College of Natural Sciences

Department of Statistics

Modeling Time-to-First Symptomatic Recovery of Major Depressive Disordered Patients; A Case Study at Jimma University Medical Center

By

Ketema Zerihun

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

July, 2021

Jimma, Ethiopia

Modeling Time-to-First Symptomatic Recovery of Major Depressive Disordered Patients; A Case Study at Jimma University Medical Center

MSc. Thesis

By

Ketema Zerihun

Main Advisor: Geremew Muleta (Ass.Prof, PhD Scholar)

Co-Advisor: Gurmessa Nugussu (MSc.)

July, 2021

Jimma, Ethiopia

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.

Ketema Zerihun Asefa

Signature: _____

Date: _____

Jimma University, Jimma, Ethiopia

Approval Sheet

As thesis research advisors, we here by certify that we have read the thesis prepared by Ketema Zerihun under our guidance, which is entitled **"Modeling Time-to-First Symptomatic Recovery of Major Depressive Disordered Patients: A Case Study at Jimma University Medical Center, Jimma, Ethiopia"**, in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and ful-fill university and department style requirements; (2) its illustrative materials including tables

Geremew Muleta (Ass.Prof, PhD Scholar)		12/08/2021
Name of Major Advisor	Signature	Date
Gurmessa Nugussu (MSc.)		12/08/2021
Name of Co-Advisor	Signature	Date
Abiy Disassa (MSc.)		12/08/2021
Name of the Internal Examiner	Signature	Date
Shibru Temesgen (PhD)		12 August 2021
Name of the External Examiner	Signature	Date
Jaleta Abdisa (MSc.)		12/08/2021
Name of the Chairperson	Signature	Date

1

Acknowledgement

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 sincere appreciation to my advisor Mr.Geremew Muleta (PhD Scholar, Ass.prof) who has the substance of a genius: for his kind support, advice,guidence, constructive comments and encouraged me to be professional and do the right thing on proposal development up to the final work of this thesis and will always be grateful for that.

Likewise, I am sincerely grateful to my Co-advisor Mr.Gurmesa Nugussu (MSc.) for his continuous advice, guidance and kind support. It is a great pleasure to express his commitment to share his experience, knowledge and his smooth contacts is also appreciable.

I would like to pass my heartfelt gratitude and pleasure to my friend Mr.Tadele Degefa who gave me moral appreciation and fruit full guidance and comment throughout my career.

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 Madda Walabu 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 Zerihun Asefa, my mother Bogalech Aschalow and to all my brothers, my sisters and all my best friends for their unconditional love, pray and supports.

Abstract

Background: Major Depressive Disorder is one of the most prevalent mental disorders and the leading cause of disability worldwide with lifetime prevalence estimates ranging from 7% to 21%. Symptoms of major depressive disorder are depressed mood, loss of interest, decreased energy, sadness, feelings of guilt or low selfworth, disturbed sleep or appetite, etc. Symptomatic recovery is a dimensional measure that refers to improvement in the magnitude of symptoms.

Objective This study aimed to model time to first symptomatic recovery of major depressive disorder in Jimma University Medical Center.

Methods: The data for this study was major depressive disorder patients under follow up at Jimma University Medical Center from Semptember 1, 2018 through August 31, 2020. Weibull, Log-logistic and Lognormal as baseline hazard functions with the Gamma and Inverse Gaussian frailty distributions were used. To select best model Akaike Information Criteria was used. Data analysis is done using R statistical software.

Results: The median first symptomatic recovery time of patients was 7 months of which about 54.1% were experienced first symptomatic recovery from major depressive disorder. The clustering effect is significant on modeling time to first symptomatic recovery from major depressive disorder. According to the result from lognormal inverse-gaussian frailty model, marital status, khat chewing, educational level, employment status, substance abuse and other cofactors were the significant factors at 5% level of significance.

Conclusion: Lognormal-inverse- Gaussian frailty model is the model that best describes time-to-first symptomatic recovery of the major depressive disorder dataset. Being educated and employed significantly shortens the time-to-first symptomatic recovery from major depressive disorder while being divorced, khat chewers, substance abuse and with other cofactors prolongs the time- to-first recovery from major depressive disorder. For those groups whose recovery time was prolonged, health professionals (physicians) should give good treatment for patients for identified stakeholders on identified risk factors.

Keywords:- Depressive Disorder, Heterogeneity, Shared Frailty Model, Symptomatic Recovery

List of Acronyms/Abbreviations

AFT:	Acceleration Failure Time
AIC:	Akaike Information Criterion
BIC:	Bayessian Information Criterion
CDC:	Centers for Disease Control and Preventation
DALYs:	Diability Adjusted Life Years
DSM:	Diagnostic and Statistical manual of mental disorders
ICD:	International Classification of Disease
JUMC:	Jimma University Medical Center
LRT:	Likelihood ratio test
MDD:	Major Depressive Disordered
MDE:	Major Depressive Episode
MRI:	Magnetic Resonance Imaging
NIMH:	The National Institute For Mental Health
OR:	Odds Ratio
PH:	Proportional Hazards
PL:	Partial Likelihood
US:	United States
WHO:	World Health Organization

Definition of Symbols

CI: Confidence Interval

Coef: Coefficients of the Model

SE: Standard Error

ϕ: Accelaration Factor

θ: Variance of Random Effect

τ: Kendall's Tau

μ: Scale

σ: Shape

*: Statistically Significant at 5% Level

Ref: Reference

Contents

Acknowledgement **iv**

Abstract **v**

List of Acronyms/Abbreviations **vi**

Definition of Symbols **vii**

1 Introduction **1**

1.1 Background of the Study 1

1.2 Statements of the Problem 5

1.3 Objectives of the Study 7

1.3.1 General Objective 7

1.3.2 Specific Objectives 7

1.4 Significance of the study 8

2 Literature Review **9**

2.1 Overview of Major Depressive Disorder 9

2.1.1 Recovery of Major Depressive Disorder 10

2.1.2 Impacts of Major Depressive Disorder 11

2.1.3 Risk Factors for Recovery of Major Depressive Disorder 11

2.2 Over view of Unobserved Heterogeneity of Frailty 16

2.3 Consequence of Ignoring Frailties 18

3 Methodology **19**

3.1 Study Area 19

3.2 Data Description 19

3.3 Study Design 20

3.4 Inclusion and Exclusion Criteria 20
3.5 Study Variables 20
3.5.1 Dependent Variable 20
3.5.2 Independent Variable 21
3.6 Statistical Methods 22
3.6.1 Survival Analysis 22
3.6.2 Descriptive Methods For Survival Data 23
3.6.3 Estimation of Survivorship Function 24
3.6.4 Median Survival Time 25
3.7 Frailty Models 26
3.7.1 Shared Frailty Model 26
3.7.2 The Frailty Distributions 28
3.7.2.1 The Gamma Frailty Distribution 29
3.7.2.2 Inverse Gaussian Frailty Distribution 30
3.7.3 Baseline Survivor and Hazard Function 31
3.7.4 Parameterization 32
3.8 Method of Parameter Estimation 33
3.9 Comparison of Models 34
3.10 Model Diagnostics 35
3.10.1 Evaluation of the Baseline Parameters 35
3.10.2 The Cox- Snell Residuals 35
3.10.3 Checking for Proportional Hazard Assumption 36
3.10.4 The Quantile - Quantile Plot 36

4 Results and Discussion **38**

4.1 Descriptive of Socio-economic and Health Related Variables 38
4.2 Survival of Significantly Different Groups 42
4.3 Univariable Analysis 43

4.4 Multivariable Analysis and Model Comparison 44
4.5 Model Diagnostics 49
4.5.1 Diagnostic plots of parametric baselines 49
4.5.2 The Cox Snell residual plots 50
4.5.3 Adequacy of accelerated failure time 50
4.6 Discussion 51
4.7 Limitations of the study 55

5 Conclusion and Recommendation **56**
5.1 Conclusion 56
5.2 Recommendation 56

References **58**

APPENDIXES **71**
Annex 1: Univariable analysis using parametric shared frailty models for MDD dataset . 71
Annex 2: Multivariable Analysis Using Parametric Shared Frailty Models For MDD Dataset 74
Annex 3: The log rank test for Survival curve for significantly different groups 79
Annex 4: Descriptive statistics for woredas of patients' 80

List of Tables

3.1 Description of Independent Variables will be used in the Analysis 21

3.2 Parametric Distributions For The Baseline Hazards 32

3.3 The Cox–Snell residuals for the baseline hazard functions considered in this study . 36

4.1 Descriptive summary of covariates with the recovery status from MDD patients. . . 40

4.2 AIC, BIC and LRT values of the models used in the study 45

4.3 Lognormal-Inverse Gaussian Multivariable Analysis 48

List of Figures

4.1 The survival functions of the significantly different groups 42
4.2 Graphical Evaluation of the weibull, loglogistic and lognormal assumptions 49
4.3 Cox-Snell residuals obtained by fitting lognormal to the MDD dataset. 50
4.4 Q-Q plots to check the adequacy of accelerated failure time model 51

1 Introduction

1.1 Background of the Study

Major Depressive Disorder is a common, often chronic and recurrent condition, marked by persistent suffering and poor overall health and with deleterious effects on psychosocial, academic, vocational and family functioning. Major Depressive Disorder is one of the most prevalent mental disorders and the leading cause of disability worldwide [1], with lifetime prevalence estimates ranging from 7% to 21% [2].

Major Depression disorder, also known as unipolar or clinical depressive disorder, is a mental illness problem that affects mood. Symptoms of major depressive disorder are depressed mood, loss of interest and enjoyment, decreased energy, sadness, anxiety, agitation or restlessness, feelings of guilt or low selfworth, disturbed sleep or appetite, etc. Major Depressive Disorder is a major cause of disability and contributor to the global burden of disease affecting more than 300 million people worldwide [3]. Despite its prevalence and disability, its neurobiological mechanisms remain incompletely understood. Magnetic resonance imaging (MRI) research has a significantly advanced understanding of brain changes associated with depression [4].

A major depressive episode is characterized by the presence of a severely depressed mood that persists for at least two weeks [5]. Episodes may be isolated or recurrent and are categorized as mild (few symptoms in excess of minimum criteria), moderate, or severe (marked impact on social or occupational functioning). An episode with psychotic features commonly referred to as psychotic depression - is automatically rated as severe. [6]. If the patient has had an episode of mania or markedly elevated mood, a diagnosis of bipolar disorder is made instead. Depression without mania is sometimes referred to as unipolar because the mood remains at one emotional state or "pole" [6].

Major Depressive Disorder has an extreme global economic burden and has been listed as the third largest cause of disease burden by the WHO since 2008, and is expected to rank the first by 2030 [7]. It is diagnosed when an individual has a persistently low or depressed mood, anhedonia or decreased interest in pleasurable activities, feelings of guilt or worthlessness, lack of energy, poor concentration, appetite changes, psychomotor retardation or agitation, sleep disturbances, or suicidal thoughts [4].

Globally depression is the second leading cause of disability, with slightly more than 4% of the world's population diagnosed with it. More than 5% of the population suffers from depression in the Middle East, North Africa, Sub-Saharan Africa, Eastern Europe and the Caribbean. The most depressed country is Afghanistan, where more than 1 in 5 people suffer from the disorder. The least depressed is Japan, with a diagnosed rate of less than 2.5% [8]. The report on Global Burden of Disease estimates the point prevalence of unipolar depressive episodes to be 1.9% for men and 3.2% for women, and the one-year prevalence has been estimated to be 5.8% for men and 9.5% for women. It is estimated that by the year 2020 if current trends for demographic and epidemiological transition continue, the burden of depression will increase to 5.7% of the total burden of disease and it would be the second leading cause of disability-adjusted life years (DALYs), second only to ischemic heart disease [9].

In Africa, there are different prevalence rates of depression. A large epidemiological study was conducted between 2002 and 2004 in a total sample of 4351 adults of South African by using the WHO composite International Diagnostic Interview and found that the prevalence of depression was 9.7% for life time and 4.9% for the 12 months prior to the interview [10]. But the other comparative study that was conducted by Amoran *et al.* [11], found that the prevalence of depression among adults in Nigeria was 5.2% which is lower than the prevalence of depression in South Africa that was conducted by Tomlinson *et al.* [10].

The prevalence of depression in Ethiopia was reported to be 5% according to the Ethiopian Federal Ministry of Health Report of 2012, and WHO survey in collaboration with Jimma University shows that the prevalence of depression in Ethiopia was 9.1% [12]. National survey of 2014 states the pooled prevalence of depression from Eight studies in Ethiopia was 11% [13]. A cross-sectional community based study in Jimma town conducted by Ermias and Samuel in 2002; 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 [14].

Major depressive disorder has been shown to impose a substantial economic burden on all levels of society. The economic burden of major depressive disorder was estimated at $210.5 billion in 2010 in the US [15]. While approximately half of this amount was due to direct medical costs, the other half was attributable to indirect costs related to absenteeism, presentism, and suicide, further underscoring the toll that MDD imposes on a patient's life. Individuals with depression lose 5.6 hours of productive time at work per week compared to 1.6 hours in non-depressed workers [16], which results in 225 million lost workdays and $36.6 billion of salary-equivalent lost productivity per year associated with depression [17]. Besides a serious medical condition, MDD can impose a heavy financial burden both on families and the society, including significant resource demands on a country's health system [15]. According to data from the WHO, approximately 121 million people worldwide are estimated to suffer from depression. This type of chronic mental disorder causes severe damage to patients, their families, and also the country.

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 [18]. 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 [19].

In considering the different factors for the first symptomatic recovery of MDD, the first symptomatic recovery time of the disease can be predicted and statistically estimated with the survival analysis. Various techniques in survival analysis, which is the statistical tool used to analyze time to first symptomatic recovery data considered in this study. Survival data is a term used for describing data that measure the time to a given event of interest [20]. In this study, the event of interest was the time to first symptomatic recovery of MDD after treatment. Proportional hazard model popularized by Cox [21], is the classical model for this kind of data. However, the correct inference based on Cox's models needs identically and independently distributed samples.

Often, subjects may be exposed to different risk levels, even after controlling for known risk factors. This is because the covariates that are relevant to the researcher are often unavailable or even unknown. Moreover, the frailty model, introduced in the statistical literature by Vaupel *et al.* [22] and discussed in details by Hougaard *et al.* [23, 24], accounts for heterogeneity in baseline. This concept is an extension of the Cox's PH model in which the hazard function depends upon an unobservable random quantity, the so-called frailty that acts multiplicatively on the hazard function.

Furthermore, study population for this study was clustered and hence clustered patients survival data may be correlated at the cluster (woreda) level. In current study, shared frailty models explored assuming that patients with in the same cluster (woreda) shares similar risk factors, which would take care of the frailty term at woreda level. This model is a conditional independence model, where the frailty is common to all individuals in a cluster, hence responsible for creating dependence between event times. This is because ignoring the full dependence among observations might lead to understated standard errors and also parameter estimates that are both biased and inconsistent [25].

Typically, the estimation of the frailty model can be parametric or semi-parametric. In the former case, parametric density is assumed for the event times, resulting in a parametric baseline hazard function. The choice of a parametric baseline hazard means that the marginal likelihood is fully parametric so that we can rely on classical maximum likelihood techniques to estimate the parameters. In the latter case, the baseline hazard is left unspecified and more complex techniques are available to approach that situation Abrahantes *et al.* [26]. Even though semi-parametric estimation offers more flexibility, the parametric estimation will be more powerful if the form of the baseline hazard is to be somehow known in advance [27].

This thesis considered parametric frailty models to investigate the relationship between different potential covariates and time to first symptomatic recovery of MDD for clustered survival data with random right censoring. The choice of distribution for the hazard is very important than the choice of frailty distribution [28]. The advantage of parametric method over the semiparametric method shows that having distribution may calculate the quantiles, simplicity and completeness are reasons for the popularity of parametric distributions [20]. Hence, in this study weibull, log-logistic and lognormal baseline hazard functions used. On the other hand, among frailty distribution we have assumed gamma and inverse Gaussian distributions to fit MDD data set. Gamma and Inverse Gaussian are the two most common choices of frailty distributions due to their mathematical tractability. For comparison of different models the AIC criteria used.

1.2 Statements of the Problem

Mental illness is a public health problem in developed as well as developing countries [29]. Mental disorders contribute to 13% of the global burden of disease, with major depression expected to be the largest contributor to this by 2030 [30]. Major Depressive Disorder is a common condition in developed countries and is a growing problem in developing countries [31]. Major depression is the most severe problem that occurs all over the world and this disorder leads to other problems

and affect an individual's life. An individual may be affected by this problem in his /her life time. It affect people's ability to participate in health-promoting behaviors, thoughts, feelings and sense of well-being. In turn, problems with physical health, such as chronic diseases, can have a serious impact on mental health and decrease a person's ability to participate in treatment and recovery [32]. In Ethiopia, mental health problems accounts for 12.45% of the burden of diseases, 12% of the people are suffering from mental health problems of which, 2% are severe cases [33]. At this time, there is no cure for major depressive disorder; however treatment can significantly decrease the associated morbidity and mortality.

Some studies have been conducted to identify covariates of major depresive disorder recovery by using logistic regression [34, 35, 36] and Cox proportional hazard models [37, 38, 39]. But Logistic regression does not account the censoring observations, that is, it does not hold for time-to-event data. Therefore, there is a restriction on the events when we use logistic models for the follow up time which loss massive information [40]. Similarly, correct inference based on Cox's models needs identically and independently distributed samples. Also in demographic applications nonparametric and semiparametric models are often used to model time to event data. In such applications, it is assumed that all heterogeneity is captured by theoretically relevant covariates [41].

