SURVIVAL ANALYSIS OF TIME TO RECOVERY OF ADIMTTED COVID-19 PATIENTS: IN HAWASSA UNVERCITY REFERAL AND COMPREHENSIV HOSPITAL TEREATMENT CENTER

MSc. THESIS

BY

AMANUEL MERDIKYOS NANA

HAWASSA, ETHIOPIA

JUNE, 2023

SURVIVAL ANALYSIS OF TIME TO RECOVERY OF ADIMTTED COVID-19 PATIENTS: IN HAWASSA UNVERCITY REFERAL AND COMPREHENSIV HOSPITAL TEREATMENT CENTER.

AMANUEL MERDIKYOS NANA

A THESIS SUBMITTED TO:

SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES

HAWASSA UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE IN STATISTICS

(SPECIALIZATION: APPLIED STATISTICS)

JUNE, 2023

HAWASSA, ETHIOPIA

APPROVAL SHEET-1

This is to certify that the thesis entitled "SURVIVAL ANALYSIS OF TIME TO RECOVERY RATE OF ADIMTTED COVID-19 PATIENTS: A CASE STUDY IN HAWASSA UNVERCITY REFERAL AND COMPREHENSIV HOSPITAL TEREATMENT CENTER **.**" submitted in partial fulfillment of the requirements for the degree of Master of Science in Statistics with Specialization of Applied Statistics of the Graduate Program of the School of Mathematical and Statistical Sciences, Hawassa University, and is a record of original research carried out by **Amanuel Merdikyos Nana Id. No. GpAStR/006/12** under my supervision and no part of the thesis has been submitted for any other degree or diploma. The assistance and the help received during the course of this investigation have been duly acknowledged. Therefore, I recommend that it would be accepted as fulfilling the thesis requirements.

____________________ ______________ ________________

Name of Main Advisor Signature Date

____________________ ______________ ________________

Name of Co-Advisor Signature Date

APPROVAL SHEET-2

We, the undersigned, members of the board of examiners of the final open defense by **Amanuel Merdikyos Nana** have read and evaluated his thesis entitled "SURVIVAL ANALYSIS OF TIME TO RECOVERY RATE OF ADIMTTED COVID-19 PATIENTS: A CASE STUDY IN HAWASSA UNVERCITY REFERAL AND COMPREHENSIV HOSPITAL TEREATMENT CENTER" and examined the candidate. This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Statistics with specialization of Applied Statistics.

____________________	______________	________________
Name of Chairperson	Signature	Date
____________________	______________	________________
Name of Main Advisor	Signature	Date
____________________	______________	________________
Name of Co-advisor	Signature	Date
____________________	______________	________________
Name of External Examiner	Signature	Date
____________________	______________	________________
Name of Internal Examiner	Signature	Date

ACKNOWLEDGEMENTS.

During my graduate studies in Hawassa University several persons and institutions collaborated directly and indirectly to my research. That is why I wish to dedicate this section to recognize their support.

First and foremost I would like to acknowledge the Hawassa University Compressive Hospital Head and staff members of Hospital to undertake this study with their cooperation and permission in using the data with special thanks for Ato Tesfahun and sister Elfinesh for their willingness to help me.

My special gratitude goes to Dr. Cheru Atsmegiorgis, my advisor and instructor, for his immense and invaluable advice and guidance that contributed to the successful realization of this study. I would like to express my sincere appreciation to my co-advisor, for his precious suggestions and comments during the entire time of the study. My sincere appreciation and thanks also go to my instructors and the rest staffs of Statistics Department at Hawassa University for their unreserved knowledge sharing and cooperation. I would like to be grateful for my sponsor Melo koza woreda Administration office for the contribution in my study.

Finally, my deepest and warm gratitude goes to my family that has been a source of pride and encouragement throughout my work. I am thankful to my father and mother for their encouragement and prayer my beloved wife Abezash Gorfu my kid Aynishet Amanuel Mehalt-Amanuel my sisters Meselch Merdikyoe, Birhanesh Merdikyos and Mebirat Merdikyos, Marta Merdikyos and beloved brother Tibebu Merdikyos, Yosef Merdikyos, who were always there to provide me with continuing motivation and encouragement, and constructive advice that helped me to complete this work.

ABSTRACT

Corona virus is one of the major pathogens that primarily target the human respiratory system, which started in Wuhan, China in December 2019, has emerged as a global health and economic security threat with an overwhelming growing incidence worldwide. When the World Health Organization (WHO) declared the disease a global public health emergency, different stakeholders stepped up efforts to convince the world that the disease is a serious problem that needs strong containment measures. The main objective of the study is to identify the determinant risk factors for the recovery of corona virus(covid-19) patients. A study population of 826 total Covid-19 Patients that had been treated at Hawassa University Comprehensive and Referral hospital from September 20, 2013 to January 20, 2014 E.C was included in the study. Descriptive statistics and Kaplan-Meier survival curves were used to estimate and compare the recovery time of corona virus (covid-19) patients among different categorical characteristics of the patients. We used survival time model to analyze the data. The Weibull regression model better fits the recovery time of corona virus (covid-19) than the exponential, log-logistic model and log-normal model. The result showed that out of a total of 826 corona virus (covid-19) patients considered total recovery are 637(77.12%) recovered from covid-19. From the result severity (HR=0.932, p-value=0.014), Co-morbidty(HR=0.89,p-value=0.038), other pains out of covid-19(HR=0.7918, P-value=0.006), shortness of breath (HR=0.83,p-value=0.025), severe headache (HR=0.843, p-value=0.034) and Age (HR=0.8948, p-value=0.000) were the significant factors for the corona virus(covid-19) patients using Weibull regression model. The model showed that the major factors that affect the recovery time of corona-virus (covid-19) and see the associations factors among patients. Patient's comorbidities have a major impact on CVID-19; So, health profession should close follow up is required for client admitted with comorbidity and create great awareness about the risk factors the corona virus (covid-19).

***Key words**: corona virus (covid-19), proportional hazard, Weibull regression, survival analysis*

Table of Contents

LIST OF ACRYNOMS

AFT:- Accelerated Failure Time

AHR :-Adjusted Hazard Ratio

AIC:- Akaiki Information Criteria

APACHE II:- Acute physiology and chronic Health Evaluation II

BIC:-Bayse Information Criteria

BMI:- High Body Mass Index

CI:-Confidence interval

COVID-19:-Corona Virus Disease 2019

CSA:- Central Statistical Agency

CT:-Computed tomography

CXR:- Chest X Ray

EC:-Ethiopia Calendar

HB:-Hemoglobin

HR:- Hazard Ratio

ICU :-Intensive care unit

IQR :-Inter quartile Range

IJID :-International Journal of infectious Disease

JCM :-Journal of Contemporary Medicine

KM :-Kaplan meiers

LDH:-Lactate Dehydrogenase

LDS:-Hospital length of stay

MERS :-Middle East respiratory syndrome

NLR:- Neutrophil lymphocyte ratio

OR:- Odd ratio

PH:-Proportional Hazards

RNS:- Ribonucleic Acid

rRT-PCR :-Real-time Reverse Transcriptase polymerase chain reaction

RTPCR:--Reverse transcriptase Polymerase chain reaction

SARS :-Severe Acute Respiratory Syndrome

WBC:- Write Blood Cell

WHO:-World Health Organization.

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

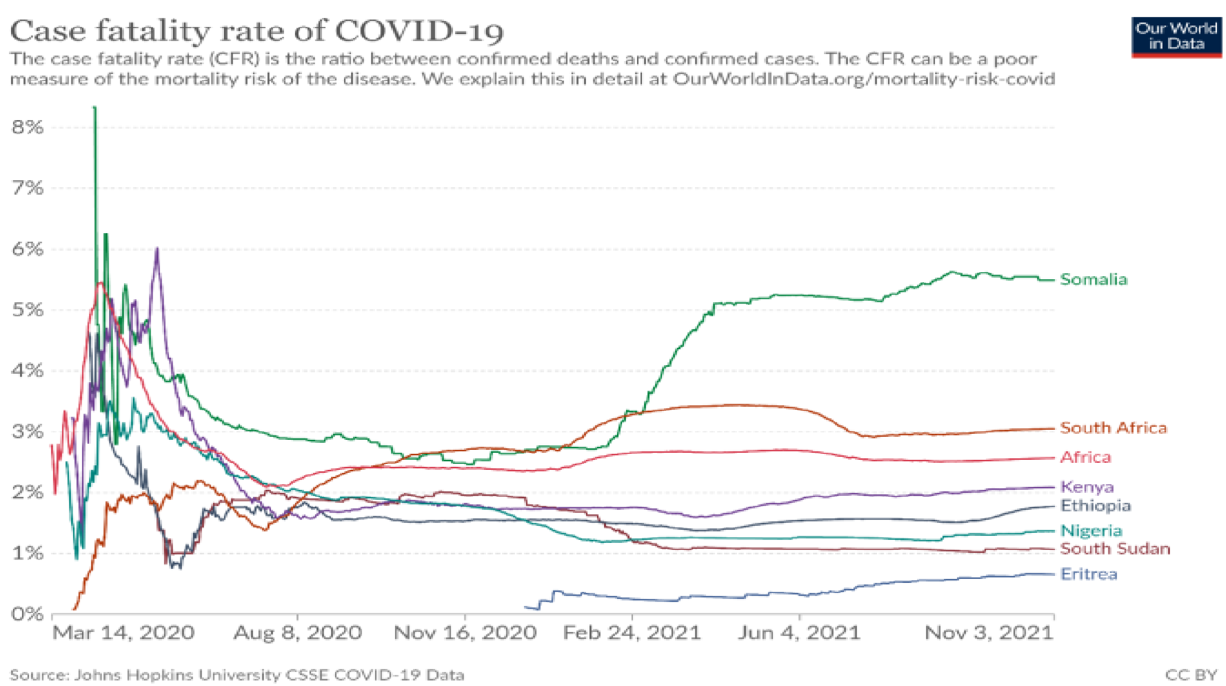
The novel coronavirus COVID-19 originally identified in December 2019 as a severe case of pneumonia in Wuhan province of China and since that has become a global pandemic, affecting greatest nations around the whole world (Wu JT, Leung K & Leung GM, 2020). Although the foundations of this disease are very similar to the severe acute respiratory syndrome (SARS) virus that took hold of Asia in 2003, it is shown to spread much more easily and there currently exists no routinely used vaccinations at the time this research was written Since the first confirmed cases were reported in China, much of the literature has focused on the outbreak in China including the transmission of the disease, the risk factors of infection, and the biological properties of the virus. However, more recent literature has started to cover an increasing number of regions outside of China.

Coronaviruses are a large group of viruses, some cause illness to human and some occur in animals. Rarely, animal coronaviruses can evolve and infect people and then may spread between people. Human coronaviruses cause routine seasonal respiratory virus infections. Other coronaviruses, like severe acute respiratory syndrome (SARS) and Middle East reparatory syndrome (MERS), can cause serious illness (Yin Y. and Wunderink R.G, 2021).

Coronavirus spreads throughout countries and was categorized as pandemics by World Health Organization (WHO) in March 2020. Coronaviruses are respiratory viruses and most commonly spread through respiratory secretion of an infected person in close proximity. The epidemiological dynamics of COVID-19 has changed dramatically over the courses of months (Pamuk S., O¨ zkan A.,et al.,2020). At this time COVID-19 has infected more than 500 million people globally, and more than 6 million people died out, and more than 450 million cases were recovered.

Africa compared to other regions, the initial wave of COVID-19 in the first half of 2020 progressed more slowly throughout. However, the second wave has hit Africa much harder, and currently shows no sign of slowing down, especially given the arrival of the highly transmissible Delta variant. Ethiopia now appears to be entering a third wave and the negative impacts of the pandemic are being further exacerbated by compounding crises, including the locust outbreak, the conflict in several parts of Ethiopia (with the worst being in the Northern regions), and extensive flooding, with significant implications for food security and poverty reduction. The COVID-19 situation in Ethiopia has evolved rapidly in the last year and a half.

Currently, Ethiopia had reported more than 470 thousand confirmed cases and more than 7500 deaths from COVID-19(OPM study,2022). While the total number of cases currently ranks Ethiopia 6th highest in Africa, cases and deaths *per population* show a very different story, with 2,994 cases per million population ranking Ethiopia at 33 out of 58 African countries, and 51 deaths per million ranking even lower at 36 out of 58 countries; Although there has been a slight increase recently, Ethiopia has experienced a lower fatality rate than the regional average throughout the pandemic. As in many countries with weak health system, the actual case and death count may be much higher than reported ones.



By comparison, South Africa that ranks first in Africa, has recorded significantly higher rates of 48,328 cases per million population, and 1,466 deaths per million; neighboring Kenya has recorded almost double the rates of Ethiopia, with 4,545 cases per million and 91 deaths per million; while Nigeria, the most populous African country, has recorded particularly low rates of 978 cases per million and 13 deaths per million. (Jin Y.-H, 2019). In many African countries, the number of cases and number of death are low compared to European and American countries, this could be due to low test capacity, underreporting, and young population. Knowing the COVID-19 fatality rate help us understand severity of the disease, identify risk factors and assess the quality of healthcare. There are difference in mortality rate among different groups due to age difference and comorbidity. And also the COVID-19 fatality rate vary across different locations may be due to populations age structure and case-mix of infected and deceased patients.

The median time to recovery from COVID-19 varies among patients and settings, in which the average recovery time from COVID-19 more than 14 days for some countries and less than 14 days for others (Bhapkar H.R., 2020). According to WHO, the recovery time is estimated to be 2 weeks for patients with mild infection and 3 to 6 weeks for those with serious illnesses (Jin Y.-H., et al., 2019). On the other hand, CDC estimated people with mild to moderate spectrum of symptom and maintain home isolation have a resolution of 3 days after the fever decreased, and there was substantial improvement in respiratory symptoms, even without use of medication(Leulseged T.W., et al.,2020).

A study conducted in treatment center found in Ethiopia showed that, the rate of prognosis from COVID-19 for asymptomatic cases was higher when compared to symptomatic COVID-19 cases, and this study reported the average recovery time of 16 days. Another study conducted in Kotebe treatment center of Ethiopia revealed the media time to recovery from COVID-19 was 19 days and it ranges from 2–71 days (Leulseged T.W., et al., 2020). Most studies indicated that severity of COVID-19 is associated with factors like age, sex, chronic underlying illness, co-infection of bacteria and fungi and clinical factors like d-dimer, ferritin, C-reactive protein and leukocytosis (Alimohamadi, Y, et al., 2020). The studies also indicated the proportion of mild, moderate and critical cases of the disease.

The overall median incubation period of COVID-19 was a little bit varying among several studies (Hu X., Wu J., 2020). However, there were instances that the incubation period specifically longer than the majority group in some cases. The incubation period also varied by age that patients with older age had a longer incubation period. Knowing the incubation period of COVID- 19 is important for disease surveillance, prevention and control measures of the disease. Although most studies suggested the 14 days incubation period, there are evidences that the incubation period extends to 24 days (Ortiz-Prado E., et al., 2019). A comprehensive literature review showed that age greater than 65 years, being hypertensive (up to 40% of patients), diabetics, obesity, cardiovascular and lung disease are possible risk factors for delayed recovery from COVID-19 . In addition, presence of clinical manifestation on admission also associated with delayed recovery from COVID-19

Leulseged T.W., et al.,2020 carries out that studies showed improved survival of covid-19 patients depends on quality of health care services like, patient management, timing of admission, understanding of disease progress and expand use of steroids.

The necessary medical care is applied only to patients with more probability to survive. Calculating the probability to survive and the effect of each feature like symptoms in our case on survival probability is done using survival analysis. Survival analysis is a model for time until a certain“event”. Time-to-event data encounters several research challenges such as censoring, symptoms (features) correlations, high-dimensionality, temporal dependencies, and difficulty in acquiring sufficient event data in a reasonable period of time (Fadnavis RA, 2019).

There are many current literature techniques for conducting this sort of survival study. Among them, the Cox Proportional Hazards Model (Cox) (Cox DR 1975) which is a Regression models that is commonly used in survival analysis .Survival analysis methods can work with specific problems, with a data type that waits for the event to occur. Cox regression is the most appropriate method to deal with this kind of data. Occasionally, the basic assumptions of the model, for example, non-proportionality for the Cox model, are not true. The choice of an appropriate model varies depending on the complexity and features that affect the suitability of the model, in model building (Grannesby JK, Borgan Q 1996). Though some scientific researches have been published regarding covid-19, many aspects of it still need more detailed valid and reliable information. This retrospective cohort study aimed at determining time to recovery of Covid-19 infected people and its predictors among Covid-19 patients admitted to treatment center of Hawass University Referral Hospital, Sidama Region Ethiopia. Hawassa Referral Hospital – this government hospital serves the population of Hawassa and extends to include surrounding areas encompassing roughly 17 million people. This hospital was started in 2003 and is attached to Hawassa University. Hospital is constructed nearby Lake Hawassa and provides curative services to the community and as a teaching referral hospital for training various health professionals. The Hospital has six wards: (Surgical, Pediatrics, Gynecology and obstetrics, Ophthalmology, Medical, and Labor wards) with 350 maximum bed capacity all together.

1.2 Statement of the Problem

The median duration of viral shedding in COVID-19 patients ranges from 8–30 days while the longest duration was reported as 47 days (Yin Y. and Wunderink R.G., 2020). Factors that are associated with prolonged duration of viral RNA shedding in COVID-19 patients include older age, a time lag from illness onset to hospital admission, diarrhea, corticosteroid treatment and lopinavir/ritonavir use. In contrast, a study reported that antiviral therapy and corticosteroid treatment were not independent factors for prolonged duration of viral shedding (Bhapkar H.R., et al., 2020). Further, a study in Japan showed that fever and time from illness onset to hospitalization were associated with increased odds of prolonged duration of viral shedding (Abrahim S.A., et al., 2020). Male sex, immunoglobulin use, Acute Physiology and Chronic Health Evaluation II (APACHE II) score, and lymphocyte count were independent factors associated with a prolonged duration of SARS-CoV-2 shedding in a study conducted among hospitalized patients with COVID-19 in Zhejiang Province, China (Leulseged T.W., et al., 2020). Some studies showed improved survival of covid-19 patients depends on quality of healthcare services like, patient management, timing of admission, understanding of disease progress and expand use of steroids (Xie J., et al.,2021). High Body Mass Index (BMI) was also reported to be associated with a longer duration of viral shedding in COVID -19 patients in a study conducted in Italy (Ortiz-Prado E., et al., 2020). A study by Xie revealed that severe illness requiring invasive mechanical ventilation was an independent factor for prolonged SARS-CoV-2 RNA shedding (Xie J., et al., 2021) while another study reported no correlation between severity and viral shedding (Jin Y.H., et al.,2020).Various factors determine the duration of the viral shedding in COVID-19 patients. Therefore, evidences that show the duration of recovery from COVID-19 in different contexts and settings are necessary to tailor appropriate treatment and prevention measures. Most of the published studies conducted on the duration of SARS-COV-2 shedding among COVID-19 patients are from China and Europe. There are few studies conducted in Africa where the epidemic seems to be different from that of other continents with regards to speed of the spread of the virus and the death toll recorded.

Hence this study aimed to investigate predictors of recovery from SARS-CoV-2 infection among COVID-19 patients admitted to Hawassa Referral Hospital, and Treatment Center, Sidama Region, Ethiopia. Understanding the factors associated with the duration of viral ribonucleic acid (RNA) shedding, the time from infection to viral RNA-negative conversion in patients with Coronavirus disease 2019 (COVID-19), is crucial in formulating preventive measures and optimizing treatment options (Yin Y. and Wunderink R.G.,2020). Centers for Disease Control and Prevention (CDC) recommends two consecutive negative results of SARS-CoV-2 by Real-time Reverse Transcriptase Polymerase Chain Reaction (rRT-PCR) in 24 hours to conclude a patient's recovery from COVID-19 (Huang C., et al., 2020).Therefore, this study aimed and planned to determine the recovery time of COVID-19 patients and factors associated with it. Among Covid-19 patients admitted to hospital length of stay (LOS) is a crucial variable. It is the time elapsed between a patient's admission to discharge from the hospital isolated and treatment center. In Hawassa Referral Hospital, the recovery rate from COVID-19 in the present analysis as the number of recovered cases per total positive coronavirus cases has emerged to be one of the major indicators of measuring the accomplishment of a country in effectively restraining the spread of the current coronavirus pandemic. Hawassa city is one of the fast growing cities in Ethiopia. It is the largest city in Sidama Region Ethiopia. There are a number of economic activities that have been undertaken in Hawassa city and the nearby areas. Both southern Sidama regional head bureaus are located in Hawassa city. Besides, Hawassa city accommodates a number of regional bureaus for international non-governmental organizations, civic societies, financial institutions and multinational companies. Many private businesses mainly focusing on hotels and tourism, coffee export business, factory product exporters like flour and textiles are actively involving in the economic activity of the country. The current study, thus, tried to investigate why the recovering time for Covid-19 patients had different or wide range understand the association between recovery time and several other socioeconomic determinants. The gap of previous studies was the data was analyzed only using descriptive statistics and non-parametric estimation. However, this study tried to use semi-parametric and parametric survival function.

The study attempted to answer the following research questions:

1. What are the determinant factors affecting the recovery time of COVID-19 patients be absolutely different in the same treatment center?
2. What is the mean recovering time of COVID-19 patients in Hawassa Referral Hospital?
3. Which survival distributions models fit the recovery time data?

1.3 Objectives of the Study

1.3.1 General objective

The general objective of the study was to identify the key socio-economic, demographic and clinical determinant factors of recovery time of COVID-19 patients in Hawassa isolation treatment center.

1.3.2 Specific objectives

The specific objectives of the study are to

- To identify the determinants of recovery time of COVID-19 in Hawassa Referral Hospital;
- To analyze the mean recovering time of covid-19 patient in a study area;
- To find out the best fit model for recovery time data;

1.4 Significance of the study

The results of this study may provide information on causes of high risk in treatment center by analyzing the impact of different variables on time of events. Specifically, to give some knowledge about determinants and risk factors of variable. The results of this study could be used as input for other studies related to recovery rate .This study could provide information to government and other concerned bodies in setting policies, strategies, and further investigation for improving the recovery time.

1.5. Scope of the study

The general objective of the study would be to identify the determinant risk factors for the recovery of corona virus (covid-19) patients using classical survival approaches. The study targets the corona virus patients found at Hawassa University Comprehensive and Referral hospital**.** The researcher used secondary data from September 20, 2013 to January 20, 2014 E.C

1.6 Definition of terms

Censoring: The survival time of an individual is said to be censored when the end point of interest has not been observed for that individual.

▪ **Right censoring:** censoring for incomplete data is right censoring where a subjects follow-up time terminates before the outcome of interest is observed.

▪ **Left censoring:** an observation is said to be left-censored if all that is known is that the individual developed the event of interest before the beginning of the study.

▪ **Interval censoring:** an observation is categorized into interval censored if it is only known that the event of interest occurs within an interval of time without the knowledge of when exactly it occurs.

▪ **Recovered (cure):** Patient that has reached the discharge criteria.

▪ **Dead**: Patient that has died whiles he/ she was in the program at this stabilization center

1.7. Organization of the paper

The paper is containing five chapters. The first chapter deals with general background of the study, the statement of the problem, objectives, significance of the research and operational definitions of the study are also addressed in this chapter. The second chapter reviews work related to the major problem being investigated. In the third chapter, the method of analysis; namely the variables considered, model fitting and model diagnostics are discussed. In addition, the assumptions of the model under consideration are discussed. Chapter four is devoted to the result and discussion of the estimated parameters of the selected model. Here, the effect of explanatory variables to the response variable, test of parameters and model-based results are described. The last chapter, Chapter five, gives conclusions and Recommendation drawn from the result of study

CHAPTER TWO

REVIEW OF LITERATURES

2.1 Literature Review of Statistical Model for Recovering Time of Covid-19 Patients in Hospital

In different literatures the recovering time of covid-19 patients was carried out using different models. (Zhuo Wang, John S. Ji,2020) carried out survival models for the research entitle Survival analysis of hospital length of stay of novel coronavirus (COVID-19) pneumonia patients in Sichuan, China. From January 16, 2020 to March 4, 2020, 538 human cases of COVID-19 infection were laboratory-confirmed, and were hospitalized for treatment. Among these, 271 (50%) were 45 years of age or above, 285 (53%) were male, 450 (84%) were considered as having mild symptoms. The median hospital length of stay was 19 days (interquartile range (IQR): 14-23, Range: 3-41). Adjusted multivariate analysis showed that longer hospital length of stay was associated with factors aged 45 and over (HR: 0.74, 95% CI: 0.60-0.91), those admitted to provincial hospital (HR: 0.73, 95% CI: 0.54-0.99), and those with serious illness (HR: 0.66, 95% CI: 0.48-0.90); living in areas with more than 5.5 healthcare workers per 1000 population (HR: 1.32, 95% CI: 1.05-1.65) was associated with shorter hospital length of stay with no gender difference.

BMC Infection Disease (BMC, 2021) was applied survival analysis for analyzing predictors of mortality in COVID-19 patients at Kinshasa Medical Center and a survival analysis Descriptive statistics consisted of calculating the mean and standard deviation for quantitative data with Gaussian distribution; the median and interquartile range (IQR) for quantitative data with non-Gaussian distribution. Proportions were used for categorical data and percentage are based on the total number of non-missing value. Pearson's Chi-square test or Fisher's exact test was used to compare the proportions. For continuous variables, the comparisons between the survivor and non-survivor groups were made using student's t-test (variables normally distributed) or Mann Whitney's test (variables not normally distributed). Kaplan Meier's method was used to describe the survival between the date of admission in KMC care and death (complete data) and the end of the study (censored data). The Log-rank test was used to compare survival curves. Factors associated with mortality in unadjusted invariable cox regression were included in a multivariable cox regression model to identify independent factors associated with mortality, the Odd ratio (OR) was calculated for each independent variable The findings demonstrate that age, respiratory rate, proteinuria, procalcitonin and AKI were significantly associated with mortality in COVID-19 patients. Additionally, increasing levels of creatinine during hospital admission were associated with an increased mortality. Mortality rate of COVID-19 patients is high, particularly in intubated patients and is associated with age, respiratory rate, procalcitonin, proteinuria and acute kidney injury.

According to (Vekaria *et al. BMC Infectious Diseases* ,2021) were used Data-driven methods entitle Hospital length of stay for COVID-19 patients On a national scale, relevant patients were identified from the COVID-19 Hospitalization in England Surveillance System (CHESS) reports. An Accelerated Failure Time (AFT) survival model and a truncation corrected method (TC), both with underlying Weibull distributions, were fitted to the data to estimate LoS from hospital admission date to an outcome (death or discharge) and from hospital admission date to Intensive Care Unit (ICU) admission date. In a second approach they fit a multi-state (MS) survival model to data directly from the Manchester University NHS Foundation Trust (MFT) and develop a planning tool that uses LoS estimates from these models to predict bed occupancy all methods produced similar overall estimates of LoS for overall hospital stay, given a patient is not admitted to ICU (8.4, 9.1 and 8.0 days for AFT, TC and MS, respectively). Estimates differ more significantly between the local and national level when considering ICU. National estimates for ICU LoS from AFT and TC were 12.4 and 13.4 days, whereas in local data the MS method produced estimates of 18.9 days.

