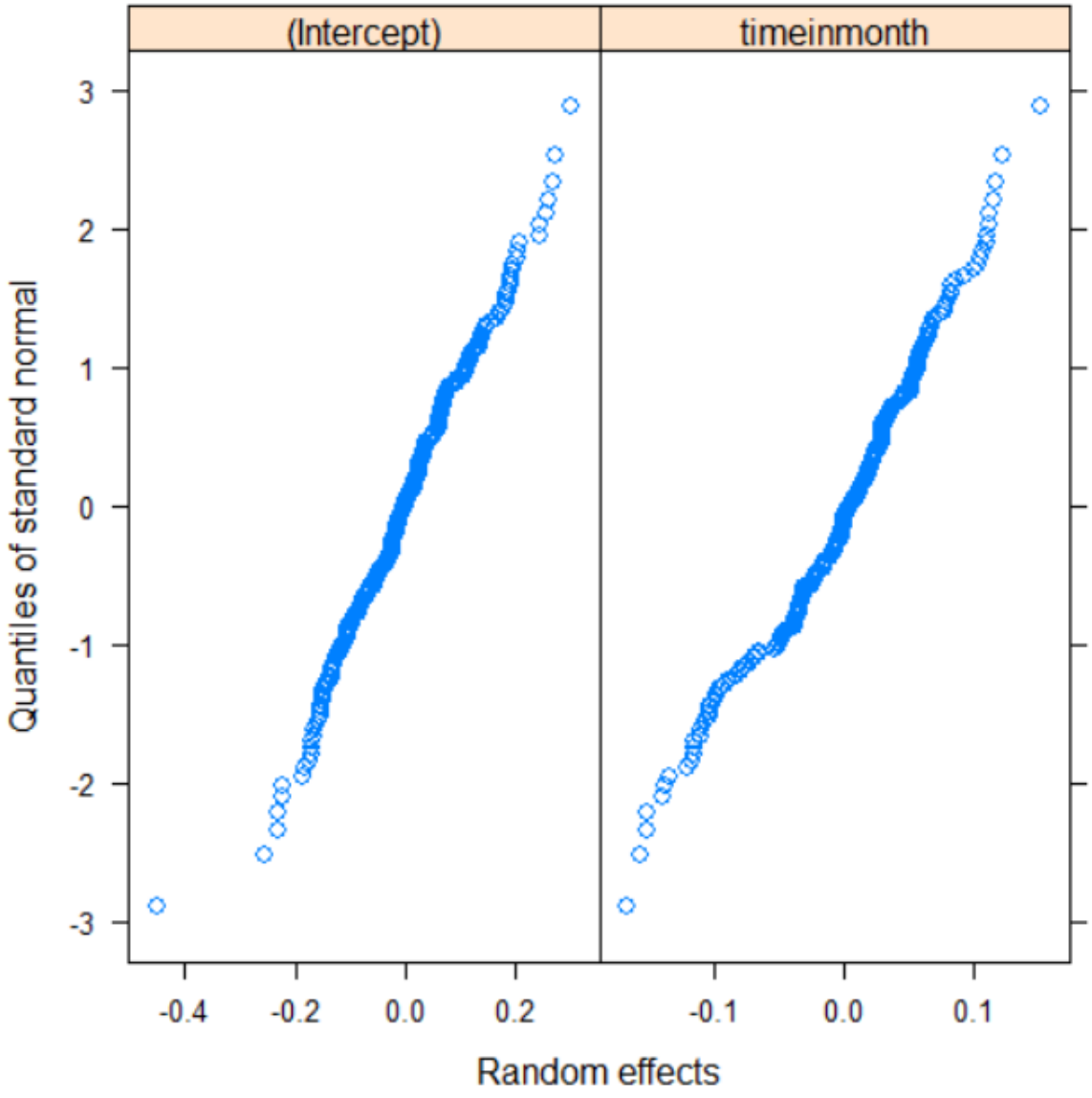

Figure 5.8: Q-Q plot to check normality of random effects In GLMM

.