Cox proportional hazards model didn't take into account any extra heterogeneity present in the data. Ignoring this heterogeneity will produce biased parameter estimates and inconsistent standard errors in survival analysis [42]. Frailty models are the survival data analog to regression models, which account for heterogeneity and random effects. A shared frailty model is a random-effects model where the frailties are common (or shared) among groups of individuals or spells and are randomly distributed across groups. Therefore, we have employed a shared frailty model to investigate the factors associated with time to first symptomatic recovery of MDD taking into account the heterogeneity present in the data.

Thus, this study addressed the following research questions:

- What are the factors that significantly affect the time to symptomatic recovery from MDD?
- Which parametric shared frailty model well describe the symptomatic recovery of MDD dataset?
- What is the estimated median first symptomatic recovery time of MDD?

1.3 Objectives of the Study

1.3.1 General Objective

The main objective of this study is to modeling time to first symptomatic recovery of the major depressive disordered patients of Jimma University Medical Center.

1.3.2 Specific Objectives

The specific objectives of this study are:-

- To determine the significant factors that affect time to first symptomatic recovery from MDD.
- To identify the parametric shared frailty model that best predicts time to the first symptomatic recovery of MDD patients well.
- To estimate the median first symptomatic recovery time.

1.4 Significance of the study

Studying the time to recovery is one way of overcoming the mental health problem in the community by addressing the effects of major depressive symptoms on the duration of recovery and significant factors of the major depressive disorder symptoms.

This study may helps physicians and researchers as a landmark for further studies related to mental disorder and others. Its also expected to give some knowledge about the determinants or risk factors of MDD and identify groups of patients, who are at higher risks to develop the disease. This have more advantage for health professionals (physicians) in order to give good treatment for identified stakeholders on identified risk factors. Provide that on model implications and arguing on choosing of the frailty distribution, the research reduces ambiguities on comparison of frailty models specifically in fitting data such as major depresive disorder.

2 Literature Review

2.1 Overview of Major Depressive Disorder

Major depressive disorder is a common mental disorder that negatively affects activities of daily living and is associated with high societal costs and functional impairment [43]. From 2013 to 2016, an estimated 8.1% of US adults 20 years and older experienced depression in any given two week period [44]. Depression can be chronic or short lasting, markedly impairing an individual's functioning at work or school or cope with daily life. In 2015 WHO estimated that 4.4% (322 million people) of the global population live with depression [1].

Major depressive disorder is the most commonly diagnosed psychiatric disorder in adults over 60 years of age [45]. Both DSM-5 and ICD-10 mark out typical (main) depressive symptoms [46]. ICD-10 defines three typical depressive symptoms (depressed mood, anhedonia, and reduced energy), two of which should be present to determine the depressive disorder diagnosis [47]. According to DSM-5, there are two main depressive symptoms: a depressed mood, and loss of interest/pleasure in activities (anhedonia). These symptoms, as well as five out of the nine more specific symptoms listed, must frequently occur for more than two weeks (to the extent in which it impairs functioning) for the diagnosis [48].

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 [49].

2.1.1 Recovery of Major Depressive Disorder

Major depressive disorder is a highly prevalent psychiatric condition that is associated with significant levels of disability, morbidity, and mortality [50]. Treatment of MDD traditionally aims to reduce depressive symptoms [51]. Consequently, the treatment is considered fully effective when complete or near-complete absence of the MDD symptoms (for a certain period of time) is achieved [52]. However, MDD is associated with major and sometimes long-lasting decreased levels of functioning and productivity. Approximately 60% of the patients with an MDD report severe or very severe functional impairment and can continue to experience (partial) impairment long after mood symptoms have been resolved [53]. Moreover, patients in remission report better functioning than those with mild depression, although their functioning is significantly worse than that found in the general population [54]. Therefore, remission of symptoms does not necessarily coincide with completely restored levels of functioning. Furthermore, MDD symptoms have differential effects on the level of functioning; depressed mood and loss of interest are strongly related to impaired functioning while weight problems, mid-nocturnal insomnia, and hypersomnia have less impact [55].

There is no commonly agreed definition of remission and recovery in MDD. Remission has been defined as a period of time in which the patient no longer meets the symptomatic criteria for the disorder or has only mild symptoms. Recovery is usually defined as sustained remission for a longer period of time. The operational criteria encompass (1) severity of symptoms assessed through symptom measurement instruments (eg, the Hamilton Depression Rating Scale; HAM-D17) [56]. and (2) duration or a certain period of time [57]. A reduction in symptom severity of $\geq 50\%$ during the course of treatment became an indicator of clinical response, that is, a clinically significant improvement [58]. A cutoff score on one of these measurements (eg, HAM-D17 $\leq$7) [59], is subsequently used to determine remission [57]. However, specific symptoms considered and symptom intensity may vary across studies [60].

2.1.2 Impacts of Major Depressive Disorder

Depression is the most widespread psychical health status in the public population [61]. People in almost all countries and cultures and of all genders, ages and experiences are affected by this common disorder named depression, where it affects 350 million people all over the world [62]. Depression impedes extraordinarily an individual's vocational power [63]. It significantly impacts on quality of life, general health and personal relationships, thereby contributing to poorer functioning at work, school life, sleeping, eating habits and within the family [64].

Major Depressive Disorder is more common in people without close interpersonal relationships, and who are divorced or separated, or widowed. No difference in the prevalence of MDD has been found among races and socioeconomic status. Individuals with MDD often have comorbid disorders such as substance use disorders, panic disorder, social anxiety disorder, and obsessive-compulsive disorder. The presence of these comorbid disorders in those diagnosed with MDD increases their risk of suicide [65]. Depression is found to be more prevalent in rural areas than in urban areas.

Major Depression disorder also often co-exists with other serious medical illnesses such as heart disease, stroke, cancer, HIV/AIDS, diabetes,asthma and parkinson's disease. Studies have shown that people who have depression in addition to another serious medical illness tend to have more severe symptoms of both depression and the medical illness, more difficulty adapting to their medical condition, and more medical costs than those who do not have co-existing depression [66].

2.1.3 Risk Factors for Recovery of Major Depressive Disorder

This section mainly covers several related literatures to the factors that are associated with the time to first symptomatic recovey of major depressive disorder.

Gender:- Gender were the significant association with depression. According to Centers for Dis-

ease Control and Preventation (CDC) mortality and morbidity a cross-sectional sample survey report in US, depression was more common in females than males [67]. Similarly, study conducted in Australia by Khawaja and Duncanson, [68], revealed that female are more depressed than male. Study in Turkey by Topba *et al.* showed that females were found to be twice more vulnerable to depression than males [69]. The other study done among Ambo University students by Adamu Birhanu indicate that being female is four times more likely to be depressed when compared with male [70]. Study conducted in Jimma by Andualem Mossie *et al.* [71], As per gender, in comparison with males, significantly larger proportion of females had depression.

Age:- According to CDC mortality and morbidity a cross-sectional sample survey report in US, people in the age group of 55 years and above had 5.21 times more odds to develop depression than people less than 25 years [67]. The possible reasons may be inability to perform daily activities, sedentary life style, and occurrence of concomitant medical illness [67]. Similarly, study conducted in Jimma by Andualem Mossie *et al.* [71], revealed that Age above 55 years is 5.9 times more likely to have depression than lower age groups. Another, cross-sectional study conducted by Hailemariam, which used multiple logistic regressions showed that the risk of depression episodes, increased as the age of depressive patient increases [12].

Marital Status:- Marital status was significantly associated with depression; An institution-based cross-sectional study was conducted among staff of Jimma University by Andualem Mossie *et al.* [71] and a cross-sectional study conducted in Debre Brehan Town by Reta *et al.* [72], widowed individuals were more likely to develop depression in comparison with the singles. The high prevalence of major depression in divorced individuals is due to both an increased risk of marital disruption in those with major depression, and also to the higher risk of this disorder in those with divorced marital status. Similary, according to L. Gu and J. Xie, Mogga and Dayess, depression was also common among marital status being divorced were significantly associated with depression [73, 74, 75].

First onset age:- Major depressive disorder can have significant effects when onset occurs in childhood and adolescence, the rate of depression increases from childhood through adolescence and into adulthood [76]. According to Ralph [77], depression is defined as a persistent experience of sad or irritable mood, loss of the ability to experience pleasure in nearly all activities. It is a serious health problem that can affect people of all ages, including children, adolescents and adults and both sex. In addition to this it affects an individual's every day activities by affecting an individual's emotion, cognition, behavior. It also includes a range of other symptoms such as change in appetite, disrupted sleep patterns, increased or diminished activity level, impaired attention and concentration [77].

Family History of mental illness:- A family history of mental illness is a significant predictor of depression. Individuals with a positive family history of depression were 2.5 times more likely to have depression. This supports the notion that genetic factors may contribute to depression in individuals [31]. Most individuals who develop major depression have a family history and one of their family member is affected by major depression. A study conducted by Khan *et al.* [78], found that out of 142 samples, 29% had a family history of depression. Based on their findings they revealed that students who had a family history of depression are 2.35 times depressed than those students who had no family history of depression. Accordingly, the odds of having depression were nearly two and a half fold higher among students who had a family member with mental illness as compared to their counterparts [79].

Chewing Khat:- According to Andualem Mossie *et al.* [71], chewing khat have significant association with depression episodes, khat chewers had more likely risk of developing depression as compared to nonchewers. Similary, according to the study conducted by Gelaw *et al.* [80] and Tekalign *et al.* [81] with WHO expert analysis [82] depression was significantly associated with khat chewing. The probability of developing depression among khat chewers is more likely

than that among nonchewers. The other study done among Ambo University students by Adamu Birhanu indicate that current use of Khat is three times more likely to be depressed [70].

Educational Status:- According to Andualem Mossie *et al.* [71], educational level have significant association with major depression, uneducated individuals were more likely to develop depression compared to educated peoples. Similarly, according to Centers for Disease Control and Preventation mortality and morbidity report in US, depression was more common in uneducated persons than educated persons [10, 67, 69], educated people have better understanding of the risk factors of depression compared to uneducated.

Employment Status:- Employment status have significant association with major depressive disorder, according to Ermias and Samuel the unmployed individuals had increased risk of having mental illness than the professionals [83]. Similary, in a study of Unemployment and Depression among emerging adults, the risk of depression is higher among the unemployed than among employed. The odds of depressions were higher for unemployed than employed emerging adults [84]. According to CDC mortality and morbidity report in US, depression was more common in unemployed persons than employed persons [67].

Religion:- Studies among adults reveal fairly consistent relationships between levels of religiosity and depressive disorders that are significant and inverse [85]. Koenig and colleagues highlight the fact that before 2000, more than 100 quantitative studies examined the relationships between religion and depression. Of 93 observational studies, two-thirds found lower rates of depressive disorder with fewer depressive symptoms in persons who were more religious [85]. In 34 studies that did not find a similar relationship, only 4 found that being religious was associated with more depression.

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 [83]. The major ethnic groups identified in the study area were Oromo 34.6%, Amhara 26.1%, Orthodox Christians account for 64.2% of the population where as 27.2% were Muslims.

Substance Abuse:- Substance abuse have shown a significant association with depression episodes. According to CDC mortality and morbidity report in US, depression was more among common substance users [73]. Study conducted in Mekelle General Prison Center by Welu *et al.* revealed that prisoners who had lifetime substance use were almost two times more likely to develop depression when compared to those who did not use substance in their life [34]. Similarly, study conducted by Covey stated that persons with major depression tend to abuse substances and have difficulties when they try to stop [86]. There are thousands of chemicals other than nicotine present in cigarette smoke, of which one or several may affect mood in the same way as a group of antidepressant medications called monoamine oxidase inhibitors or does. These medications effectively increase levels of specific neurotransmitters involved in the regulation of mood. Smoking, therefore, may be one way for depressed individuals to alleviate depressive symptoms [86].

A study conducted in Borena semi-nomadic community in southern Ethiopia [87], revealed the life time prevalence of all psychiatric disorders including substance abuse was 21.6%. The mental disorder excluding substance abuse was 14.6% among which neurotic and somatoform disorders were the most frequent disorders with life time prevalence of 14%.

Other Cofactors:- Other factors associated with depression is having chronic medical illness, comorbid medical illnesses (like, hypertension, epilepsy, HIV/AIDS, Diabetic mellitus renal disease) were significantly associated with depression [12]. The possible explanation may be that medical illness can cause tremendous life changes which may limit mobility and independency, interferes with doing enjoyable activities, and consequently decreases self-confidence that results

in depressive symptoms [12]. Similarly, according to study conducted by Egede LE the odds of major depression are high in individuals with chronic medical conditions, and major depression is associated with significant increases in utilization, lost productivity and functional disability. [88]. Studies have shown that people who have depression in addition to another serious medical illness tend to have more severe symptoms of both depression and the medical illness, more difficulty adapting to their medical condition, and more medical costs than those who do not have co-existing depression [66].

2.2 Over view of Unobserved Heterogeneity of Frailty

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and unobserved heterogeneity. In its simplest form, a frailty is an unobserved random factor that modifies multiplicatively the hazard function of an individual or a group or cluster of individuals [89]. Models constructed in terms of group-level frailties are sometimes referred to as 'shared' frailty models because observations within a subgroup share unmeasured 'risk factors' that prompt them to exit earlier than other subgroups. Frailty models [90], are increasingly popular for analyzing clustered survival data, where frailties or random effects often enter into the baseline hazard multiplicatively to model the correlation among observations within the same cluster [91]. We should be aware that, neither theory nor data typically provides much guidance for choosing a specific distribution from which to draw the frailty term and the parameter estimates can be very sensitive to the assumed parametric form [92].

The choice of the frailty distribution is often governed by the problem at hand in terms of the model implications [93]. The gamma distribution has been widely applied as a mixture distribution [94, 23]. From a computational and analytical point of view, it fits very well as a mixture distribution to failure data [95], used several distributions for frailty, including Gamma and Inverse Gaussian distributions and claimed that these two distributions are relevant and mathematically tractable as a frailty distribution for a heterogeneous population between groups. The most com-

mon reason for using the gamma distribution is its mathematical convenience. This is due to the simplicity of the derivative of the Laplace transform, meaning that traditional maximum likelihood procedures can be used for parameter estimation [96, 97]. Its flexible shape is another reason given for selection of the gamma distribution as the frailty distribution [25, 98]. Although it may be the most commonly used frailty distribution for the mathematical reasons, [99], emphasized that there are no biological reasons for choosing the gamma distribution.

The inverse Gaussian (inverse normal) distribution was introduced as a frailty distribution alternative to the gamma distribution by [100], and was used, for example, by [101, 20]. Similar to the gamma frailty model, simple closed-form expressions exist for the unconditional survival and hazard functions, this makes the model attractive. The particular interest in the multivariate case is the association between related event times. Indeed, different dependence structures result from different frailty distributions In particular; gamma frailties typically generate very strong dependence at late times and the inverse gaussian frailties at mid times [99]. The choice of a family of frailty distributions should therefore be accompanied by an assessment of fit. It is natural to consider the mean of the frailty variable conditionally on the observed filtration, which should fluctuate around one (1) [102].

Estimation of the frailty model can be parametric or semi-parametric. In the parametric case, a parametric density is assumed for the event times, resulting in a parametric baseline hazard function. Estimation is then conducted by maximizing the marginal log-likelihood [103]. In the semi-parametric case, the baseline hazard is left unspecified and more complex techniques are available to approach that situation [26]. The parametric estimation will be more powerful if the form of the baseline hazard is somehow known in advance [27]. Although the baseline hazard function may be modeled parametrically, some have argued that the parameters of the frailty distribution may be sensitive to the choice of distribution for the hazard, and that the choice of distribution for the hazard may, in fact, be more important than the choice of frailty distribution [104]. But, inference for

Cox frailty models is usually complicated because of issues surrounding the infinite-dimensional nuisance parameters which may arise in the estimation of the baseline hazard function, thereby violating the usual regular estimating assumptions [105].

2.3 Consequence of Ignoring Frailties

Ignoring the existence of heterogeneity will produce incorrect estimation of parameters and their standard errors in survival analysis [106], ignoring heterogeneity overestimates life expectancy based on their study on estimating life expectancy in a heterogeneous population [107], showed that when heterogeneity is ignored, it caused underestimation of covariate effects in his study of unemployment rates. Frailty models are used to make adjustments for over dispersion/under dispersion. When unobserved or unmeasured effects are ignored, the estimates of survival may be misleading [108], showed that ignoring frailty leads to regression coefficient estimates biased towards zero by an amount depending on the distribution and the variability of the frailty terms.

3 Methodology

3.1 Study Area

The study was conducted at Jimma University Medical Center 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, the capital city of Ethiopia. It has latitude and longitude of $7^0$40'N $36^0$50'E. Jimma has relatively cool tropical monsoon climate. The temperatures are in comfortable range, with the daily mean staying between 20^0C and 25^0C year-round. Jimma is the birth place of coffee and it represents about 11.8% of Ethiopians total coffee.

Jimma University Medical Center is one of the oldest Public Hospitals in Ethiopia. It was established in 1938 G.C by Italian invaders for service of their soldiers. After the withdrawal of the colonial occupants, it has been governed by the name of "Ras Desta Damtew Hospital" and later "Jimma Hospital" During Dergue Regime and Currently JUMC. 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 illpatients *www.hindawi.com/journals/psychiatry/2020/87395*.

3.2 Data Description

This data is secondary data recorded at the hospital in which patient's registry date to the event time or censoring time. Therefore data has extracted from the patient's card which contains epidemiological, laboratory and clinical information of MDD patient's card and information sheet after identification of patients who have admitted and follow up from September 01, 2018 to August 31, 2020. The event for this study was first symptomatic recovery, otherwise censored. Though, the information for the censored or truncated subjects is not completely known. MDD patients, who do not experience symptomatic recovery during the study time, lost and dropped before symp-

tomatic recovery, are considered as censored. The total number of patients considered in the study was 366 who were patient from all woredas. Woredas that contribute single patients were omitted. Therefore, a total of 366 MDD patients were considered in this study.