Furthermore, International Journal of Infectious Diseases (IJID, 2020) described survival analysis of all critically ill patients with COVID-19 admitted to the main hospital in Mogadishu, Somalia, 30 March–12 June 2020 Of the 131 patients admitted to the hospital with COVID-19, 52 (40%) died and 79 (60%) survived. The main factors associated with the risk of in-hospital death were age ≥60 years {survival probability on day 21 was 0.789 [95% confidence interval (CI) 0.658–0.874] in patients aged <60 years vs 0.339 (95% CI 0.205–0.478) in patients aged ≥60 years), cardiovascular disease [survival probability 0.478 (95% CI 0.332–0.610) in patients with cardiovascular disease vs 0.719 (95% CI 0.601–0.807) in patients without cardiovascular disease] and non-invasive ventilation on admission (patients who were not ventilated but received oxygen were significantly more likely to survive than patients who were ventilated; $P<0.001$).

Considering the risk factors mentioned in this study – age ≥60 years, presence of cardiovascular disease and use of non-invasive ventilation is critical when dealing with patients with severe COVID-19 in an environment where a trained and skilled health workforce to manage patients in high-dependency units is limited and critical care services are rudimentary. This study confirms the importance of high levels of hospital emergency preparedness including decision-making in managing critical care.

The findings highlight the importance of the critical aspect of deciding whether or not to ventilate critical patients in a low-resource setting when advanced levels of critical care are not available, and instead optimizing the use of available healthcare resources in these settings, such as the use of medical oxygen to improve the probability of survival. Thus, the study results have important policy implications by highlighting the value of available, accessible and affordable low-cost interventions in a fragile and resource-poor setting to inform case management for severe acute respiratory diseases.

The scaling up of availability of medical oxygen in such a setting will also promote improved access to care for childhood pneumonia and other respiratory diseases, and result in improved outcome in terms of lives saved and deaths averted as lower respiratory infections are the third leading cause of death and second leading cause of disability-adjusted life-years for both sexes in Somalia (GBD 2019 Universal Health Coverage Collaborators, 2020).

International Journal of Population Data Science (IJPDS , 2021) carried out Survival Analysis for the research entitle Length of Stay in ICU of Covid-19 Patients in England, March - May 2020 used Accelerated Failure Time survival models with Weibull and log-normal distributional assumptions to investigate the effect of predictors, which are known to be associated with poor Covid-19 outcomes, on the LoS in ICU Patients admitted before 25 March had significantly longer LoS in ICU (mean = 18.4 days, median = 12), controlling for age, sex, whether the patient received Extracorporeal Membrane Oxygenation, and a co-morbid risk factors score, compared with the period after 7 April (mean = 15.4, median = 10). The periods of admission reflected the changes in the ICU admission policy in England. Patients aged 50-65 had the longest LoS, while higher co-morbid risk factors score led to shorter LoS. Sex and ethnicity were not associated with ICU

LoS.The skew of the predicted LoS suggests that a mean LoS, as compared with median, might be better suited as a measure used to assess and plan ICU beds capacity. This is important for the ongoing second and any future waves of Covid-19 cases and potential pressure on the ICU resources. Also, changes in the ICU admission policy are likely to be confounded with improvements in clinical knowledge of Covid-19.

Journal of Contemporary Medicine (JCM, 2021) using simple linear regression and proportional hazard model entitled of Factors Impacting Length of Hospital Stay of COVID-19 Inpatient Surviving ICU patients have longer hospital stay time than non-surviving ICU patients, which in turn longer than non-ICU patients. Older age with in certain range is always correlated with a longer hospital stay. The factors with strongest signal in our dataset are C-reactive protein (CRP), hemoglobin (HGB) and calcium level: increased CRP level, decreased HGB level and calcium level are associated with longer hospital stay, independent from the contribution from surviving status. Other potentially stay time-impacting factors include d-dimer, urea, glucose, white blood cell(WBC) count, neutrophil, but these signals may not be robust against log transformation of the stay time, or confounding with age. We also observed that glucose is more important than HbA1C or diabetes status in its influence on hospital stay time. Almost all factors we collected contribute to a faster/slower mortality or discharge rate, in particular C-reactive protein and hemoglobin. Measurement of the associated factors in COVID-19 patients could be used for a better hospital bed management.

Zhuo Wang, Yuanyuan Liu (2022) carried out the study on survival analysis entitle “What are the risk factors of hospital length of stay in the novel coronavirus pneumonia (COVID-19) patients?” estimated the relationship between LoS and the possibly determinant factors, including demographic characteristics of confirmed patients, individual treatment behavior, local medical resources and hospital grade.

The Kaplan-Meier method and the Cox Proportional Hazards Model were applied for single factor and multi-factor survival analysis. From January 16, 2020 to March 4, 2020, 538 human cases of COVID-19 infection were laboratory-confirmed, and were hospitalized for treatment, including 271 (50%) patients aged 45, 285 (53%) males, and 450 patients (84%) with mild symptoms. The median LoS was 19 (interquartile range (IQR): 14–23, range: 3–41) days. Univariate analysis showed that age and clinical grade were strongly related to LoS ($P<0.01$).

Adjusted multivariate analysis showed that the longer LoS was associated with those aged > 45 (Hazard ratio (HR): 0.74 95% confidence interval (CI): 0.60–0.91), admission to provincial hospital (HR: 0.73, 95% CI: 0.54–0.99), and severe illness (HR: 0.66, 95% CI: 0.48–0.90). By contrast, the shorter LoS was linked with residential areas with more than 5.5 healthcare workers per 1,000 population (HR: 1.32, 95% CI: 1.05–1.65). Neither gender factor nor time interval from illness onset to diagnosis showed significant impact on LoS.Understanding COVID-19 patients' hospital LoS and its risk factors is critical for governments' efficient allocation of resources in respective regions. In areas with older and more vulnerable population and in want of primary medical resources, early reserving and strengthening of the construction of multi-level medical institutions are strongly suggested to cope with COVID-19 outbreaks.

Akancha Singh, Aparajita Chattopadhyay (2021) used a linear regression model entitle COVID-19 recovery rate and its association with development. The positive role of the health system, better economy, urbanization, and good governance. A good investment in the health system has proven the success of case recovery by furnishing easy access to health centers, good quality of care, and handling emergency health conditions by mobilizing health resources.

(IHME COVID-19 health service utilization forecasting team, Murray, C.J. et.al, 2020) used that Modelling studies predicting bed occupancy published so far have broadly relied on very few sources of information for LoS estimates, which were often derived from very different settings. Estimates for LoS can be obtained from a variety of studies, but are often an incidental result rather than a study's primary outcome, and typically only summary statistics are reported. In general, LoS distributions are right-skewed due to a minority of patients with long hospital stays and are often modelled using gamma, log-normal or Weibull distributions (although log-normal is less preferred due to its heavier tails).A particular challenge is how to synthesize appropriate LoS distributions from a range of relevant sources in similar settings, capturing the variation both within and between them. Incorporating the uncertainty and stochasticity in parameters using a distribution, rather than fixed point estimates (such as the mean over all studies), allows for more realistic model predictions.

To identify the evidence on LoS for COVID-19 patients worldwide. they also presented a method for generating LoS summary distributions by combining information from different summary statistics (mean and medians) reported in multiple studies, and accounting for differences in sample sizes. In doing this work, we aim to inform the efforts of modelers and policy makers to better anticipate healthcare needs during the evolving COVID-19 pandemic.

Overall summary distributions were created for general hospitalization LoS and for ICU LoS. We included studies in the estimation of these summary distributions if they reported both the sample size along with either the median and interquartile range or the mean and standard deviation. If no measure of variation was provided (either IQR or standard deviation), the point estimates were included in figures but excluded from these summary results. A Weibull distribution was fitted to the summary data from each grouping (by country setting and general/ICU classification) in the appropriate studies.

According to Seyed Alinaghi *et al*., 2021 cross-sectional study was implemented in the COVID-19 clinic of a teaching and referral university hospital in Tehran. Patients with the highly suggestive symptoms who had computed tomography (CT) imaging results with typical findings of COVID-19 or positive results of reverse transcriptase-polymerase chain reaction (RTPCR) were enrolled in the study. Inpatient and outpatient COVID-19 participants were followed up by regular visits or phone calls, and the recovery period was recorded. Results: A total of 478 patients were enrolled. The mean age of patients was 54.11±5.65 years, and 44.2% were female. The median time to recovery was 13.5 days (IQR: 9). Although in the bivariate analysis, multiple factors, including hypertension, fever, diabetes mellitus, gender, and admission location, significantly contributed to prolonging the recovery period, in multivariate analysis, only dyspnea had a significant association with this variable (p=0.02, the adjusted OR of 2.05; 95% CI 1.12–3.75).Conclusion. This study supports that dyspnea is a predictor of recovery time. It seems like optimal management of the comorbidities plays the most crucial role in recovery from COVID-19.

2.2 Empirical Studies on Recovering Time of Covid-19 Patients in Ethiopia

In Ethiopia, multiple interventions were implemented to prevent the COVID-19 pandemic before it causes a substantial impairment to the community. The government declared a state of emergency, and established a COVID-19 taskforce at national who were informing disease prevention measures, deliver regular situational updates, and organize massive awareness creation efforts using diverse social and mass media stages. The COVID-19 catastrophe sends a solid message to resistant health systems that can only be realized with a committed health workforce. Protecting everyone requires urgently addressing shortages of health workers, updating infection prevention measures, investing in capacity building, and warranting safe working environments (Chen Y, Tong X, Wang J, Huang W, Yin S, Huang R, 2020).

According to Demisu Zenbaba et al., 2021 Socio-demographic-, knowledge-, and health facility related factors to compliance towards COVID-19 prevention measures were identified using an ordinary logistic regression model. Health professionals who working in general hospitals were 45% times less likely to have good compliance towards COVID-19 prevention measures than health professionals working in referral hospitals (AOR = 0.55; 95% CI 0.38, 0.79). The odds of having good compliance towards COVID-19 preventive measures were 2 times more likely among health professionals with 3–6 service years than health professionals who had ≤ 2 service years (AOR = 2.10; 95% CI 1.35, 3.21). Health professionals who have a good knowledge regarding COVID-19 preventive measures were 1.80 more likely to have good compliance regarding COVID-19 preventive measures than their counterparts (AOR = 1.80; 95% CI 1.14, 2.89). Similarly, the odds of having good compliance were nearly 3 times more likely (Misganu Endriyas ,Aknaw Kawza, Abraham Alano, Mamush Hussen,Endashaw Shibru,2020).

After running multivariate binary logistic regression, sex, educational status, family size and overall knowledge about COVID-19 were associated with practicing COVID-19 prevention measures. Women were 44% more likely to practice COVID-19 prevention measures as compared with male respondents (AOR at 95%CI: 1.44 (1.08 to 1.92)). Those who attended primary and secondary education, and had certificate and above education were 89%, 74% and 89% more likely to practice COVID-19 prevention measures, respectively, as compared with those who never attended education. Respondents from larger (≥6 members) household were 39% more likely to practice COVID-19 prevention measures as compared with those who had fewer number of family members (≤5 members) (AOR at 95%CI: 1.39 (1.04 to 1.84)). Respondents who have medium-level and high-level knowledge scores were 2.04 and 2.47 times more likely to practice COVID-19 prevention measures as compared with respondents with poor knowledge scores (AOR at 95%CI: 2.04 (1.51 to 2.75) and 2.47 (1.80 to 3.40), respectively)

According to Seyed Alinaghi *et al*., 2021 the mean age of patients was 54.11±5.65 years, and 44.2% were female. The median time to recovery was 13.5 days (IQR: 9). Although in the bivariate analysis, multiple factors, including hypertension, fever, diabetes mellitus, gender, and admission location, significantly contributed to prolonging the recovery period, in multivariate analysis, only dyspnea had a significant association with this variable (p=0.02, the adjusted OR of 2.05; 95% CI 1.12–3.75). Wubedle Zelalem Temesgan et al., 2021 carried out a community-based cross-sectional study was conducted from July 1st to 30th, 2021, in Gondar city. In this study, adherence towards COVID-19 preventive practice in pregnant women is low. Hence, it is important to strengthen women's awareness about COVID-19 through different media and health education. In addition, empowering women to attain ANC and special consideration should be given to women who had no formal education.

Tinsae Abeya Geleta, Berhanu Senbeta Deriba , Kemal Jemal , 2022 carried out an institutional-based cross-sectional study was conducted from May 5/2020 to June 5/2020 in public hospitals in the North Shoa zone, Ethiopia. Data were collected using a structured questionnaire and study participants were recruited using a simple random sampling technique. The data were checked for completeness and entered into the Epi Data manager version 4.4.1 and transferred to SPSS version 23 for analysis purposes. Bivariate and multivariate logistic regression was computed and a significant association was declared with a p-value less than 0.05.This study finding revealed that the level of knowledge, attitude, and practice towards COVID-19 prevention among hypertension patients was low. Therefore, increasing knowledge, attitude, and practice on COVID-19 among hypertension patients requires a coordinated effort from the government, non-government, and health professionals.

Gemechu Churiso, Kuma Diriba , Henok Girma, Soressa Tafere ,2021 carried out a retrospective study design was conducted in 220 patients confirmed by real time polymerase chain reaction and admitted to Dilla University Referral Hospital treatment center from September 2020 to July 2021. Descriptive statistics were used for clinical features, and median time to recovery was computed by using Kaplan–Meier. They found that COVID-19 patients frequently show cough, shortness of breath, fever, headache, easy fatigue and joint pain. Median time to recovery was 5 days. Having a normal body temperature, normal breathing rate, and severe disease status had statistically significant association with median recovery time. So, close follow up is required for client admitted with severe disease. Tadesse Tolossa et al,.2021 carried out a hospital-based retrospective cohort study conducted among 263 adult patients admitted with COVID-19 in WURH treatment center from March 29, 2020 through September 30, 2020. A Cox proportional hazard regression model was fitted to determine factors associated with recovery time. A variable with P-value ≤ 0.25 at bivariable Cox regression analysis were selected for multivariable Cox proportional model. Multivariable Cox regression model with 95% CI and Adjusted Hazard Ratio (AHR) was used to identify a significant predictor of time to recovery from COVID-19 at P-value < 0.05.the median recovery time of patients with COVID-19 cases was long, and factors such as older age group, presence of fever, and comorbidity was an independent predictors of delayed recovery from COVID-19.

2.3 Empirical Studies on Recovering Time of Covid-19 Patients in Other Countries

One particular issue in the COVID-19 pandemic is that, due to the fast transmission of virus by asymptomatic carriers, the number of patients can increase very fast in a short period of time. Even if only a percentage of them need medical care, the hospitals can be flooded with patients requiring beds, and an even smaller percentage needs Intensive Care Unit (ICU) and ventilators. The number of both the regular beds and ICU beds in a hospital are limited, and during peak time for the number of infected, a hospital can be in a crisis. To manage the crisis, it is very important to know what patient's information can potentially be used to predict their hospital stay length. With that information, it is possible to anticipate if the numbers of beds are sufficient. A hospitalized patient may have two outcomes: mortality/death and discharge.

Although there is a third possibility at the time of data collection, i.e., the patient is still in the hospital, it is less common for COVID-19 patients. A COVID-19 patient may be either discharged or die in the matter of days. If the two outcomes are treated separately, the time-to-event data is usually dealt with by the competing-cause or competing-risk survival analysis. For the issue of hospital stay, the two types of events are equivalent: a bed will become available whether the patient is discharged or has died. Therefore, the competing-cause survival analysis can become a regular survival analysis if there are right censored data. If there are no right-censored data, the analysis strategy is even simpler: it is a simple analysis relating a factor, which can either be categorical/discrete or continuous, and the hospital stay time.

The goal of study was to determine factors influencing the hospital stay time. In this paper, we aim at finding which factors are associated with long hospitalization stay time for COVID-19 inpatients. This question is related, but not identical, to the question of which factors are associated with severity of COVID-19 disease. More specifically, the associated factor should cause severity of the disease, but not too severe so that the patient dies earlier. Comparing that two associated factors, CRP and HGB, there had been reports linking CRP level to severity of COVID-19 disease, and discussion on how anemia affects the quality of life in elder COVID-19 patients these results are not directly on the hospital stay time, they show potential mechanisms linking a factor with the stay time. As we have argued already, that the same mechanism can either cause a patient to stay longer in the hospital or cause another patient to stay less timed-dimer, glucose, urea, calcium, creatinine, potassium, sodium, WBC, neutrophile, lymphocyte, NEU/LYM ratio, BUN/CR ratio, PLT/LYM ratio.[21-34] The fact that every single one of the factors we found to be significantly associated with rate of mortality/discharge.

Leclerc QJ, Fuller NM, Keogh RH, et al, .2021. Among publications on the topic of hospital stay time for COVID-19 patients, some of them only address the overall statistics or difference between countries, on bed types and not on factors which might impact the stay time.it was reported that median hospital stay is 14 days in China, whereas it is 5 days outside of China.

Rees EM, Nightingale ES, Jafari Y, et al, 2020 entitled “COVID-19 length of hospital stay a systematic review and data synthesis the paper estimated that majority of ICU hospital stay time is between 7 to 11 days.(Lane EA, Barrett DJ, Casey M, et al,2020) These are comparable to our hospital stay values.For studies of association between factors and hospital stay time, although the types of data varies (e.g. some only include non-severe COVID-19 patients), or the analysis plan is different (e.g., hospital stay time is binarized into longer or shorter than 14 days), the study do find some interested comparisons.

Liu X, Zhou H, Zhou Y, Wu X, Zhao Y, Lu Y, etal., 2020 entitled Risk factors associated with disease severity and length of hospital stay in COVID-19 patients It was observed that patients with low lymphocyte count (lymphopenia) stay in hospital longer with a log-rank test p-value of 0.027 for Kaplan-Meier curve. Other analyses of hospital stay time applying the survival analysis techniques can be found in the literature and Factors potentially associated with hospital stay time, e.g, obesity, which may have a nonlinear impact on COVID-19 severity, the prescription of drugs. (Şirin Çetin1, Ayse Ulgen 2, Hakan Şıvgın 3, Wentian,et,al.,2021)carry out A Study On Factors Impacting Length of Hospital Stay of COVID-19 Inpatient find that the median length of hospital stay for the study sample was 5.5 days (IQR 9 days), with a range of 1–35 days and a mean of 7.7 days (SD 6.9). Study in Vietnam, the median length of hospital stay was 21 (range 16–34) days (Thai et al., 2020),while in the USA and the UK, the length of stay was shorter with an average of 7–8 days (Bhatraju et al., 2020; Bialek et al., 2020;Docherty et al., 2020).

Two-thirds of patients were male (69%). The mean age of male patients was 58.5 years vs 56.9 years for females. These findings are consistent with studies in Vietnam (Thai et al., 2020) and China (Yang et al., 2020) which suggest a gender difference in hospitalized patients with COVID-19. The signs and symptoms by age group and sex in the study do not differ from those reported in studies outside of Africa (Bhatraju et al., 2020; Bialek et al., 2020; Docherty et al., 2020; Salinas-Escudero et al., 2020; Thai et al.,2020; Wu and McGoogan, 2020; Wu et al., 2020); however, there are no comparable data from African settings, especially subSaharan African countries. The main found of the study contributes to an understanding of the risk factors for in-hospital death from COVID-19 in Somalia. Risk factors for in-hospital death from COVID-19 have not been studied or published from sub-Saharan countries previously.

In addition, the study has used findings from survival analyses to assess which interventions have a higher probability of improved clinical outcome of patients with severe COVID-19 in resource-poor settings by early use of medical oxygen. Considering the risk factors mentioned in this study – age ≥60 years, presence of cardiovascular disease and use of non-invasive ventilation – is critical when dealing with patients with severe COVID-19 in an environment where a trained and skilled health workforce to manage patients in high-dependency units is limited and critical care services are rudimentary. This study confirms the importance of high levels of hospital emergency preparedness including decision-making in managing critical care. The findings highlight the importance of the critical aspect of deciding whether or not to ventilate critical patients in a low-resource setting when advanced levels of critical care are not available, and instead optimizing the use of available healthcare resources in these settings, such as the use of medical oxygen to improve the probability of survival. Thus, the study results have important policy implications by highlighting the value of available, accessible and affordable low-cost interventions in a fragile and resource-poor setting to inform case management for severe acute respiratory diseases. .The scaling up of availability of medical oxygen in such a setting will also promote improved access to care for childhood pneumonia and other respiratory diseases, and result in improved outcome in terms of lives saved and deaths averted as lower respiratory infections are the third leading cause of death and second leading cause of disability-adjusted life-years for both sexes in Somalia. According to Rosenbaum,L,Facing et.,2020 carried out research entitle Covid-19 in Italy | Ethics, Logistics, and Therapeutics on the Epidemic's Front Line. Understanding and predicting hospital bed demand (as well as associated staff or equipment requirements) provides crucial evidence for decision-making and contingency planning. Predicting demand for hospital services requires an estimate of the number of patients requiring hospitalization, and an estimate of how long each person will require hospital care. It is possible to model the rate of hospitalization in many settings based on estimated epidemic curves. However, estimating length of stay (LoS) in hospitals requires observation of individual patient pathways.COVID-19 presents at varying levels of severity. Hospital care can vary from general ward based care to high dependency units with oxygen support to intensive care where patients may be intubated for mechanical ventilation. The LoS is likely to depend on the level of care required, as well as the geographic setting due to varying COVID-19 care guidelines. Some hospitals in China were initially used as isolation settings. As knowledge of effective treatments changes, the pathways, staff, beds and equipment required are also likely to affect the duration and level of care needed. Moreover, patient characteristics - such as age and comorbidities - impact disease severity and are likely to influence LoS. If differences are significant then capacity planning may need to account for these characteristics to provide accurate predictions of the number of beds required at each level of care.

The median duration of hospitalization ranged from 5 to 29 days. There was no observed trend with respect to when the study was conducted. Estimates for LoS amongst patients who died in hospital were generally shorter than those for patients who were discharged alive, with medians between 4 and 21 days compared to 4 and 53 days, respectively. Where median LoS was lower for those discharged alive in 6 out of 8 studies that reported both outcomes. In studies that reported general hospital LoS by disease severity there was a trend towards more severe cases having longer LoS. However, the definition of different levels of severity was inconsistent between studies so it is not possible to draw any confident conclusion. Visual inspection of the study estimates suggested some evidence of a difference between general hospital LoS reported within and outside China, but studies outside China were too few. However, LoS reported within the ISARIC report in particular (which includes contributed data from 25 countries, but with the majority of patients from the UK) gave a median and IQR (4 days (1 - 9)) substantially lower than the weighted mean from the studies from China (15.3 days) Shi, C &Wang, C. et al., 2020).

Among patients discharged alive there appears to be little difference in average LoS between studies with the youngest and oldest patients, but the longest estimates came from studies with average age in the upper end of the range with average age of 68 and 69, respectively; . The LoS estimates which included non-survivors tended to come from studies with older populations, as is to be expected given the well-documented, age-dependent fatality rate .Estimated summary hospital LoS distributions for studies from China and studies outside China are shown in Fig. 4. The median and IQR for general hospital was estimated to be 14 (10-19) for China and 5 (3-9) excluding China. This was also repeated for ICU LoS, with a median and IQR 8 (5-13) for China and 7 (4-11) outside China. Studies from China which had complete follow-up with respect to general hospital LoS were compared with studies with incomplete follow-up (Wang, Shi *et al.*, 2020).

Understanding how long patients hospitalized with COVID-19 remain in hospital is critical for planning and predicting bed occupancy as well as associated staff and equipment needs. This review found that hospital LoS observations for COVID-19 patients published in the literature to date varied from less than a week to nearly two months. Stay in intensive care was shorter and less variable, with studies reporting medians of one to three weeks. Where LoS was reported according to discharge status, stay was found to be shorter for those who died than for those discharged alive; however, this difference was only apparent in terms of overall stay and not stay in ICU (no statistical comparison was made). With respect to practical implications, knowledge of a difference between survivors and non-survivors is of less use since the outcome will not be known in advance in order to influence decision-making.

The studies found that LoS is often not the primary measure of interest in studies which report it, however it is an important parameter when it comes to forecasting bed occupancy during an outbreak. By conducting this review we have systematically gathered a range of published estimates, providing a source from which researchers and decision makers can obtain estimates specific to their population of interest (e.g. with respect to comorbidities) and allowing comparison of LoS between several different populations and settings. There have been numerous previous studies which have aimed to forecast the number of hospital beds required for COVID-19 patients [16{22,54]. Many of these studies published so far have used point estimates, only originating from one study which often does not reflect the context of interest. In particular, many used estimates from (Zhou *et al, 2020).*

The LoS estimate is a critical parameter within a bed forecasting model, and as such any model is likely to be very sensitive to the value or distribution being assumed, with huge implications for policy and planning. This review has highlighted several potential sources of variation in LoS, and identified common issues and biases which influence each individual estimate. This gives a motivation for considering a wider range of values than can be obtained in a single study, aiming instead to capture the overall distribution of LoS across avariety of possible patient trajectories. Here we have included estimates of the overall LoS distribution in two settings (China/Other) for which we obtained sufficient data. It is preferable to use data from the setting for which you are trying to forecast bed occupancy (as was done by the IHME COVID-19 health service utilization forecasting team , however data on completed patient stays will often not be available until well after the onset of the epidemic. Furthermore, LMICs may have reduced capacity for surveillance and monitoring in order to obtain these data. In such cases, where countries are in the early stages of an outbreak, it would be better to use a conservative (i.e broad) distribution of LoS from another setting.

This review summarized the available literature to provide estimates of LoS for general admission and ICU which can be applied for planning and preparedness for SARS-CoV-2. We found substantial differences between China and other settings in terms of total hospital stay, but little evidence for an impact on LoS of time of study, age or disease severity. We present summary distributions which can be used within models making predictions about bed requirements, and suggest that this may be a more robust and realistic way to characterize LoS than relying on summary data from just one setting or hospital. The majority of the data presented in this review comes from China and, as more data become available, it will be important to update this with setting-specific LoS estimates. Understanding the duration of hospitalization of COVID-19 patients is critical for providing insights as to when hospitals will reach capacity, as well predicting associated staff or equipment requirements.

A few people are asymptomatic but they may also transmit the virus, although relatively less than symptomatic people. Everyone is at risk for getting COVID-19 if they are exposed to the virus and some may require hospitalization and intensive care. World Health Organization (WHO) stated that almost 85% of COVID-19 cases have mild to moderate symptoms and illness, 10% to 15% people have severe symptoms and out of these severe symptoms only 5% may require ICU care. According to WHO, the median time from onset to clinical recovery for mild cases is approximately 2weeks and 3 to 6weeks for patients with severe or critical disease. The incidence of COVID-19 is still rising worldwide, though better controlled in India. Hence, it is essential to study the factors influencing the length of stay so that the pathological mechanisms underlying these factors can be dealt with to give effective interventions or treatment. Since the cases have been rapidly increasing it is important to address the length of stay, by which the availability of beds in the hospital, especially in the Intensive Care Unit (ICU) can be addressed. Though clinical presentations have been analyzed with respect to recovery time, comorbidities, and certain laboratory parameters such as creatinine, Lactate dehydrogenase (LDH), Ferritin, hemoglobin, neutrophil-lymphocyte ratio (NLR), and D-dimer have not been addressed in earlier studies as a whole to model recovery time.