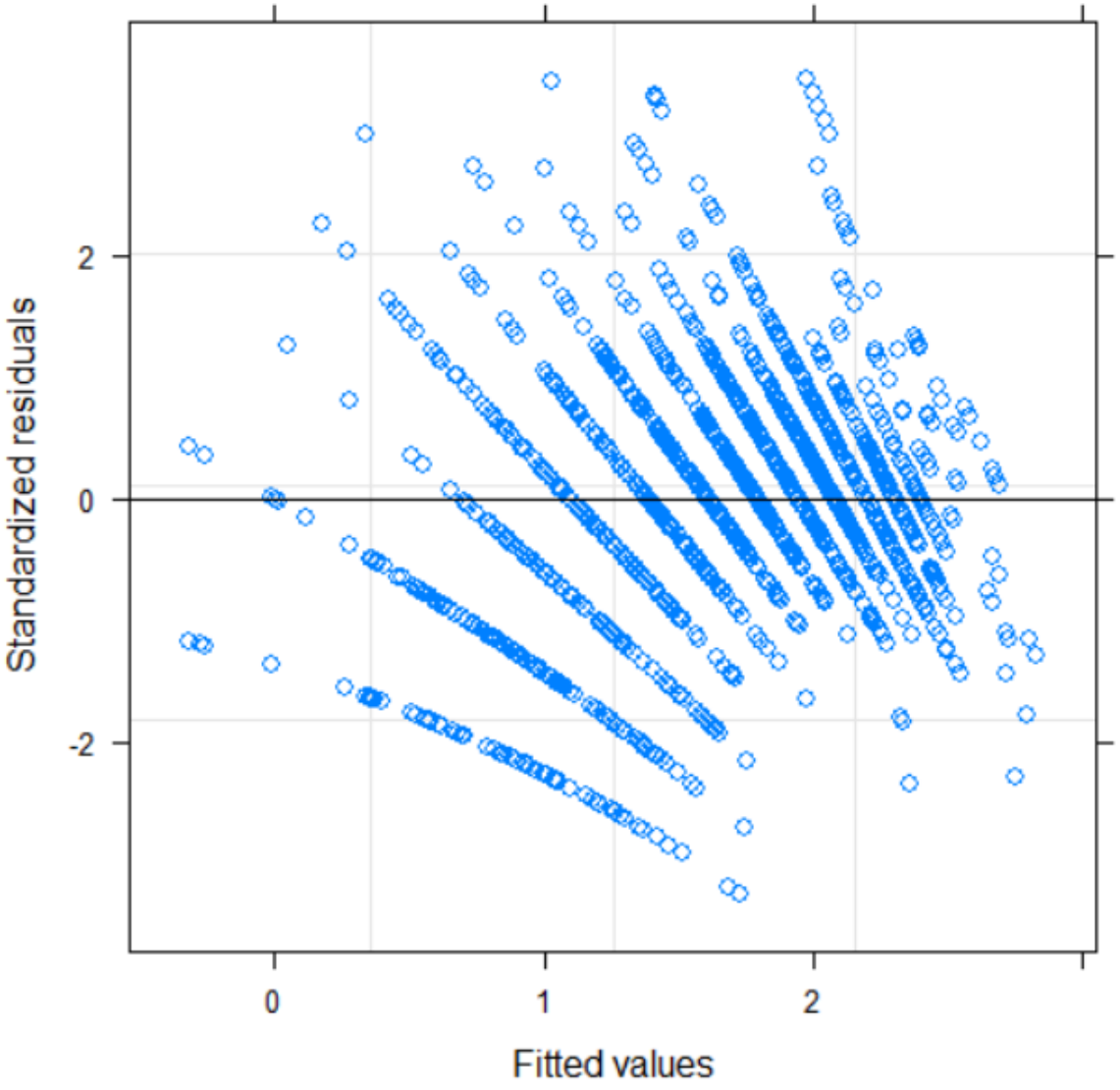

Figure 5.9: plot of standard residuals with fitted values of GLMM

.