3.3 Study Design

The study is a retrospective cohort type admitted follow up of all major depressive disorder patients, who have followed at least three visits from September 1, 2018 to August 31, 2020 in Jimma University Medical Center is included.

3.4 Inclusion and Exclusion Criteria

In this study, all MDD patients who are registered, diagnosed for MDD and under follow up in the JUMC has been included. Children less than 12 ages were excluded, because depression is less common 1-2% in pre-puberty children when compared with adult which is 20% [109]. On other hand all patients who are diagnosed for MDD but not started follow up were excluded and in addition to this for analysis, woredas that contribute single patient were omitted, since the shared frailty model should be done on at least two patients in a woreda.

3.5 Study Variables

3.5.1 Dependent Variable

The response variable considered in this study was the time to first symptomatic recovery from major depressive disorder in between registry time (September 01, 2018) to study ends (August 31, 2020), which is the length of time in months to get the first symptomatic recovery. Time to first symptomatic recovery means the time until patients is free from any diagnostic criteria for major depressive disorder at least for six month duration according to DSM-IV. It is recorded on patients' card when all symptoms are improved or totally removed respectively.

3.5.2 Independent Variable

The candidate covariates used in the research and coding of categories are described in the following table where woreda is considered as a clustering variable in all frailty models.

Table 3.1: Description of Independent Variables will be used in the Analysis

Variables	Variables Description	Categories and Coding
Gender	Gender of Patients	0 = Male, 1 = Female
Age	Age of the Patients (in years)	1 = 13-19, 2 = 20-25, 3 = 26-49, 4 = $\geq$50
Event of Relapse	Relapse history of patients	0 = No, 1 = Yes
Frist onset age	Frist onset age of patients	1 = Childhood, 2 = Adolescence age 3 = Adult age & above
Educational Level	Educational Level of patients	0 = Uneducated, 1 = Educated
Other Cofactors	Other Cofactors	0 = No, 1 = Yes
Marital Status	Marital Status of patients	1 = Single, 2 = Married, 3 = Widowed, 4 = divorced
Family History	Family History of mental illness	0 = No, 1 = Yes
Substance abuse	Patient whether use alcoholic drinks	0 = No 1 = Yes
Religion	Religion of patients	1 = Orthodox, 2 = Muslims, 3 = Protestant, 4 = Other's
Chewing Khat	Habit of Chewing Khat	0 = No, 1 = Yes
Employment	Employment status of patients	0 = No, 1 = Yes
Ethnicity	Ethnicity of patients	1 = Oromo, 2 = Amhara, 3 = Other's

3.6 Statistical Methods

3.6.1 Survival Analysis

Survival analysis is an important statistical technique used to describe and model time-to-event data. It is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. By time we mean, years, months, weeks, or days from the beginning of follow-up of an individual until an event occurs. The use of survival analysis, as opposed to the use of other statistical methods, is most important when some subjects are lost to follow up or when the period of observation is finite and certain patients may not experience the event of interest over the study period. In this latter case one cannot have complete information for such individuals. These incomplete observations are referred to as being censored.

A common characteristic of survival data is censoring, truncation, or combination of censoring and truncation. In essence, censoring occurs when we have some information about individual survival time, but we don't know the survival time exactly. There are three categories of censoring, [20], such as right censoring, left censoring and interval censoring. Right censoring is the most common form of censoring, where a subject's follow up time terminates before the outcome of interest is observed. This type of censoring is commonly recognized survival analysis and also considered in this study. An observation is said to be left censored if individuals developed the event of interest prior to the beginning of the study. And, interval censoring is when event of interest occurs within an interval of time without the knowledge of when exactly happened. The data used for this study is right censored data and the censored was:

- Patients who die because of depression or other disease before mental disorder.
- Patients who dropout or referred to other hospital.
- At end time of the study, patients who are in the study but not develop mental disorder.

3.6.2 Descriptive Methods For Survival Data

An initial step in the analysis of a set of survival data is to present numerical or graphical summaries of the survival times in a particular group. In summarizing survival data, the two common functions applied are the survivor function and the hazard function [110].

Survival Function:- The basic quantity employed to describe time-to-event phenomena is the survival function, the probability of an individual surviving or being event-free beyond time t (experiencing the event after time t). Moreover, the distribution of survival time is characterized by three functions: survivorship function, probability density function, and hazard function.

Let T be a random variable associated with the survival times, t be the realization of the random variable T and f(t) be the underlying probability density function of the survival time t. The cumulative hazard function H(t), which represents the probability that a subject selected at random will have a survival time (in this case, survival time to return) less than some stated value t, is given by:

$$F(t) = P(T \leq t) = \int_0^t f(u)du, t \geq 0$$

The survival function is defined as the probability that the survival time is greater or equal to t.

$$S(t) = P(T > t) = 1 - F(t), t \geq 0$$
$$S(t) = 1 - F_T(t)$$

Theoretically, as t ranges from 0 to infinity, the survivor function can be graphed as a smooth curve. This survival function gives the probability of surviving or being event free beyond time t. Because Survival functions $(S(t))$ is probability, it is characterized by: Survival function, $S(t)$ is characterized that, they are non-increasing function. Actually, at the time, $t = 0; S(t) = S(0) = 1$. That is, at the start of the study no one has experienced the event yet, the probability of surviving past time 0 is one (1). However, as time $t \to \infty; S(t) \to 0$ that is, theoretically, if the study period

increased without limit, eventually nobody would survive, so the survivor curve must eventually converge to zero.

Hazard Function:- The hazard function is a measure of the probability of failure during a very small interval, assuming that the individual has survived at the beginning of the interval. The hazard function describes the concept of the risk of an outcome (e.g., death, failure, hospitalization, recovery) in an interval after time t, conditional on the subject having survived to time t. It is the probability that an individual dies somewhere between $t + \Delta t$, divided by the probability that the individual survived beyond time t. hazard function h(t) can be formulated as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t+\Delta t / T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$
$$h(t) = \frac{f(t)}{S(t)}$$

The survival and cumulative hazard functions can be given in terms of the hazard function as:

$$h(t) = \int_o^t h(u)du$$

Using the above expressions the hazard function h(t) can also be given as:

$$H(t) = -\frac{\log S(t)}{dt} = \frac{dH(t)}{dt}$$

3.6.3 Estimation of Survivorship Function

In survival analysis, it is always a good idea to present numerical or graphical summaries of the survival times for the individuals. In general, survival data are conveniently summarized through estimates of the survival function and hazard function. This method is non-parametric or distribution free, since they require no specific assumptions to be made about the underlying distribution of the survival times [110]. In this study, the other estimators of the survivor function the Kaplan-Meier (KM) for estimation survival function and log-rank test for comparison between two or more groups of categorical covariates were used. The Kaplan-Meier estimator of the survivorship function [111], also called product limit estimator, is the estimator used by most software packages. This estimator incorporates information from all of the observations available, both uncensored

and censored, by considering survival to any point in time as a series of steps defined by the observed survival and censored times.

Suppose we have a sample of independent observations, their survival times denoted by $t_1, t_2, t_3, ..., t_n$ and indicators of censoring denoting by $\delta_1, \delta_2, \delta_3, ..., \delta_n$ where

$$\delta_i = \begin{cases} 1, & \text{if the first symptomatic occur} \\ 0, & \text{otherwise} \end{cases}$$

Thus, the survival data are denoted by $t_i, \delta_i; i = 1, 2, 3, .., n$. The first step to obtain the Kaplan-Meier estimator of the survival function is to order the survival times as $t_1, t_2, t_3, ..., t_n$. Assume that among the n observations $m \leq n$ event occurred at distinct m times. The main quantity of interest is the probability that an event would not occur by time $t : S(t) = P(T > t)$. Kaplan and Meier [111] develop an estimator for the survival function .

$$\hat{S}_{KM}(t) = \prod_{t_i \leq t} (\frac{n_i - d_i}{n_i})^{\delta_i} = \prod_{t_i \leq t} (1 - \frac{d_i}{n_i})^{\delta_i}$$

Where, d_i is number of patients experienced event at t_i and n_i is number of patients at risk before t_i.

The log-rank test which is used for comparison of the survival curves of two or more categorical covariates also applied. Log-rank test is first proposed by Breslow, and it gives information on the significance of difference for the survival of two or more groups of patients [112].

3.6.4 Median Survival Time

Median survival time is the time beyond which 50% of the individuals in the population under study are expected to survive and is given by that value t(50) which is such that $S\{t(50)\} = 0.5$. Due to the fact that the non-parametric estimates s(t) are step functions, it will not usually be possible to realize an estimated survival time that makes the survival function exactly equal to 0.5. Instead, the estimated median survival time, is defined to be the smallest observed survival time

for which the value of the estimated survival function is less than 0.5.
In mathematical terms,

$$\hat{t}(50) = \min\{\frac{t_i}{\hat{S}(t_j)} < 0.5\}$$

Where t_i is the observed survival time for the i^{th} individual, i = $1, 2, ..., n$ and t_j is the j^{th} ordered death time, j = $1, 2, ..., r$

3.7 Frailty Models

The notion of frailty provides a convenient way to introduce random effect, association and unobserved heterogeneity into models for survival data. Frailty models are the survival data analog to regression models, which account for heterogeneity and random effects. For example, in many cases it is impossible to measure all relevant covariates related to the disease of interest, sometimes because of economic reasons, sometimes the importance of some covariates is still unknown. The frailty approach is a statistical modeling concept which aims to account for heterogeneity, caused by unmeasured covariates. In statistical terms, a frailty model is a random effect model for time-to-event data, where the random effect (the frailty) has a multiplicative effect on the baseline hazard function. This random effect explains the dependence in the sense that had we known the frailty, the events would be independent. In other words, the life times are conditional independent, given the frailty. A frailty is a latent multiplicative effect on the hazard function and is assumed to have unit mean and variance q, which is estimated along with the other model parameters.

3.7.1 Shared Frailty Model

A shared frailty model is a random effects model where the frailties are common (or shared) among groups of individuals or spells and are randomly distributed across groups. A natural extension of the univariate frailty model would be a multivariate survival model where individuals are allowed to share the same frailty value.

Many statistical models and methods proposed to model failure time data assume that the observations are statistically independent of each other. However, this does not hold in many applications. Shared frailty model is a conditional model in which frailty is common to all subjects in a cluster. The shared frailty model is responsible for creating dependence between event times. It is also known as a mixture model because the frailties in each cluster are assumed to be random. It assumes that, the given frailty, all event times in a cluster are independent. Shared frailty model was introduced by Clayton (1978) without using the notion frailty and extensively studied in [113, 114].

Frailty models are the extensions of the proportional hazards model which is best known as the Cox model [21],the most popular model in survival analysis. Normally, in most clinical application, survival analysis implicitly assumes a homogeneous population of individuals to be studied. This means that all individuals sampled in that study are subject in principle under the same risk (e.g., risk of death, risk of disease recurrence). In many applications, the study population cannot be assumed to be homogeneous, but must be considered as a heterogeneous sample i.e., a mixture of individuals with different hazards. Multivariate frailty model is an extension of the univariate frailty model which allows the individuals in the same cluster to share the same frailty value. When frailty is shared, dependence between individuals who share frailties is generated.

Conditional on the random term, called the frailty denoted by u_i, the survival times in cluster i($1 \leq$ i $\leq$ n) are assumed to be independent, the proportional hazard frailty model assumes

$$h_{ij}(t/X_{ij}, u_i) = \exp(\beta' X_{ij} + u_i) h_o(t)$$

where u_i the random term of all the subjects in cluster.

Where as an alternative if the proportional hazards assumption does not hold is the accelerated failure time frailty model which assumes

$$h_{ij}(t/X_{ij}, u_i) = \exp(\beta' X_{ij} + u_i) h_o(exp(\beta' X_{ij} + u_i)t)$$

Where i indicates the i^{th} cluster and j indicates the j^{th} individual for the i^{th} cluster, $h_0(.)$ is the baseline hazard, u_i the random term of all the subjects in cluster i, X_{ij} the vector of covariates for subject j in cluster i, and β the vector of regression coefficients.

If we let Z = $\exp(u_i)$, in the thesis assumed that Z has the gamma or the inverse gaussian distribution so that the hazard function depends upon this frailty that acts multiplicatively on it. If the number of subjects n_i is 1 for all groups, the univariate frailty model is obtained [24], otherwise the model is called the shared frailty model [96, 114], because all subjects in the same cluster share the same frailty value z_i.

The main assumption of a shared frailty model is that all individuals in cluster i share the same value of frailty Z_i (i = 1, ..., n), and this is why the model is called the shared frailty model. The lifetimes are assumed to be conditionally independent with respect to the shared (common) frailty. This shared frailty is the cause of dependence between lifetimes within the clusters.

3.7.2 The Frailty Distributions

The frailty z_i is an unobservable realization of a random variable Z with probability density function $f(.)$ which is the frailty distribution. Since z_i multiplies the hazard function, Z has to be non-negative. Another constraint is further needed for identifiability reasons, similar to the zero mean constraint of a random effect in a standard linear mixed model. More specifically, the mean of Z is typically restricted to unity when possible (i.e., when $E(Z)$ exists) in order to separate the baseline hazard from the overall level of the random frailties.

The choice of the frailty distribution is of crucial importance to arrive at a good description of the dependence structure present in the data. Therefore, the choice of the frailty distribution is even more important as the choice of the distribution of the random effect(s) in mixed models. Different distributions have been proposed for the frailty term. In this study, we used Gamma and

inverse Gaussian frailty distributions. In both cases, as a single heterogeneity parameter (denoted by θ) shows the degree of independence.

3.7.2.1 The Gamma Frailty Distribution

The gamma distribution is very-well known and has simple densities. It is the most common distribution used for describing frailty. Even though gamma models have closed form expressions for survival and hazard functions, from a computational view, it fits well to frailty data and it is easy to derive the closed form expressions for unconditional survival and hazard functions. The gamma distribution has been widely applied as a mixture distribution for example [104, 115]. From a computational and analytical point of view, it fits very well to failure data. It is widely used due to mathematical tractability [24]. To make the model identifiable, we restrict that expectation of the frailty equals one and variance be finite, so that only one parameter needs to be estimated. Thus, the distribution of frailty Z is the one parameter gamma distribution. Under the restriction, the corresponding density function and Laplace transformation of gamma distribution:-

$$f_z(Z_i) = \frac{Z_i^{(1/\theta)-1}\exp(-Z_i/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \ , \ \theta > 0$$

Where $\Gamma(.)$ is the gamma function, it corresponds to a Gamma distribution $\mathrm{Gam}(\mu, \theta)$ with μ fixed to 1 for identifiability. Its variance is then θ, with Laplace transform

$$\mathrm{L(u)} = (1 + u/\theta)^{-\theta} \ , \ \theta > 0$$

Note that if $\theta > 0$, there is heterogeneity. So the large values of θ reflect a greater degree of heterogeneity among groups and a stronger association within groups.
The conditional survival function of the gamma frailty distribution is given by [103].

$$S_o(t) = [1 - \theta \ln(S(t))]^{-1/\theta} \ , \ \theta > 0$$

And the conditional hazard function is given by:

$$h_0(t) = h(t)[1 - \theta \ln(S(t))]^{-1}$$

Where S(t) and h(t) are the survival and the hazard functions of the baseline distributions. Larger variance indicates a stronger association within groups. For the Gamma distribution, the Kendall's Tau, which measures the association between any two event times from the same cluster in the multivariate case, can be compute by:

$$\tau = \frac{\theta}{\theta+2}, \text{ where } \varepsilon(0,1), \text{ With } SE(\tau) = \frac{2SE(\theta)}{(\theta+2)^2}, \text{ Where } \tau\varepsilon(0,1)$$

3.7.2.2 Inverse Gaussian Frailty Distribution

The inverse Gaussian (inverse normal) distribution was introduced as a frailty distribution alternative to the gamma distribution by [100]. Similar to the gamma frailty model, simple closed-form expressions exist for the unconditional survival and hazard functions, this makes the model attractive. In particular, Inverse Gaussian frailties generate stronger dependence at mid time. As an alternative to the gamma distribution. The probability density function of an inverse Gaussian shared distributed random variable with parameter $\theta > 0$ is given by;

$$f_z(z_i) = (\frac{1}{2\pi\theta})^{1/2} Z_i^{-3/2} \exp(\frac{-(z_i-1)^2}{2\theta z_i}),\ \theta > 0,\ \text{z} > 0$$

For identifiability, we assume z has expected value equal to one and variance θ.
The Laplace transformation of the inverse Gaussian distribution is:-

$$L(s) = \exp(\frac{1}{\theta}(1-\sqrt{1+2\theta s})),\ \theta > 0,\ \text{s} > 0$$

For the inverse Gaussian frailty distribution the conditional survival function is given by:

$$S_0(t) = \exp\{\frac{1}{\theta}(1-[1-2\theta\ln\{S(t)\}]^{1/2}\},\ \theta > 0$$

And For the inverse Gaussian frailty distribution the conditional hazard function is given by:

$$h_0(t) = h(t)[1-2\theta\ln(S(t))]^{-1/2},\ \theta > 0$$

Where S(t) and h(t) are the survival and the hazard functions of the baseline distributions. With multivariate data, an Inverse Gaussian distributed frailty yields a Kendall's Tau given by,

$$\tau = \frac{1}{2} - \frac{1}{\theta} + 2\frac{\exp(2/\theta)}{\theta^2}\int_{2/\theta}^{\infty}\frac{\exp(-u)}{u}du \text{ , where } \tau \in (0, \frac{1}{2})$$

3.7.3 Baseline Survivor and Hazard Function

As in the proportional hazards model, parametric or non-parametric forms of baseline hazard can be assumed in frailty models. If non-parametric form is assumed for $h(t)$, then semi parametric proportional hazards model is considered and the estimates are usually obtained by using Expectation-Maximization (EM) algorithm.