Clinical symptoms like fever, cough, sore throat, breathlessness, myalgia, loss of smell and taste, vomiting, and loose stool were included. Laboratory parameters such as oxygen saturation (SpO2), creatinine, lactate dehydrogenase (LDH), ferritin, hemoglobin (Hb),total leucocyte count (TC), differential count (neutrophil, lymphocyte counts, and neutrophil lymphocyte ratio-NL ratio derived), and D-dimer were collected. Survival analysis is a form of data analysis in which the outcome variable is time till the event occurs.

Most important factors that influenced the LOS were oxygen saturation, presence of more than 2 comorbidities and certain laboratory parameters such as LDH, ferritin, D-dimer, neutrophil-lymphocyte ratio.. In early studies from India, the mean length of hospital stay was seen to be as long as 17 days, IQR 15 to 20days (where LOS was influenced significantly only by the presence of SARI or a travel history) 13 to 24 days (16-34days at 95% CI; the length of hospital stay in this study was influenced only by gender and age).

According to Barman MP, Rahman T, Bora K, Borgohain C.,2020 in entitle COVID-19 pandemic and its recovery time of patients in India: a pilot study A study done in Belgium reported that the length of stay ranged from 3 to 10.4 days and it significantly increased with the age of the patient.

(Faes C, Abrams S, Van Beckhoven D, Meyfroidt G, Vlieghe E, Hens N;et al.2020)A study which was done among 538 confirmed patients between January and March 2020 in China found that the median hospital stay was 19 days with interquartile range of 14 to 23 days and it was influenced by age, serious illness, and density of health care workers.

(Ji JS, Liu Y, Liu R, et al.2020) In a study from Vietnam on 251 patients, the median duration of hospitalization was seen to be 16 days and they found that age, residence of the patient, and the source of infection influenced the LOS significantly.

Rees EM, Nightingale ES, Jafari Y, et al ,2020) which analyzed 52 studies from China and the rest of the world between 24th December 2019 and 16th April 2020, it was found that the length of hospital stay was much longer for patients in China when compared to the rest of the world [14 days (Inter Quartile Range (IQR) 10-19) for China and 5days (IQR3-9) for the rest of the world] which is probably attributable to the fact that China was the first to face the pandemic and the rest of the world learnt from them. As can be seen, there have not been many studies analyzing the factors affecting hospital LOS which could help us in making contingency plans to build infrastructure during the pandemic. We have not studied the impact of treatment modalities on the length of stay as it will be too elaborate from a contingency planning point of view.

The only means to prevent spread of COVID-19 are universal masking, social distancing, hand hygiene, and vaccination but getting even 50% compliance in community driven initiatives is a far-fetched dream. Hence, we may have to wait for vaccination as the next best option. But, at the rate at which new strains are emerging in this small, very connected world, the need of the hour is to devise a ready planner such that if need be, more infrastructure can be created or readied in time so that patients don’t suffer. The estimated length of stay of these patients is needed so that we can model bed-occupancy and make contingency plans. Such studies from various parts of the world soon so that we can manage the growing number of patients more efficiently.

It's being more than a year since the first case of COVID-19 reported in Wuhan, China. Since then, the pandemic has affected millions of lives around the globe. It has interrupted routine healthcare delivery and challenged healthcare infrastructure of even developed countries. Length of hospital stay (LOS) for a disease is a vital estimate for healthcare logistics planning.

According to TsaiPF,ChenYY,et al.,2020 decreased LOS is reported to be associated with lowered risk of hospital acquired infection, reduction in financial burden for treatment among the patients, higher bed turnover rate of the hospitals (increases bed availability for the other patients) and vice versa can be told for increased LOS.

According to (Baek H,Cho M,Kim S,Hwang H,Song MYoo Set al.,2018 For COVID-19 the reported associates of LOS are age, gender, nutritional status, presenting symptoms (ie, fever, breathlessness, fatigue, anorexia etc.), co-morbidities (ie, hypertension, heart disease, diabetes etc.), vitals (ie, respiratory rate, blood pressure etc.) laboratory parameters [ie, d-dimer C-reactive protein (CRP), leucocyte count, lactate dehydrogenase (LDH), aspartate aminotransferase (AST) etc.], radio graphical parameters (ie, chest X-ray, CT scan findings etc.), and medications (ie, tocilizumab, ACEI, ARB, metformin etc.). Although in terms of healthcare logistic planning the factors that could predict LOS of COVID-19 patients early bears additional importance. Patient characteristics and vital signs at the time of admission are being already linked with the level of care requirement, disease progression and mortality among COVID-19 patients by some prior studies.

According to Sands KE,Wenzel RP, McLean LE,et al.,2020 Although considering specifically for LOS, no such prior attempts have been made in Indian context. Early identification of factors influencing LOS of COVID-19 patients may help policymakers to plan healthcare logistics (man, money, and material) for the disease accordingly. On extensive literature search we could only retrieve 2 prior Indian studies which have explored LOS among COVID-19 patients.(Neeraj Agarwal, Bijit Biswas et al.2012) entitled Early Determinants of Length of Hospital Stay: A Case Control Survival Analysis among COVID-19 Patients admitted in a Tertiary Healthcare Facility of East India The median age of the studied COVID-19 patients was 55years with interquartile range (IQR) of 40 to 65years.Overall, we found the median LOS for the survivors to be 8 days [interquartile range (IQR): 7-10days] while the same for the non-survivors was 6 days [IQR: 2-11 days]. Survivors had more probability of higher LOS compared to non-survivors. In univariate cox-proportional hazard model; higher age, number of co-morbidities and travelling distance to seek healthcare, self-reporting to the healthcare facility, complaint of breathlessness, deranged RR, SpO2, PR, SBP,DBP, GCS, and SOFA score on admission increased LOS while complaint of sore throat on admission decreased the same. Similarly, among the reported co-morbidities, COPD/asthma was found be associated with slowest recovery followed by CKD, diabetes, hypertension and ischemic heart disease. The median days since symptomatic prior to admission was 6 days with IQR of 4 to 8 days. The duration of symptoms prior to admission was not found to be associated with LOS [HR: 0.85, 95% CI: (0.63-1.16)] In the multivariable cox-proportional hazard model; travel distance (>16 km) [adjusted hazard ratio (aHR):0.69, 95% CI: (0.50-0.95)], mode of transport to the hospital (ambulance) [aHR: 0.62, 95% CI: (0.45-0.85)],breathlessness (yes) [aHR: 0.56, 95% CI: (0.40-0.77)], number of co-morbidities (1-2) [aHR: 0.66, 95% CI: (0.47-0.93)] (≥3) [aHR: 0.16, 95% CI: (0.04-0.65)], COPD/asthma (yes) [aHR: 0.11, 95% CI: (0.01-0.79)], DBP (<60/≥90) [aHR:0.55, 95% CI: (0.35-0.86)] and qSOFA score (≥2) [aHR: 0.33, 95% CI: (0.12-0.92)] were the significant attributes affecting LOS of the COVID-19 patients.

According to Tadess ,Tolossa Dejene Seyoum Geber Emiru Merdassa,Atomass,Getahun Fentensa et al,2021 carried out study entitle Time to recovery from COVID-19 and its predictors among patients admitted to treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia: Survival analysis of retrospective cohort study found the prolonged recovery time from coronavirus disease. The study revealed that older age, fever at admission, and having at least one comorbid condition as a poor prognostic factors of novel coronavirus disease. Thus, elders and individuals with comorbidity has to get due attention to prevent infection by the virus. Moreover, elders and patients with comorbidity should get priority in management of coronavirus disease in order to enhance good clinical outcome.

2.4. Conceptual frame work of the study

A conceptual framework in research is a textual or visual depiction of the anticipated relationship between the variables under study. It depicts what a researcher expects to find through their research and maps out the steps that must be carried out through the course of the study. As the conceptual frame work illustrates a researcher's understanding of how variables connect, it can be used to identify the key variables that need to be investigated. It acts as a map that provides researchers with a shape and structure to their research, helping them carry out their study more effectively.

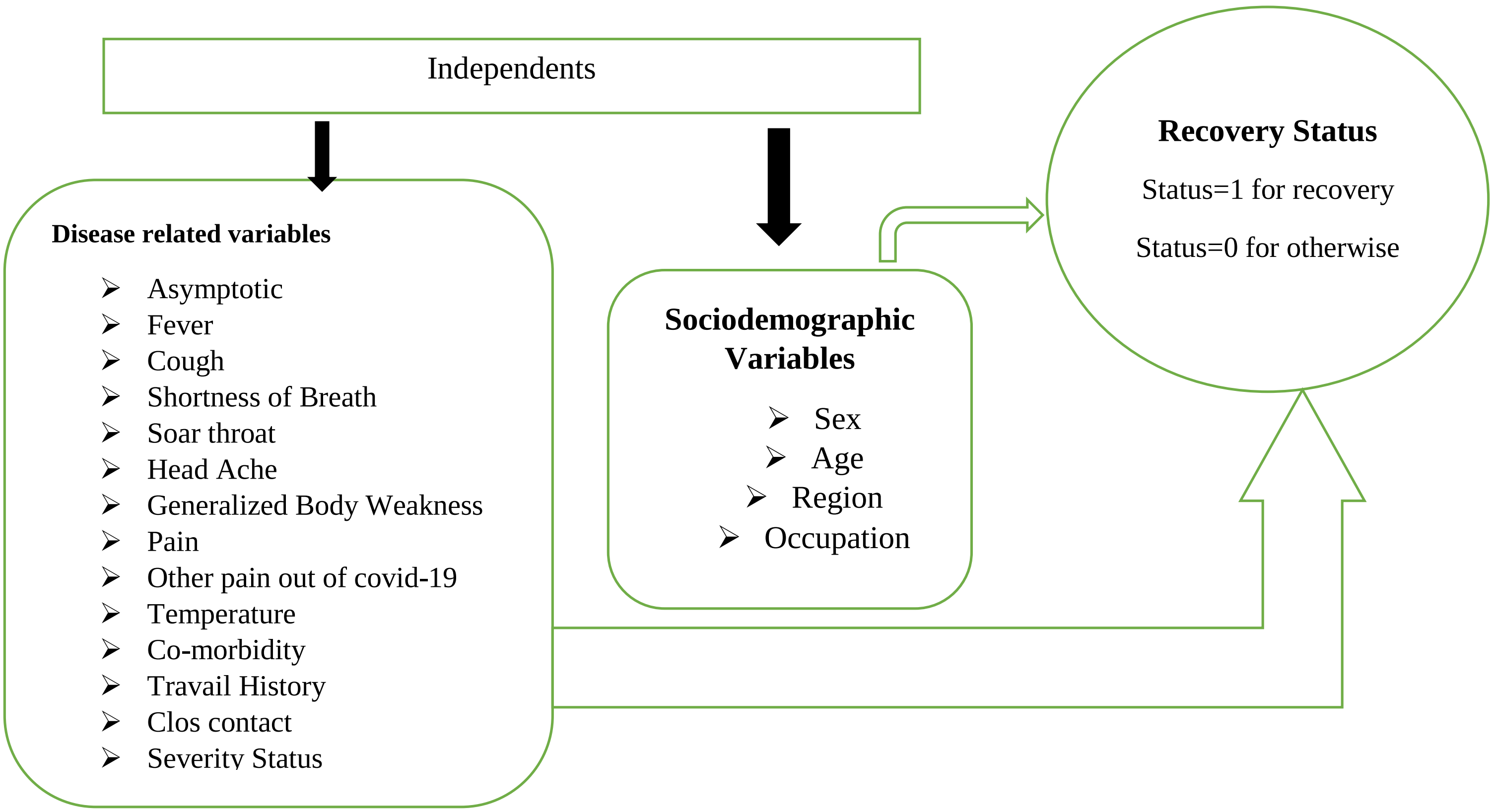


Figure 2.1 conceptual frame work for the predictors.

CHAPTER THREE

MATERIALS AND METHODOLOGY

In this chapter study area, source or nature of data and methodology part for analysis of survival analysis in the covid-19 disease to determine the main factors for recovery time of coronavirus and associated factors, Cox PH model, Survival Weibull Regression model, approaches, model comparison, model diagnosis, and survival analysis were presented.

3.1 Description of Study Area

This study was carried out at Hawassa Referral Hospital Hwassa city, 275km away from Addis Abeba. Geographically, Hawasssa lies between $07^0 05'$ north,$38^0 29'$ east and its altitude was 1708 meters above sea leabel(Med Liberary.org).Beased on report from the Central statistical Agency of Ethiopia, Hawassa city had an estimated total population of 159,013 out of which 81,984 were male and 770729 femals(CSA,2010).

The study participants were all COVID-19 patients. Who were tested positive for COVID-19 by using rRT-PCR test and admitted to treatment center from September 20, 2013 to January 20, 2014 E.c with a definite outcome (event or censored) and whose chart is available during the data collection period. Patients with incomplete outcome variable were excluded from the analysis.

3.2 Study Population

Based on a retrospective cohort study, information of 826 total Covid-19 Patients that had been treated at Hawassa University Comprehensive and Referral hospital from September 20, 2013 to January 20, 2014 E.C was be assigned for this study and those patients come to hospital for treatment more than ones times must be excluded from the study because of the patients id number is already recorded in the documents for reducing redundancy case.

3.3 Sample Size Determination

In conducting researches that require taking a sample, we always have the stage of deciding the sample size. The decision is important because taking too large sample implies waste of resources while too small sample reduces the usefulness of the results. In order to have an optimum sample size, there are a number of issues/points one has to take into account. Some of the issues are: objective of the research, design of the research, cost constraint, degree of precision required for generalization and etc. Based on the above information, there are several formulas developed for sample size calculation that conform to different research situations. Accordingly, the sample size determination formula (Cochran, 1977) is adopted for this study.

$$n = \frac{\frac{z^2 p(1-p)}{d^2}}{1 + \frac{1}{N}[\frac{z^2 p(1-p)}{d^2} - 1]}$$

Where n=the sample size needed, N=the total population size,Z is the upper $\frac{\alpha}{2}$ points of standard normal distribution with $\alpha = 0.05$ significant level. Suppose the maximum allowable difference between the maximum likelihood estimate and the unknown population parameter, denoted by d, desired to be 0.021. The specification of d must be small to have a good precision. The parameter p represents proportion of recovery due to covid-19 disease. Few previous studies describe the proportion of recovery due to covid-19 in Ethiopia .in this study the estimated proportion of the recovery rate from covid-19 disease to be=0.07(Acording to Leulseged T.W,2020).

A retrospective cohort study design was be used to retrieve relevant information from the medical records of corona virus (covid-19) patients to address the objective of the study

3.4 Inclusion and Exclusion Criteria

In our study all Covid-19 patients who were recorded in the medical record room of Hawassa University Referral and Comprehensive Hospital and those cards which have the vital data for the research were included in the study. All patients whose age was recorded in their treatment card were included in our study without any restriction on the age of patients .However the study could exclude those patients who were incomplete information and transferred to other health treatment center and those patients follows treatment ah Referral hospital was discharged and come again in study time for Covid-19 Treatment.

3.5.1 The Response Variables

The dependent or response variable is the waiting time until the occurrence of the event (recovery) (status has 0 for censored observation, 1 for patients who recovered during the study period for our study total recovery are 646(78.2%) and censored death 181(21.8%).

3.5.2 Predictor (Explanatory) (Independent) variables

The predictor variables in survival data analysis are called covariate. These Covariates can be categorical or continuous. The predictor (covariate) variables which assumed to influence the survival of Covid-19 patients included in literature and in our present study are:

Sex	Head Ache
Age	Generalized Body Weakness
Region	Pain
Occupation	Temperature
Asymptotic	Others Pains out of covid-19
Fever	Under comorbidity and Number of comorbidity
Cough	Travel History in the last 14 days
Shortness of breath	Closed Contact with Confirmed
Soar throat	Severity Status

3.6 Statistical Model

3.6.1 Survival Data Analysis

What makes survival analysis so Special is that we cannot use ordinary statistical models due to censoring. In the study of survival data one has to wait for the event to occur. When the study ends and the analysis begins. We commonly note that the event of interest has occurred for some individuals and for some others not. We then have two types of data complete and incomplete data. The latter is called censoring in survival and event history analysis. Therefore, survival data almost always incomplete. The statistical terminological for such data is censoring. Now Days, survival analysis techniques have become important tools for analyzing the data belonging to field to medicine, engineering, marketing etc. Although names medical research give it the name of survival analysis. Survival analysis deals with models methods and is used for analyzing data of lifes.one of the common uses of survival analysis in clinical trial is the comparison of survival times of different treatments in some fatal diseases.in our study on Covid-19 disease patients

Survival analysis is different from the other procedures due to following reasons: In Survival analysis response variable is always time.Staggered entries are more common in medical research.

By staggered entries we mean than all individuals in the study do not have the same entrance time. This does not affect the survival analysis as the analysis deals with the length of the observation time and not based on the same entrance. The assumption of normality is not hold in survival analysis as survival data are generally skewed. The commonly used distributions is survival analysis are exponential, Weibull, lognormal, gamma, log-logistic etc.

Concept of censoring which may affect the hazard function. Time dependent covariates. Censoring is common in survival analysis and it is consider as an important feature of survival data.
A data sample is said to be censored when values of the variable are not observed for some of the items in the sample. In medical studies the actual time of death of same subject may not be noted for many reasons e.g they move away or the allocated time for the study elapses prior the events. One of the main reasons of censoring is the limiting duration of study period (Gross Aj,et at,1975).

Survival analysis well suited for such data which are very common in medical health situation can occur due to the following reasons

- ❖ When an individual's survive beyond the study period or the individual does not experience the event.
- ❖ Lost to follow-up that is an individual may drop out transfer to other place etc.
- ❖ Deaths Due to other cases different from that/those specified in the study.

The most common encountered from of a censored observation is one in which observation begins at the defined time, say t=0, and terminates before the outcome of interest is observed. Since the incomplete nature of the observation occurs in the right tail of the time axis such observations are said to be right censoring. In our study data was right censored the other mechanism that can lead to incomplete observation of time is truncation. Truncated observation is one which is incomplete due to a section process inherent in the study design.

Survival methods have been developed for the analysis of Survival data. Some of these are Descriptive Statistics which include life tables, survival distribution and Kaplan --Meier times from a sample. Nonparametric test are available for comparing the survival experience between two or more groups. The most common and widely used of these test is the log-rank test Generalized Wilcox on test and Peto-prentice test. The Multivariate Method uses Cox-proportional hazards model.it is considered as the most interesting survival modeling in the interest of examining the relationship between survival and one or more predictors. Covariates may be categorical or continuous. In addition the model has the capability of including both time-dependent and time-independent variables.

3.6.1.1 Non Parametric Method

The simple and easy method which is free of assumptions is the nonparametric method.

Nonparametric methods are often very easy and simple to understand as compared to parametric methods. Furthermore, nonparametric analysis are more widely used in situations, where there is doubt about the exact from of distribution. In the nonparametric methods, the most popular and commonly used method is the Kaplan-Meier method (Kaplan El,Meire Pl,1985).

3.6.1.1.1 Kaplan-Meire method

It is also called one-sample nonparametric method.it is a very popular method.it is a very popular for its advantage such as it does not require any assumption for its hazard function of proportional hazard.it simply takes in to account with the empirical probability of surviving over certain time but it assumes the value of survival function between successive distinct observation is constant it is widely applied for studying survival function of population. And also it is very commonly used to compare the survival functions among groups. The method does not take into account of covariates so it is mainly descriptive. For better presentation the survival function usually shows as stepwise reduction plot.

The Kaplan-Meier estimator of the survivorship function (Kaplan and Meier, 1985), also called the product the product limit estimator is probably the most popular approach because it

- ✓ Is used by most software packages
- ✓ Incorporates information from all of the observation available both uncensored and censored
- ✓ It based on individual observations, so it is more precise than the life table estimator.

  The Kaplan-Meire estimator of the survivorship function (or survival probability) $s(t) = P(T \geq t)$ is defined as

  $\hat{S}_{(t)} = \prod_{t_{(i)} \leq t} (1 - \frac{d_i}{n_i})$……………………………………………………3.1

With the convention that

$$\hat{S}_{(t)} = 1\ if\ t \leq t(i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.2$$

Where $n_j = number\ at\ risk\ of\ recoviring$

$$d_{(j)} = observed\ number\ of\ deaths$$

$$t_{(i)} = rank - ordered\ survival\ times$$

Some properties of Kaplan estimator:

- The KM estimator is step function which only jumps at uncensored observations. Different jumps are random are dependent on censored observations.
- When the largest observation is censored the KM estimator does not converge to zero at infinity and is often taken as undefined
- The variance of the product-limit estimator is estimated by Green-wood's formula

3.6.1.1.2 Comparison of survival function

When comparing groups of subjects, it is always a good idea to be begun with a graphical display of the data in each group. The figure in general shows if the pattern of one survivorship function lying above another which means the group defined by the upper curve lived longer, or had a more favorable survival experience, that the group defined by the lower curve. Now the statistical question is whether the observed difference seen in the figure is significant. The general form of this test statistics is given by

$$Q= \frac{[\sum_{i=1}^{m} w_i(d_{1i}-\hat{e}_{1i}]^2}{\sum_{i=1}^{m} w_i{}^2\hat{v}_{1i}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.3$$

Where:

$e_{1i} = \frac{n_{1i} \times d_i}{n_i}$ is expected number of failures corresponding in group 1 at time t_i

$\hat{v}_{1i} = \frac{n_{1i} n_{2i} d_i (n_i - d_i)}{n_i{}^2 (n_i - 1)}$ is the variance of the number of failures in group 1 at time t_i

n1i is the number at risk at observed survival time t(i) in group 1

n2i is the number at risk at observed survival time t(i) in group 2

d1i is the number of observed recover in group 1 d2i   is the number of observed recover in group 2

ni   is the total number at risk

di   is the total number of births at time t(i)

The contribution to test statistic depends on which of the various tests is used. But each may be expressed in the form of a ratio of weighted sums over the observed survival times. Under the null hypothesis assuming that the two survivorship functions are the same, and that the censoring experience is independent of group, and that the total number of observed events and the sum of the expected number of events is large, Q follows a chi-square distribution with one degree of freedom. The above can be used to compare more than two groups (Hosmer & Lemeshow, 1999). The log rank test which is a special case of Q is used in this study.

3.6.1.1.3 Log Rank Test

Log-Rank method is one of the non-parametrical statistical tests that have been proposed to answer whether there are differences of survival time between groups (Hosmer 2008). It is constructed by computing the observed and expected number of event of interest in one of the groups at each observed time to event of interest. The calculation of each test is based on a contingency table of groups by status at each observed survival time. These methods are appropriate for survival data that deal with a single (first) time to event but not for recurrent events of both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. In addition, Log-Rank method is one of the non-parametric statistical tests that have been proposed to answer whether there are differences in survival time between groups (Hosmer 2008). It is constructed by computing the observed and expected number of event of interest in one of the groups at each observed time to event of interest. The calculation of each test is based on a contingency table of groups by status at each observed survival time. These methods are appropriate for survival data that deal with a single (first) time to event but not for recurrent events.

The log rank test, developed by Mantel and Haenszel, is a non-parametric test for comparing two or more independent survival curves. Since it is a non-parametric test, no assumption about the distributional form of the data is required. This test is however most powerful when used for no overlapping survival curves. The log rank test is based on weight equal to one, i.e *wi* =1. Its statistic becomes:

$$Q_{LR} = \frac{[\sum_{i=1}^{m}(d_{1i}-\hat{e}_{1i}]^2}{\sum_{i=1}^{m}\hat{v}_{1i}} \quad \text{……}3.4$$

3.6.1.1.4 The Generalized Wilcoxon Test

Gehan (1965) and Breslow (1974) generalized the Wilcoxon rank sum test to allow for censored data. This test uses weights equal to the number of subjects at risk at each survival time, i.e

$w_i = n_i$, and is called Wilcoxon or generalized Wilcoxon test in most software packages. Thus the Wilcoxon test can be defined as:

$$Q_{LR} = \frac{[\sum_{i=1}^{m} n_i(d_{1i}-\hat{e}_{1i})]^2}{\sum_{i=1}^{m} n_i{}^2\hat{v}_{1i}} \quad \text{……}3.5$$

We can think of this as testing whether the two survival curves are identical or not (Survival Analysis). Moreover,

H_O: *no statistical signficant different* H_1: *there is statistical signficant different*

The Survival curves will be estimated by the Kaplan-Meier process for each group and then we use the log rank test to compare the group's Survivals.

3.6.1.2. Semi-parametric methods

3.6.1.2.1 Cox regression Model

One of the most popular types of regression models used in survival analysis is the cox proportional hazard model (cox, 1972).

3.6.1.2.1.1. Cox's Proportional hazard Model

Cox's semi-parametric model is widely used in survival analysis to model the effect of covariates on hazard rates. Coxi's proportional hazards model assumes a parametric from for th effect of the covariates, but it allows an unspecified from for the underling survivor function. The partial likelihood of the model also allows time-dependent covariates. A covariate is time dependent if its value for model for any individual can change over time. The validity of the proportional hazards model can be testing for interaction between time-dependent covariates and the response time.

The cox Proportional Hazard Model is a multiple regression method used to evaluate the effect of multiple covariates on the survival.(Cox,D.R. and Oakes,D,1984).Proposed a semi-parametric model for the hazard function that allows the addition of covariates, while keeping the baseline hazards unspecified and can take only positive values. With this parameterization the cox hazard function is:

$$\lambda(t, x, \beta) = \lambda_0(t)e^{\beta' x} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.6$$

Where $\lambda_0(t)$ is the baseline hazard function that characterizes how the hazard function change as a function of survival time, $\lambda(t, x, \beta)$ represent the hazard function at time t with covariates x=$\left(\left(x_{1,}x_2 \ldots\ldots\ldots . x_p\right)', \beta = (\beta_1, \beta_2, \ldots\ldots\ldots \beta_p\right)$ is a column vectors of regression parameters

$e^{\beta' x}$ and characterizes how the hazard function changes as a function of subject covariates.