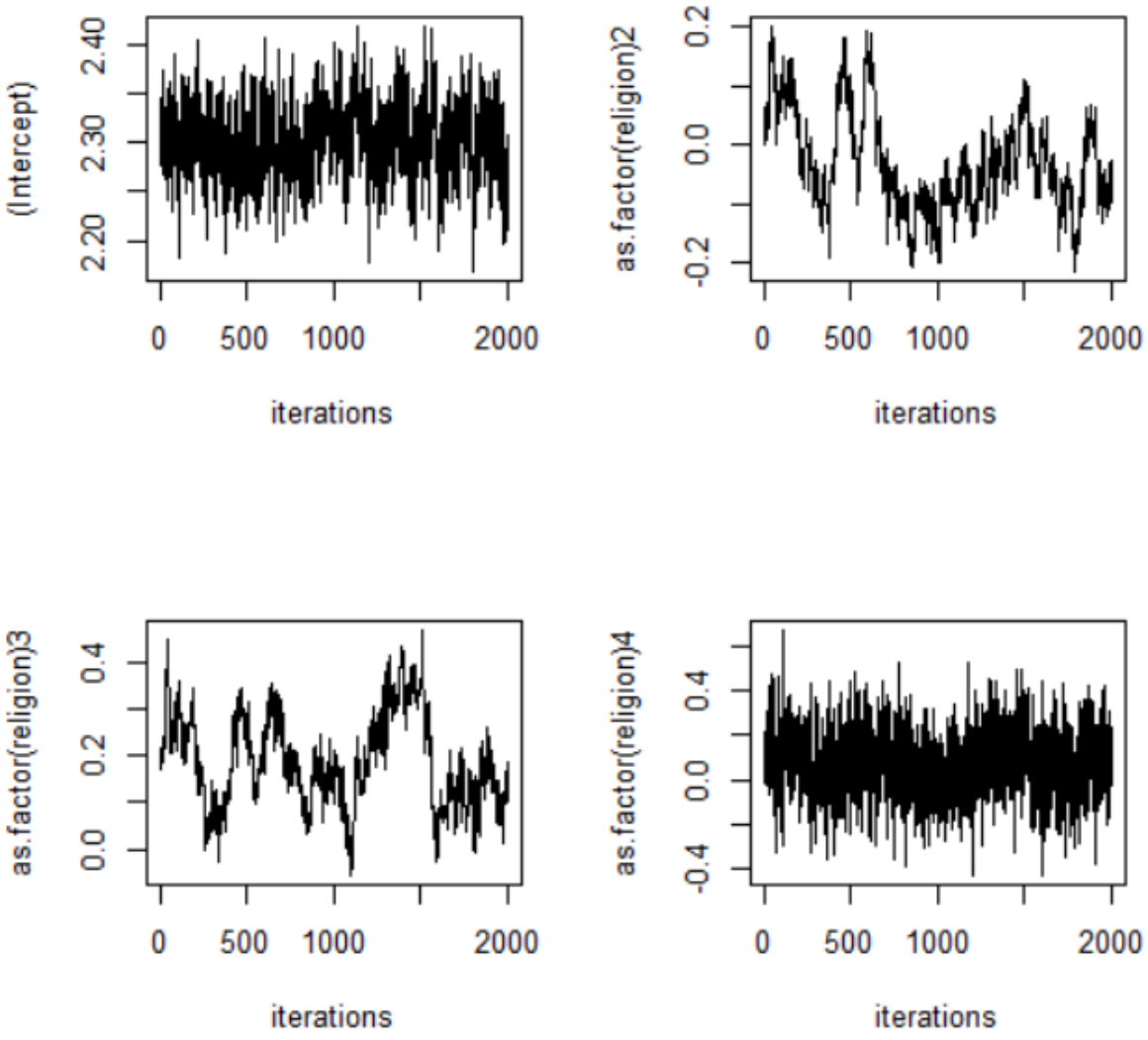

Figure 5.10: trace plot of time series iteration of gibbs sampling to check convergence

.