The survival time T is assumed to follow a distribution with density function $f(t)$, then the survival function is given by $S(t) = P(T > t) = \int_t^\infty f(u)du$
The hazard function is a measure of the probability of failure during a very small interval, assuming that the individual has survived at the beginning of the interval. It is defined as:-

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{-d}{dt}S(t)}{S(t)}$$

The relationship between the survival and the hazard function is given by $S(t) = \exp(-\int_0^\infty h(u)du)$. Under the parametric approach, the baseline hazard function is defined as a parametric function and the vector of its parameters, say ψ, is estimated together with the regression coefficients and the frailty parameter(s).

The cumulative hazard function is given by $H(t) = \int_o^t h(u)du$. Specifying one of the four functions f(t), S(t), h(t) or H(t) specifies the other three functions. The parameter λ is reparameterized in terms of predictor variables and the regression parameters. Typically for parametric models, the shape parameter ρ is held fixed. where $\lambda, \rho, \sigma > 0$, PH is Proportional hazards, AFT is accelerated failure time and $\Phi(z)$ denoted the standard normal cumulative distribution. In this research the following distributions are considered.

Table 3.2: Parametric Distributions For The Baseline Hazards

Distribution	f(t)	s(t)	h(t)	H(t)	Parameter Space
Weibull	$\rho\lambda t^{\rho-1}\exp(-\lambda t^{\rho})$	$\exp(-\lambda t^{\rho})$	$\rho\lambda t^{\rho-1}$	λt^{ρ}	$\lambda,\rho>0$
Log-Logistic	$\frac{\lambda\rho t^{\rho-1}}{(1+\lambda t^{\rho})^2}$	$\frac{1}{1+\lambda t^{\rho}}$	$\frac{\lambda\rho t^{\rho-1}}{1+\lambda t^{\rho}}$	$\ln[1+(\frac{t}{\lambda})^{\rho}]$	$\lambda\in\Re,\lambda>0$
Log-normal	$\lambda\exp(-\lambda t)$	$\exp(-\lambda t)$	λ	λt	$\lambda>0$

3.7.4 Parameterization

When we say proportional hazards (PH) it means that the hazard function of a group is proportional to the hazard function of the other group, i.e., the hazard ratio is constant over time [20]. The hazard ratio is hence given by HR = $\exp(\beta^{'}X_{ij})$ is the hazard ratio (HR). Where $\beta^{'}=\beta_1,\beta_2,\beta_3,.....,\beta_p$ is a vector of regression coefficients and X_{ij} is the vector of covariates for subject j in cluster i. On the other hand, the accelerated failure-time (AFT) model describes stretching out or contraction of survival time as a function of predictor variables. The acceleration factor which is usually denoted by ϕ is given by $\exp(\alpha^{'}X_{ij})$ where $\alpha^{'}=\alpha_1,\alpha_2,\alpha_3,.....,\alpha_p$ is a vector of regression coefficients in case of AFT model. For the weibull, log logistic and log-normal survival model, the relationship between α and β is given by.

For weibull, $\beta_j=-\alpha_j\rho$, where ρ is the shape parameter and hence, HR = $\exp(\beta)$ = $\exp(-\alpha_j\rho)$ is the hazard ratio of the j^{th} group with the reference groups. On the other hand for log-logistic distribution, since the log-logistic model is a proportional odds (PO) model, i.e. it has constant OR for two groups. Therefore $\beta_j=-\alpha_j\rho$, where ρ is the shape parameter and OR = $\exp(\beta)=\exp(-\alpha_j\rho)$ indicates the failure odds ratio of the j^{th} group with the reference groups. Lastly, for log-normal, has shape similar to the log-logistic distribution accommodates an AFT model (as log-logistic), but is not a proportional odds model.

3.8 Method of Parameter Estimation

Estimation of the frailty model can be parametric or semi-parametric. In the former case, a parametric density is assumed for the event times, resulting in a parametric baseline hazard function. Estimation is then conducted by maximizing the marginal log-likelihood [27]. In the second case, the baseline hazard is left unspecified and more complex techniques are available to approach that situation. Even though semi-parametric estimation offers more flexibility, the parametric estimation will be more powerful if the form of the baseline hazard is somehow known in advance.

Frailty models account for the clustering present in grouped event time data. For right-censored clustered survival data, the observation for subject $j \in J_i = \{1, \dot{...}, n_i\}$ from cluster $i \in I = \{1, \dot{...}, s\}$ is the couple (y_{ij}, δ_{ij}), where $y_{ij} = min(t_{ij}, c_{ij})$ is the minimum between the survival time tij and the censoring time c_{ij}, and where $\delta_{ij} = I(t_{ij} \leq c_{ij})$ is the event indicator. When covariate information are been collected the observation will be $(y_{ij}, \delta_{ij}, X_{ij})$, where X_{ij} denote the vector of covariates for the ij^{th} observation. In the parametric setting, estimation is based on the marginal likelihood in which the frailties have been integrated out by averaging the conditional likelihood with respect to the frailty distribution. Under assumptions of non-informative right-censoring and of independence between the censoring time and the survival time random variables, given the covariate information, the marginal log-likelihood of the observed data can be written as.

$$l_{marg}(\psi, \beta, \theta; Z, X) = \sum_{i=1}^{s} \{[\sum_{j=1}^{ni} \delta_{ij}(\log(h_0(y_{ij})) + X_{ij}^T \beta)] + \log[(-1)^{di} L^d([\sum_{j=1}^{ni} H_o(y_{ij}) \exp(X_{ij}^T)])]\}$$

Where $d_i = \sum_{j=1}^{ni} \delta_{ij}$ is the number of events in the i^{th} cluster, and $L^q(.)$ the q^{th} derivative of the Laplace transform of the frailty distribution defined as

$L(s) = E[\exp(-Zs)] = \int_0^{\infty}(-Z_i s) f(Z_i) dz_i$, $S \geq 0$ and
$L(s) = E(-q) \int_0^{\infty}(Z^q \exp(-Zs)) f(Zi) dzi$, $q \geq 0$

Where Ψ represents a vector of parameters of the baseline hazard function, β the vector of regression coefficients and θ the variance of the random effect. The estimates of Ψ, β, θ are obtained by maximizing the marginal log-likelihood of the above. This can be done if one is able to compute higher order derivatives $L^q(.)$ of the Laplace transform up to $q = max\dot{(}d_1, d_2, d_3, ..., d_s)$. Symbolic differentiation is performed in R, but is impractical here; mainly because this is very time consuming [27].

3.9 Comparison of Models

Model comparison and selection are among the most common problems of statistical practice, with numerous procedures for choosing among a set of models [116]. There are several methods of model selection. The most commonly used methods include information criteria. One of the most commonly used model selection criteria is Akaike Information Criterion (AIC). A data-driven model selection method such as an adapted version of Akaike's information criterion AIC is used to find the truncation point of the series [117]. In some circumstances, it might be useful to easily obtain AIC value for a series of candidate models [27]. In this study, we used the AIC criteria to compare various candidates of parametric frailty models. The model with the smallest AIC value is considered a better fit. For comparing models that are non-nested type, the Akaike's information criterion (AIC) which is defined as:

$$AIC = -2\log(L) + 2(k + c + 1)$$

Where k is the number of covariates, c the number of model specific distributional parameters. The preferred model is the one with the lowest values of the AIC. Manipulation of the comparison was done using the R software.

3.10 Model Diagnostics

3.10.1 Evaluation of the Baseline Parameters

The graphical methods can be used to check if a parametric distribution fits the observed data or not. Appropriateness of assumed distributions baseline hazard function is evaluated as follows:

Model with the weibull baseline has a property that the $\log(-log(S(t))$ is linear with the log of time, where $\hat{S}(t) = \exp(-\lambda t^{\rho})$. Hence, $\log(-\log(\hat{S}(t)) = \log(\lambda) + \rho \log(t)$. This property allows a graphical evaluation of the appropriateness of a Weibull model by plotting $\log(-\log(\hat{S}(t)))$ versus $\log(t)$ where $\hat{S}(t)$ is Kaplan-Meier survival estimate [118].

For log-normal baseline plot of $\Phi^{-1}\{1-\exp(-H(t))\} = \Phi^{-1}\{1-\hat{S}(t)\}$ versus log time (t) should be linear, if the log-normal distribution is appropriate.

For log logistic baseline plot $\log(\frac{1-\hat{S}(t)}{\hat{S}(t)})$ versus $\log(t)$. This should be linear with slope ρ. The log-failure odd versus log time of the log-logistic model is linear. Where log failure odds can be written as: $\log(\frac{1-\hat{S}(t)}{\hat{S}(t)}) = \log(\lambda\rho t^{\rho}) = \log(\lambda) + \rho\log(t)$. Where $\hat{S}(t)$ is Kaplan-Meier survival estimate [118].

3.10.2 The Cox- Snell Residuals

The Cox-Snell residuals method can be applied to any parametric model and the residual plots can be used to check the goodness of fit of the model. For the parametric regression problem, analogs of the semi parametric residual plots can be made with a redefinition of the various residuals to incorporate the parametric form of the baseline hazard rates [20]. The first such residual is the Cox–Snell residual that provides a check of the overall fit of the model. The Cox–Snell residual, r_j, is defined by: $r_j = \hat{H}(T_j/Xj)$ where $\hat{H}$ is the cumulative hazard function of the fitted model. If the model fits the data, then the $r_j^{'}s$ should have a standard (λ= 1) exponential distribution, so that

a hazard plot of r_j versus the Nelson–Aalen estimator of the cumulative hazard of the $r_j^{'}s$ should be a straight line with slope one.

Table 3.3: The Cox–Snell residuals for the baseline hazard functions considered in this study

Distribution for Baseline Hazard Function	Cox-Snell residuals (r_j)
Weibull	$\hat{\lambda}\exp(\hat{\beta}'X_j)t^{\rho}$
Log-Logistic	$\ln(\frac{\hat{1}}{1+\hat{\lambda}\exp(\beta'X_j)t_j^p})$
Log-normal	$\ln(\frac{1}{1+\rho t\exp(\beta'Xj)})$

3.10.3 Checking for Proportional Hazard Assumption

Let $\hat{H}_0(t)$ is the cumulative baseline hazard function, and K is the number of disjoint categories. To check the proportionality assumption we could plot $\ln(\hat{H}_{10}(t)), \ln(\hat{H}_{20}(t)), ..., \ln(\hat{H}_{k0}(t))$ versus t. If the assumption holds, then, these should be approximately parallel and the $\ln(\hat{H}_{g0}(t))$ and $\ln(\hat{H}_{h0}(t))$ should give a crude estimate of the factor needed to obtain constant vertical separation between $\hat{H}_{h0}(t)$ from $\hat{H}_0(t)$. An alternative approach is to plot $\ln(\hat{H}_{g0}(t)) - \ln(\hat{H}_{10}(t))$ versus t for g = 2... k. If the proportional hazards model holds, each curve should be roughly constant [20].

3.10.4 The Quantile - Quantile Plot

A quantile-quantile or q-q plot is made to check if the accelerated failure time model provides an adequate fit to the data. The plot is based on the fact that, for the accelerated failure-time model,

$$S_1(t) = S_0(\phi t)$$

Where S_0 and S_1 are the survival functions in the two groups and ϕ is the acceleration factor. Let t_{0p} and t_{1p} be the p^{th} percentiles of groups 0 and 1, respectively, that is

$$t_{kp} = S_k^{-1}(1-p), k = 0, 1.$$

Using the relation $S_1(t) = S_0(\phi t)$ we must have $S_0(t_o p) = 1 - p = S_1(t_1 p) = S_0(\phi t_{1p})$ for all t. If the accelerated failure time model holds, $t_{op} = \phi t_{1p}$. To check this assumption we compute the Kaplan–Meier estimators of the two groups and estimate the percentiles t_{1p}, t_{op}, for various values of p. If we plot the estimated percentile in group 0 versus the estimated percentile in group 1 (i.e., plot the points t_{1p}, t_{op}, for various values of p), the graph should be a straight line through the origin, if the accelerated failure time model holds. If the curve is linear, a crude estimate of the acceleration factor q is given by the slope of the line [20].

4 Results and Discussion

4.1 Descriptive of Socio-economic and Health Related Variables

The data for this study consists of 366 patients who were major depressive disorder patients under psychiatric follow up at Jimma University Medical Center, from September 1, 2018 to August 31, 2020 were considered. Of all 366 MDD patients during the time period, 198(54.1%) were experienced the event (first symptomatic recovery from MDD) whereas 168(45.9%) of them were censored (Table 4.1). The estimated median symptomatic recovery time for MDD patients was found to be 7 months.

Among the total number of MDD patients, 187(51.1%) were males and 179(48.9%) were females; of which 41.2% of males and 67.6% of females were experienced symptomatic recovey. The median symptomatic recovery time of male and female were 11 and 9 months, respectively.

Regarding to age of patients, about 37(10.1%) of patients were aged between 13-19, 98(26.8%) were aged between 20-25, 144(39.3%) were aged between 26-49 and about 87(23.8%) were aged 50 and above; of which 59.5%, 64.3%, 63.9% and 24.1% were recovered, respectively. The median recovery time of MDD patients who aged 13-19, 20-25, 26-49 and 50 & above were 6, 7, 6 and 13 months, respectively. This result seems to indicate that majority of MDD patients were found between 26-49 age group, and patients who aged 50 and above took longer period of time to first symptomatic recovery than the other age groups.

Relatively among the woreda, Jimma Town experienced highest MDD 18.1% followed by Kersa woreda 8.4% and Gomma woreda 6.5%. Setema woreda was experienced the lowest MDD 2.4% followed by Limmu Seka woreda 3%. Setema woreda has also the highest median time to first symptomatioc recovery from MDD (13 months) among the other woreda. (see Annex 4).

Regarding marital status of patients, about 110(30.1%) were single, 136(37.2%) were married, 62(16.9%) were widowed and 58(15.8%) were divorced. Among those who were experienced event of interest, symptomatic recovery in our case, it showed that 72(65.5%), 98(72.1%), 12(19.4%) and 16(27.6%) were single, married, widowed and divorced in marital status categories. The median recovery time of MDD patients single, married, widowed and divorced were 6, 5, 21 and 19 months respectively.

Based on the results of this study, it shown that most of the patients join to the JUMC at their adolescent age. Out of 366 patients 20(5.5%) were childhood, 198(54.1%) patients were join at adolescent age while remaining 148(40.4%) were join at adult age and above. Patients who were adolescent age seems that they took shorter time to first symptomatic recovery from major depressive disorder (7 months).

Patients who had family history of mental illness seems that they stay longer time 12 months than who had no family history of mental illness 6 month to experience recovery. Out of total, 190(51.9%) had no family history of mental illness while 176(48.1%) had a family history of mental illness. From those who has family history of mental illness 75(42.6%) were recoverd and 101(57.4%) were censored.

Regarding khat chewing of patients, 230(62.9%) were non chewing khat while 136(37.1%) were khat chewers. The median recovering time of patients of non chewing khat and chewing khat are 5 and 13 months respectively, and of which 77% and 15.7% recovered respectively.

It also observed that from total patients, only 215(58.7%) of patients had educated educational level while 151(41.3%), of patient's had uneducated educational level. Patients' with uneducated educational level stayed 13 months, of which 29.3% were recovered and educated educational level stayed 9 months of which 89.4%. Regarding employment of MDD patients, from total of MDD

Table 4.1: Descriptive summary of covariates with the recovery status from MDD patients.

Variable	Categories	Recovery Status		Total(%)	Median Time (in months)	95% CI
		Censored(%)	Events(%)			
Gender	Male	110(58.8%)	77(41.2%)	187(51.1%)	11	(9, 18)
	Female	58(32.4%)	121(67.6%)	179(48.9%)	9	(5, 13)
Age	13-19	15(40.5%)	22(59.5%)	37(10.1%)	6	(5, 20)
	20-25	35(35.7%)	63(64.3%)	98(26.8%)	7	(5, 8)
	26-49	52(36.1%)	92(63.9%)	144(39.3%)	6	(5, 9)
	≥ 50	66(75.9%)	21(24.1%)	87(23.8%)	13	(12, 14)
Marital Status	Single	38(34.5%)	72(65.5%)	110(30.1%)	6	(5, 9)
	Married	38(27.9%)	98(72.1%)	136(37.2%)	5	(4, 7)
	Widowed	50(80.6%)	12(19.4%)	62(16.9%)	21	(20, 22)
	Divorced	42(72.4%)	16(27.6%)	58(15.8%)	19	(13, 22)
First Onset Age	Childhood	8(40%)	12(60%)	20(5.5%)	11	(5, 12)
	Adolescent	78(39.4%)	120(60.6%)	198(54.1%)	7	(6, 9)
	Adult	82(55.4%)	66(44.6%)	148(40.4%)	9	(7, 18)
Family History	No	67(35.3%)	123(64.7%)	190(51.9%)	6	(5, 7)
	Yes	101(57.4%)	75(42.6%)	176(48.1%)	12	(9, 18)
Chewing Khat	No	53(23%)	177(77%)	230(62.9%)	5	(4, 6)
	Yes	115(84.6%)	21(15.4%)	136(37.1%)	13	(11, 20)
Educational Level	Uneducated	152(70.7%)	63(29.3%)	215(58.7%)	13	(12, 15)
	Educated	16(10.6%)	135(89.4%)	151(41.3%)	9	(3, 15)
Employment	No	105(60.3%)	69(39.7%)	174(47.5%)	9	(8 19)
	Yes	63(32.8%)	129(67.2%)	192(52.5%)	6	(4, 7)
Religion	Orthodox	47(49.5%)	48(50.5%)	95(25.9%)	9	(6, 18)
	Muslims	90(43.7%)	116(56.3%)	206(56.3%)	10	(6, 11)
	Protestant	24(52.2%)	22(47.8%)	46(12.6%)	7	(6, 8)
	Other's	7(36.8%)	12(63.2%)	19(5.2%)	6	(4, 7)
Ethnicity	Oromo	112(46.3%)	130(53.7%)	242(66.1%)	8	(6, 11)
	Amhara	42(49.4%)	43(50.6%)	85(23.2%)	7	(6, 18)
	Others	14(35.9%)	25(64.1%)	39(10.7%)	6	(5, 13)
Substance Abuse	No	61(32.1%)	129(67.9%)	190(51.9%)	5	(4, 6)
	Yes	107(60.8%)	69(39.2%)	176(48.1%)	12	(10, 18)
Other Cofactors	No	40(19.8%)	162(80.2%)	202(55.2%)	5	(4, 6)
	Yes	128(78%)	36(22%)	164(44.8%)	10	(9, 22)
Event of Relapse	No	64(33.7%)	126(66.3%)	190(51.9%)	6	(5, 7)
	Yes	104(59.1%)	72(40.9%)	176(48.1%)	13	(5, 17)

patients, about 174(47.5%) of the patients has unemployed while 192(52.5%) has employed. The median symptomatic recovering time of the MDD patient has unemployed and has employed 9 and 6 months respectively, and of which 39.7% and 67.2% were recovered respectively.