T is the failure time. The survival time of each member of the sample is assumed to follow own hazard function.in such a case the above model can equivalently be written as:

$$\lambda_i(t, x_i, \beta) = \lambda_o(t)\exp\left(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}\right) i = 1,2,,,,,,,,,,,,, n \ldots\ldots\ldots\ldots .3.7$$

Where n is total number of TB patients included in the study,$X = (x_1, x_2, \ldots\ldots . x_p)'$ is a column vector of measured covariates for the i^{th} individual (patient) which are expected to affect the survival probability. The proportional hazards estimation method computes a coefficient for each predictor variable that indicates the direction and degree of flexing that the predictor has on survival. The proportional hazard is the most popular regression method for analysis of censored survival data. The popularity is because

- It allows flexible choice of covariate (we can accommodate time varying, time independent, continuous and discrete covariates).
- It is fairly easy to fit. Standard software packages are programmed to handle proportionality hazards model such as R, SPSS, SAS, STATA and S-PLUS etc.
- Dose not make any assumption about the underlying survival distribution (does not require the knowledge of the shape of the survival distribution)
- Dose not require estimation of the baseline hazard rate,$\lambda_o(t)$, to estimate the regression parameters. The cox proportional hazard model most is formulated as the hazard function which measures the risk to death or failure at time t

3.6.1.2.1.2. Proportional Hazards Assumption

- The baseline hazard $\lambda_o(t)$ depends on t, but not on covariate$(x_1, \ldots\ldots, x_p)$.
- The hazard ratio,i.e, $\exp(\beta' x)$, depends on the covariates $X = (x_1, x_2, \ldots.x_3).\ but\ not\ on\ time\ t$
- The covariate x_i,do not depends on time t:

Assumption (2) is what led us to call this a proportional hazards model. To express this mathematically. Consider two distinct values of the covariate X say x_1 and x_2:

Then the hazard ratio becomes:

$$\frac{\lambda(t,x_1,\beta)}{\lambda(t,x_2,\beta)} = \frac{\lambda_0(t)e^{\beta;x_1}}{\lambda_0(t)e^{\beta' x_2}} = \frac{e^{\beta' x_1}}{e^{\beta' x_2}} = e^{\beta'(x_1 - x_2)}$$

In independent of time t .This shows that the ratio of the hazard functions for two individuals with different covariate dose not vary with time.

3.6.1.2.2 Parametric Estimation

In cox proportional hazards model we can estimate the vector of parameters β with out having any assumption about the hazard $\lambda_o(t)$.As a consequence this model is more flexible and estimate of the parameter can be obtained easily. Conspired n independent individuals the data that we need for the Cox proportional hazard model is represented by $(t_i, \delta_i, x_i), i = 1,2, \ldots n$

Where t_i =the survival time for the i^{th} individual δ_i=an indicator of censoring for the i^{th} individual given by 0 for censored and 1 for event/death x_i =a vector of covariates for individual i $(x_{i1}, x_{i2} \ldots \ldots. x_{ip})$

The full likelihood for right censored data can be constructed as

$$L(\beta) = \prod_{i=1}^{n} \lambda((t_i, x_i, \beta)^{\delta_i} S(t_i, x_i, \beta) \ldots \ldots \ldots \ldots \ldots 3.9$$

Where $\lambda(t_i, x_i, \beta) = \lambda_o(t_1)e^{\beta' x_i}$ is the hazard function for individual i:

$S(t_i, x_i, \beta)$=$(s_o\ (t_i)\exp(\beta' x)$ is the survival function for the individual i:

It follow that

$$L(\beta) = \prod_{i=1}^{n} (\lambda_o(t) \exp(\beta'^{x_i})^{\delta_i} s_o(t_i) \exp(\beta'^{x_i})) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..3.10$$

The full maximum likelihood of β can be obtained by differentiating the right hand side of equation(3.10) with respect to the component of β and the base line hazard $\lambda_o(t)$.

This implies that unless we explicitly specify the base line hazard,$\lambda_o(t)$.we can not obtain the maximum likelihood estimators for the full likelihood. To avoid the specification of the base line hazard (cox,D.R. 1975) proposed a partial likelihood approach that treats the baseline hazard as a

Nuisance parameter and removed it from the estimating equation:

3.6.1.2.2.1 Partial likelihood

Instead of construction a full likelihood we consider the probability that an individual experiences an event time t_i given that an event occurred at the time.

Let R_i denote the set of individual at risk time to t_i,no ties.The probability that individual i with covariates x_i is the one who experience the event at time t is equal to the p(individual i has experiences an event at time $t(i)$ one event at time $t(i)$

$$\frac{\lambda(t,x_i)}{\sum_{j\epsilon R_{t(i)}} \lambda(t,x_j)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots3.11$$

And under the proportional hazards assumption on using (3.11),the ratio

$$\frac{\lambda_o(t)\exp(\beta' x_i)}{\sum_{j\epsilon R_{t(i)}} \lambda_o(t)\exp(\beta' x_j)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..3.12$$

Shows the contribution to the partial likelihood at each death time $t_{(i)}$ by the individuals with covariate x_i in the risk set $R_{t(i)}$.where$R_{t(i)}$ is the over all subjects in the risk set at time $t_{(i)}$ By eliminating the base line hazards function in the numerator equation (3.12) becomes

$\frac{\exp(\beta' x_i)}{\sum_{j\epsilon R_{t(i)}} \exp(\beta' x_j)}$ Thus the partial likelihood is the product over all failure time $t_{(i)}$ for i=1,2,…,m of the conditional probability(3.13) to give the partial likelihood

$$L_p(\beta) = \prod_{i=1}^{m} \frac{\exp(\beta' x_i)}{\sum_{j\epsilon R_{t(i)}} \exp(\beta' x_j)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..3.14$$

The product is over the m distinct ordered survival time and x_i denotes the value covariate for the subject with ordered survival time $t_{(i)}$. The log partial likelihood function is

$L_p(\beta) = \sum_{i=1}^{m}[\beta'^{x_j} - \ln(\sum_{j\epsilon R_{t(i)}} \exp(\beta'^{x_j}))]$……………………………………………3.15

We obtain the maximum partial likelihood by differentiating the right hand side of equation(3.15) with respect to the component of β,setting the derivative equal to zero and solving t=for the unknown parameters.

The partial likelihood derivate above is valid when there are no ties in the data set. But in most real situation tied survival times are more likely to occur.in addition to the possibility of more than one censored observations at a time .To handle this real-world fact, partial likelihood algorithms have been adopted to handle ties.

There are three approaches in common to estimate regression parameter when there are ties

The most popular and easy approach is Breslow's approximation.

3.6.1.2.3 The Breslow Approximation

This approximation is proposed by Breslow and peto by modifying the partial likelihood takes the following form:

$L_B(\beta) = \prod_{i=1}^{m} \frac{\exp(\beta' S_i}{\sum_{j\epsilon R_{t(i)}} \exp(\beta' x_j)^{dj}}$…………………………….3.16

Where S_i the sum of covariates over d_i subject at time $t(i)$ $d(i)$ the number of death occurred at time t(i) Now the partial log likelihood of (3.16) is given as

$$l_B(\beta) = \sum_{i=1}^{m} \beta'^{s_i} - d_i ln \sum_{j\epsilon R_{t(i)}} \exp(\beta'^{x_i}) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots .3.17$$

We obtain the Breslow maximum partial likelihood estimator adjusted for observation by differentiating equation (3.17) with respect to the component of $\beta\ and$ setting the derove equal to zero and solving for the unknown parameters.

3.6.1.2.5 Model Development

3.6.1.2.5.1 Selection of Covariates

There are three kind of method selection of beast subset of the covariates i.e. purposeful selection stepwise selection best subset selection .But for this particular study purposeful selection is used:

Purposeful Selection of Covariates

Step1 it begins by including covariates that are statistical significant (p-value<20-50% modest level of significant.

Step 2 Include covariates that are considered more important.

Step 3 use these covariates (those selected out step1 and 2)

Step 4 Select covariates that are statistical significant (p-value<0.05) using Wald-test

Step 5 Retain some covariates that are "important "even though they are not significant and checking also for potential confounders.

Step 6 In addition to these for deleted (remove) covariate check whether their contribution is statistical significant or not using particular likelihood ratio test.

This selection processes provide preliminary subsets of covariates. These are considered as main effects.

3.6.1.2.5.2 Assessment of Model Adequacy

The use of diagnosis procedure for model checking diagnosis is an essential part of the modeling process. There are different commonly used model diagnoses to evaluate whether the appropriate functional form for a covariate is used in the model to assess the fitted model.

The methods for assessment of a fitted proportional hazards model are essential the same as for other regression .in general requirements for model assessment are:

- Methods for tasting the assumption of proportional hazards
- Subject-specific diagnostic statistics that extended the notation of leverage and influence to the proportional hazards model
- Overall Summary measures of goodness of fit

3.6.1.2.5.3 Checking for Proportionality Assumption

In order to use the Cox model we must check the assumption of whether the effects of covariates on hazard ratio remains constant over time .This is a critical assumption of proportional hazards model and must be checking for each covariate. Different studies suggest that several test and graphical techniques can be used to assess proportionality assumptions in fitting the Cox Model. The Grambsch-themeau test of non-proportionality uses partial residuals for the test of proportional hazards assumption. In order to use this test for the i^{th} covariate (Grambsch and Themeau, 1994) propose a time-varying coefficient as:

$$\beta_j(t) = \beta_j + \gamma_i g_i(t) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 3.18$$

Where $\beta_j(t)\ is$ time varying coefficient, β_j is constant $g_i(t)$ is some specified function of time usually $g_i(t) = \ln(t)$ becomes:

$$\lambda\left(t, x_i, \beta_j(t)\right) - \lambda_0(t)\exp(\beta_j(t)x)$$

Substitute$\beta_j(t) = \beta_j + \gamma_i g_i(t)$ gives:

$$=\lambda_0(t)\exp\left(\beta_j\right) + \gamma_i lnt)x = \lambda_0(t)\exp\left(\beta_j\right)x + \gamma_i(lnt)x \dots\dots\dots\dots 3.19$$

This looks like the proportional hazards model where the interaction term,xln(t) is include in the model in addition to the main effect x_i to test the significant of the interaction terms $x_i\ln(t)$ that is $H_o: \gamma = 0$ against$H_1: \gamma \neq 0$ we can use likelihood-based test like Wald test

If $\gamma = 0$ is reject $\beta_j s$ are not time varying coefficients and hence the proportional hazards assumption is satisfied. If $\gamma = 0$ is rejected then the proportional hazards assumption is not satisfied and we have to look for another model.

The schoenfeld graphical technique can be used to assess cox model assumptions the technique is based on individual contribution to the log-partial likelihood and measures the difference between the covariate for the i^{th} individual and a weighted average of the covariate over the risk set at the time the i^{th} individual event(Schoenfeld,1982).For greater diagnostic power the selected schoenflied residuals are the scaling can be done on the variance of the i^{th} subject Schoenfeld residuals.

To check the proportional assumption for each covariate we plot the schoenfeld residual on the Y-axis against log of survival time on the X-axis .if the proportional hazards assumption is satisfied the distribution of residual over time is random that is does show a particular trend and the smoothed plot called Locally Weighted polynomial regression (Lowes) line Summarizing the residuals should be a straight line and close to the horizontal reference line. Otherwise a plot of scaled residuals for a given covariate may reveal a violation of the proportional hazard assumption (Arjas, E.1988).

3.6.1.2.5.4 Identification of Influential and Poorly Fit Subject

Another important aspect of model evaluation is a thorough examination of regression diagnostic statistics to identify which if any observations:

- Have an usual configuration of covariates
- Exert an undue influence on the estimate of the parameters, and
- Have an undue influence on the fit of the model

Statistics similar to those used in linear and logistic regression are available to perform these tasks with a fitted proportional hazards model. There are some different in types of statistics used in linear and logistic regression and proportional hazards regression but the essential ideas are the same in all the three setting.

Leverage is a diagnostic statistics that measure how 'unusual' the value of the covariate of the coverage are for an individual .in linear and logistic regression leverage is the distance of the value of the covariate for a subject to the covariates. Leverage is not easily defined nor does it have the same nice properties in multiple risk set and thus may be present in multiple terms in the partial likelihood.

The score process residual for the i^{th} subject on the k^{th} covariate may be expressed as:

$$L_{ik} = \sum_{i=1}^{n} x_{ik} - \bar{x}_{wjk})dM_i(t_i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.20$$

Is a weighted average of the distance of the value,x_{ik},to the risk set means $\bar{x}_{wjk}$,where the weights are the change in martingale residuals $dM_i(t_i)$ defines as:

$$dM_i(t_i) = dN_I(t_i) - Y_i(t_i)\exp(x_i'\beta)\,\lambda_o\big(t_j\big) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.21$$

Where $dN_i(t_i)$ is the change in the count function for the i^{th} subject at time t_j always equal to zero for censored subjects and one for non-censored subjects at actual observed survival time. The function $Y_I(t_i)$ is called the risk process and defined as zero if $t_i \leq t_j$ and one if

$t_i \geq t_{j,} \lambda_0(t_{j)}$ is the value of $\frac{\delta_i}{\sum_{j \epsilon R(t)} \exp(x_j' \beta)}$ evaluated at $t_{j,}$ .The net effect is that for continuous covariates the score residuals have the linear regression leverage property that the further the value is from the mean the larger the scores residual is but 'large' may be either positive or negative. Thus the score residual are sometimes referred to as the leverage or partial leverage residuals. We plot score residual against each continuous covariates to observe if there is individuals far away from the mean.

3.6.1.2.5.5 Overall Goodness of Fit

As in regression analysis some measure analogues to R^2 may be of interest as a measure of model performance. There is not a single simple easy to calculate useful easy to interpret measure for a proportional hazards regression model. In particular al measures depend on the proportion value that are censored. A perfectly adequate model may have what at face value seems like a terribly low R^2 due to a high percent of censored data (Hosmer and Lemeshow, 1998).We use R^2 as it is the easiest and beast one to use and it is defined as:

$$R_p{}^2 = 1 - \exp(\frac{2}{N}((LL_o - LL_\beta)\text{..............................}3.22$$

Where N is the total number of observation in the model LL_o is the Log partial likelihood for model zero. LL_β is the log partial likelihood for the fitted model with p covariates.

3.6.1.3 Parametric Model

None parametric Survival model makes no assumption about the shape of the survival curve. Does no control for covariate and it requires categorical predictors. Ox PH model allows to estimates and make inference about the parameters without assuming any distribution for the survival time. However when the PH assumption is not tenable this model will not be suitable.in this section we will introduce parameter model in which specific probability distribution is assumed for the survival time. The basis of this method was to avoid having to specify the hazard function completely .However there may be setting in which the distribution of the survival time is specific parametric distribution that justifies the use of a fully parametric model to better address the goal of the analysis(Lee,E.and Wang,J. 2003).Many models using different distributions have been developed. Some of most common Survival models are:

3.6.1.3.1 The Exponential Regression Model

For the time to event data and skewed to the right with distribution of the time is exponatial the time of survival for a single covariate which is called accelerated failure time expressed as

$$T = \exp(\beta_0 + \beta_1 x + \in) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..3.23$$

Where $\{t \sim Exp(\alpha)\}$ and ∈ be error component

The exponential model is the simplest parametric model and assumes a constant risk or hazard over time. The survival function may be obtained by expressing in terms of time as

$$S(t, x, \beta) = \exp\left(\frac{-t}{e^{\beta+\beta_1 x}}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 3.24$$

And the hazard function of the exponential regression model is

$$h\big(t, x, \beta) = (\lambda_o(t) e^{-(\beta_0+\beta_1 x)}\big) \dots\dots\dots\dots\dots\dots\dots\dots 3.25$$

The exponential regression model for the k covariates and individual Covid-19 patients is expressed as:

$$h_i(t, x_i, \beta) = \lambda_o(t) \exp(\beta_0 + \beta_1 x_{1i} + \cdots .. \beta_k x_{ik}) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..3.26$$

3.6.1.3.2 The Weibull Regression Model

Survival time t is a positive random variable with Weibull density function can be expressed as:

$$f(t; \mu, \alpha) = \frac{\alpha}{\mu}(\frac{t}{\mu})^{\alpha-1} \exp(\frac{-t}{\mu})^{\mu} \dots\dots\dots\dots\dots\dots\dots\dots\dots..3.27$$

Where $\mu > 0\ and, \alpha > 0$

The baseline hazard function of the distribution becomes

$$h(t; \mu, \alpha) = \frac{\alpha}{\mu}(\frac{t}{\mu})^{\alpha-1} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.3.28$$

Parametrization the Weibull distribution using $\lambda = \mu^{-\alpha}$ then $h_o(t) = \lambda\alpha t^{\alpha-1}$ will be the baseline hazard function. Now incorporate covariates x in the hazard function the Weibull Regression model become

$$h(t; x, \beta) = \lambda\alpha t^{\alpha-1} \exp(x\beta) \dots \dots \dots \dots \dots \dots \dots \dots \dots .3.29$$

This yields the following survivor function $S(t) = \exp[-(\frac{t}{\mu})^{\alpha}]$ and the cumulative hazard function becomes $H(t) = (\frac{t}{\mu})^{\alpha}$

3.6.1.3.3 The Log-Logistic Regression Model

Single covariate log-logistic accelerated failure time may be expressed as

$$\ln(t) = \beta_0 + \beta_1 x + \sigma\varepsilon \dots \dots \dots \dots \dots \dots \dots \dots \dots 3.30$$

Where σ is the scale parameter and ε is the residual (unexplained) variation in the transformed survival times (David-Collett,D,1994).

The survivorship function for the model is $S(t, x, \beta, \sigma) = [1 + \exp(z)]^{-1} \dots \dots \dots \dots \dots \dots \dots 3.31$

Where z is the standard log-time outcome variable that is $z = \frac{y-\beta_0-\beta_1 x}{\sigma}$ and $y = \ln(t)$

The odds of a survival time evaluated at x=0 and x=1 is

$$OR(x = 0, x = 1) = \frac{\exp\{-(\frac{y-\beta_0-\beta_1 x1}{\sigma}}{\exp\{-(\frac{y-\beta_0-\beta_1 x0}{\sigma})} = \exp(\frac{\beta_1}{\sigma})$$

This is independent of time

3.6.1.3.4 The Log-Normal Regression Model

The log-normal model may takes censored time dependent variable that allows the hazard rate to increase and decrease. The log-normal model assume that $\varepsilon \sim N(0,1). Let\ h(t)$ be the hazard function of T for(3.23) when$\beta = 0$ i.e $\beta_0 = \beta_1 = \beta_p = 0$

Then it can be shown that $h(t)$ has the following functional form

$$h(t) = \frac{\phi(\frac{\log(t)}{\sigma})}{[1 - \phi\left(\frac{\log(t)}{\sigma}\right)]} \dots\dots\dots\dots\dots\dots\dots .3.33$$

Where $\phi(t) = \frac{1}{\sqrt{2\pi}}\exp(\frac{-t^2}{2})$ be the probability density function and $\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}}\exp(\frac{-u^2}{2})du$ is the cumulative density function. Obviously we no longer have a proportional hazard model. If the baseline hazard function is desired it can be obtained from equation (3.33) by setting x=0 .The survival function $s\left(\frac{t}{x}\right)$ at any covariate x can be expressed as

$$s\left({}^{l}/_{x}\right) = \emptyset\left[\beta_0{}^* + \beta_1{}^* x_1 + \cdots + \beta_p{}^* x_p - \alpha \log(t)\right] \dots\dots\dots\dots\dots\dots .3.34$$

Where $\alpha = \frac{1}{\sigma^*}, \beta^*{}_j = \frac{\beta_j}{\sigma}$ for j=0,1,….,p.this is the survival model with intercept depending with t.

3.6.1.4 Model Selection

To select the model that can predict the survival time of Covid-19 patients we have two methods. The first is graphical approach (shoenfld, D, 1984). For this method the Cox-snell residual plot is the common one. It is used to determine how well a specified distribution fits to the observed data. This plot will be approximately linear if the specific theoretical distribution is the correct model. Easy fit displays the reference diagonal line along which the graph point should fall along with the goodness of fit tests; the distribution plot can be helpful to determine the beast fitted model the fundamental difference of the approaches is that it is quite subjective to came to conclusion while the goodness of fit test are 'exact' in the sense that the result do not depend on the researcher(provided) that the tests are preformed correctly).using the plot method is consider to be more empirical approach in model selection. (Akaikie, 1974) proposed an information criteria (AIC) statistic to compare different models and/or models with different of parameters. For model the value is computed as

$$AIC = -2\log(likelihood) + 2(\rho + 1 + s) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 3.35$$

Where ρ denotes the number of covariates in the model without including the constant terms and s is the number of parameters minus one i.e s=0 for the exponential regression model and s=1 for Weibull, Log logistic and Log normal regression models. According to the criterion a model with small AIC value will be consider a better fit.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Descriptive Statistics

The descriptive statistics was presented in the Table 4.1 below. In this study, a sample of 826 Covid-19 patients was considered. The medical cards of those patients were reviewed, out of which 36.2% were female and 63.8% were male. Among those patients 77.6% and 22.4% were recovered and censored respectively. A Recover proportion of females which is 29.9% seems lower than males 47.7%. There are 43.2% and 56.8% patients were No Sympotomic and Sympotomic, the Recovery proportion of Non Sympotomic Patients were 33.4% lower than Sympotomic patients 44.2%, respectively.

Regarding to Age, which showed that 9.9% were under 20 years, 12.1% were [20-30 years old] 18.6% were[30-40 years old],13.7% were[40-50 year old],14.5%were[50-60 year olds],17.7% were [60-70 year],8.4%were [70-80] and the rest 8.7%were above 80 years old. The Recovered proportion of under20 year,[20-30],[30-40],[40-50],[50-60],[60-70] and greater than 70 year of patients were 74 (8.96%), 90 (10.9%), 132 (15.98%), 90 (10.9%), 92 (11.14%), 100 (12.11%), and 63 (7.63%), respectively.

There are 31.84%, 0.12%, 11.74%, and 56.3% patients from Oromia, Amhara, SNNPR, and Sidama regions, respectively. Out of the entire subjects integrated in this study, 58.7% of the patients (had) Fever whereas 41.3% had no fever. The Recover proportion was higher for those fever 44.92%, while lower for those non fever patients 32.69%. The sample data also revealed that 65.6% patients were non other pain without covid-19 and 34.4% were other pains without covid-19. The Recover proportion of other pain without covid-19 215(26.03%) was lower than non-other pain without covid-19 424(51.57%).

The outcome in Table 4.1 also revealed that, 38.7% patients were diagnosed with Comorbidity and 61.3% patients were diagnosed with non-comorbidity were more likely to Recovered (50.37%) compared to with comorbidity(26.88%).

There are 83.01% patients who had no travel history for last 14 days, 16.95% had travel history last 14 days.80.02% of the patients had no visit health facility before on set , 19.98% had visit the health facility before on set. Among those the Recover proportion was higher for non-visitor 60.17% and lower for health facility before on set (17.47%).

From all patients included in this study, the highest proportion 69.7% of the patients had Caught, followed by those who had no caught 30.3%, while the Recovery proportion of patients who had caught were 19.03%, followed by those who had no caught which was 26.51%, while the higher proportion of recovery 50.64% were accounted for patients who had no caught problem.

There are 61.9% patients who had Shortens of breath problem and 38.1% of the patients had no shortens of breath. The recovery proportion of patents that have adverse shortens of breath was 43.34%, while lower for no shortness of breathing problem 34.26%. Out of the total sample, 31.4% patients had sore throat problem, 68.6% had no sore throat problem. Regarding sore throat condition, the recover proportion was higher which 54.48% is for those patients who had no sore throat followed by those who had sore throat problem 23.12%.

The proportions of patients who had General body weakness were 57.6%, whereas 42.4% had no general body weakness .The death proportion were higher for those patients who had general body weakness 40.56%, while smaller no general body weakness 37.05%.

There are 76.3% patients who had no clos contact history with the confirmed case, 32.7% had close contact with the confirmed case Recovery proportion for no close contact with confirmed case was higher than 51.7% that of the patients had close contact with confirmed case 26.03%.

The severity status were 21.4%,25.5%,23.8%,11.41%,14.41%,3.6% no severity, mild,moderate,sevier,critical,Asymptotic and the proportionality of recovery follow as 15.5%,22.64%,21.19%,9.99%,7.47%,2.78% no severity, mild, moderate, sevier, critical and Asymptotic.

Table 4.1 Summary Result of Covid-19 Recover event vs Risk Factor

No.	Factors	Strata	Value	Total	Percentage	Improved (Events)	Death (Censors)
1	sex	1	Female(0)	299	36.20	244	55
		2	Male(1)	527	63.80	393	134
2	Age	1	<20(0)	82	9.9	74	8
		2	20-30(1)	100	12.1	89	11
		3	30-40(2)	154	18.6	130	24
		4	40-50(3)	113	13.7	90	23
		5	50-60(4)	120	14.5	92	28
		6	60-70(5)	146	17.7	100	46
		7	70-80(6)	111	13.4	62	49
3	Sympotomic	1	No(0)	357	43.22	273	84
		2	Yes(1)	469	56.78	364	105
4	Fever	1	No(0)	341	41.28	267	74
		2	Yes(1)	485	58.72	370	115
5	Head Ache	1	No(0)	470	56.90	371	99
		2	Yes(1)	356	43.10	266	90
6	other pain out of covid-19 pain	1	No(0)	542	65.62	424	118
		2	Yes(1)	284	34.38	213	71
7	Comorbidity	1	No(0)	506	61.26	415	91
		2	Yes(1)	320	38.74	222	98
8	number of comorbidity	1	No(0)	438	53.03	367	71
		2	One(1)	278	33.66	204	74
		3	Two(2)	92	11.14	55	37
		4	Greater than 2	18	2.18	11	7
9	Travel History last 14 days	1	No(0)	686	83.05	526	160
		2	Yes(1)	140	16.95	111	29
10	Health facility visit before on set	1	No(0)	661	80.02	511	150
		2	Yes(1)	165	19.98	126	39
11	Nationality	1	Ethiopia(0)	824	99.76	635	189
		2	Other Country(1)	2	0.24	2	0
12	Region	1	Sidama(0)	465	56.30	348	117
		2	SNNPR(1)	97	11.74	78	19
		3	Oromia(2)	263	31.84	210	53

	Factors	Strata	Value	Total	Percentage	Improved (Events)	Death (Censors)
13	Generalize Body Weakness	1	No(0)	350	42.37	303	47
		2	Yes(1)	476	57.63	334	142
14	pain	1	No(0)	248	30.02	209	39
		2	Yes(1)	578	69.98	428	150
15	Close Contact With Confirmed	1	No(0)	556	76.08	423	133
		2	Yes(1)	270	79.26	214	56
16	Severity	1	No Severity(0)	173	20.94	123	50
		2	Mild(1)	211	25.54	184	27
		3	Moderate(2)	197	23.85	174	23
		4	Sever(3)	96	11.62	64	32
		5	Critical(4)	119	14.41	67	52
		6	Asymptotic(5)	30	3.63	25	5
17	Occupation	0	unemployed	425	51.45	324	101
		1	Student	78	9.44	62	16
		2	merchant	30	3.63	20	10
		3	civil servant	77	9.32	62	15
		4	Others	216	26.15	169	47
18	Caught	1	No(0)	250	30.30	220	30
		2	Yes(1)	575	69.7	416	159
19	shortens of Breath	1	No(0)	315	38.14	281	34
		2	Yes(1)	511	61.86	356	155
20	Sore throat	1	No(0)	567	68.64	446	121
		2	Yes(1)	259	31.36	191	68

4.1.1 Comparison of Survival Experience of COVID-19 Patients

In order to investigate if there is significant difference between the survivals of a patient by gender, Kaplan-Meier survivor estimates for the two gender groups are plotted in Figure P1.1 in appendix 1. This Figure shows that males had slightly higher survival until the 20 to 60 Days compared with females whereas; females survive better before 20 days five. But the difference in survival were not supported by Statistical tests, since log-rank (mantelcox) test and breslow (generalized wilcoxon) test in Table 4.2 shows that there is insignificant difference between male and female with respect to survival time. Comparing the survivor functions between different Severity status of among covid-19 patients, Kaplan Meier survivor estimates for the six groups are plotted in Figure A1.1 in appendix 1. This Figure shows that patients with critical severity status had slightly higher survival compared sever, no Sympotomic, modernet, mild and asymptotic patients until 30 days than sever had slightly high survival probability camper to other severity status after 30 days . Statistical test is made by using log-rank (mantel-cox) test and breslow (generalized wilcoxon) test this shows that there is significant difference between covid-19 patients those severity status was Critical ,sever, no Sympotomic, modernet, mild and asymptotic with respect to survival time. Among different Age groups, patients had the age category below 40 year higher survival time and it is also statistically significant ($p<.000$). These variables log-rank test for survival difference were all highly significance. The information presented above is summarized in Table 4.2 and Figure A1.1 in Appendix 1 below. Among the close contact with confirmed case category patients who had non-contact history with the confirmed case was higher survival time particularly until 30 days followed by close contact with confirmed case.