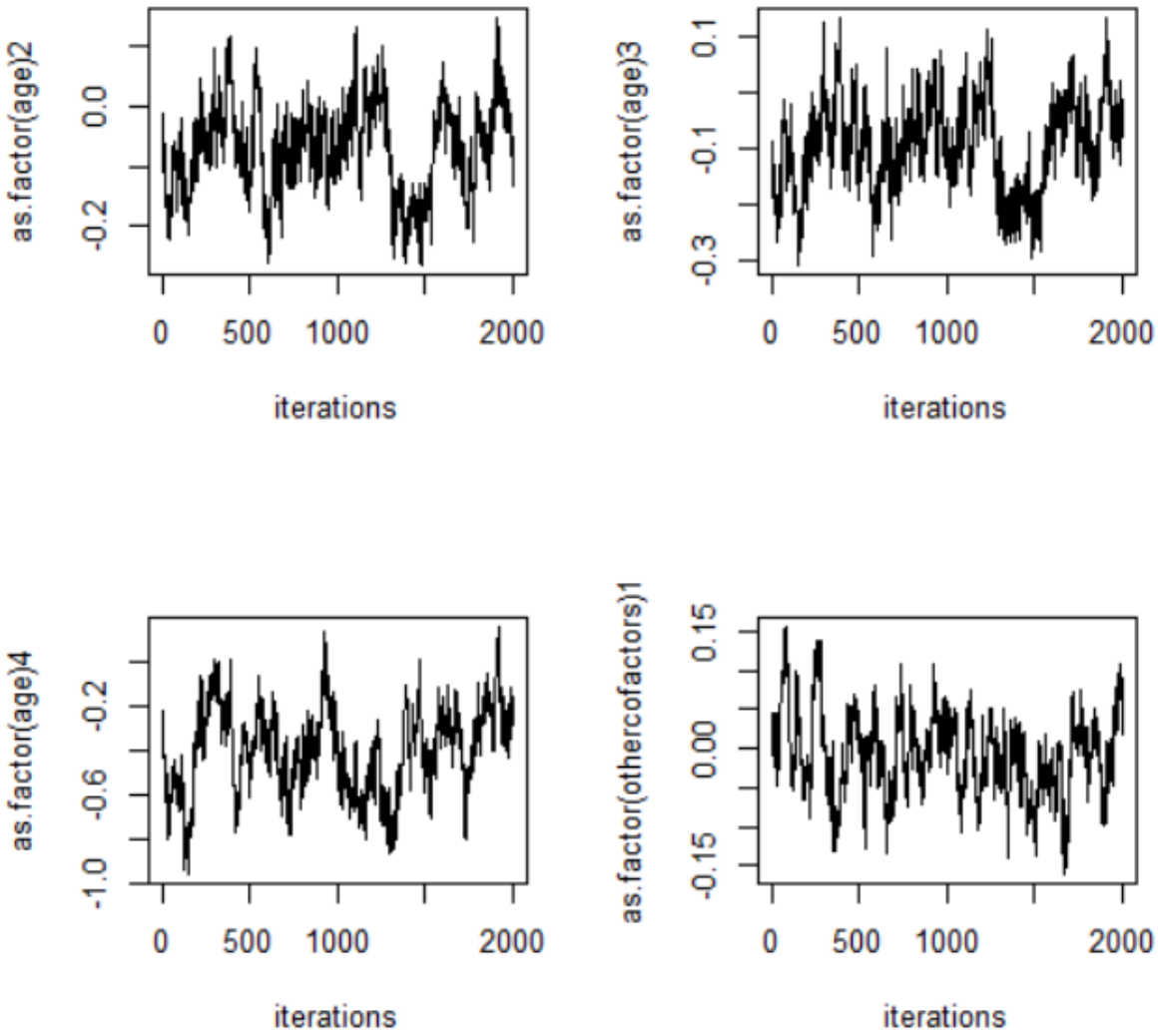

Figure 5.11: trace plot of time series iteration of gibbs sampling to check convergence

.