Regarding to ethnicity of patients, the major ethnic groups identified in the study area were Oromo 242(66.1%), of which 130(53.7%) of them have recovered, 85(23.2%) of them are Amhara of which 43(50.6%) have recovered and the other ethnicities are 39(10.7%) of which 25(64.1%) of them have recovered. Orthodox account for 95(25.9%) of the population, muslims account for 206(56.3%) of the population, protestant account for 46(12.6%) where as 19(5.2%) were other's religion of patients.

Regarding to substance abuse of patients, about 176(48.1%) of patients were substance abused while, 190(51.9%) were not substance abused. Among patients who were experienced recovery of MDD 69(39.2%) were substance abused while, 129(67.9%) were not substance abused. The median recovering time for substance abused and not substance abused patients were 12 and 5 months respectively.

Patients who have other cofactors seems that they stay longer time 10 months than who had no other cofactors 5 to experience recovery. Out of total, 202(55.2%) had no other cofactors while 164(44.8%) had a other cofactors. From those who have other cofactors 36(22%) were recoverd and 128(78%) were censored.

Lastly, regarding event of relapse, 190(51.9%) had not event of relapse while, 176(48.1%) had event of relapse. Among patients who were experienced first symptomatic recovery from MDD 126(66.3%) had not event of relapse while, 72(40.9%) had event of relapse. The median recovering time for who had event of relapse and who had not event of relapse patients were 6 and 13 months respectively.

4.2 Survival of Significantly Different Groups

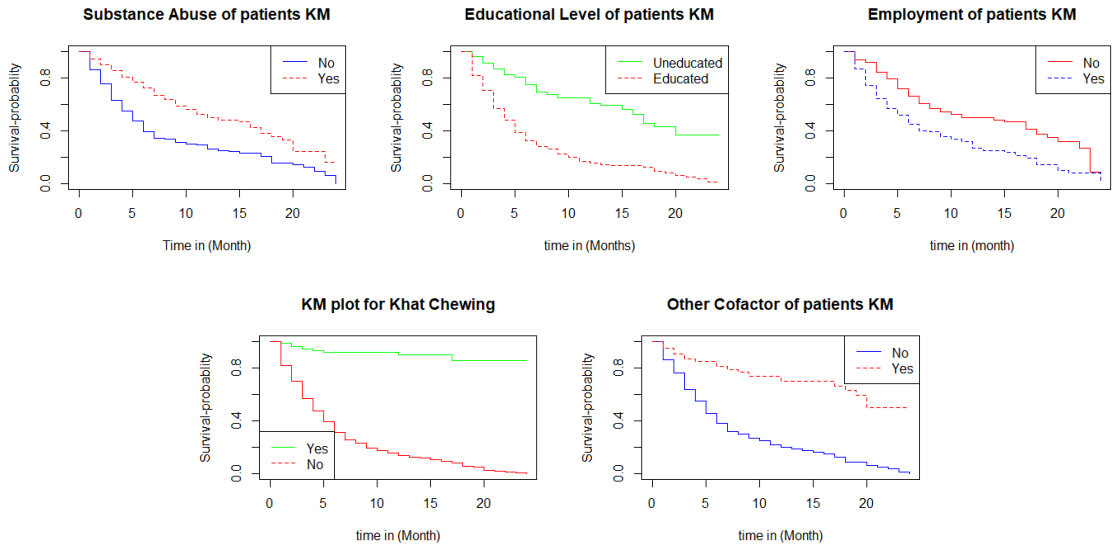

Figure 4.1: The survival functions of the significantly different groups

The survival time to first symptomatic recovery for patients whose abused substance is greater than those patients who were didn't abused substance. This indicates that the probability of prolonging symptomatic recovery time at a given specific time is greater for patients whose abused substance (Figure 4.1). The result of log rank test also revealed that difference is significant at 5% level of significance (p = < 0.001)(see Annex 3).

The survival time to first symptomatic recovery for patients who were educated is less than those patients who were uneducated (Figure 4.1). This indicates that educated patients had shorter symptomatic recovery time than uneducated patients. The result of log rank test also revealed that difference is significant at 5% level of significance (p = < 0.001) (see Annex 3).

The survival time to first symptomatic recovery for patients who were employed is less than those patients who were unemployed (Figure 4.1). This indicates that employed patients had shorter

symptomatic recovery time than unemployed patients. The result of log rank test also revealed that difference is significant at 5% level of significance (p = < 0.001) (see Annex 3).

The survival time to first symptomatic recovery for patients whose khat chewers is greater than those patients who were non chewers (Figure 4.1). This means, the probability of prolonging symptomatic recovery time at a given time for the khat chewers group of patients is greater as compared to the nonchewers group. The result of log rank test also revealed that difference is significant at 5% level of significance (p = < 0.0001)(see Annex 3).

The survival time to symptomatic recovery for patients whose had other cofactors is greater than those patients who hadn't other cofactor (Figure 4.1). This indicates that probability of prolonging symptomatic recovery time a given time for the who had other cofactor group of patients is greater as compared to the who hadn't other cofactor. The result of log rank test also revealed that difference is significant at 5% level of significance (p = < 0.0001) (see Annex 3).

4.3 Univariable Analysis

In the univariable analysis, covariates with p-value less than or equal to 25% were considered for multivariable analysis. Then, the multivariable models were fitted including all the potential covariates that were significant at 25% level of significance at the univariate level (see Annex 1). From the univariable analysis we observed that the covariate gender of patients, age, marital status, family history of mental illness, khat chewing, educational level, employment status, substance abuse, other cofactor and event of relapse history of patients were significant in the entire models used. This indicates that they are important prognostic factor for the time to first symptomatic recovery of MDD patients.

However, first onset age, religion and ethnicty of MDD patients were not a significant factor for the first symptomatic recovery of MDD according to all the candidate models (i.e, Weibull-Gamma,

Weibull-Inverse Gaussian, Loglogistic-Gamma, Loglogistic-Inverse-Gaussian, Lognormal-Gamma and Lognormal-Inverse Gaussian). Therefore, based on this result, it is better to ignore these covariates and shall do our multivariable analysis using the significant factors. Hence, the effects of the gender of patients, age, marital status, family history, khat chewing, educational level, employment, substance abuse, other cofactor and event of relapse history of patients on the time of first symptomatic recovery from MDD shall better be interpreted using the multivariable analysis.

4.4 Multivariable Analysis and Model Comparison

For time-to-first symptomatic recovery from MDD, the multivariable survival models of the Weibull, Loglogistic and Lognormal for the baseline hazard function; and the Gamma and the Inverse Gaussian frailty distributions were fitted again by assuming all the significant covariates in the univariable analysis at 25% level of significance. The output of the Lognormal-Inverse-Gaussian multivariable frailty model is presented in Table 4.3; and the output of the other multivariable frailty models were similarly drawn (see Annex 2).

The variance of the random effect or frailty θ is significant for all baseline frailty models at 5% level of significance. It is highest when we assume the inverse gaussian frailty distribution $(\theta = 0.21)$ followed by the gamma distribution $(\theta = 0.172)$ with the lognormal baseline hazard function. The Kendall's tau τ is used to measure the dependence within the clusters (woredas) and it is higher for the higher variance of random effect θ values. Accordingly, the dependence within the clusters for the lognormal-inverse-gaussian frailty model ($\tau = 0.081$) is the maximum followed by the lognormal gamma frailty model ($\tau = 0.079$). This indicates that within group correlation on times-to-first symptomatic recovery of MDD within the clusters (woredas).

The most commonly used methods include information and likelihood based criteria. For shared frailty models information based criteria is used while for the nested frailty model likelihood ratio test is used. Thefore, to compare the Gamma and Inverse Gaussian shared frailty models with

Weibull, Loglogistic and Lognormal hazard functions, this study used information criteria. The most commonly used model selection are the Akaike Information criterion (AIC) and Bayesian Information criterion (BIC). The model with the smallest AIC value is considered a better fit. The AIC value of the Lognormal-Inverse-Gaussian model that i.e. 1172.549 is the minimum from all the other AIC values of the models which indicates that it is the most efficient model to describe the major depresive disorder dataset among the different parametric shared frailty models (Table 4.2)

Table 4.2: AIC, BIC and LRT values of the models used in the study

Baseline hazard function	Frailty distribution	AIC	BIC	LRT	τ
Weibull	Gamma	1197.6797	1316.999	232	0.026
	Inverse-Gaussian	1192.313	1282.455	223	0.027
Loglogistic	Gamma	1183.948	1302.879	235	0.015
	Inverse-Gaussian	1174.908	1244.216	218	0.014
Log-normal	Gamma	1181.907	1300.928	225	0.079
	Inverse-Gaussian	1172.549	1244.136	210	0.081

The Lognormal-Inverse-Gaussian frailty model result showed that the marital status of patients, khat chewing, educational level, employment, substance abuse and other cofactors were significant at 5% level of significance (Table 4.3). This indicates that they were the contributing factor for the first symptomatic recovery of MDD patients. However, according to this model the gender, family history of mental illness of patients, event of relapse and age of patient's has no significant effect on the first symptomatic recovery of MDD patients.

The result of this study suggested that marital status of patients had a significant effect on the first symptomatic recovery status of MDD patients; the acceleration factor of divorced patients was 1.858 times higher than those who were single (ϕ = 1.858), which means that the survival time of single patients was reduced by 85.8% when compared with those who had divorced mari-

tal status. Therefore, patients with divorced marital status had prolonged time to first symptomatic recovery from MDD by a factor of 1.858 than the patients with single marital status.

Similarly, khat chewing had a significant effect on the first symptomatic recovery status of MDD patients; the acceleration factor of patients who chew khat was 2.466 times that of patients who did not chew khat ($\phi = 2.466, 95\% CI = 2.125, 2.807$), which means that patients who chew khat had longer symptomatic recovery time from MDD by a factor of 2.466 than those who did not chew khat. Therefore, khat chewers were less likely to had first symptomatic recovery than their counterparts.

Looking at the effect of education status, after adjusting other confounding variables, the acceleration factor of being recovered of patients with educated status was 0.596 times less than the factor of those with uneducated status ($\phi = 0.596, CI : 0.323, 0.867$); this indicates that the symptomatic recovery time of patients who were educated was reduced by 40.4% when compared with patients who were uneducated. Therefore, symptomatic recovery time was shorter for educated patients.

Employment status was another covariate which had a significant impact on the symptomatic recovery time of patients; the acceleration factor for being recovered of patients who were employed was 0.658 times less than that of patients who were unemployed ($\phi = 0.658, CI : 0.406, 0.911$), indicating that the symptomatic recovery time of patients who employed was reduced by 34.2% when compared with patients who were unemployed. Therefore, employed patients had shorter symptomatic recovery time than unemployed patients.

Furthermore, holding other covariates constant and accounting for frailty, substance abuse had also a significant effect on the symptomatic recovery time of mental patients. The acceleration factor of being recovered of mental patients who were substance abused was 1.487 times the factor of those who were not substance abused ($\phi = 1.487, CI : 1.224, 1.749$), which means that the

survival time of substance abused patients was 48.7% less when compared with non-substance abused patients. Thus, substance abused patients had longer symptomatic recovery time than the non-substance abused patients.

Looking at the effect of other co-factors, after adjusting other covariates, patients who had other co-factors were found to be associated with high survival time, whose acceleration factor was 1.633 times that of the patients without other co-factors ($\phi = 1.633, 95\%CI = 1.337, 1.929$), which means that the symptomatic recovery of patients who had other co-factors was decreased by about 63.3% when compared to patients without other co-factors and the decrements could be as low as 33.7% and as high as 92.9%. Thus, patients with others co-factors were less likely to have first symptomatic recovery than their counterparts.

The value of the shape parameter in the lognormal-inverse Gaussian frailty model is σ = 3.56. This value is greater than unity that indicates the shape of hazard function is unimodal, i.e. it increases up to some time and then decreases. The variability (heterogeneity) in the population of clusters (woredas) estimated by our best model is θ = 0.21, and the dependence within clusters is about τ = 8.1%.

Table 4.3: Lognormal-Inverse Gaussian Multivariable Analysis

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender	Male	Ref		1		
	Female	-0.149	0.129	0.862	[0.609 1.114]	0.25
Age of patients (in years)	13-19	Ref		1		
	20-25	-0.057	0.215	0.944	[0.523 1.366]	0.79
	26-49	-0.031	0.211	0.969	[0.556 1.383]	0.88
	≥ 50	0.38	0.255	1.471	[0.962 1.962]	0.13
Martital status	Single	Ref		1		
	Married	-0.212	0.143	0.808	[0.529 1.089]	0.14
	Widowed	0.3215	0.240	1.379	[0.909 1.849]	0.18
	Divorced	0.6195	0.230	1.858	[1.407 2.309]	0.0071 *
Family History	No	Ref		1		
	Yes	0.1419	0.132	1.1523	[0.894 1.411]	0.28
Chewing Khat	No	Ref		1		
	Yes	0.9028	0.174	2.466	[2.125 2.807]	$\leq$0.001 ***
Educational Level	Uneducated	Ref		1		
	Educated	-0.517	0.138	0.596	[0.323 0.867]	$\leq$0.001**
Employment	No	Ref		1		
	Yes	-0.4179	0.129	0.658	[0.406 0.911]	0.0012**
Substance Abuse	No	Ref		1		
	Yes	0.3966	0.134	1.487	[1.224 1.749]	0.003 **
Other Cofactors	No	Ref		1		
	Yes	0.4905	0.151	1.633	[1.337 1.929]	0.0011 ***
Event of Relapse	No	Ref		1		
	Yes	0.2058	0.132	1.228	[0.969 1.487]	0.12
	$\theta = 0.21$	$\tau = 0.081$	μ=0.000131	$\sigma = 3.56$	AIC = 1172.549	

4.5 Model Diagnostics

4.5.1 Diagnostic plots of parametric baselines

The final step in the model assessment is to see the overall goodness of fit. Therefore, it is desirable to determine whether a fitted parametric model adequately describes the data or not. To check the adequacy of our baseline hazard: weibull is plotted by $\log(-\log(\hat{S}(t)))$ with the logarithm of time of the study; the log-logistic is plotted by log odds of failure or $\log(\frac{1-\hat{S}(t)}{\hat{S}(t)})$ with the logarithm of time and the log-normal is plotted by the qnorm(1-survival) or $\Phi^{-1}[1-\hat{S}(t)]$ with the logarithm of time (Figure 4.2). The plot of lognormal is more linear than the other plots. The patterns suggests that the lognormal hazard function is appropriate in the model.

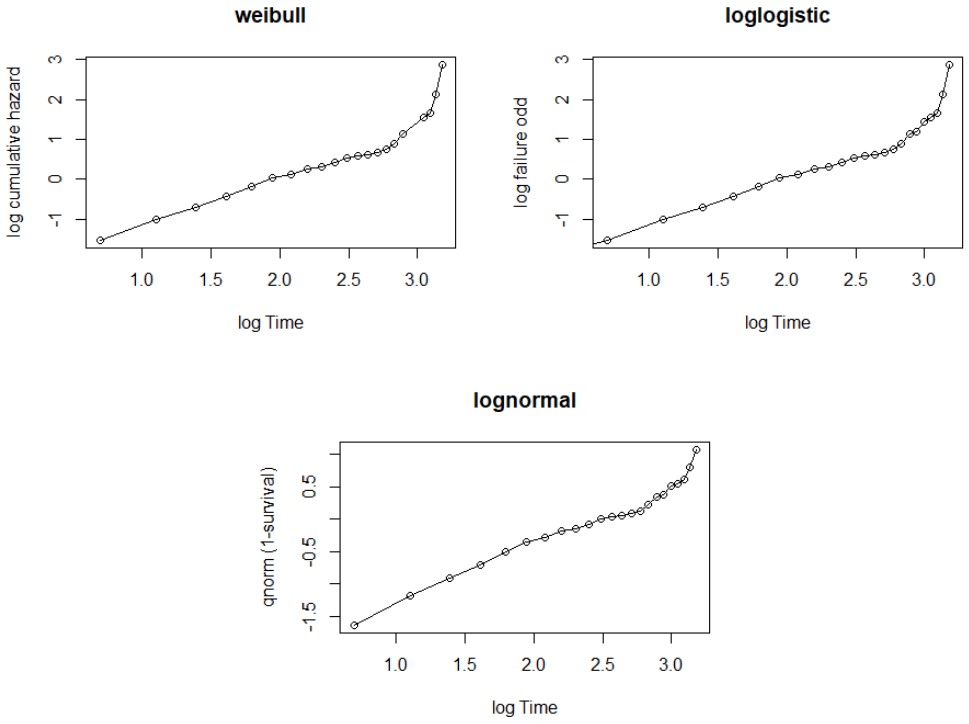

Figure 4.2: Graphical Evaluation of the weibull, loglogistic and lognormal assumptions

4.5.2 The Cox Snell residual plots

The Cox-Snell residuals are one way to investigate how well the model fits the data. In this case we used the Cox-Snell residuals to check the overall goodness of fit for different parametric models. The Cox- Snell residuals obtained from fitting the lognormal model to our data via maximum likelihood estimation. By comparing with Weibull and Loglogistic, this plot shows tha the line to the Cox-Snell residuals of the lognormal models were nearest to the line through the origin, again indicating that this model describes the MDD dataset well.