But Statistical test using log-rank test shows that there is no significant difference among the close contact with confirmed case and non-contact with confirmed case groups. Kaplan-Meier survivor estimate for Comorbidity of covid-19 patient the Survival probability of length of stay time in hospital was high for covid-19 patient with comorbidity problem and the survival probability of staying in hospital was low for without comorbidity. The statistical test or log-rank test revels that there are significant different for the length of staying time in hospital for covid-19 patient with comorbidity case and non-comorbidity problems. Similarly, Kaplan-Meier survivor estimates for other pain with out-covid-19 the survival probability plotted in Figure A1.1 in Appendix 1 of staying time in hospital were long for covid-19 patients with other pains and the log rank test indicated that there is significant different between this two category. The information presented above is summarized in Table 4.2 and Figure A1.1. and in Appendix 1 below.Similar analysis is performed to investigate difference in the survival time among the patients with respect to shortness of breath and headache problems. Results from the Kaplan-Meier curves in Figure A1.1 in Appendix 1 shows that the curves indicating that survival time were different for these groups. Log-rank test also show that there is statistically significance on the difference in the survival of a patient.

Table 4.2: Comparison of survival experience on Covid-19 patients using demographic, health and risk behavior variables (at Hawass Referral Hospital, during September 2013- January 2014)

Variables	Observed Events	Total	Percent of events	Df	Log rank test		Berswel(Wlcoxon)Test	
					x^2	p-value	x^2	p-value
Age <20	74	82	9.9	6	41.90	0.0000	44.33	0.0000
20-30	89	100	12.1					
30-40	130	154	18.6					
40-50	90	113	13.7					
50-60	92	120	14.5					
60-70	100	146	17.7					
70 and above	62	111	13.4					
Severity no symptom	123	173	20.94	5	30.91	0.0000	27.77	0.0000
mild	184	211	25.54					
moderate	174	197	23.85					
sever	64	96	11.62					
critical	67	119	14.41					
Asymptotic	25	30	3.63					
Contact. H No	423	556	76.08	1	0.30	0.5817	0.00	0.9885
Yes	214	270	79.26					
Health. vs no	511	661	80.02	1	1.56	0.2116	1.08	0.2978
yes	126	165	19.98					
Traveling Hx no	526	686	83.05	1	1.59	0.2075	5.10	0.0239
yes	111	140	16.95					
Comorbidity no	415	506	61.26	1	12.57	0.0004	14.10	0.0002

yes	222	320	38.74					
No of Comorbidity				3	7.37	0.0611	4.91	0.1785
No comorbidity	367	438	53.03					
one comorbidity	204	278	33.66					
two co- morbidity	55	92	11.14					
more than two	11	18	2.18					
Pain no	209	248	30.02	1	0.61	0.4335	0.08	0.7801
yes	428	578	69.98					
Other Pain no	424	542	65.62	1	8.48	0.0036	4.34	0.0371
yes	213	284	34.38					
Body weaken no	303	350	42.37	1	303	334	1.27	0.2604
yes	334	476	57.63					

Variables	Observed events	Total	Percent of events	Df	Log rank test		Berswel(Wlcoxon)Test	
					x^2	p-value	x^2	p-value
Headache no	371	470	56.90	1	5.05	0.0246	3.83	0.0503
yes	266	356	43.10					
Sore throat no	446	567	68.64	1	0.56	0.4534	0.50	0.4810
yes	191	259	31.36					
S Breath no	281	315	38.14	1	5.69	0.0171		0.0947
yes	356	511	61.86				2.79	
Cough no	220	250	30.30		1.96	0.1620	1.48	0.2231
yes	416	575	69.70					
Fever no	267	341	41.28	1	3.68	0.0550	4.55	0.0329
yes	370	485	58.72					
Symptomatic no	273	357	43.22	1	0.01	0.9434	0.48	0.4877
yes	364	469	56.78					
Region Amahara	1	1	0.12	3	0.45	0.9294	0.58	0.9014
Oromia	210	263	31.84					
SNNPR	78	97	11.74					
Sidama	348	465	56.30					
Nationalty Ethiopia	635	824	99.76	1	0.49	0.4831	1.37	0.2411
other country	2	2	0.24					
Sex female	244	299	36.20	1	0.16	0.6926	0.03	0.8548
male	393	527	63.80					
Occupation Unemploy	324	425	51.45	4	7.62	0.1066	6.45	0.1679

Student	62	78	9.44					
Merchant	20	30	3.63					
Civil Serv.	62	77	9.32					
	169	216	26.15					

4.2 Cox proportional Hazards Regression Model

4.2.1 Single Covariate Analysis

Single covariate Cox proportional hazards model analysis is an appropriate procedure that is used to screen out potentially important variables before directly included in the multivariate model. The relationship between each covariates and survival time of covid-19 patients are presented in Table 4.3. As can be seen from this Table, survival of the patients is significantly related with severity, Comorbidity ,number of co-morbidity, other pain without covid-19,Head ache Shortness of breathing , age,. But the covariate like sex, Occupations, Region, Health Facility Visit before onset, Close contact with confirmed case, Travel history last 14 days, Pains, General Body weakens, Caught and Sore throat are not statistically significant at 5% significant level.

Table 4.3: Uni-variate analysis of Cox proportional hazards on the time to event of covid-19 patients (at Hawassa university referral Hospital, during 2013-2014)

Variables	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]
severity	.9322217	.026595	-2.46	0.014	.881527 .9858319
Contact	1.044743	.0880696	0.52	0.604	.8856353 1.232435
Hvisit	1.124576	.1120882	1.18	0.239	.9250137 1.367191
Traveling Hx	.8827047	.0928785	-1.19	0.236	.7182104 1.084874
Comorbidity	.757429	.0631937	-3.33	0.001	.6431682 .8919886
Co-morbidity	.8902167	.0499655	-2.07	0.038	.7974805 .9937368
pain	.9395309	.079454	-0.74	0.461	.7960253 1.108907
Other Pain	.7918651	.0675072	-2.74	0.006	.6700165 .9358729
Body weakness	.9455134	.0754728	-0.70	0.483	.8085801 1.105636
Headache	.843223	.067988	-2.11	0.034	.7199646 .9875833

Sore throat	.9403254	.0818874	-0.71	0.480	.7927787 1.115332
S Breath	.8355351	.0668725	-2.25	0.025	.7142303 .9774422
cough	.8957872	.0748345	-1.32	0.188	.7604928 1.055151
Fever	.8646044	.0695989	-1.81	0.071	.7384098 1.012366
Symptomatic	.99464	.0799361	-0.07	0.947	.8496841 1.164325
Region	.9790278	.0424708	-0.49	0.625	.8992271 1.06591
Nationality	.6310305	.4471309	-0.65	0.516	.1573665 2.530395
sex	.9700089	.0793137	-0.37	0.710	.8263735 1.13861
Temperature	1.002327	.0013269	1.76	0.079	.9997298 1.004931
Occupation	.961608	.0217931	-1.73	0.084	.919829 1.005285
age	.8948325	.0190395	-5.22	0.000	.8582832 .9329383

4.3 Multivariable Covariates Analysis

One problem of single covariate approach is that it ignores the possibility that a collection of variables, each of which is weakly associated with the status, can become an important predictor of the status when taken together. It is for this reason that we used p-value of 0.25 for selection of variables that are candidates for the multiple covariate analysis from single covariate findings. Results presented in Table 4.4 indicate the parameter estimates of coefficients be seen from Table 4.3 the covariates comorbidity, number of comorbidity, other pain without covid-19, Head ache Shortness of breathing, age passed the first filtration of variables for multiple covariate analysis and then forward Stepwise (Conditional LR) variable selection method was used to select the important variables to be included in Cox proportional hazards model. In order to decide whether or not a variable is significant, the p-value associated with each parameter has been estimated and variables that have p-value less than 0.05 are considered as important variables and hence, are included in the study. Survival of covid-19 patients was significantly related with comorbidity, number of comorbidity, other pain without covid-19, Headache, severity, traveling history of last 14 days and age.

Table 4.4: Results of multivariable Cox PH Model time-to-recovery from COVID-19

Covariates	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]
Occupation Unemployer ®					
Students	**1.043005**	**.1631862**	**0.27**	**0.788**	**.7675544 1.417305**
Merchant	**.8484427**	**.2051527**	**-0.68**	**0.497**	**.5282053 1.362832**
Civil Servant	**.9256503**	**.1374426**	**-0.52**	**0.603**	**.6919244 1.238327**
Others	**.8685706**	**.091269**	**-1.34**	**0.180**	**.7069054 1.067208**
severity No severity®					
mild	**1.380436**	**.1708187**	**2.61**	**0.009**	**1.083145 1.759326**
moderet	**1.18427**	**.146921**	**1.36**	**0.173**	**.9286465 1.510258**
sever	**.8885356**	**.1438726**	**-0.73**	**0.465**	**.6469153 1.2204**
crtical	**.7828052**	**.123525**	**-1.55**	**0.121**	**.5745607 1.066526**
Assymptotic	**1.232401**	**.2841902**	**0.91**	**0.365**	**.7842711 1.936592**
Hvisit No ®					
yes	**1.61218**	**.1979198**	**3.89**	**0.000**	**1.267405 2.050745**
TravelingHx No ®					
yes	**.5743962**	**.085234**	**-3.74**	**0.000**	**.4294402 .7682815**
comorbidty No ®					
yes	**.6230674**	**.083512**	**-3.53**	**0.000**	**.4791212 .8102606**
Nuembr of morbidity NoComorbodity ®					
one comorbodity	**1.434926**	**.2017101**	**2.57**	**0.010**	**1.089368 1.890099**
two co morbodity	**1.299057**	**.2368521**	**1.44**	**0.151**	**.9087239 1.857054**
More than two	**1.738234**	**.5832944**	**1.65**	**0.099**	**.9004774 3.355394**
Other_Pain No ®					
yes	**.7032371**	**.0664086**	**-3.73**	**0.000**	**.5844138 .8462196**
HeadAche No ®					
yes	**.801149**	**.072865**	**-2.44**	**0.015**	**.6703414 .957482**
SBreath No ®					
yes	**1.010042**	**.0930455**	**0.11**	**0.914**	**.8431921 1.209909**
cough No ®					
yes	**1.030722**	**.0984472**	**0.32**	**0.751**	**.8547531 1.242917**
Fever No ®					
yes	**.9339943**	**.0822626**	**-0.78**	**0.438**	**.7859116 1.109979**
age <20 ®					
20-30	**1.285365**	**.2072104**	**1.56**	**0.119**	**.9371447 1.762977**
30-40	**1.144618**	**.1732915**	**0.89**	**0.372**	**.8507297 1.540032**

40-50	**.8670399**	**.1425252**	**-0.87**	**0.385**	**.6282283 1.196632**
50-60	**.7489048**	**.1235871**	**-1.75**	**0.080**	**.5419486 1.034892**
60-70	**.7049835**	**.1145863**	**-2.15**	**0.031**	**.5126568 .969463**
70 and above	**.7638391**	**.1395943**	**-1.47**	**0.140**	**.5338777 1.092854**

SE: Standard Error **Df:** Degree of freedom **CI:** Confidence Interval **R**= Reference group

4.3.1 Assessment of Model Adequacy

4.3.1.1 Graphical Assessment Proportional Hazard Assumption

To check the PH assumption for covariates included in the fitted model, we used the ln (-ln (survival probability)) plot In principle the schoenfeld residuals are independent of time .a plot that shows a non-random pattern against time is evidence of violation of the PH assumption

4.3.1.1.2 Test for checking PH assumptions Use global test

Table 4.5: Test of proportional-hazards assumption based on Schoenfeld residuals

	Chi-square	Df	P-value
Global test	12.04	19	0.884

The global test indicates that the PH assumption fulfilled (table 4.5)

From the output above the test is not statistically significant for each of the covariates,and the global test is also not statistically significant.Therefore ,we can assume the proportional hazards

4.3.1.2 Assessment of Model Adequacy

To check the PH assumption for covariates included in the fitted model, we used the ln (-ln (survival probability)) plot versus ln survival time figures 4.10 in appendix suggested that two lines corresponding to ln[-ln[S(t)]] versus ln(time) are not distributed parallel, the distance is changing over time and cross each other which suggests violation from PH assumption for covariates severity, comorbidity, number of comorbidity, other pain without covid-19 age. The formal test applied to the model presented in Table 4.5 shows the time-dependent covariates (interaction of covariates with logarithm of time) were not significant for severity, comorbidity, number of comorbidity, other pain without covid-19 age which justifies the proportional hazard assumption holds at 5% level of significance. But newly added covariates were statistically significant which indicates that proportional hazard assumptions were not satisfied.

PH assumption the predictor variables their interaction with log time for time-to-Recovery

Table 4.5 PH assumption the predictor variables their interaction with log time for time-to-Recovery

Covariates	HR	Std.Err.	z	P>z	[95%Conf. Interval]
Newage	.8269834	.0651365	-2.41	0.016	.7086843 .9650299
Hvisit	1.325108	.5890551	0.63	0.527	.5544515 3.166934
TravelingHx	.2115995	.1212628	-2.71	0.007	.0688194 .650607
Other_Pain	.8952005	.2773846	-0.36	0.721	.4877184 1.643128
HeadAche	.6094231	.1890857	-1.60	0.110	.3317543 1.119493
severity	1.101146	.1097841	0.97	0.334	.9056919 1.338781
comorbidty	.6066266	.233617	-1.30	0.194	.2851784 1.290406
Co_morbidity	1.13678	.2713271	0.54	0.591	.7120516 1.814854
tvc					
Newage	1.035415	.0348806	1.03	0.302	.969258 1.106087
Hvisit	1.051731	.2006818	0.26	0.792	.7235788 1.528703
TravelingHx	1.504469	.359391	1.71	0.087	.9419897 2.402814
Other_Pain	.9269515	.1201979	-0.58	0.559	.7189217 1.195178
HeadAche	1.108482	.1468911	0.78	0.437	.8549316 1.43723
severity	.9408148	.0399346	-1.44	0.151	.8657118 1.022433
comorbidty	1.095737	.1750804	0.57	0.567	.8011216 1.498697
Co_morbidity	.9883188	.0992866	-0.12	0.907	.8116807 1.203397

Note: Variables in tvc equation interacted with ln(_t).

The plot of Schoenfeld residuals in Appendix figure 4.10 shows that distributes in nonsystematic way about the reference line (without definite increment or decrement), but the Lowess curve connecting the values of the smoothed residuals is slightly upward and down ward (not horizontally) for same covariates i.e. Schoenfeld Residuals do not give straight forward answer, however they might suggest violation from PH assumption for those covariates. Thus, the researcher doubts the accuracy of the PH assumption and considers the parametric model for this data set.

4.4 Parametric Regression Modeling for the Recovery Time of Covid-19 Patients

The common applicable criterion to select the model is the Akaikie information criterion (AIC) statistic proposed by Akaikie (1316.232). From Table 4.8 the Weibull regression model is the least AIC value which shows that the data of Recovering time of covid-19 patients fits for the Weibull regression model.

Table 4.6: AIC and BIC value of parametric Model for time-to-recovery

Baseline Distribution	AIC	BIC
Exponentials	**1687.568**	**1725.165**
Gompertez	**1480.35**	**1522.623**
Loglogistic	**1344.786**	**1387.06**
Weibull	**1316.232***	**1358.512***
Lognormal	**1441.45**	**1483.725**

4.4.1 Multivariate Analysis of Weibull Regression Model

In order to decide whether or not a variable is significant, the p-value associated with each parameter has been estimated and variables that have p-value less than or equal to 0.05 are considered as important variables and hence, are included in the study. Results presented in Table 4.7 indicate the parameter estimates of coefficients $\widehat{B_\iota}$ for the covariates in the final Weibull regression model along with the associated standard error covariate effects(γ_j) significance level, hazard ratio and 95% confidence interval for the hazard ratio. Survival time of covid-19 patients were significantly related with Severity, Comorbidity, Other_Pain , HeadAche, Age The Wald statistics for the parameter estimates indicate that at least one of the parameters in each covariate is significantly different from zero at 0.05 levels of significance. The formal tests are applied to the model adequacy and the results are displayed in section. From the Weibull regression model, after fixing other coefficients, the hazard rate for Severity patients were 0.87 times that of patients who had Mild Sever patients, whereas the hazard rate decreases for patients who had no severity symptoms. The hazard rate of patients who had comorbidity was staying time 1.26 times more than that who did not had comorbidity. Considering status of patents, keeping other confounding covariates constant, the hazard rate for patents who had otherpainwithoutcovid-19 is 1.24 times greater than non-other pain without covid-19.

Table 4.7: Parameter estimates, standard errors and the hazard ratios in the final Weibull regression model (at Hawass Referral Hospital, during 2013-2014)

	Coffi.	Std. Err.	Z	P>|z|	Haz. Ratio	95% CI for HR
Severity						
Mild	.3264644	.1689908	2.68	0.007	1.386059	1.091444 1.760199
Moderet	.1339984	.140327	1.09	0.275	1.143391	.898934 1.454326
Sever	-.0723883	.1471774	-0.46	0.647	.9301696	.6821495 1.268366
Crtical	-.2463373	.1226847	-1.57	0.117	.7816585	.5746684 1.063205
Assymptotic	.3054374	.3036688	1.37	0.172	1.357219	.8753847 2.104266
Hvisit						
Yes	.5082829	.2014198	4.20	0.000	1.662434	1.311032 2.108025
TravelingHx						
Yes	-.640271	.0728463	-4.63	0.000	.5271496	.402075 .6911314
comorbidty						
Yes	-.5681494	.0732555	-4.39	0.000	.566573	.4397431 .7299829
Co_morbidity						
one comorbidity	.4268915	.2035902	3.21	0.001	1.532486	1.181177 1.988284
two co morbidity	.3273931	.2450887	1.85	0.064	1.387347	.9813224 1.961365
morethan two	.6436124	.62844	1.95	0.051	1.903344	.9964841 3.635501
Other_Pain						
Yes	-.4675198	.0579452	-5.06	0.000	.6265543	.5226821 .751069
HeadAche						
Yes	-.2346949	.0717008	-2.59	0.010	.7908121	.6620597 .9446033
Newage						
20-30	.2592005	.2055328	1.63	0.102	1.295894	.9496555 1.768368
30-40	.0522361	.1551102	0.35	0.723	1.053624	.7895425 1.406035
40-50	-.220926	.1280067	-1.38	0.166	.801776	.5863479 1.096354
50-60	-.3494256	.1134637	-2.17	0.030	.705093	.5143643 .9665446
60-70	-.4207173	.1046997	-2.64	0.008	.6565757	.4803408 .8974705
70 and above	-.3756402	.1222438	-2.11	0.035	.6868494	.4845815 .9735456

_cons	-4.947399	.0015296	-22.97	0.000	.0071019	.0046563 .0108319
/ln_p |	.6930414	.0291055	23.81	0.000	.6930414	.6359956 .7500871
p |	1.999788	.0582049			1.999788	1.888902 2.117184
1/p |	.5000529	.0145543			.5000529	.4723254 .5294081

4.4.1.1 Assessment of Adequacy of the Weibull Regression Model

From the likelihood ratio test Table 4.8 below, it can be seen that the model is significant and in using the log likelihood values of the null model and the full model it can be seen that the model has a significant improvement after the covariates are added in the model.

Table 4.8: The likelihood ratio and significance of the Weibull regression model ((at Hawass Referral Hospital, during 2013-2041)

Loglik(intercept only)	Loglik (model)	Chi-square	Df	P	Scale	intercept
-694.06302	-632.34112	123.44	19	0.0000	0.0015296	-4.947399

Using the regression model of equation (3.36) and with the parameters found, the survival time of diabetic patients have Weibull distribution, which can be expressed as: $\mathrm{time} \sim \mathrm{weibull}(\alpha, \mu)$ with parameter $\mu = \exp(\mathrm{intercept}) = \exp(-4.947399\ ) = 0.0071$ and $\alpha = \frac{1}{\mathrm{scale}} = \frac{1}{0.0015296} = 653.76$ .By substituting the parameters in the final Weibull model with substitution of $\lambda = \mu^{-\alpha} = 0.0071^{-653.76} = 7.1 * 10^{-656}$ then

$$h_o(t) = \lambda\alpha t^{\alpha-1} = \ 7.1 * 10^{-656} * 654 * t^{653} = 4.643 * 10^{-653} * t^{653}$$

The Weibull hazard regression model that predicts the recovery time of covid-19 patients with identical data settings were:

$$h(t, x, \beta) = 4.643 * 10^{-653} * t^{653} * \exp(\beta)$$

In parametric settings, except for exponential regression models the base line function is not proportional for all subjects as a case of Cox regression model. For the Weibull regression model the base line hazard will vary with $h_o(t) = \lambda\alpha t^{\alpha-1}$ so the base line hazard function of covid-19 patients at Hawassa Referral Hospital, was with formula of (3.46) in every increase in time measured in years:

$$h_o(t) = \lambda\alpha t^{\alpha-1} = 4.643 * 10^{-653} * t^{653}$$

The interpretations of the hazard from Weibull outputs

The length of staying time in hospital acovid-19 patients with mild sever was 1.386 times higher than a patients no severity one. The hazard ratio for age groups above 60 year old were increased the staying time by 0.705 times higher than a patients below 20 years. Finally the patients with comorbidity has higher hospital staying time compered to non-comorbidity patients.

4.5 Discussion on the Results

The objective of the study was to identify significant risk factors that affect time-to-recovery from Covid-19 patients in Hawassa University Comprehensive and referral hospital. For determining the risk factors for covid-19 patients and modeling the survival time, a total of 826 patients were included in the study out of which 77.6% were recovered and the rest 22.4% were censored.

In a study done in Kuyha COVID-19 Isolation and Treatment Centre, Mekelle University, North Ethiopia, 81.7% were recovered and the rest 18.3% were censored, which is almost similar what was observed by the current study.the log-rank test revealed that Age, Severity status, comorbidity, pains out of covid-19 Head Ache and shortness of Birthing were significant difference in recovery time. This result is in line with other studies (Zhuo Wang, Yuanyuan Liu (2022).

The main aim of the study was to modeling time-to-recovery from Covid-19 using appropriate survival models. The comparison of distributions of the models was done using the AIC criteria, where a model with minimum AIC is accepted to be the best (Munda, 2012). In this study, Weibull regression model which had AIC value is 1338.35 was the most appropriate model to describe the Covid-19 data set. This study showed that there was heterogeneity between the ages on the timing of recovery among patients. While time of study was concluded in other studies as a covariate that does not impact LOS, we found that each week after the onset of pandemic is associated with decreased LOS. This could be explained by the fact that caregivers have become more experienced and proficient at managing the disease. There were also newer technologies with quicker turn-around time for tests, resulting in a shorter pre-discharge holding time in the hospital. Analysis also shows that the age of hospitalized COVID-19 patients' decreases since the pandemic started. Hence, the decrease of LOS is also consistent with the age association with LOS discussed above.

Assuming patients coming from the same age catagory share similar risk factors related to covid-19. This finding is similar to finding of ((Vekaria *et al. BMC Infectious Diseases, 2021*).

One of the factors that affect time-to-recovery from covid-19 is the Severity status of the patients. This study shows that the accelerated factor a patient with non-severity Sympotomic is fast as compared to those whose mild sever. This indicates that higher severity status shorter the chance of recovery as compared to non-sever patients. The result is in accordance with the earlier studies (Zhuo Wang, John S. Ji,2020) Comorbidity status is an important predictor for the recovery of covid-19 patient with [Φ=1.26, 95%CI: 0.648, 1.05], which suggested that the survival time of comorbidity patients had prolonged time-to-recovery from covid-19 than Non-comorbidity patients. The most prevalent comorbidity in population was hypertension and next prevalent comorbidities were fluid and electrolytes disorders, and obesity. This result is in line with the study by (International Journal of Population Data Science (IJPDS, 2021).The other pains out of covid-19 is another prognostic factor that significantly predicts the recovery time of covid-19 patient at [Φ=1.24, 95%CI: 1.13, 1.37]. The study revealed that the survival time for patients living with other pain is increased by a factor of 1.24 compared to the living non-other pains. The result is similar to earlier study (Biadgilign, S. et al). Acording to Noel George, Naresh K. Tyagi, Jang Bahadur Prasad COVID-19 pandemic and its average recovery time in Indian states

The average time of recovery from the disease in India ranges from 5 days to 60 days. However, Hospital studies by Manashet al. (2020) showed that COVID-19 patients in India have an estimated recovery time of 25 days (CI = 16 days–34 days). The recovery period estimated for male and female patients and patients' belonging to different age groups is also statistically significant. This result is consistent with (Noel George *, Naresh K. Tyagi, Jang Bahadur Prasad).This study pointed out that the median time to recovery from SARS COV-2 was 12 days. This is similar to the studies done in Eka Kotebe General Hospital, Ethiopia in which viral clearance lasted for 19 days (Abrahim S.A., et al., 2020) and 16 days (Leulseged T.W., et al., 2020). This might be due to relative similarity in care and treatment given for the patients in both study areas. Moreover, it is also is consistent with the previous study findings from Israel (20–21) days (Voinsky I., Baristaite G., and Gurwitz D., 2020). However, the median recovery time was lower in the previous studies done in Guangzhou Eighth People's Hospital, China (12 days) (Chen X., et al., 2020), University of California San Diego Health (7 days) (Daniels L.B., et al.,2020), Zhejiang University and the Shenzhen Third People's Hospital, China (15 days) (Xu K., et al.,2020) , and in Singapore (12 days) (Young B.E., et al.,2020) . The possible reason for the observed discrepancy between the studies might be due to variation in sample size, study setting, socioeconomic characteristics, and the severity of the disease. Evidences are showing the severe the disease condition, the longer the duration of viral RNAclearance (Xu J., et al., 2020).The present study found older age as independent predictor of delayed recovery time from coronavirus disease. This is consistent with previous study findings from Guangzhou, China(Chen X., et al., 2020), Korea (Das A. and Gopalan S.S.,2020), Wuhan Pulmonary Hospital, China(Du R.-H., et al.,2020) , Shenzhen, China(Xu K., et al.,2020) , three hospitals in Wuhan, China(Xu J., et al.,2020) , and Qingdao, China (Hu X., et al.,2020). This might be attributed to older age related severity progression of COVID-19 cases which in turn leads to either death or delayed duration of viral clearance in elderly patients (Wu J., et al., 2020). Moreover, it might be due to the fact that older age is not without comorbid conditions which are among the major risk factor of lower recovery rate form coronavirus disease and even death related to COVID-19. Besides, older age is associated with degeneration of pulmonary function and compromised immunity that contributes for severe COVID-19 cases and poor clinical outcomes. The current study has also demonstrated that patients with comorbid condition had 44%lower odds of recovery rate from coronavirus disease compared to their counterparts. Similarly, existing evidences are supporting the present study finding, for instance, the study donein Italy (Boari G.E., et al.,2020), Fairfield General Hospital, Bury, UK (Chinnadurai R., et al.,2020), Wuhan Pulmonary Hospital, China(Du R.-H., et al.,2020), Jin Yin-tan Hospital and Tongji Hospital (Ruan Q., et al.,), and Wuhan, China (Xu K., et al.,2020) claim comorbid conditions majorly cardiovascular diseases attributed to the delayed duration of recovery from SARS COV-2 cases. Furthermore, our study has also identified absence of fever as a good prognostic factor of COVID-19 cases. This is in line with study conducted in Eka Kotebe treatment center of Ethiopia, in which presence of clinical manifestation on admission prolong the time of recovery from COVID-19(Leulseged T.W., et al., 2020). This is also supported by the study finding from Changsha, China (Qi L., et al., 2020). This could be due to the fact that the function of respiratory system is dependent on body temperature variations (Rubini A., 2011). This can be explained that an increment in body temperature results in increment in respiratory rate which increases the pulmonary work load eventually leading pulmonary insufficiency and lower recovery rate ((Rubini A.,2011).