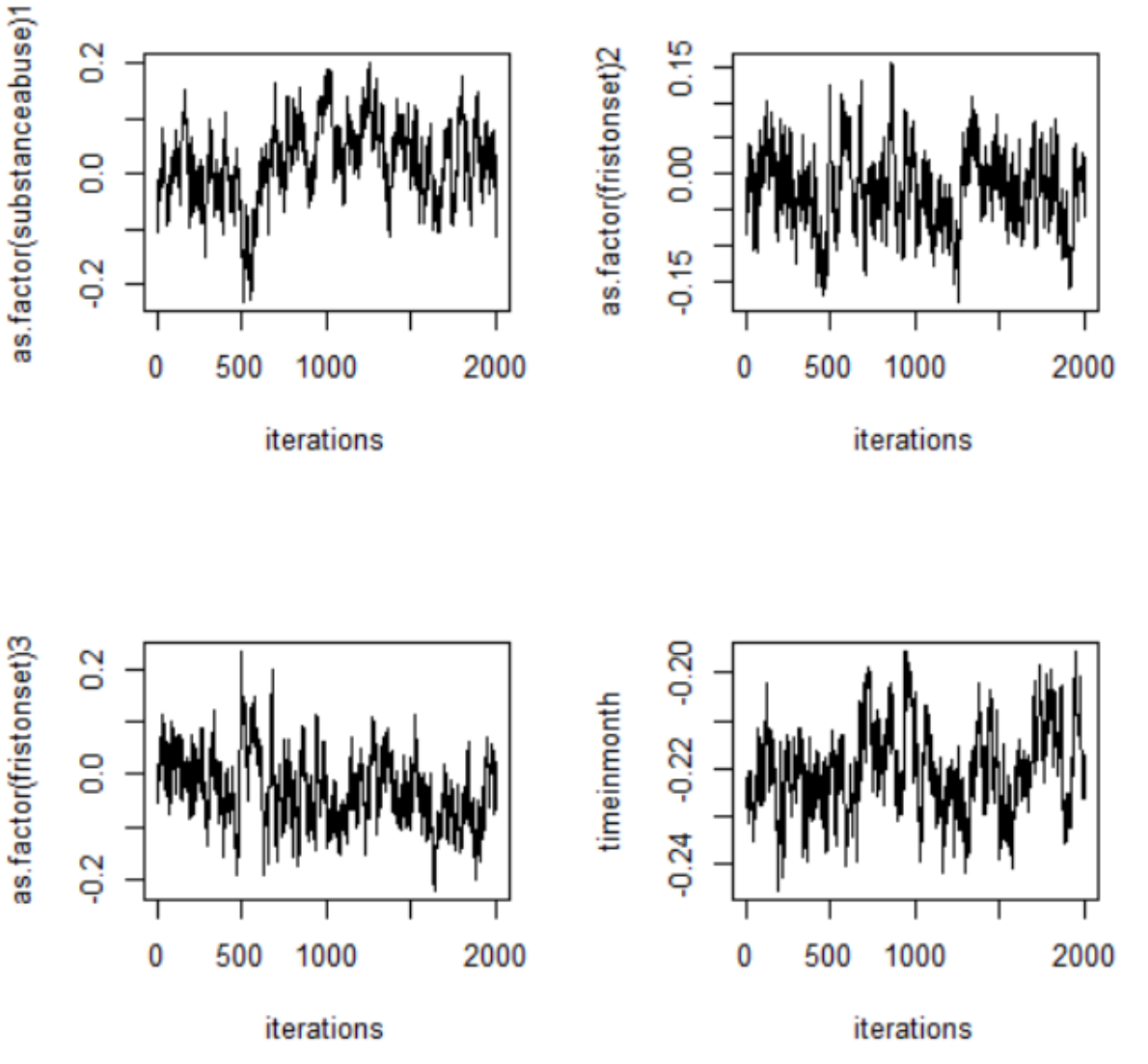

Figure 5.12: trace plot of time series iteration of gibbs sampling to check convergence

.