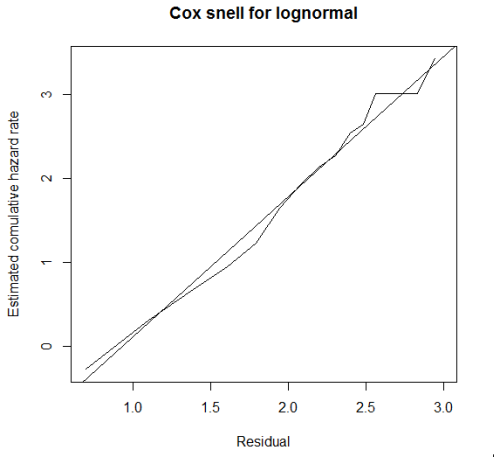

Figure 4.3: Cox-Snell residuals obtained by fitting lognormal to the MDD dataset.

4.5.3 Adequacy of accelerated failure time

A quantile-quantile or q-q plot is made to check if the accelerated failure time provided an adequate fit to the data using two different groups of population. We shall graphically check the adequacy of the accelerated failure-time model by comparing the significantly different educational level (educated, uneducted), employment status (employed, unemployed), marital status (single, married, divorced, widowed), chewing khat (yes, no), other cofactor (yes, no) and substance abuse (yes, no) (Figure 4.4). The figures appear to be approximately linear for all covariates.

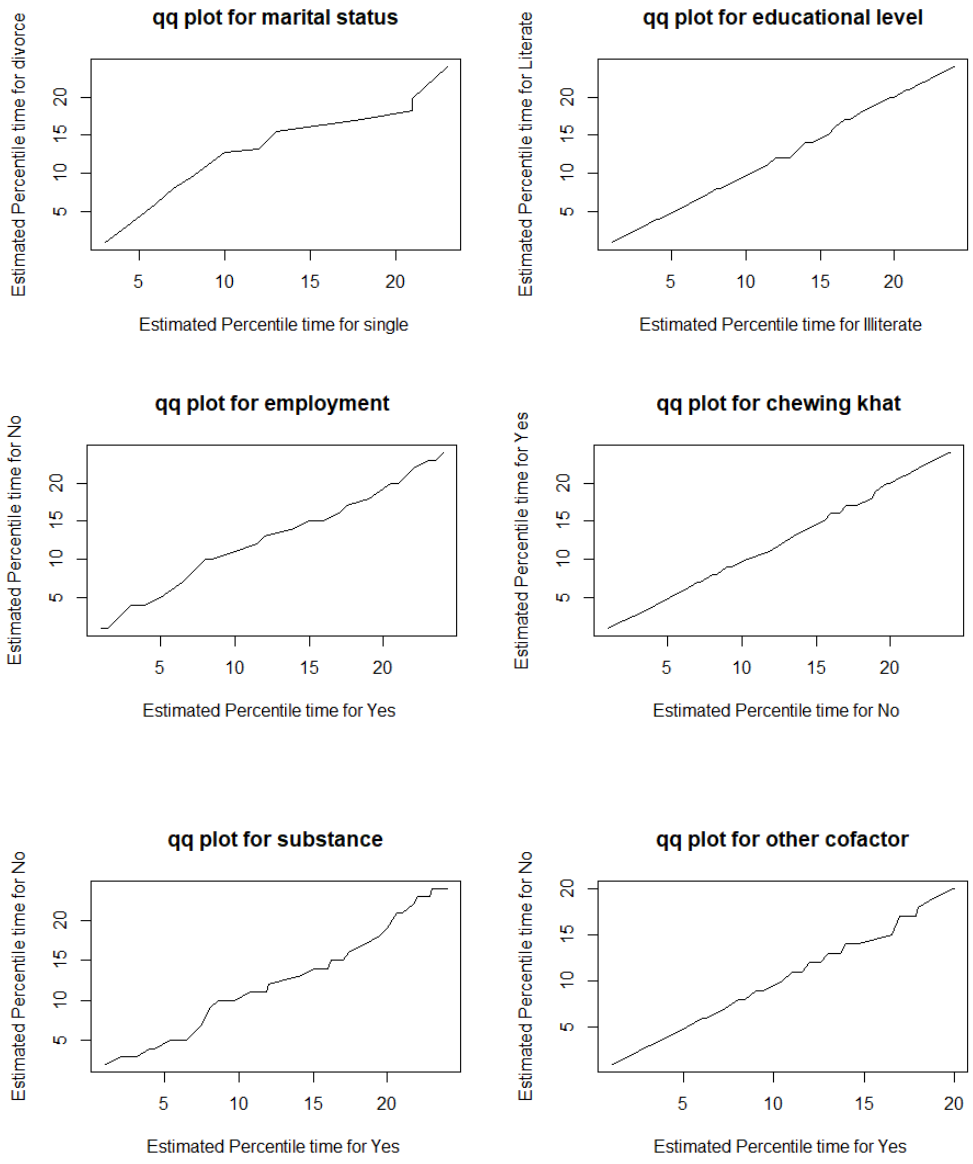

Figure 4.4: Q-Q plots to check the adequacy of accelerated failure time model

4.6 Discussion

The findings of this study revealed that being educated and employed status significantly shorten/decelerate the time-to-first symptomatic recovery from major depressive disoreder, while being divorced, khat

chewers, abused substance and with other cofactors accelerates time-to-first symptomatic recovery among major depressive disoreder patients in Jimma University Medical Center.

In this study, of 366 MDD individuals under psychiatric follow up during the time period, $198(54.1\%)$ of them were faced first symptomatic recovery whereas $168(45.9\%)$ of them were censored. This result seems to imply that the majority of MDD patients, who were selected as a sample for study, were achieved first symptomatic recovery. In complement with this, the study done by Novic *et al.* [119], retorted that among the patients considered, 52.1% were achieved recovery and 47.9% did not achieved recovery.

Lognormal-Inverse-Gaussian shared frailty model having minimum AIC value selected as best fit the MDD data set. Lognormal really shines for skewed distributions, large variances (i.e, data with a large standard deviation), and all-positive values. Additionally, if we were to take the natural log of each random variable and its result is a normal distribution, then the Lognormal is the best fit.

This research also showed that there was a clustering (frailty) effect on modeling time-to-first symptomatic recovery from MDD which might be due to the heterogeneity in woreda from which the patients came-from which means that patients' coming from the same woreda share similar risk factors related to MDD. Clusters with minimum median time have smaller frailties, so that these clusters are predicted to have a high hazard and more probable to first symptomatic recovery [23]. These nuisance terms modify the hazard function, so that the hazard function should be evaluated conditionally on this effect. Woredas frail more are more likely to symptomatic recovery than the less frail woredas (since the event is positive).

Results of this study showed that educational level of patients was significantly influenced time to first symptomatic recovery from MDD. Patients who were educated had more likely symptomatic recovery time from MDD by a factor of 0.596 than those who were uneducated. The result

is similar with studies conducted by Novic *et al.* [119]. They suggested that patients achieving recovery were more likely to have a higher level of education (p = 0.0075). Additionally, the likelihood of achieving recovery was negatively associated with lower levels of education [120]. The result is similar with studies conducted in South Africa and Turkey which described that the possible explanation for this could be the fact that individuals with low socioeconomic and educational status were given less value to their self-esteem and live a stressful life as compared with educated individuals. In addition, educated people have better understanding of the risk factors of depression compared to uneducateds [10, 67, 69].

The findings of this study implied that marital status of patients significantly influenced time to first symptomatic recovery from MDD. Acceleration factor of ϕ = 1.858 indicates that patients with divorced marital status had prolonged time to symptomatic recovery from major depressive disorder as compared to the patients with single marital status. Study conducted by Andualem Mossie and Novic *et al.* [71, 119] also supports this fact. This study is also consistent with the study conducted by L.Gu and J.Xie, Mogga and Deyessa [73, 74, 75] which stated that depression has shown a significant association with marital status. Widowed individuals were five times more likely to develop depression as compared to single individuals. This could be due to the fact that the loss of a spouse or lovers has been identified as one of the most stressful life events, requiring more psychological therapy compared to many others.

The result of this study also revealed that the khat chewing is another risk factor for the first symptomatic recovery from MDD. The finding was supported by studies done in Jimma University and Ambo University [71, 70]. The probability of developing depression among khat chewers is more likely than that among nonchewers. This result is consistent with the study conducted in Jimma University by Gelaw *et al.* [80] and Tekalign *et al.* [81] with WHO expert analysis [82] which revealed that depression was significantly associated with khat chewing. The probability of developing depression episodes among khat chewers is tenfold higher than that among nonchewers.

The results of this study suggested that employment status severely impacted patients recovery. For patient those who were employed, the acceleration factor was less than one (ϕ = 0.658) which means that patients who were employed took shorter time to recovery as compared to unemployed. The finding was supported by a study done in United States [67]. This finding is consistent with a study done by Ermias Mekonnen & Samuel Esayas. They reported that Unemployed individuals had shown increased risk of having mental illness than professionals [83]. In a study of Unemployment and Depression Among Emerging Adults, the risk of depression is higher among the unemployed than among the employed. In bivariate analyses, depression was more likely among unemployed emerging adults compared with employed emerging adults (p<.001). The odds of depression were about 3 times higher for unemployed than employed emerging adults [84].

Substance abuse has been identified as prognostic factor for first symptomatic recovery from MDD. Substance abused patients were less likely to recover from MDD than that of non-substance abused patients ($\phi = 1.487$). This result is supported by study done in mekelle General Prison Center [34] and study conducted in Borena smi-nomadic community [87]. In addition, this finding is in agreement with the study done by Dierker [121] who retorted that substance use increases the risk of major depressive disorder. Similary, study conducted by Covey stated that persons with major depression tend to abuse substances and have difficulties when they try to stop. There are thousands of chemicals other than nicotine present in cigarette smoke, of which one or several may affect mood in the same way as a group of antidepressant medications called monoamine oxidase inhibitors or does. These medications effectively increase levels of specific neurotransmitters involved in the regulation of mood. Smoking, therefore, may be one way for depressed individuals to alleviate depressive symptoms [86].

Furthermore, the study findings showed that other cofactor of patients significantly influenced time to symptomatic recovery from MDD. Patients those who were with other cofactor had longer

recovery time by a factor of 1.633 than the patients with no other cofactor. The result is similar with study conducted by Leonard which suggested that the 12-month prevalence and odds of major depression are high in individuals with chronic medical conditions, and major depression is associated with significant increases in utilization, lost productivity and functional disability [88]. Similarly, the result was supported by other studies done in Malaysia and Ethiopian [12, 122].

This study also showed that there was a clustering (frailty) effect on modeling time to first symptomatic recovery of MDD which might be due to the heterogeneity with in woredas from which the patients came from. Assuming patients coming from the same woreda share similar risk factors related to MDD, indicating that it was important considering the clustering effect in modeling the hazard function. The heterogeneity in the woredas was significant and estimated to be $\theta = 0.21$, and the dependence within clusters is about $\tau = 0.081(8.1\%)$. These values were the maximum among the variance of the random effects and the Kendall's tau of all the models. This result consolidate the idea that larger values of q, indicates that there is a higher degree of heterogeneity among groups and strong association within groups [123].

4.7 Limitations of the study

However, the thesis is not done without limitation. Though prognostic factors for recovery of MDD are many, the research is limited only to the thirthen covariates. This is because the patient's card consists more of the characteristics that are not related to the recovery of MDD and some relevant covariates like economic status, social relationship, loneliness and low wealth index has not been recorded. Since these are the expected risk factors from many literature's and lack of prior research studies on the topic.

5 Conclusion and Recommendation

5.1 Conclusion

This study used survival time of Major Depressive Disorder patients data set for those patients who were received treatments and under follow up from September 1, 2018 through August 31, 2020, with the aim of modeling the determinants of time to first symptomatic recovery of MDD patients in Jimma University Medical Center.

The Lognormal-Inverse-Gaussian frailty model is the model that best described the time to first symptomatic recovery from Major Depressive Disorder patients data set. The result of lognormal-inverse-gaussian shared frailty model showed that marital status, khat chewing, employment status, educational level, substance abuse and other cofactor were found to be statistically significant risk factors for first symptomatic recovery of major depressive disorder patients.

According to the study, the median first symptomatic recovery time from major depressive disorde was 7 months. There is a frailty (clustering) effect on the time-to-first symptomatic recovery from major depressive disorders' that arises due to heterogeneity in between the woredas.

5.2 Recommendation

Based on the conducted study, different significant factors were identified for the first symptomatic recovery of major depressive disorder patients. The following recommendations were made for the ministry of health, policy makers, psychiatrists or clinicians and the public at large.

- Good treatments have to be given for the society on the risk of divorce of marriage by giving the treatments that start from the marriage dissolution case.
- Uneducated patients should have to get emphasis and good treatments in accordance with their understanding level in order to recover from their illness.

- Patients with khat chewing habits were less likely to recover from the disorder. Therefore, it has to be needed to work against such bad habits by considering it as one of the treatment parts.
- The researcher recommended that all the concerned bodies should take their parts in helping major depressive disorder patients who are substance abused.
- The governments and non-government sectors will be devote huge efforts in order to reduce the effects of unemployment.
- The health workers, especially psychiatrists, should be cautious when a patient is under a follow up of a mental case with who had other cofactors.
- The researcher suggested that if those identified risk factors could be well managed, many mental health problems are preventable and that there is considerable scope for increasing interventions that reduce the incidence of people developing major depressive disorder problems and increase the potential for sustained recovery after illness.
- The additional information of the patients history such as economic status, drug side effects, loneliness, poor social support and severity of MDD using charting the MDD illness should be recorded in patients card and included in the further studies.

References

[1] Depression, W. H. O. (2017). Other common mental disorders: global health estimates. Geneva: World Health Organization, 1-24.

[2] Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. Annual review of public health, 34, 119-138.

[3] World Health Organization. (2016). Health and Social Effects of Nonmedical Cannabis Use (The). World Health Organization.

[4] Malhi, G. S., & Mann, J. J. (2018). Course and prognosis. Lancet, 392, 2299-312.

[5] Edition, F. (2013). Diagnostic and statistical manual of mental disorders. Am Psychiatric Assoc, 21.

[6] Parker, G. F. (2014). DSM-5 and psychotic and mood disorders. Journal of the American Academy of Psychiatry and the Law Online, 42(2), 182-190.

[7] Dadi, A. F., Miller, E. R., Bisetegn, T. A., & Mwanri, L. (2020). Global burden of antenatal depression and its association with adverse birth outcomes: an umbrella review. BMC Public Health, 20(1), 1-16.

[8] Dewey, C. (2013). A stunning map of depression rates around the world. The Washington Post, 7.

[9] Ezzati, M., Vander Hoorn, S., Lopez, A. D., Danaei, G., Rodgers, A., Mathers, C. D., & Murray, C. J. (2006). Comparative quantification of mortality and burden of disease attributable to selected risk factors. Global burden of disease and risk factors, 2, 241-396.

[10] Tomlinson, M., Grimsrud, A. T., Stein, D. J., Williams, D. R., & Myer, L. (2009). The epidemiology of major depression in South Africa: results from the South African Stress and Health study: mental health. South African Medical Journal, 99(5), 368-373.

[11] Amoran, O., Lawoyin, T., & Lasebikan, V. (2007). Prevalence of depression among adults in Oyo State, Nigeria: A comparative study of rural and urban communities. Australian Journal of Rural Health, 15(3), 211-215.

[12] Hailemariam, S., Tessema, F., Asefa, M., Tadesse, H., & Tenkolu, G. (2012). The prevalence of depression and associated factors in Ethiopia: findings from the National Health Survey. International journal of mental health systems, 6(1), 1-11.

[13] Bedaso, A., Kediro, G., & Yeneabat, T. (2018). Factors associated with depression among prisoners in southern Ethiopia: a cross-sectional study. BMC research notes, 11(1), 1-6.

[14] Hanlon, C., Medhin, G., Alem, A., Araya, M., Abdulahi, A., Hughes, M., ... & Prince, M. (2008). Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. Journal of affective disorders, 108(3), 251-262.

[15] Greenberg, P. E., Fournier, A. A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). The Journal of clinical psychiatry, 76(2), 0-0.

[16] Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., & Morganstein, D. (2003). Cost of lost productive work time among US workers with depression. Jama, 289(23), 3135-3144.

[17] Kessler, R. C., Akiskal, H. S., Ames, M., Birnbaum, H., Greenberg, P., . A, R. M., ... & Wang, P. S. (2006). Prevalence and effects of mood disorders on work performance in a nationally representative sample of US workers. American journal of psychiatry, 163(9), 1561-1568.

[18] Zergaw, A. (2008). Economic burden of schizophrenia and bipolar disorder in rural Ethiopia. Addis Ababa University. Ethiopia.

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

[20] Klein, J. P., & Moeschberger, M. L. (2003). Survival analysis: techniques for censored and truncated data (Vol. 1230). New York: Springer.

[21] Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2) , 187-202

[22] Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.

[23] Goethals, K., Janssen, P., & Duchateau, L. (2008). Frailty models and copulas: similarities and differences. Journal of Applied Statistics, 35(9), 1071-1079.

[24] Wienke, A. (2011). Frailty models in survival analysis. CRC press.

[25] Sastry, N. (1997). A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil. Journal of the American Statistical Association, 92(438), 426-435.