CHAPTER FIVE

Conclusions And Recommendations

5.1 Conclusions

This study used survival times of covid-19 patients' dataset of those patients who started their covid-19 treatment from Based on a retrospective cohort study ,information of 826 total Covid-19 Patients that had been treated at Hawassa university Comprehensive and Referral hospital from September 20,2013 to January 20,2014 E.c with the aim of modeling time-to-recovery from covid-19 in HUCRH. A total of 826 patients were included in the study out of which 77.7% were cured and the rest 22.4% were censored. In assessing the significant risk factors the Log-Rank test revealed that, Age, severity status, comorbodity, pain out of covid-19, headache, and shortness of Birth had significant survival probability difference for time-to-recovery from covid-19.

The Cox regression analysis showed that the major factors that affect the survival of covid-19 patients are Age, mild severity status, comorbidity, other pain out of covid-19, head ache were less likely to survive. The result of this study also indicated that survival probability of a patient is not statistically different among groups classified by sex, occupation, Body temperature, regional effect contact history of previous patients.

To predict and model the survival time of covid-19 patients, various parametric regression models were applied. Among which the Weibull regression survival model is better fits to predict the survival time of the disease for the data of covid-19 patients of Hawassa University Comprehensive and Referral Hospital than the other parametric models. By means of stepwise selection the covariates that are selected by the model are Age, mild severity status, comorbidity, other pain out of covid-19, head ache.

5.2 Recommendations

Based on the result of the study different factors are identified for the recovery of covid-19 Patients. The following recommendations are made for health policy makers, clinicians and the public at large.

According to the results of this study the main predictive factors for the survival time of covid-19 patients are more of clinical variables. So, health workers should be cautious when a patient has high severity status, comorbidity, other pain out of covid-19, head ache. The government and concerned bodies should work on perception about the disease and its risk factors, so that patients should be well informed about the disease, early diagnose to minimize the length of hospital staying time.

Future studies also need to assess the level of awareness, treatment and control of these risk factors. The economic and social consequences of covid-19 and other chronic diseases should also receive due attention in future research, as these diseases involve lifelong medical care and social support with significant socioeconomic burden to the individual and the society at large.

The Ministry of Health need to equip health centers with necessary materials and human power to perform caesarian section at health center level. Patients with comorbidity have prolonged time-to-recovery. So special attention should be given to comorbidity patients to shorter time-to-recovery from their disease. Hawassa University Medical Center need to improve public and professional awareness, early detection and prompt treatment using feasible, effective regimens and include detailed patients characteristics in the covid-19 registry data. Thus, elders and individuals with comorbidity has to get due attention to prevent infection by the virus. Moreover, elders and patients with comorbidity should get priority in management of coronavirus disease in order to enhance good clinical outcome.

5.3 Limitation of the Study

This study had some limitations: the first is that the study used data from single hospital. Thus, the findings of this study should be interpreted very carefully when they are inferred to the national level. The second limitation is lack of published literature on our country related to the survival time of covid-19; the references are more of other countries outcome. Finally as different literature pointed out, there are different factors that are assumed to have impacts on the survival of covid-19 patients unregistered outcome and incomplete baseline data were excluded from the analysis, the reviewed records might lack very important variables that could influence recovery rate from coronavirus disease.

REFERENCES

Abrahim S.A., et al., Time to recovery and its predictors among adults hospitalized with COVID-19:A prospective cohort study in Ethiopia. PloS one, 2020. 15(12): p. e0244269. https://doi.org/10.1371/journal.pone.0244269 PMID: 33378367.

Alimohamadi, Y, Sepandi M, Taghdir M, Hosamirudsari H.2021. Determine the most common clinical symptoms in COVID-19 infection (Review). International Journal of Molecular Medicine. 47(6): 100 doi:10.3892/ijmm.

Agegnehu B, Abera M, Azene T, Behailu T, Adherence with COVID-19 Preventive Measures and Associated Factors among Residents of Derashe District, Southern Ethiopia. This article was published in the following Dove Press journal: Patient Preference and Adherence downloaded from https://www. dovepress.com/ by 197.156.93.185 on 14-May-2021

Boari G.E., et al., Prognostic factors and predictors of outcome in patients with COVID-19 and related pneumonia: a retrospective cohort study. Bioscience Reports, 2020. 40(12). https://doi.org/10.1042/ BSR20203455 PMID: 33201172

Chen Y, Tong X, Wang J, Huang W, Yin S, Huang R, et al. High SARS-CoV-2 antibody prevalence among healthcare workers exposed to COVID-19 patients. J Infect. 2020;81(3):420 6. https://doi.org/10.1016/j.jinf.2020.05.067.

Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R. Features, evaluation, and treatment of coronavirus (COVID-19).StatPearls. URL: https://www.statpearls.com/ArticleLibrary/viewarticle/52171 [accessed 2021-05-03]

Chen X., et al., Associations of clinical characteristics and treatment regimens with the duration of viral RNA shedding in patients with COVID-19. International Journal of Infectious Diseases 2020. 98: p. 252–260. https://doi.org/10.1016/j.ijid.2020.06.091 PMID: 32619760

Chinnadurai R., et al., Older age and frailty are the chief predictors of mortality in COVID-19 patients admitted to an acute medical unit in a secondary care setting-a cohort study. BMC geriatrics, 2020. 20 (1): p. 1–11.

Clinical Infectious Diseases, 2020.

Ditekemena J, Doumbia S, Ebrahim SH. COVID-19's final frontier: The central Africa region. Travel Med Infect Dis 2020; 37:101694. https://doi.org/10.1016/j.tmaid.2020.101694 PMID: 32360410

Das A. and Gopalan S.S., Epidemiology of CoVID-19 and predictors of recovery in the Republic of Korea. medRxiv, 2020. https://doi.org/10.1155/2020/7291698 PMID: 32774918

Daniels L.B., et al., Relation of statin use prior to admission to severity and recovery among COVID-19 inpatients. The American journal of cardiology, 2020. 136: p. 149–155. https://doi.org/10.1016/j. amjcard.2020.09.012 PMID: 32946859

Du R.-H., et al., Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. European Respiratory Journal, 2020. 55(5).

Fadnavis RA (2019) Application of machine learning for survival analysis- a review. IOSR J Eng (IOSRJEN) 9(5):56–60Cox DR (1975) Partial likelihood. Biometrika 62:269–762

(Gemechu Churiso, Kuma Diriba , Henok Girma, Soressa Tafere ,2021). Clinical Features and Time to Recovery of Admitted COVID-19 Cases at Dilla University Referral Hospital Treatment Center,South Ethiopia.

Grqnnesby JK, Borgan Q (1996) A method for checking regression models in survival analysis based on the risk score. Lifetime Data Anal 2:315–328

Hu X., et al., Factors associated with negative conversion of viral RNA in patients hospitalized with COVID-19. Science of the Total Environment, 2020: p. 138812. https://doi.org/10.1016/j.scitotenv. 2020.138812 PMID: 32335406

Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, Open COVID-19 Data Working Group, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020 May 01;368(6490):493-497 [FREE Full text] [CrossRef] [Medline]

Leulseged T.W., et al., Determinants of Time to Convalescence among COVID-19 Patients at Millennium COVID-19 Care Center in Ethiopia: A prospective cohort study. medRxiv, 2020.

(Misganu Endriyas ,Aknaw Kawza, Abraham Alano, Mamush Hussen,Endashaw Shibru,2020) COVID-19 prevention practices in urban setting during early introduction of the disease: results from community survey in SNNP Region, Ethiopia

OPM study OPM study:www.opml.co.uk/projects/building-resilience-in-ethiopia

Qi L., et al., Factors associated with duration of viral shedding in adults with COVID-19 outside of Wuhan, China: A retrospective cohort study. International Journal of Infectious Diseases, 2020. https:// doi.org/10.1016/j.ijid.2020.05.045 PMID: 32425636

Ruan Q., et al., Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive care medicine, 2020. 46(5): p. 846–848. https://doi.org/10. 1007/s00134-020-05991-x PMID: 32125452

Rubini A., The effect of body warming on respiratory mechanics in rats. Respiratory physiology & neurobiology, 2011. 175(2): p. 255–260. https://doi.org/10.1016/j.resp.2010.11.012 PMID: 21112417

(SeyedAlinaghi et al .,2021)Predictors of the prolonged recovery period in COVID-19 patients: a cross-sectional study

Shewasinad Yehualashet S, Asefa KK, Mekonnen AG, Gemeda BN, Shiferaw WS, Aynalem YA, et al.(2021) Predictors of adherence to COVID-19 prevention measure among communities in North Shoa Zone, Ethiopia based on health belief model: A cross-sectional study. PLoS ONE 16(1): e0246006.https://doi.org/10.1371/journal.pone.0246006 PMID: 33481962

(Salata C, Calistri A, Parolin C, Palù G (2019) Coronaviruses: a paradigm of new emerging zoonotic diseases. Pathog Dis. https://doi.org/10.1093/femspd/ftaa006

(Tadesse Tolossa et al,.2021) Time to recovery from COVID-19 and its predictors among patients admitted to treatment center of Wollega University Referral Hospital (WURH), Western Ethiopia: Survival analysis of retrospective cohort study

(Tinsae Abeya Geleta , Berhanu Senbeta Deriba , Kemal Jemal , 2022) Evaluation of the Knowledge, Attitude, and Practice of COVID-19 Prevention Methods Among Hypertensive Patients in North Shoa, Ethiopia

Voinsky I., Baristaite G., and Gurwitz D., Effects of age and sex on recovery from COVID19:Analysis of 5,769 Israeli patients. Journal of Infection, 2020.
Wang Z, He T, Zhu L, Sheng H, Huang S, Hu J. Active quarantine measures are the primary means to reduce the fatality rate of COVID-19. Bull World Health Organ. 2020:1–12.

World Health Organization. Statement on the second meeting of the International health regulations(2005) emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), 2020. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-internationalhealthregulations-(2005)-emergency-committee-regarding-the-outbreak-ofnovel-coronavirus-(2019-ncov) [Accessed 24 Jul 2020].
World meter, COVID-19 update, Coronavirus Update (Live): 119,723,984 Cases and 2,653,796 Deaths from COVID-19 Virus Pandemic—World meter (worldometers.info) <https://www.worldometers.info/coronavirus/>.

World Health Organization (2020) Director-General's remarks at the media briefng on 2019-nCoV on 11 February 2020, https://www.who.int/dg/speeches/detail/who-director-general-s-remar ks-at-the-media-briefing-on-2019-ncov-on-11-february-2020Accessed from 1 Apr 2020

(Wubedle Zelalem Temesgan etal.,2021) Adherence to COVID-19 preventive practice and associated factors among pregnant
women in Gondar city, northwest Ethiopia, 2021: Community-based cross-sectional study

Wu J., et al., Early antiviral treatment contributes to alleviate the severity and improve the prognosis of patients with novel coronavirus disease (COVID-19). Journal of internal medicine, 2020.

Wu JT, Leung K, Leung GM. 2020. Now casting and forecasting the potential domestic and international spread of the 2019nCoV outbreak originating in Wuhan, China: a modelling study. Lancet.:1-3.

Xu J., et al., Clinical course and predictors of 60-day mortality in 239 critically ill patients with COVID-19: a multicenter retrospective study from Wuhan, China. Critical Care, 2020. 24(1): p. 1–11. https://doi. org/10.1186/s13054-019-2683-3 PMID: 31898531

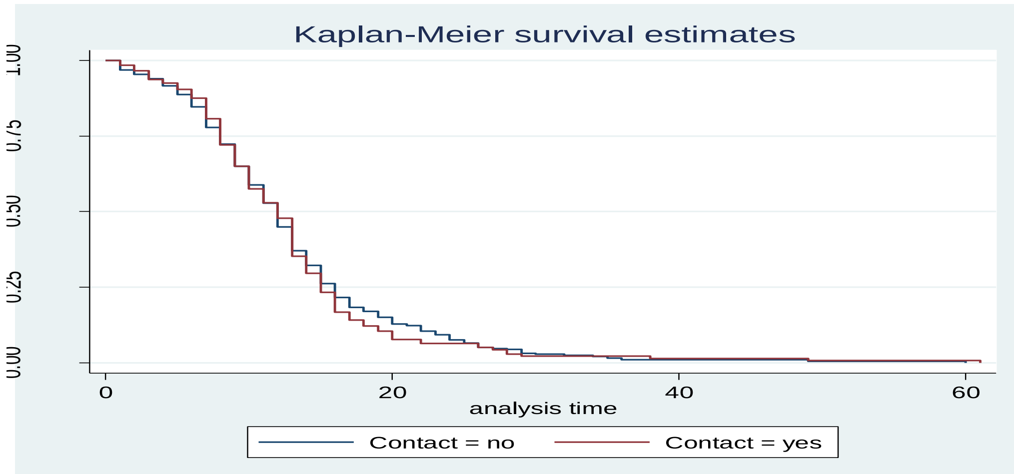
Xu K., et al., Factors associated with prolonged viral RNA shedding in patients with COVID-

Young B.E., et al., Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. Jama, 2020. 323(15): p. 1488–1494. https://doi.org/10.1001/jama.2020.3204 PMID:32125362

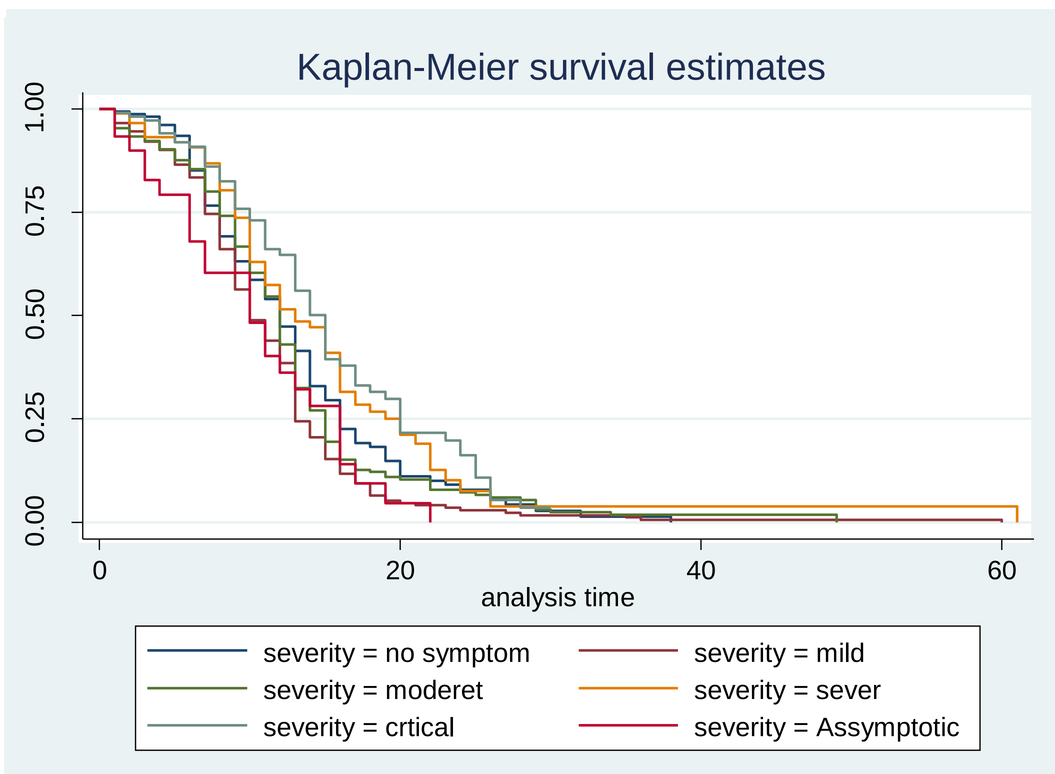
Yin Y. and Wunderink R.G.2021. The transformative Research During the Covid-19 Pandemic. A perspective. Department of Pediatrics, Division of Neonatology, Medical College of Wisconsin, Milwaukee, WI, United States. Zenbaba et al. Tropical Medicine and Health (2021) 49:30 Compliance towards infection prevention measures among health professionals in public hospitals, southeast Ethiopia: a cross-sectional study with implications of COVID-19 prevention (https://doi.org/10.1186/s41182-021-00318-y).

APPENDICES

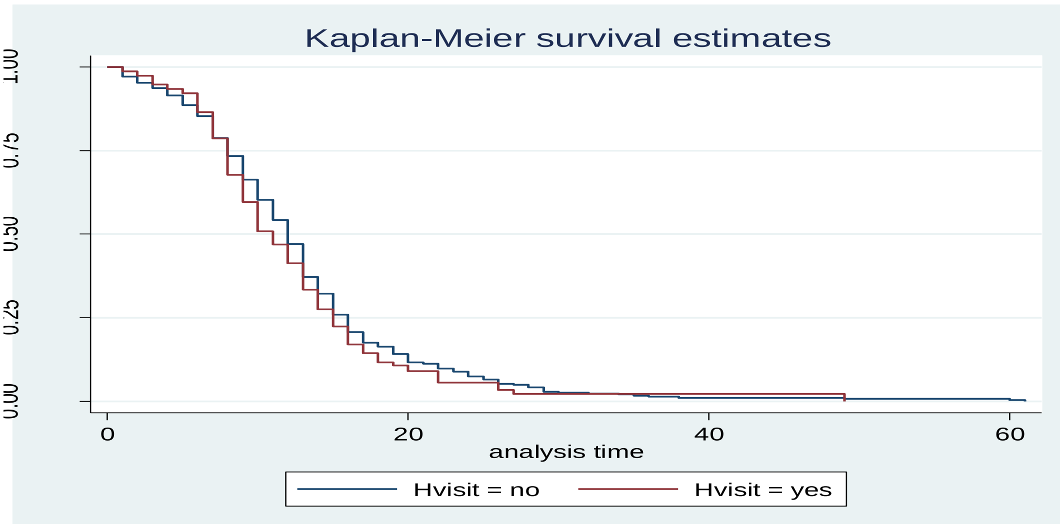
Appendix 1: The Kaplan-Meier survival function estimates



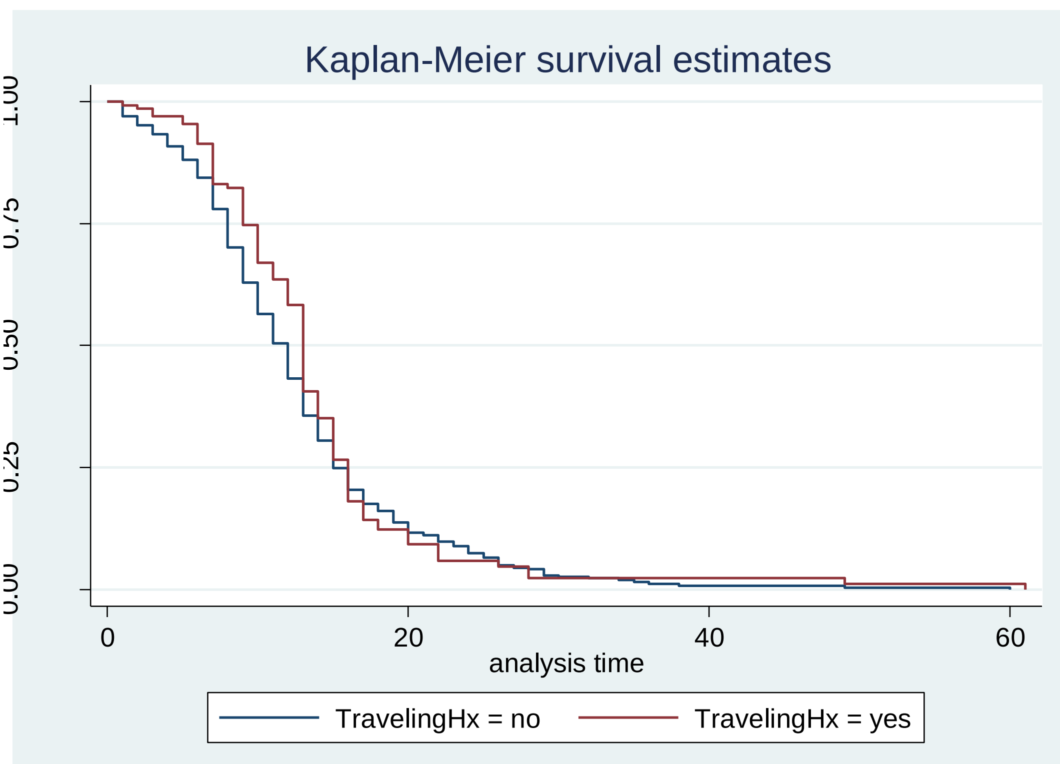
A) KM estimates of survival for the variable Clos contact with confirmed case



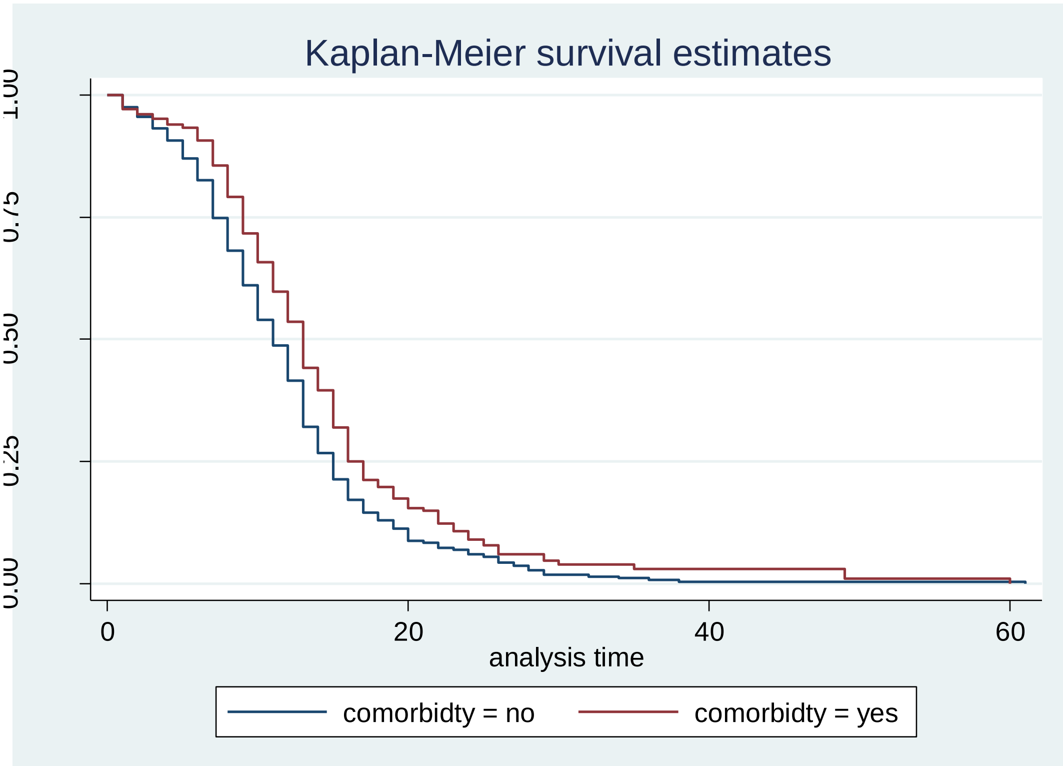
B) KM estimates of survival for the variable Severity



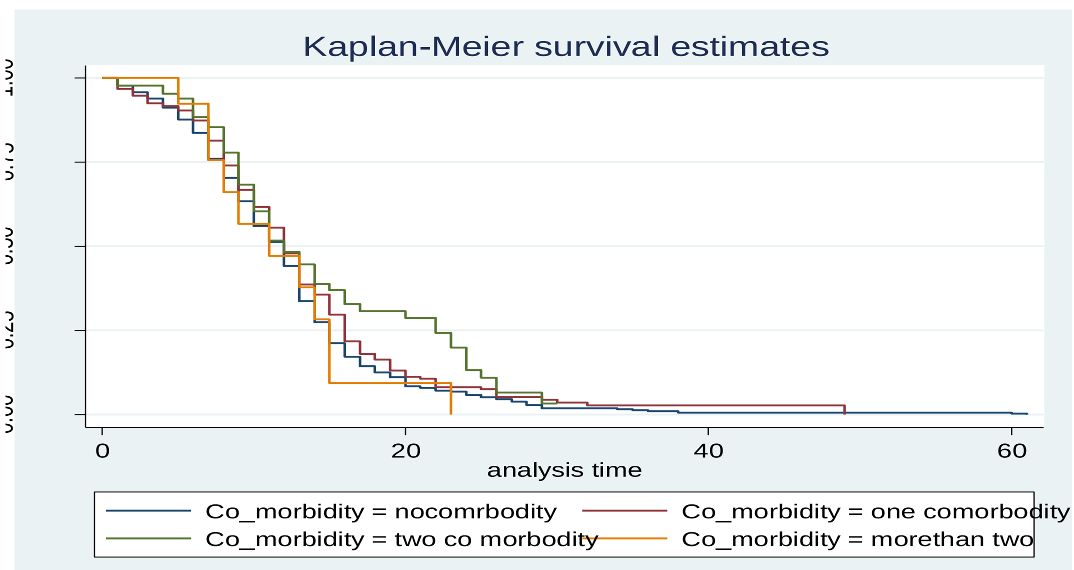
C) KM estimates of survival for the variable Visiting health facility before on set