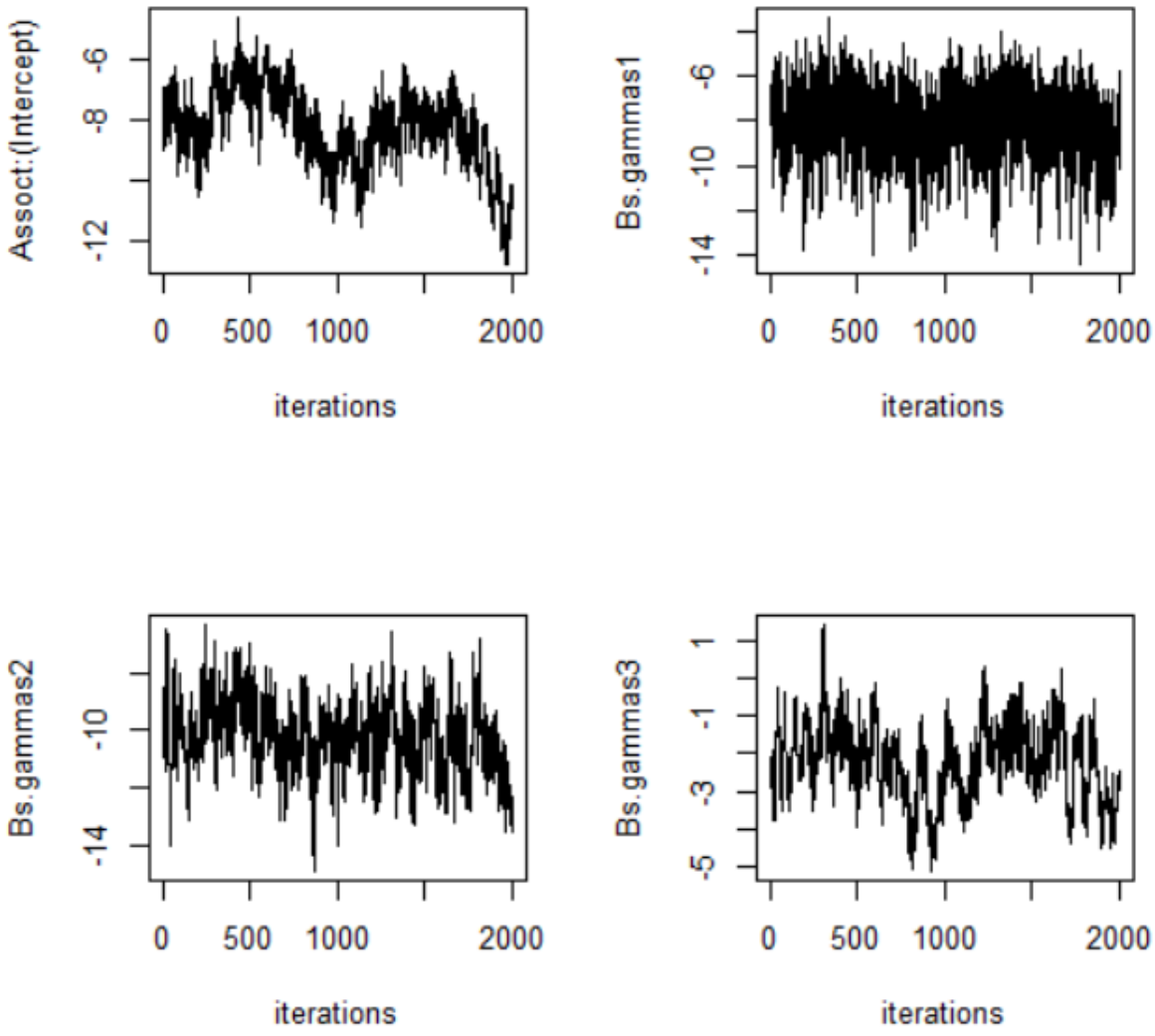

Figure 5.13: trace plot of time series iteration of gibbs sampling to check convergence

.