[26] Abrahantes, J. C., Legrand, C., Burzykowski, T., Janssen, P., Ducrocq, V., & Duchateau, L. (2007). Comparison of different estimation procedures for proportional hazards model with random effects. Computational statistics & data analysis, 51(8), 3913-3930.

[27] Munda, M., Rotolo, F., & Legrand, C. (2012). parfm: Parametric frailty models in R. Journal of Statistical Software, 51(11), 1-20.

[28] Ha, I. D., & Lee, Y. (2003). Estimating frailty models via Poisson hierarchical generalized linear models. Journal of Computational and Graphical Statistics, 12(3), 663-681.

[29] World Health Organization. (2001). The World Health Report 2001: Mental health: new understanding, new hope.

[30] World Health Organization. (2008). The global burden of disease: 2004 update. World Health Organization.

[31] Moledina, S. M., Bhimji, K. M., & Manji, K. P. (2018). Prevalence and associated factors of depression in an Asian community in Dar es Salaam, Tanzania. Psychiatry journal, 2018.

[32] Lando, J., Williams, S. M., Sturgis, S., & Williams, B. (2006). A logic model for the integration of mental health into chronic disease prevention and health promotion. Preventing Chronic Disease, 3(2).

[33] Uznanski, A., & Roos, J. L. (1997). The situation of mental health services of the World Health Organisation, African region, in the early 1990s.

[34] Welu, S. G., Aregawi, D. H., Gebreslassie, H. T., & Kidanu, K. G. (2021). Prevalence and Associated Factors of Depressive Disorder among Prisoners in Mekelle General Prison Center, Tigray, Ethiopia: A Cross-Sectional Study Design. Depression Research and Treatment, 2021.

[35] Molla, G. L., Sebhat, H. M., Hussen, Z. N., Mekonen, A. B., Mersha, W. F., & Yimer, T. M. (2016). Depression among Ethiopian adults: cross-sectional study. Psychiatry journal, 2016.

[36] Duko, B., Erdado, M., & Ebrahim, J. (2019). Prevalence and factors associated with depression among hospital admitted patients in South Ethiopia: cross sectional study. BMC research notes, 12(1), 1-4.

[37] Luppa, M., Luck, T., Ritschel, F., Angermeyer, M. C., Villringer, A., & Riedel-Heller, S. G. (2013). Depression and incident dementia. An 8-year population-based prospective study. PloS one, 8(3), e59246.

[38] Kim, G. E., Jo, M. W., & Shin, Y. W. (2020). Increased prevalence of depression in South Korea from 2002 to 2013. Scientific reports, 10(1), 1-9.

[39] Lye, M. S., Tey, Y. Y., Tor, Y. S., Shahabudin, A. F., Ibrahim, N., Ling, K. H., ... & Abdul Razak, N. A. (2020). Predictors of recurrence of major depressive disorder. PloS one, 15(3), e0230363.

[40] Lin, D. Y., Wei, L. J., & Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika, 80(3), 557-572.

[41] Trussell, J., & Rodriguez, G. (1990). Heterogeneity in demographic research. Convergent issues in genetics and demography, 111-132.

[42] Wienke, A. (2010). Frailty models in survival analysis. CRC press.

[43] Egede, L. E., Bishu, K. G., Walker, R. J., & Dismuke, C. E. (2016). Impact of diagnosed depression on healthcare costs in adults with and without diabetes: United States, 2004–2011. Journal of affective disorders, 195, 119-126.

[44] Brody, D. J., Pratt, L. A., & Hughes, J. P. (2018). Prevalence of depression among adults aged 20 and over: United States, 2013-2016.

[45] Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E. K., & Lanctôt, K. L. (2010). A meta-analysis of cytokines in major depression. Biological psychiatry, 67(5), 446-457.

[46] Gruenberg, A. M., Goldstein, R. D., & Pincus, H. A. (2005). Classification of depression: research and diagnostic criteria: DSM-IV and ICD-10. Biology of Depression, 11, 43.

[47] For the Revision, I. A. G. (2011). A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. World Psychiatry, 10(2), 86..

[48] Latendresse, G., Elmore, C., & Deneris, A. (2017). Selective serotonin Reuptake Inhibitors as First- Line Antidepressant Therapy for Perinatal Depression. Journal of midwifery & women's health, 62(3), 317-328.

[49] de Barros Pellegrinelli, K., de O. Costa, L. F., Silval, K. I. D., Dias, V. V., Roso, M. C., Bandeira, M., ... & Moreno, R. A. (2013). Efficacy of psychoeducation on symptomatic and functional recovery in bipolar disorder. Acta Psychiatrica Scandinavica, 127(2), 153-158.

[50] Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., ... & Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS medicine, 10(11), e1001547.

[51] Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. H., Frank, E., ... & Schatzberg, A. F. (2006). Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology, 31(9), 1841-1853.

[52] Thase, M. E. (2003). Effectiveness of antidepressants: comparative remission rates. Journal of Clinical Psychiatry, 64(1), 3-7.

[53] Sheehan, D. V., Harnett-Sheehan, K., Spann, M. E., Thompson, H. F., & Prakash, A. (2011). Assessing remission in major depressive disorder and generalized anxiety disorder clinical trials with the discan metric of the Sheehan disability scale. International Clinical Psychopharmacology, 26(2), 75-83.

[54] Sacchetti, E., Frank, E., Siracusano, A., Racagni, G., Vita, A., & Turrina, C. (2015). Functional impairment in patients with major depression in clinical remission: results from the VIVAL-D-Rem, a nationwide, naturalistic, cross-sectional survey. International clinical psychopharmacology, 30(3), 129-141.

[55] Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. PloS one, 9(2), e90311.

[56] Hamilton, M. (1960). A rating scale for depression. Jourrial qf. Neurology, Neurosurgery and Psychiatry, 12, 56, 62.

[57] Keller, M. B. (2004). Remission versus response: the new gold standard of antidepressant care. Journal of Clinical Psychiatry, 53-59.

[58] Cusin, C., Yang, H., Yeung, A., & Fava, M. (2010). Chapter 2. Rating Scales for Depression. Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health, Current Clinical Psychiatry.

[59] Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., ... & Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. Archives of general psychiatry, 48(9), 851-855.

[60] Fava, G.A., Ruini, C. and Belaise, C., 2007. The concept of recovery in major depression. Psychological Medicine, 37(3), p.307.

[61] Sinyor, M., Rezmovitz, J., & Zaretsky, A. (2016). Screen all for depression.

[62] World Health Organization. (2012). Public health action for the prevention of suicide: a framework.

[63] Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., & Morganstein, D. (2003). Cost of lost productive work time among US workers with depression. Jama, 289(23), 3135-3144.

[64] Aminu, A. S., Chandrasekaran, V., & Nair, S. (2017). Depression among patients with diabetes: a community-based study in South India. Journal of Medical Sciences, 37(6), 237.

[65] Lyness, J. M., Niculescu, A., Tu, X., Reynolds III, C. F., & Caine, E. D. (2006). The relationship of medical comorbidity and depression in older, primary care patients. Psychosomatics, 47(5), 435-439.

[66] Cassano, P., & Fava, M. (2002). Depression and public health: an overview. Journal of psychosomatic research, 53(4), 849-857.

[67] Centers for Disease Control and Prevention (CDC. (2010). Current depression among adults—United States, 2006 and 2008. MMWR. Morbidity and mortality weekly report, 59(38), 1229-1235.

[68] Khawaja, N. G., & Duncanson, K. (2008). Using the university student depression inventory to investigate the effect of demographic variables on Students' depression. Journal of Psychologists and Counsellors in Schools, 18(2), 195-209.

[69] Topbaş, M., Yaris, F., Can, G., Kapuca, M., & Sayar, K. (2004). The Effect of Sociodemographic Factors and Medical Status on Depressive Symptoms of the Inhabitants of Nursing Home and Community-Dwelling Elders in Trabzon-Turkey. Klinik Psikofarmakoloji Bulteni, 14(2).

[70] Birhanu, A., & Hassein, K. (2016). Prevalence and factors associated to depression among Ambo university students, Ambo, West Ethiopia. Prevalence, 25.

[71] Yeshaw, Y., & Mossie, A. (2017). Depression, anxiety, stress, and their associated factors among Jimma University staff, Jimma, Southwest Ethiopia, 2016: a cross-sectional study. Neuropsychiatric disease and treatment, 13, 2803.

[72] Reta, Y., Getachew, R., Bahiru, M., Kale, B., & Workie, K. (2019). Depressive disorder and its associated factors among prisoners in Debre Berhan Town, North Showa, Ethiopia. bioRxiv, 703223.

[73] Gu, L., Xie, J., Long, J., Chen, Q., Chen, Q., Pan, R., ... & Su, L. (2013). Epidemiology of major depressive disorder in mainland china: a systematic review. PloS one, 8(6), e65356.

[74] Mogga, S., Prince, M., Alem, A., Kebede, D., Stewart, R., Glozier, N., & Hotopf, M. (2006). Outcome of major depression in Ethiopia: population-based study. The British Journal of Psychiatry, 189(3), 241-246.

[75] Deyessa, N., Berhane, Y., Alem, A., Hogberg, U., & Kullgren, G. (2008). Depression among women in rural Ethiopia as related to socioeconomic factors: a community-based study on women in reproductive age groups. Scandinavian journal of public health, 36(6), 589-597.

[76] Avenevoli, S., Baio, J., Bitsko, R. H., Blumberg, S. J., Brody, D. J., Crosby, A., ... & Visser, S. N. (2013). Mental health surveillance among children–United States, 2005-2011.

[77] Ralph, E. (2004). Depression in children and adolescents: Information for parents and educators. National Association Of School Psychologists: Social/Emotional Development. Nova South-Eastern University.

[78] Khan, M. S., Mahmood, S., Badshah, A., Ali, S. U., & Jamal, Y. (2006). Prevalence of depression, anxiety and their associated factors among medical students in Karachi, Pakistan. Journal-Pakistan Medical Association, 56(12), 583.

[79] Ahmed, G., Negash, A., Kerebih, H., Alemu, D., & Tesfaye, Y. (2020). Prevalence and associated factors of depression among Jimma University students. A cross-sectional study. International journal of mental health systems, 14(1), 1-10.

[80] Gelaw, Y., & Haile-Amlak, A. (2004). Khat chewing and its socio-demographic correlates among the staff of Jimma University. Ethiopian Journal of Health Development, 18(3), 179-184.

[81] Damena, T., Mossie, A., & Tesfaye, M. (2011). Khat chewing and mental distress: a community based study, in jimma city, southwestern ethiopia. Ethiopian journal of health sciences, 21(1), 37-46.

[82] World Health Organization. (2004). Promoting mental health: Concepts, emerging evidence, practice: Summary report. World Health Organization.

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

[84] McGee, R. E., & Thompson, N. J. (2015). Peer reviewed: unemployment and depression among emerging adults in 12 states, behavioral risk factor surveillance system, 2010. Preventing chronic disease, 12.

[85] Koenig, H. G., Mccullough, M. E., & Larson, D. B. (2001). The manual of religion and health.

[86] Covey, L. S., Glassman, A. H., & Stetner, F. (1998). Cigarette smoking and major depression. Journal of addictive diseases, 17(1), 35-46.

[87] Beyero, T., Alem, A., Kebede, D., Shibire, T., Desta, M., & Deyessa, N. (2004). Mental disorders among the Borana semi-nomadic community in Southern Ethiopia. World Psychiatry, 3(2), 110.

[88] Egede, L. E. (2007). Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. General hospital psychiatry, 29(5), 409-416.

[89] Wintrebert, C. M. A. (2007). Statistical modelling of repeated and multivariate survival data (Doctoral dissertation, Leiden University).

[90] Clayton, D., & Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. Journal of the Royal Statistical Society: Series A (General), 148(2), 82-108.

[91] Li, Y., & Lin, X. (2000). Covariate measurement errors in frailty models for clustered survival data. Biometrika, 87(4), 849-866.

[92] Economou, P., & Caroni, C. (2005). Graphical tests for the assumption of gamma and inverse Gaussian frailty distributions. Lifetime Data Analysis, 11(4), 565-582.

[93] Nguti, R. W. (2003). Random effects survival models applied to animal breeding data. LUC.

[94] Hougaard, P. (2012). Analysis of multivariate survival data. Springer Science & Business Media.

[95] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. Biometrika, 73(2), 387-396.

[96] Ripatti, S., & Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. Biometrics, 56(4), 1016-1022.

[97] Locatelli, I., Lichtenstein, P., & Yashin, A. I. (2003). A Bayesian correlated frailty model applied to Swedish breast cancer data. Università commerciale Luigi Bocconi.

[98] Manda, S. O. (2001). A comparison of methods for analysing a nested frailty model to child survival in Malawi. Australian & New Zealand Journal of Statistics, 43(1), 7-16.

[99] Hougaard, P. (1995). Frailty models for survival data. Lifetime data analysis, 1(3), 255-273.

[100] Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. Biometrika, 71(1), 75-83.

[101] Manton, K. G., Stallard, E., & Vaupel, J. W. (1986). Alternative models for the heterogeneity of mortality risks among the aged. Journal of the American Statistical Association, 81(395), 635-644.

[102] Pipper, C. B., & Martinussen, T. (2004). An estimating equation for parametric shared frailty models with marginal additive hazards. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(1), 207-220.

[103] Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. The Stata Journal, 2(1), 22-44.

[104] dos Santos, D. M., Davies, R. B., & Francis, B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. Journal of statistical planning and inference, 47(1-2), 111-127.

[105] Ha, I. D., & Lee, Y. (2005). Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. Biometrika, 92(3), 717-723.

[106] Keyfitz, N., & Littman, G. (1979). Mortality in a heterogeneous population. Population studies, 33(2), 333-342.

[107] Lancaster, T. (1990). The econometric analysis of transition data (No. 17). Cambridge university press.

[108] Henderson, R., & Oman, P. (1999). Effect of frailty on marginal regression estimates in survival analysis. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61(2), 367-379.

[109] French, P. (2006). IACAPAP textbook of child and adolescent mental health. The Lancet.

[110] Hosmer, D. W., & Lemeshow, S. (1999). Applied survival analysis: Time-to-event (Vol. 317). Wiley-Interscience.

[111] Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481.

[112] Breslow, N. E. (1975). Analysis of survival data under the proportional hazards model. International Statistical Review/Revue Internationale de Statistique, 45-57.

[113] Therneau, T. M., & Grambsch, P. M. (2000). The cox model. In Modeling survival data: extending the Cox model (pp. 39-77). Springer, New York, NY.

[114] Duchateau, L., & Janssen, P. (2004). Penalized partial likelihood for frailties and smoothing splines in time to first insemination models for dairy cows. Biometrics, 60(3), 608-614.

[115] Hougaard, P. (2000). Analysis of Multivariate Survival Data Springer-Verlag New York Google Scholar.

[116] Kadane, J. B., & Lazar, N. A. (2004). Methods and criteria for model selection. Journal of the American statistical Association, 99(465), 279-290.

[117] Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.

[118] Dätwyler, C., & Stucki, T. (2011). Parametric survival models.

[119] Novick, D., Montgomery, W., Vorstenbosch, E., Moneta, M. V., Dueñas, H., & Haro, J. M. (2017). Recovery in patients with major depressive disorder (MDD): results of a 6-month, multinational, observational study. Patient preference and adherence, 11, 1859.

[120] Viinamaki, H., Haatainen, K., Honkalampi, K., Tanskanen, A., Koivumaa-Honkanen, H., Antikainen, R., ... & Hintikka, J. (2006). Which factors are important predictors of non-recovery from major depression? A 2-year prospective observational study. Nordic Journal of Psychiatry, 60(5), 410-416.

[121] Dierker, L. C., Avenevoli, S., Merikangas, K. R., Flaherty, B. P., & Stolar, M. (2001). Association between psychiatric disorders and the progression of tobacco use behaviors. Journal of the American Academy of Child & Adolescent Psychiatry, 40(10), 1159-1167.

[122] Kader Maideen, S. F., Mohd. Sidik, S., Rampal, L., & Mukhtar, F. (2014). Prevalence, associated factors and predictors of depression among adults in the community of Selangor, Malaysia. PloS one, 9(4), e95395.

[123] Abdulkarimova, U. (2013). Frailty Models for Modelling Heterogeneity (Doctoral dissertation).