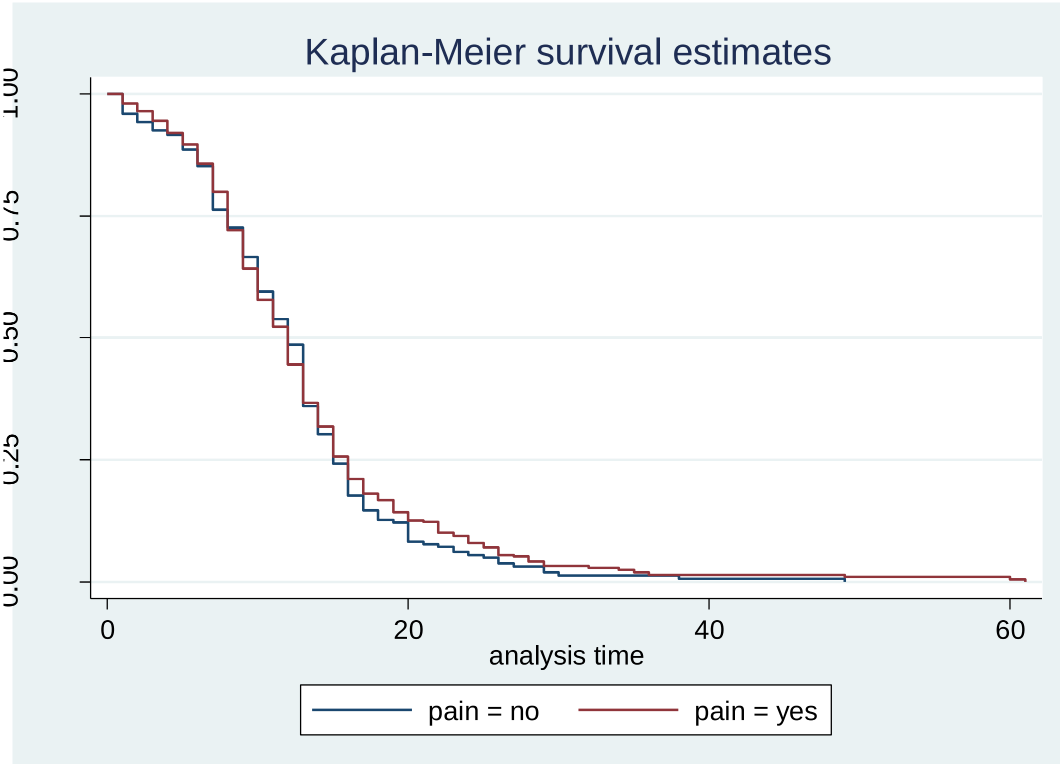
D) KM estimates of survival for the variable traveling history of last 14 days



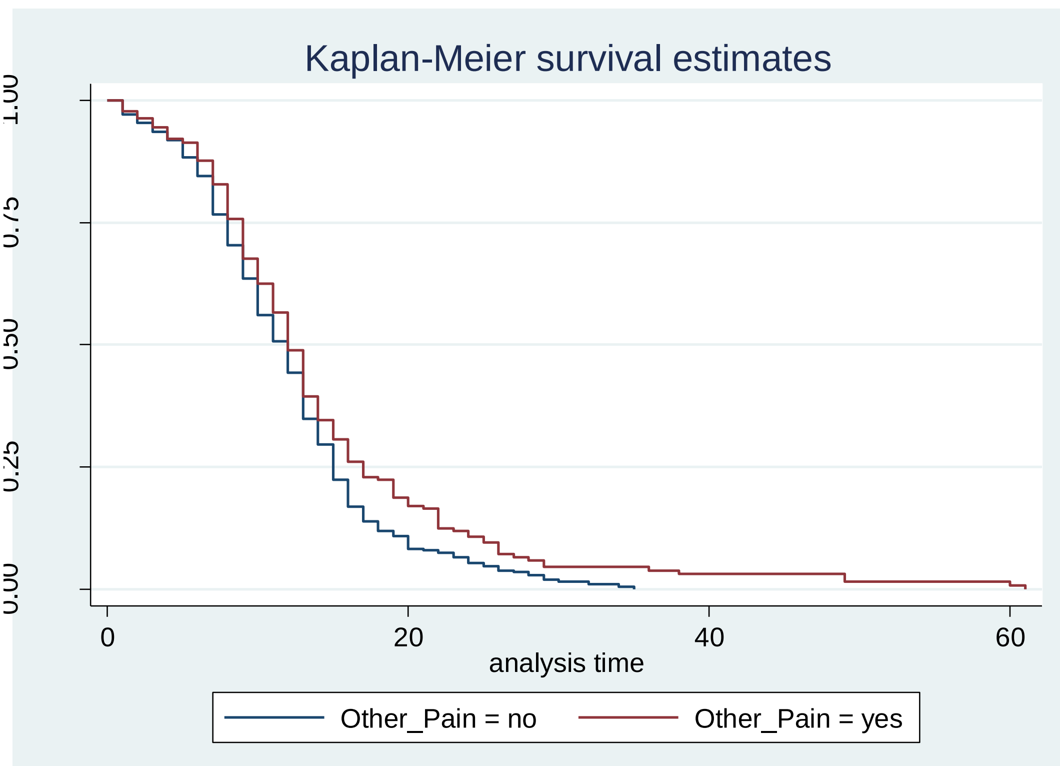
E) KM estimates of survival for comorbidity



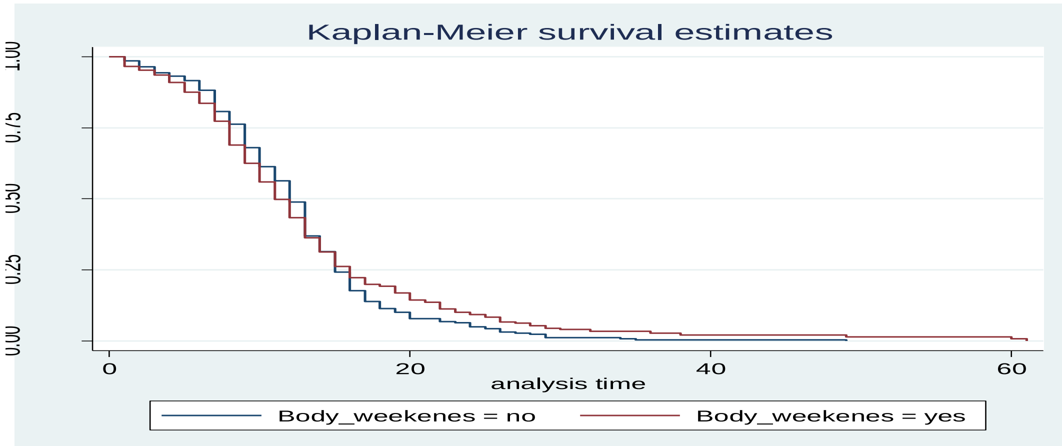
F) KM estimates of survival for Number of comorbidity



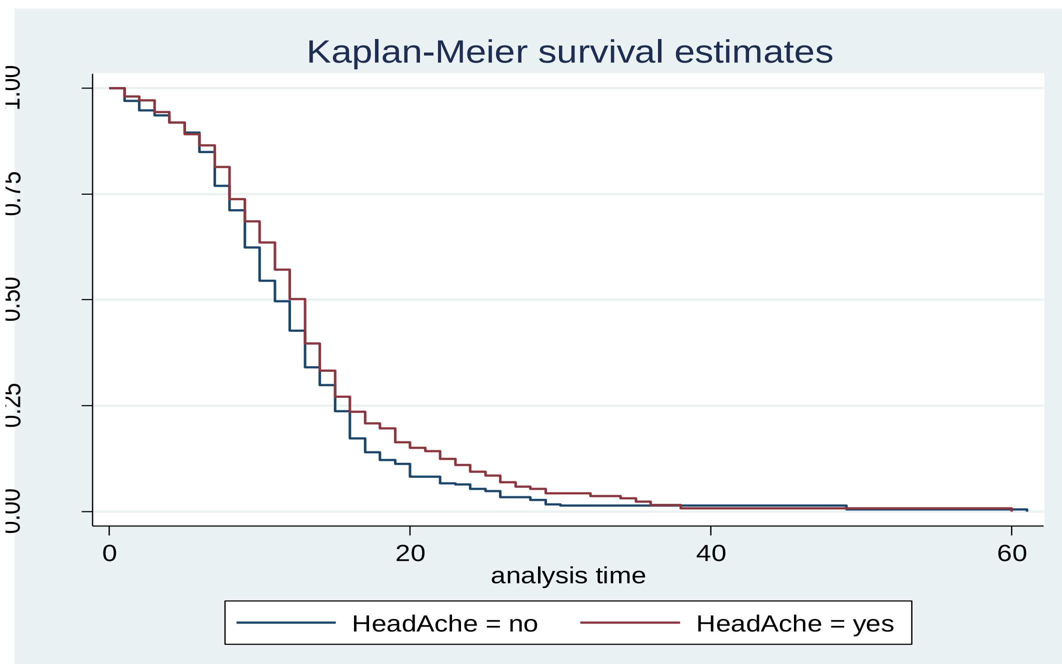
G) KM estimates of survival for pain



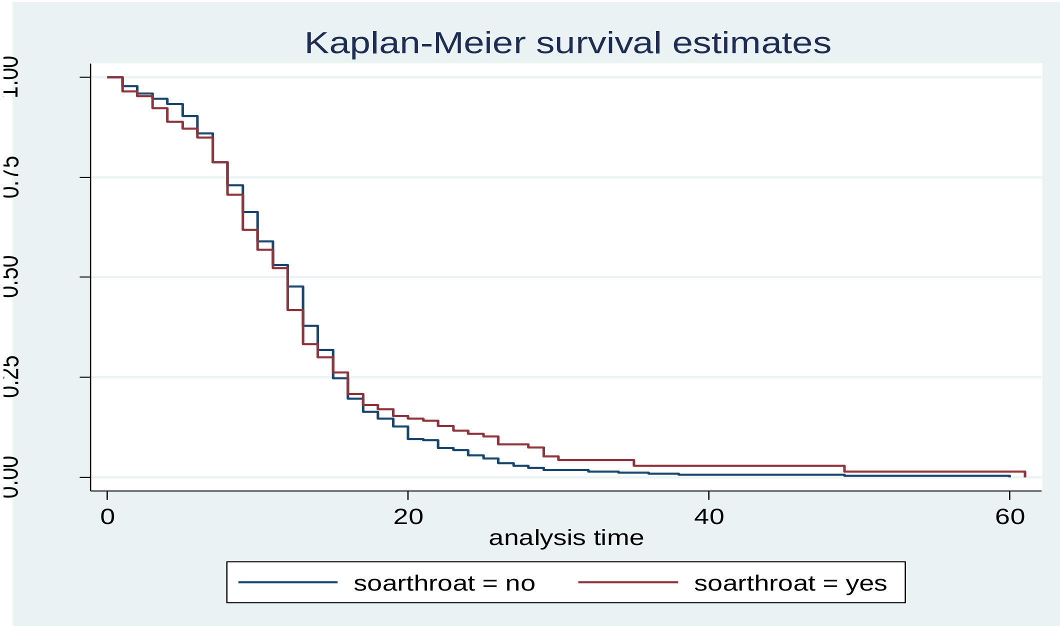
H) KM estimates of survival for the variable other pains with out covid-19



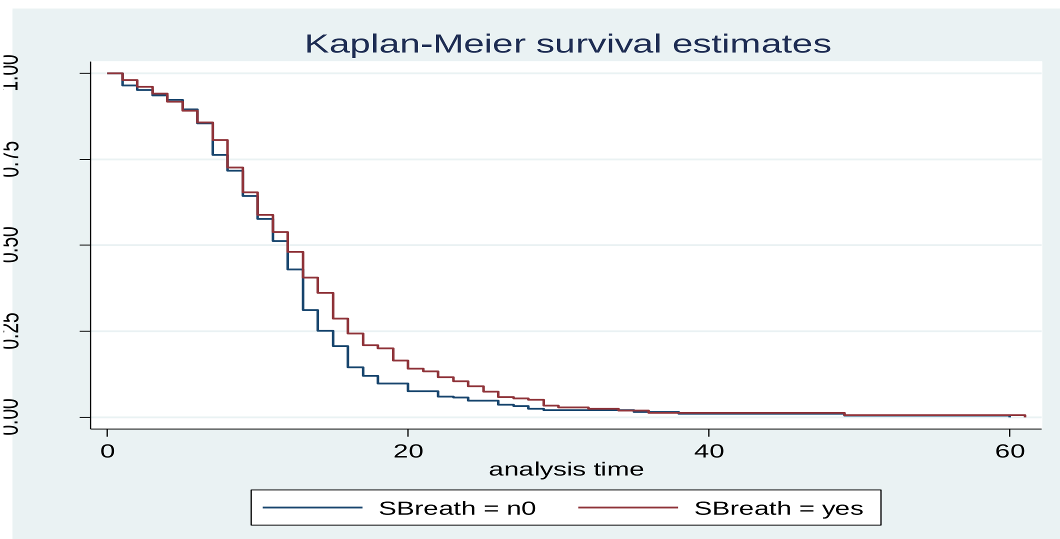
I) KM estimates of survival for the variable General Body weakness



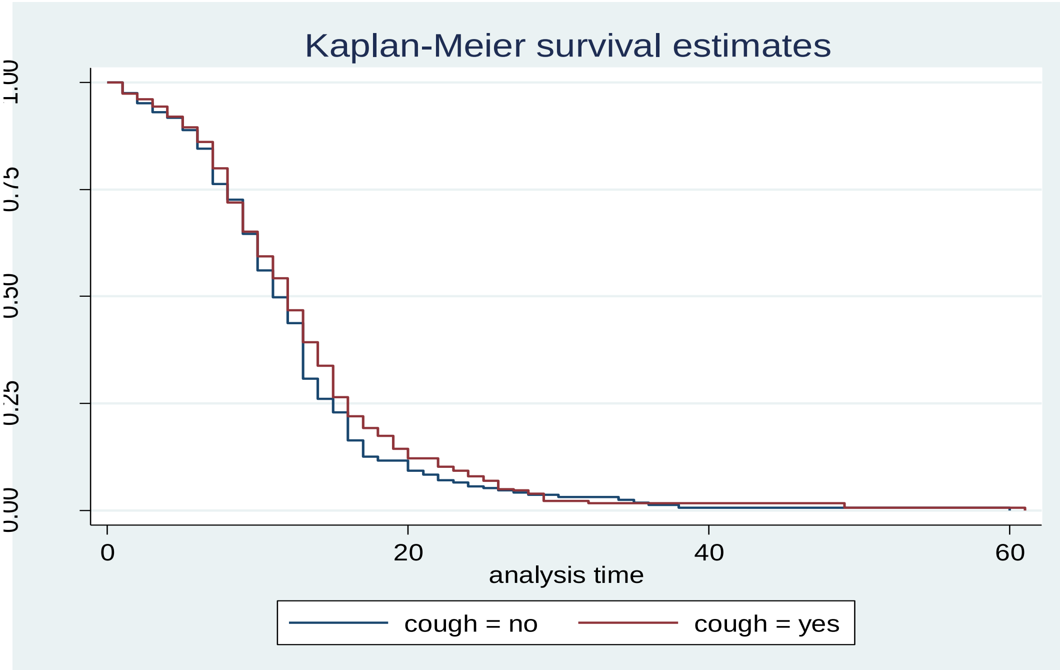
J) KM estimates of survival for the variable Head Ache



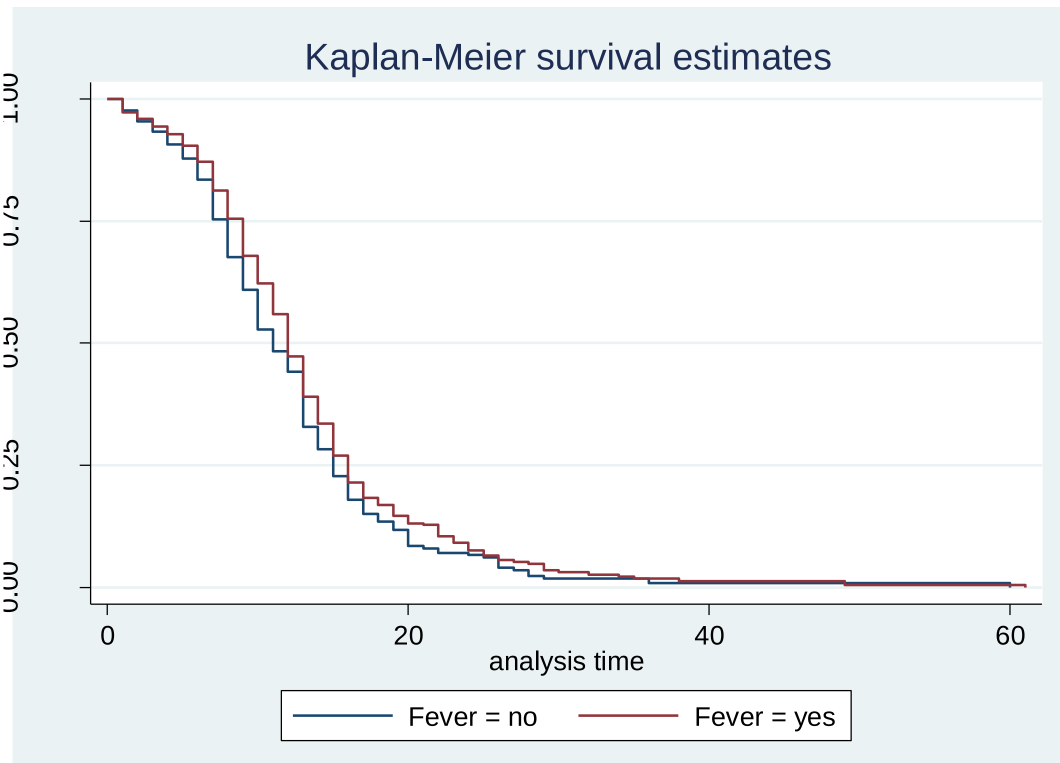
K) KM estimates of survival for the variable Soar throat of patients



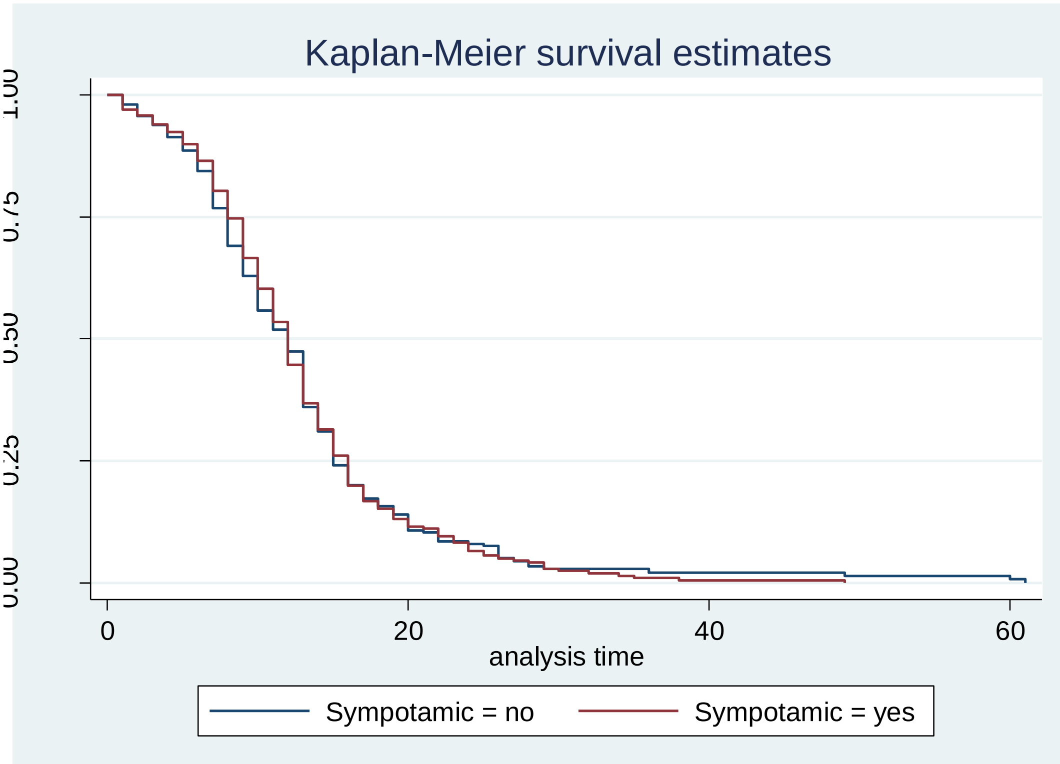
L) KM estimates of survival for the variable Shortness of breath



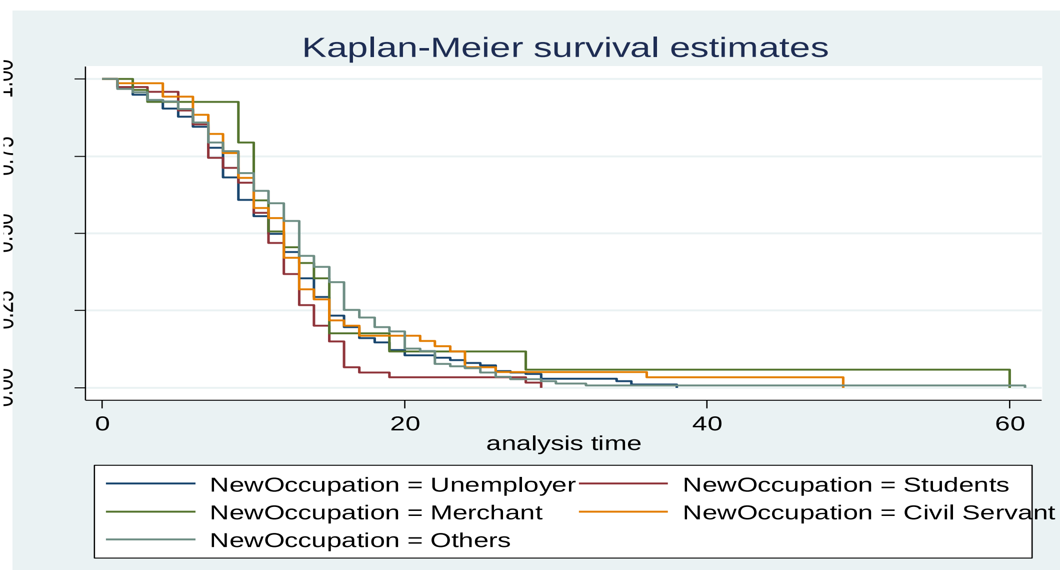
M) KM estimates of survival for the variable Cough



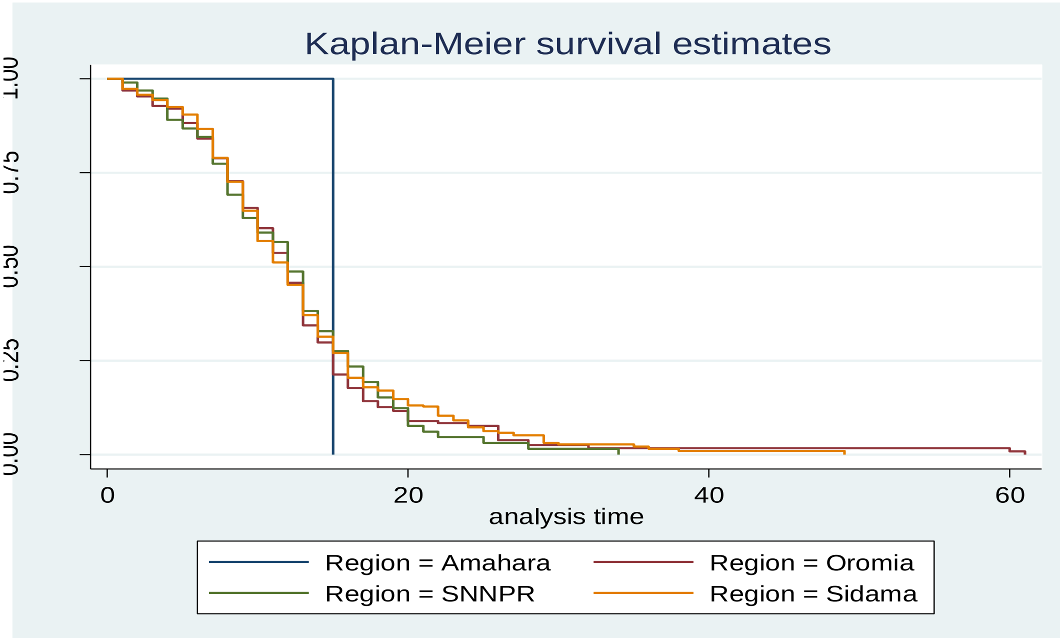
N) KM estimates of survival for the variable Fever



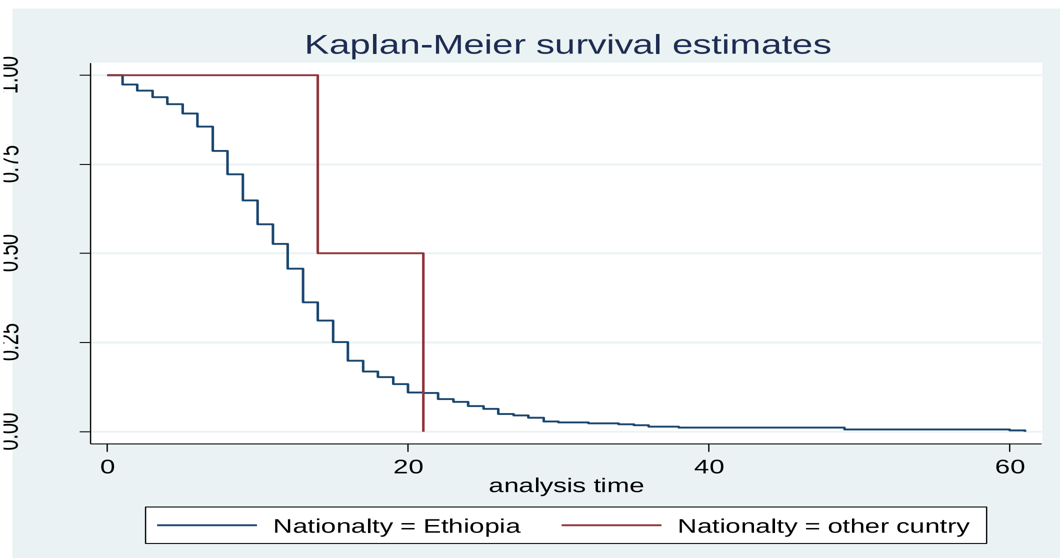
O) KM estimates of survival for the variable Sympotomic



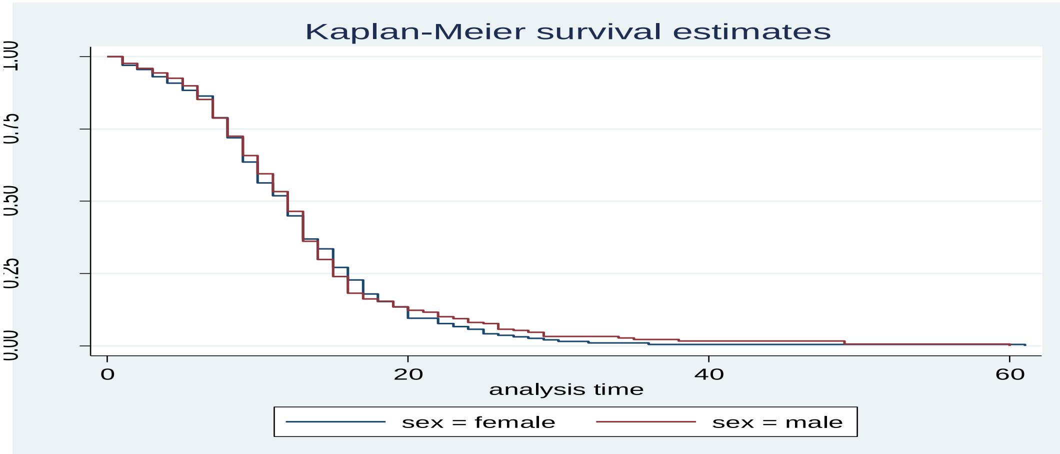
P) KM estimates of survival for the variable Occupation



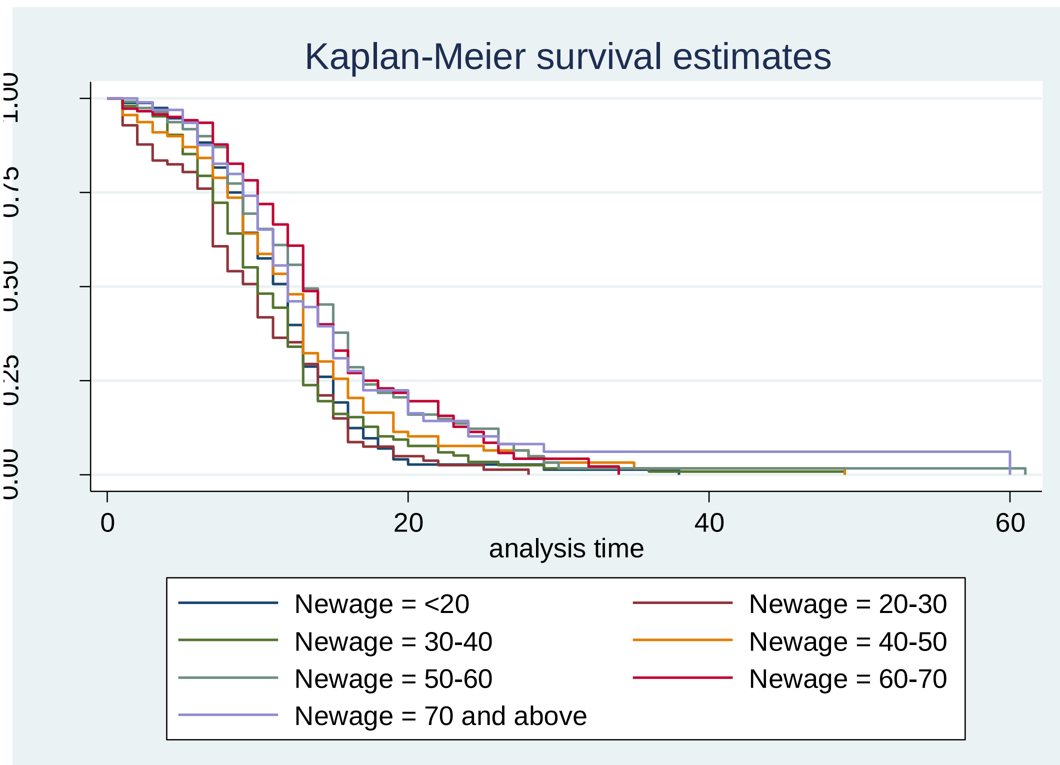
O) KM estimates of survival for the variable Region



P) KM estimates of survival for the variable Nationality



Q) KM estimates of survival for the variable Sex



R) KM estimate of survival for the variable Age.