APPENDIXES

Annex 1: Univariable Analysis Using Parametric Shared Frailty Models For MDD Dataset

A. Weibull-Gamma and Weibull-Inverse-Gaussian Univariable Analysis

Variable	Category	Weibull-Gamma				Weibull-Inverse-Gaussian			
		Coef	SE	ϕ	p-value	Coef	SE	ϕ	p-value
Gender	Female	-0.38	0.134	0.68	0.004	-0.384	0.13	0.68	0.004
Age	20-25	-0.08	0.219	0.92	0.71	-0.09	0.22	0.91	0.68
	26-49	-0.174	0.21	0.84	0.41	-0.175	0.212	0.83	0.41
	$\geq$50	0.754	0.272	2.12	0.006	0.752	0.27	2.12	0.006
Marital Status	Married	-0.148	0.141	0.86	0.3	-0.171	0.14	0.84	0.22
	Widowed	0.976	0.281	2.65	$\leq$ 0.001	0.997	0.284	2.71	$\leq$0.001
	Divorced	0.845	0.248	2.33	$\leq$ 0.001	0.859	0.25	2.36	$\leq$0.001
First Onset Age	Adolescent	-0.1613	0.270	0.85	0.55	-0.164	0.27	0.85	0.54
	Adult Above	-0.0915	0.28	0.91	0.74	-0.095	0.28	0.91	0.73
Family History	Yes	0.557	0.130	1.74	$\leq$ 0.001	0.55	0.13	1.73	$\leq$ 0.001
Chewing Khat	Yes	1.56	0.21	4.76	$\leq$ 0.001	1.58	0.213	4.85	$\leq$ 0.001
Educational Level	Educated	-1.08	0.144	0.34	$\leq$ 0.001	-1.1	0.145	0.33	$\leq$ 0.001
Employments	Yes	-0.59	0.135	0.55	0.001	-0.59	0.136	0.6	$\leq$ 0.001
Religion	Muslims	-0.1347	0.154	0.87	0.38	-0.1337	0.154	0.87	0.39
	Protestant	0.006	0.228	1.01	0.98	0.011	0.23	1.01	0.96
	Other's	-0.482	0.289	0.62	0.095	-0.481	0.289	0.62	0.97
Ethnicity	Amahara	0.084	0.1567	1.1	0.59	0.0824	0.157	1.1	0.6
	Others	-0.138	0.194	0.87	0.48	-0.1439	0.195	0.87	0.46
Substance abuse	Yes	0.71	0.133	2.02	0.001	0.683	0.134	1.97	$\leq$.001
Other Cofactors	Yes	1.25	0.164	3.45	$\leq$0.001	1.25	0.166	3.5	$\leq$.001
Event of Relapse	Yes	0.513	0.132	1.67	$\leq$0.001	0.52	0.13	1.68	$\leq$.001

B. Loglogistic-Gamma and Loglogistic-Inverse- Gaussian Univariable Analysis

Variable	Category	Weibull-Gamma				Weibull-Inverse-Gaussian			
		Coef	SE	ϕ	p-value	Coef	SE	ϕ	p-value
Gender	Female	-0.454	0.142	0.63	0.0014	-0.45	0.142	0.64	0.0015
Age	20-25	-0.141	0.244	0.87	0.56	-0.145	0.245	0.86	0.56
	26-49	-0.177	0.233	0.83	0.45	-0.179	0.23	0.84	0.44
	$\geq$50	0.742	0.273	0.48	0.0065	0.743	0.274	2.1	0.007
Marital Status	Married	-0.192	0.15	0.83	0.21	-0.199	0.155	0.82	0.2
	Widowed	0.962	0.258	2.61	$\leq$ 0.001	0.977	0.261	2.65	0.0002
	Divorced	0.853	0.239	2.34	$\leq$ 0.001	0.88	0.24	2.4	0.00024
First Onset Age	Adolescent	-0.215	0.297	0.81	0.47	-0.214	0.297	0.81	0.47
	Adult Above	-0.16	0.305	0.85	0.6	-0.159	0.31	0.85	0.6
Family History	Yes	0.539	0.14	1.71	$\leq$ 0.001	0.537	0.138	1.71	$\leq$ 0.001
Chewing Khat	Yes	1.59	0.174	4.9	$\leq$ 0.001	1.61	0.1754	5	$\leq$ 0.001
Educational Level	Educated	-1.1	0.139	0.33	$\leq$ 0.001	-1.14	0.14	0.33	$\leq$ 0.001
Employments	Yes	-0.652	0.141	0.52	$\leq$ 0.001	-0.655	0.141	0.52	$\leq$ 0.001
Religion	Muslims	-0.156	0.169	0.86	0.36	-0.1545	0.169	0.86	0.36
	Protestant	-0.053	0.25	0.95	0.83	-0.0527	0.246	0.95	0.83
	Other's	-0.279	0.308	0.76	0.37	-0.2732	0.307	0.76	0.37
Ethnicity	Amahara	0.096	0.170	1.10	0.57	0.0963	0.170	1.10	0.57
	Others	0.0389	0.215	1.04	0.86	0.0399	0.215	1.04	0.85
Substance abuse	Yes	0.827	0.135	2.28	$\leq$ 0.001	0.811	0.135	2.25	$\leq$ 0.001
Other Cofactors	Yes	1.22	0.157	3.39	$\leq$ 0.0001	1.22	0.1573	3.39	$\leq$ 0.001
Event of Relapse	Yes	0.564	0.139	3.06	$\leq$ 0.001	0.563	0.14	1.76	$\leq$ 0.001

C. Log-normal-Gamma and Log-normal-Inverse-Gaussian Univariable Analysis

Variable	Category	Weibull-Gamma				Weibull-Inverse-Gaussian			
		Coef	SE	ϕ	p-value	Coef	SE	ϕ	p-value
Gender	Female	-0.443	0.137	0.64	0.001	-0.442	0.137	0.64	0.0012
Age	20-25	-0.071	0.236	0.93	0.76	-0.0834	0.237	0.92	0.72
	26-49	-0.145	0.226	0.86	0.52	-0.1504	0.226	0.86	0.51
	$\geq$50	0.7535	0.26	2.12	0.0037	0.7518	0.261	2.12	0.004
Marital Status	Married	-0.214	0.150	0.81	0.15	-0.212	0.151	0.81	0.16
	Widowed	0.859	0.239	2.31	$\leq$ 0.001	0.872	0.24	2.39	$\leq$ 0.001
	Divorced	0.807	0.229	2.24	$\leq$ 0.001	0.850	0.231	2.34	$\leq$ 0.001
First Onset Age	Adolescent	-0.235	0.294	0.78	0.42	-0.234	0.294	0.79	0.43
	Adult Above	-0.195	0.301	0.82	0.52	-0.193	0.301	0.83	0.52
Family History	Yes	0.481	0.134	1.6	$\leq$ 0.001	0.478	0.134	1.61	$\leq$ 0.001
Chewing Khat	Yes	1.5	0.167	4.5	$\leq$ 0.001	1.53	0.168	4.6	$\leq$ 0.0001
Educational Level	Educated	-1.06	0.136	0.35	$\leq$ 0.001	-1.08	0.137	0.33	$\leq$ 0.0001
Employments	Yes	-0.61	0.137	0.54	<.001	-0.615	0.137	0.541	$\leq$ 0.001
Religion	Muslims	-0.14	0.162	0.87	0.39	-0.1392	0.16	0.87	0.39
	Protestant	-0.055	0.234	0.95	0.81	-0.0543	0.233	0.95	0.82
	Other's	-0.233	0.323	0.79	0.47	-0.2299	0.323	0.794	0.48
Ethnicity	Amahara	0.116	0.164	1.12	0.48	0.118	0.164	1.12	0.47
	Others	0.056	0.219	1.06	0.8	0.055	0.219	1.06	0.8
Substance abuse	Yes	0.754	0.133	2.12	$\leq$ 0.001	0.742	0.134	2.09	$\leq$ 0.001
Other Cofactors	Yes	1.08	0.148	3.0	$\leq$ 0.001	1.08	0.149	3.0	$\leq$ 0.001
Event of Relapse	Yes	0.518	0.135	1.68	$\leq$ 0.001	0.521	0.136	1.68	$\leq$ 0.001

Annex 2: Multivariable Analysis Using Parametric Frailty Shared Models For MDD Dataset

A. Weibull-Gamma Multivariable Shared Frailty Model

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender	Male	Ref		1		
	Female	-0.0882	0.132	0.915	[0.657 1.174]	0.50
Age of patients (in years)	13-19	Ref		1		
	20-25	0.2077	0.208	1.231	[0.823 1.638]	0.32
	26-49	0.2244	0.211	1.251	[0.838 1.665]	0.29
	≥ 50	0.5444	0.254	1.723	[0.226 2.137]	0.11
Martital status	Single	Ref		1		
	Married	-0.1662	0.143	0.847	[0.819 0.875]	0.25
	Widowed	0.3896	0.267	1.476	[0.953 1.999]	0.14
	Divorced	0.5064	0.245	1.659	[1.179 2.139]	0.038 *
Family History	No	Ref		1		
	Yes	0.0985	0.139	1.1035	[0.831 1.376]	0.48
Chewing Khat	No	Ref		1		
	Yes	1.035	0.203	2.815	[2.47 3.213]	$\leq$0.001 ***
Educational Level	Uneducated	Ref		1		
	Educated	-0.404	0.146	0.667	[0.382 0.954]	0.0056 **
Employment	No	Ref		1		
	Yes	-0.363	0.130	0.695	[0.441 0.950]	0.0053 **
Substance Abuse	No	Ref		1		
	Yes	0.4584	0.141	1.581	[1.298 1.851]	0.0011 **
Other Cofactors	No	Ref		1		
	Yes	0.625	0.164	1.869	[1.547 2.189]	0.00014 ***
Event of Relapse	No	Ref		1		
	Yes	0.251	0.133	1.286	[0.025 1.546]	0.059
$\theta = 0.054$	$\tau = 0.026$	AIC = 1197.679				

B. Weibull-Inverse-Gaussian Multivariable Shared Frailty Model

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender of patients	Male	Ref		1		
	Female	-0.0967	0.130	0.907	[0.653 1.163]	0.46
Age of patients (in years)	13-19	Ref		1		
	20-25	0.0957	0.206	1.10	[0.697 1.504]	0.64
	26-49	0.1490	0.209	1.161	[0.751 1.573]	0.48
	≥ 50	0.4434	0.253	1.558	[0.058 1.654]	0.08
Martital status of patients	Single	Ref		1		
	Married	-0.1297	0.141	0.878	[0.602 1.155]	0.36
	Widowed	0.4519	0.262	1.571	[0.289 1.622]	0.084
	Divorced	0.5027	0.237	1.653	[1.189 2.118]	0.034 *
Family History of patients	No	Ref		1		
	Yes	0.0200	0.137	1.02	[0.752 1.289]	0.88
Chewing Khat	No	Ref	1			
	Yes	0.9955	0.201	2.706	[2.312 3.100]	$\leq$0.001**
Educational Level	Uneducated	Ref		1		
	Educated	-0.4454	0.143	0.641	[0.360 0.921]	0.001 **
Employment	No	Ref		1		
	Yes	-0.3482	0.129	0.706	[0.453 0.959]	0.0018 **
Substance Abuse	No	Ref		1		
	Yes	0.4021	0.134	1.495	[1.232 1.758]	0.0027 **
Other Cofactors	No	Ref		1		
	Yes	0.6127	0.164	1.845	[1.524 2.167]	$\leq$0.001**
Event of Relapse History	No	Ref		1		
	Yes	0.2166	0.130	1.242	[0.987 1.497]	0.097
$\theta = 0.058$	$\tau = 0.027$	AIC = 1192.313				

C. Log-log logistic-Gamma Multivariable Shared Frailty Model

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender of patients	Male	Ref		1		
	Female	-0.1720	0.127	0.842	[0.592 1.090]	0.18
Age of patients (in years)	13-19	Ref		1		
	20-25	0.0649	0.213	1.067	[0.649 1.484]	0.76
	26-49	0.1289	0.207	1.137	[0.653 1.543]	0.53
	≥ 50	0.4988	0.247	1.648	[0.446 1.700]	0.163
Martital status of patients	Single	Ref		1		
	Married	-0.3196	0.143	0.726	[0.663 1.675]	0.335
	Widowed	0.1272	0.241	1.136	[0.663 1.608]	0.60
	Divorced	0.4985	0.230	1.646	[1.195 2.097]	0.030 *
Family History of patients	No	Ref		1		
	Yes	0.2087	0.132	1.232	[0.973 1.491]	0.11
Chewing Khat	No	Ref		1		
	Yes	0.9674	0.180	2.631	[2.278 2.984]	$\leq$0.001***
Educational Level	Uneducated	Ref		1		
	Educated	-0.4475	0.142	0.639	[0.361 0.917]	0.0017 ***
Employment	No	Ref		1		
	Yes	-0.4062	0.133	0.666	[0.350 0.871]	0.0023 **
Substance Abuse	No	Ref		1		
	Yes	0.4929	0.139	1.637	[1.365 1.909]	0.00039 **
Other Cofactors	No	Ref		1		
	Yes	0.4903	0.156	1.633	[1.327 1.939]	0.0017***
Event of Relapse History	No	Ref		1		
	Yes	0.2547	0.130	1.291	[0.571 1.588]	0.147
$\theta = 0.030$	$\tau = 0.015$	AIC = 1183.948				

D. Log logistic-Inverse-Gaussian Multivariable Shared Frailty Model

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender	Male	Ref		1		
	Female	-0.1967	0.129	0.821	[0.647 1.575]	0.13
Age of patients (in years)	13-19	Ref		1		
	20-25	-0.0371	0.214	0.964	[0.544 1.383]	0.86
	26-49	0.064	0.211	1.066	[0.652 1.479]	0.76
	$\geq$ 50	0.403	0.253	1.496	[0.172 1.164]	0.11
Martital status	Single	Ref		1		
	Married	-0.303	0.147	0.738	[0.453 1.029]	0.339
	Widowed	0.1996	0.247	1.221	[0.737 1.705]	0.42
	Divorced	0.555	0.231	1.742	[1.289 2.195]	0.016 *
Family History	No	Ref		1		
	Yes	0.1783	0.135	1.195	[0.931 1.459]	0.19
Chewing Khat	No	Ref	1			
	Yes	0.9537	0.181	2.595	[2.240 2.950]	$\leq$0.001 ***
Educational Level	Uneducated	Ref		1		
	Educated	-0.5139	0.144	0.598	[0.316 0.880]	$\leq$0.001 ***
Employment	No	Ref		1		
	Yes	-0.4849	0.131	0.616	[0.359 0.872]	$\leq$0.001 **
Substance Abuse	No	Ref		1		
	Yes	0.4612	0.136	1.586	[1.319 1.852]	$\leq$0.001 **
Other Cofactors	No	Ref		1		
	Yes	0.5068	0.156	1.599	[1.354 1.966]	0.0012 ***
Event of Relapse	No	Ref		1		
	Yes	0.2488	0.132	1.282	[0.572 1.809]	0.06
θ = 0.030	τ = 0.014	AIC = 1174.908				

E. Lognormal-Gamma Multivariable Shared Frailty Model

Covariates	Category	Coef	S.E	ϕ	95% CI	p-value
Gender of patients	Male	Ref		1		
	Female	-0.1487	0.128	0.862	[0.611 1.113]	0.25
Age of patients (in years)	13-19	Ref		1		
	20-25	0.0449	0.216	1.046	[0.623 1.469]	0.84
	26-49	0.0276	0.209	1.028	[0.618 1.438]	0.90
	≥ 50	0.4637	0.251	1.589	[0.098 2.082]	0.065
Martital status of patients	Single	Ref		1		
	Married	-0.2421	0.141	0.785	[0.509 1.061]	0.086
	Widowed	0.2504	0.237	1.284	[0.820 1.749]	0.29
	Divorced	0.5500	0.232	1.733	[1.278 2.188]	0.018 **
Family History of patients	No	Ref		1		
	Yes	0.1773	0.130	1.194	[0.939 1.449]	0.17
Chewing Khat	No	Ref		1		
	Yes	0.9057	0.172	2.474	[2.136 2.811]	$\leq$0.001 ***
Educational Level	Uneducated	Ref		1		
	Educated	-0.4835	0.138	0.62	[0.346 0.887]	$\leq$0.001 **
Employment	No	Ref		1		
	Yes	-0.3510	0.131	0.704	[0.447 0.961]	0.0074 **
Substance Abuse	No	Ref		1		
	Yes	0.4290	0.135	1.536	[1.271 1.800]	0.0014 **
Other Cofactors	No	Ref		1		
	Yes	0.4850	0.149	1.624	[1.332 1.916]	0.0012 ***
Event of Relapse History	No	Ref		1		
	Yes	0.2234	0.130	1.25	[0.995 1.505]	0.086
$\theta = 0.172$	$\tau = 0.079$	AIC = 1181.907				

Annex 3: The log rank test for Survival curve for significantly different groups

Variables	Category	N	Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{(O-E)^2}{V}$	χ^2	df	Sig.
Marital Status	Single	110	72	62.4	1.48	2.36			
	Married	136	98	69.3	11.92	20.16			
	Widowed	62	12	31.1	11.75	15.30			
	Divorced	58	16	35.2	10.48	14.04	39.1	3	<.0001
Khat Chewing	No	230	177	115.8	32.3	85.7			
	Yes	136	21	82.2	45.5	85.7	85.7	1	<0.0001
Educational Status	Illiterate	215	63	120.5	27.5	77.1			
	Literate	151	135	77.5	42.7	77.1	77.1	1	<.0001
Employment Status	No	174	69	99.8	9.51	21.1			
	Yes	192	129	98.2	9.66	21.1	21.1	1	<.001
Substance Abuse	No	190	129	94.9	12.2	25.9			
	Yes	176	69	103.1	11.3	25.9	25.9	1	<.001
Other Cofactors	No	202	162	103.9	32.5	75.6			
	Yes	164	36	94.1	35.9	75.6	75.6	1	<.0001

Annex 4: Descriptive statistics for woredas of patients'

	Category	Censored	Events	Total	Median Time (in months)
Woreda	Jimma	35(53%)	31(47%)	66(18.1%)	6
	Agaro	13(72.2%)	5(27.8%)	18(4.9%)	10
	Kersa	17(54.8%)	14(45.2%)	31(8.4%)	5
	Dedo	6(37.5%)	10(62.5%)	16(4.3%)	11
	Omo Nada	11(55%)	9(45%)	20(5.5%)	9
	Limmu Kosa	5(31.2%)	11(68.8%)	16(4.4%)	7
	Seka Chokorsa	10(47.6%)	11(52.4%)	21(5.7%)	11
	Tiro Afeta	5(29.4%)	12(70.6%)	17(4.6%)	4
	Setema	3(33.3%)	6(66.7%)	9(2.4%)	13
	Gumay	7(58.3%)	5(41.7%)	12(3.3%)	7
	Gera	7(38.9%)	11(61.1%)	18(4.9%)	7
	Shabe Sombo	7(38.9%)	11(61.1%)	18(4.9%)	7
	Sokoru	11(47.8%)	12(52.2%)	23(6.3%)	11
	Mana	10(47.6%)	11(52.4%)	21(5.7%)	4
	Gomma	10(43.5%)	13(56.5%)	23(6.2%)	7
	Sigimo	5(35.7%)	9(64.3%)	14(3.9%)	6
	Limmu Seka	2(18.2%)	9(81.8%)	11(3.0%)	2
	Chora Botor	4(33.3%)	8(66.7%)	12(3.3%)	9