Appendix 2: plots for model assessment

Figure4.10 To check the PH assumption for covariates included in the fitted model, we used the ln (-ln (survival probability)) plot

For All Varibel graphical tests

stphplot, by(severity)

. stphplot, by(Contact)

. stphplot, by(Hvisit)

. stphplot, by(TravelingHx)

. stphplot, by(comorbidty)

. stphplot, by(Co_morbidity)

. stphplot, by(pain)

. stphplot, by(Other_Pain)

. stphplot, by(Body_weekenes)

. stphplot, by(HeadAche)

. stphplot, by(soarthroat)

. stphplot, by(SBreath)

. stphplot, by(cough)

. stphplot, by(Fever)

. stphplot, by(Sympotamic)

. stphplot, by(Region)

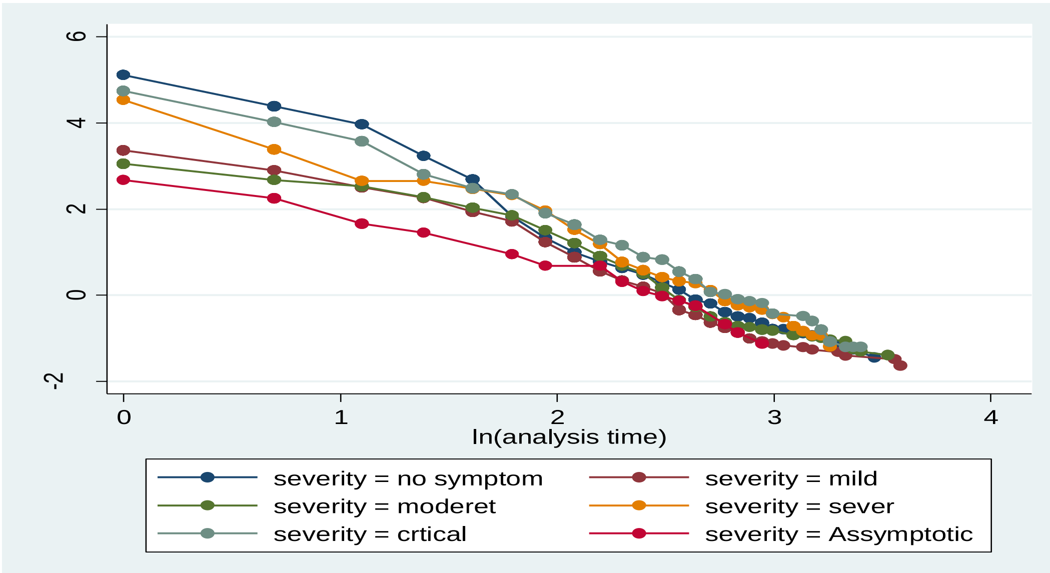
. stphplot, by(Nationalty)

. stphplot, by(sex)

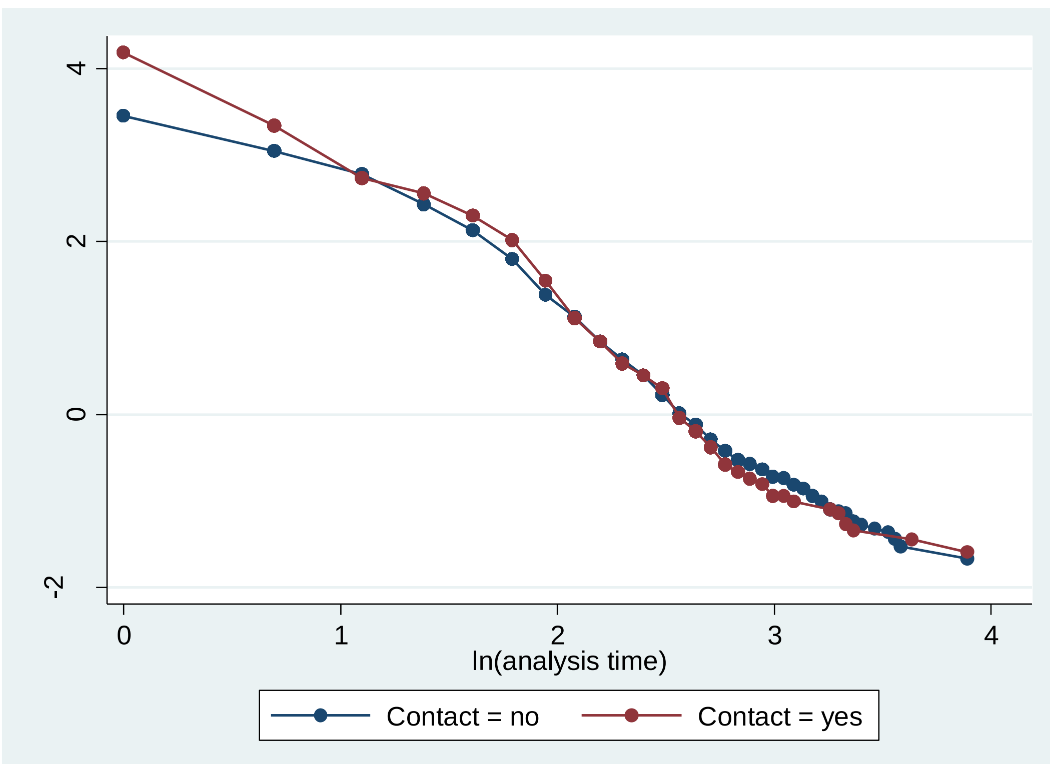
. stphplot, by(Temperture)

. stphplot, by(NewOccupation)

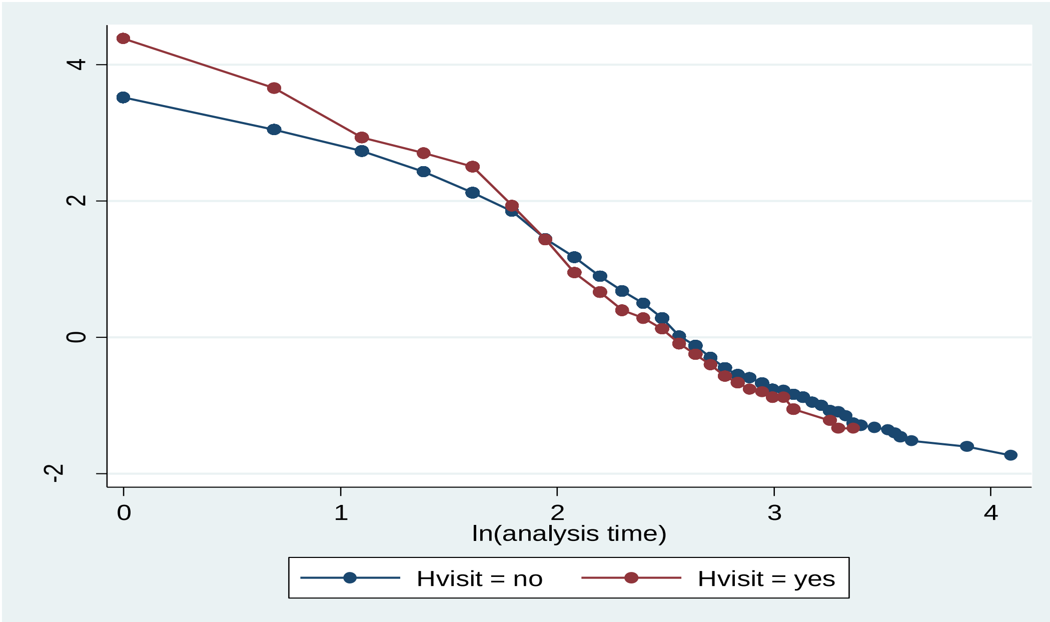
. stphplot, by(Newage)



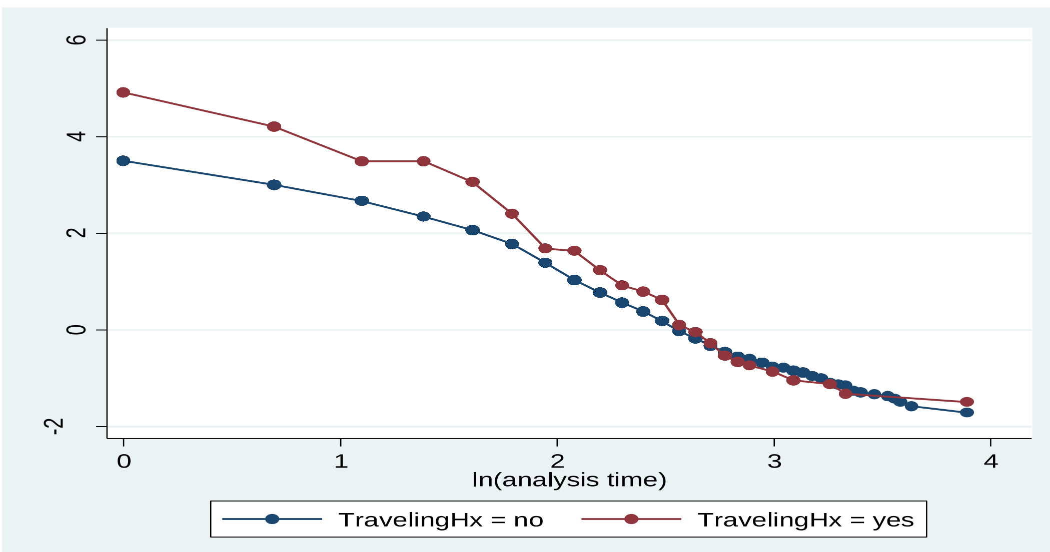
A) The plot of severity status to check the validity of the PH assumption



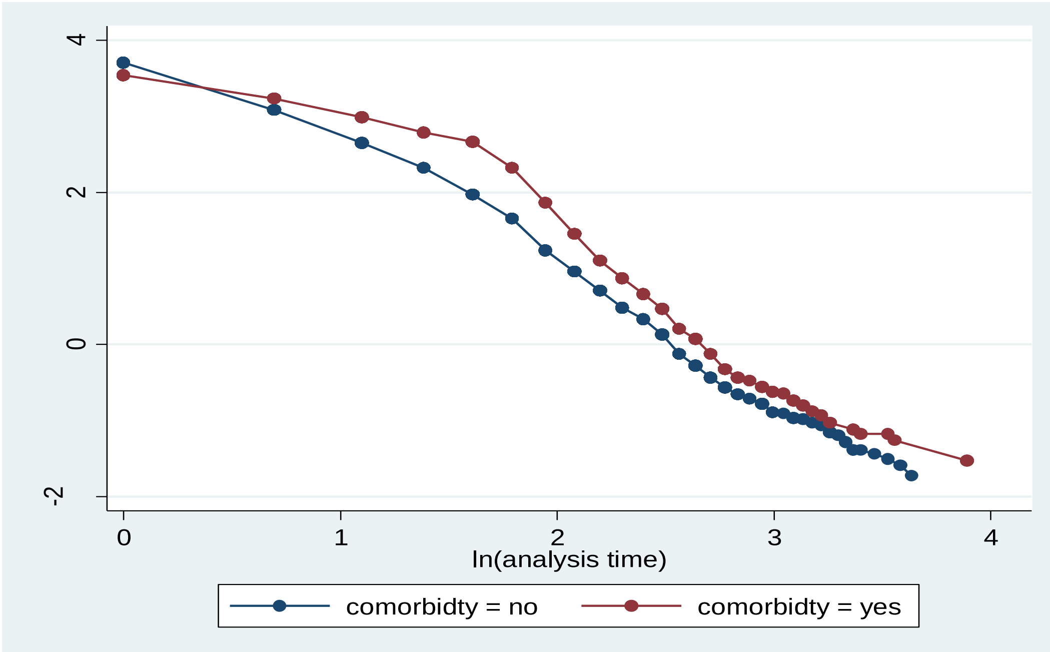
B) The plot of contact history with confirmed case to check the validity of the PH assumption



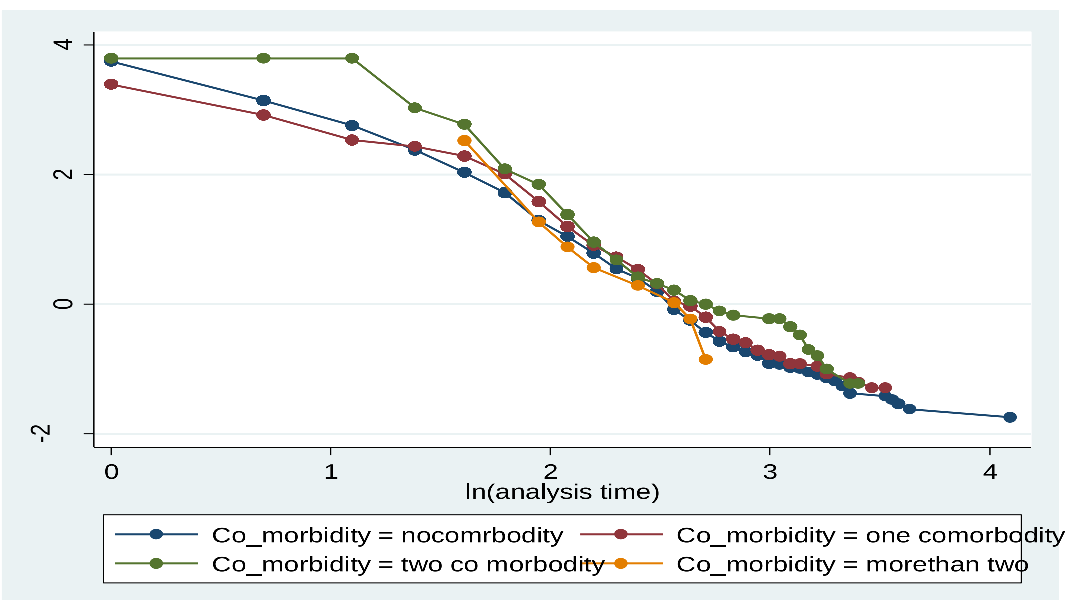
C) The plot of visiting Health facility before onset to check the validity of the PH assumption



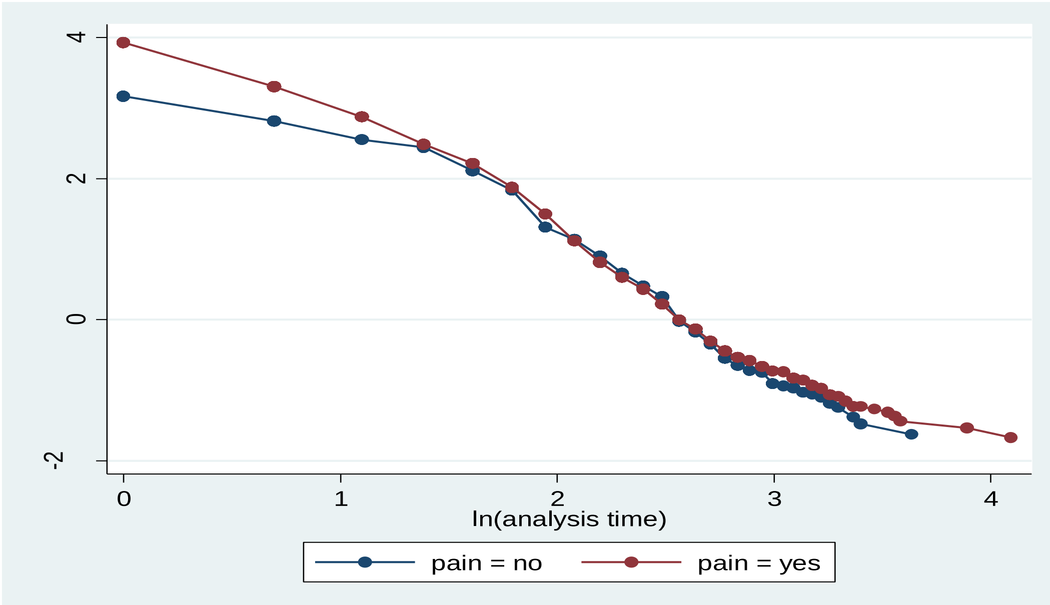
D) The plot of travailing history of last 14 days to check the validity of the PH assumption



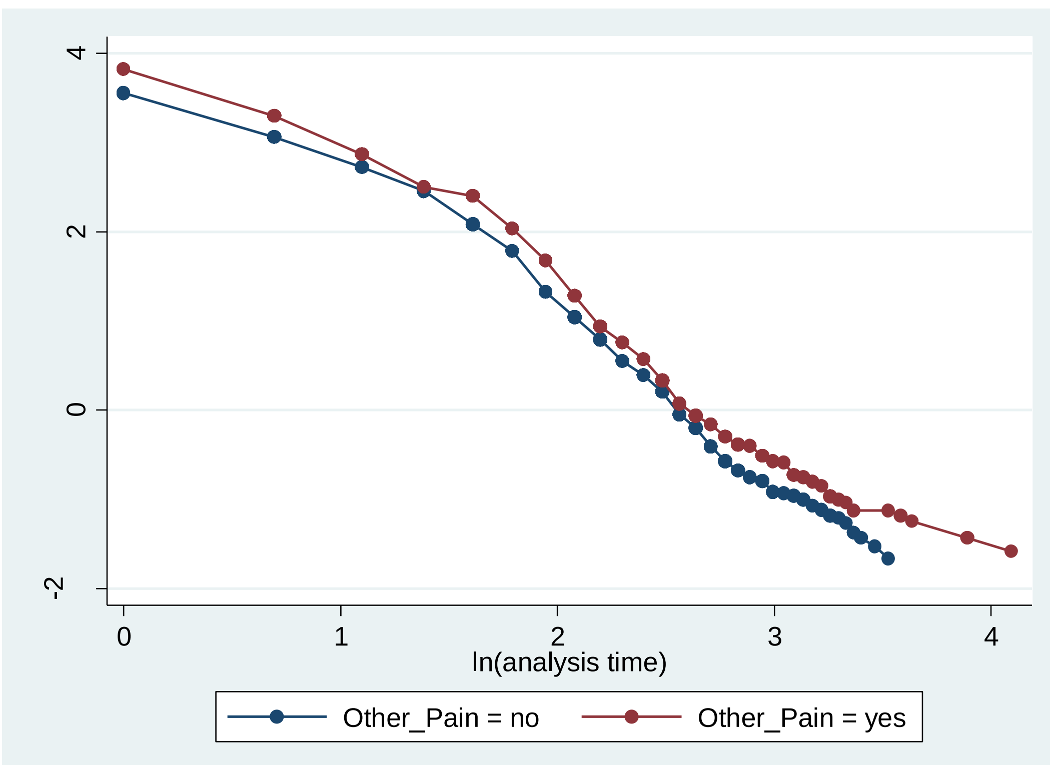
E) The plot of comorbidity to check the validity of the PH assumption



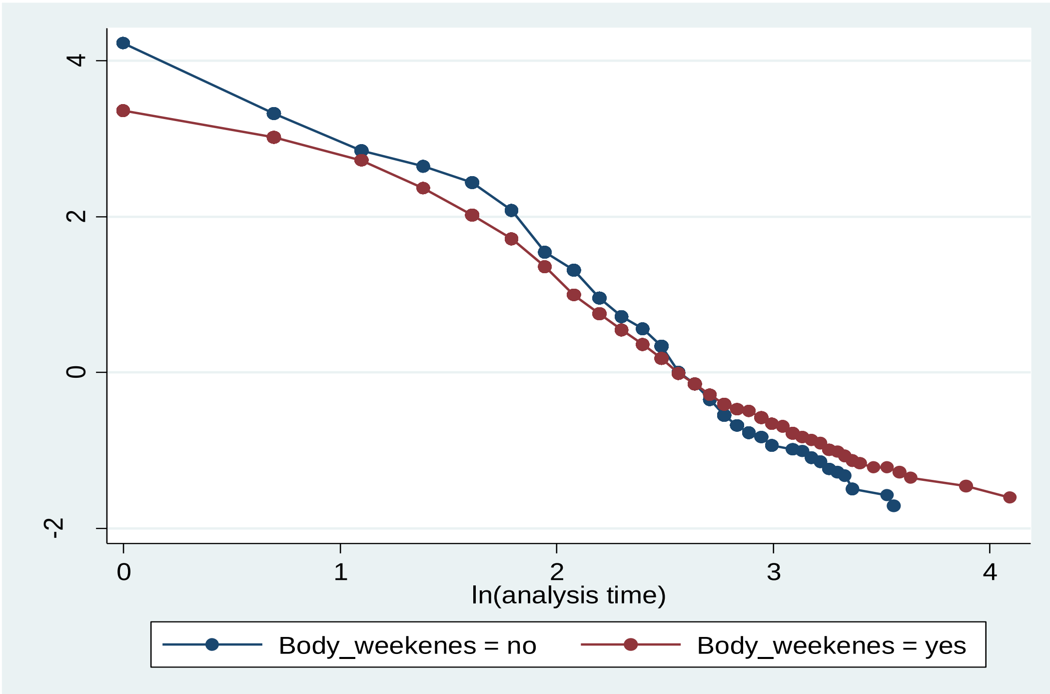
F) The plot of number of comorbidity to check the validity of the PH assumption



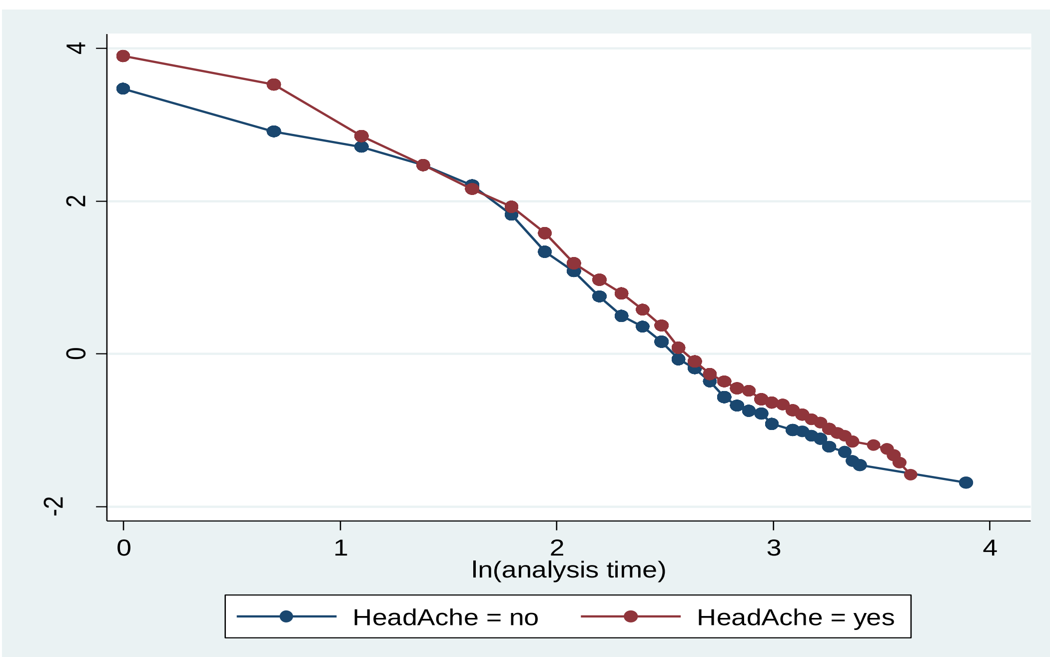
G) The plot of pain to check the validity of the PH assumption



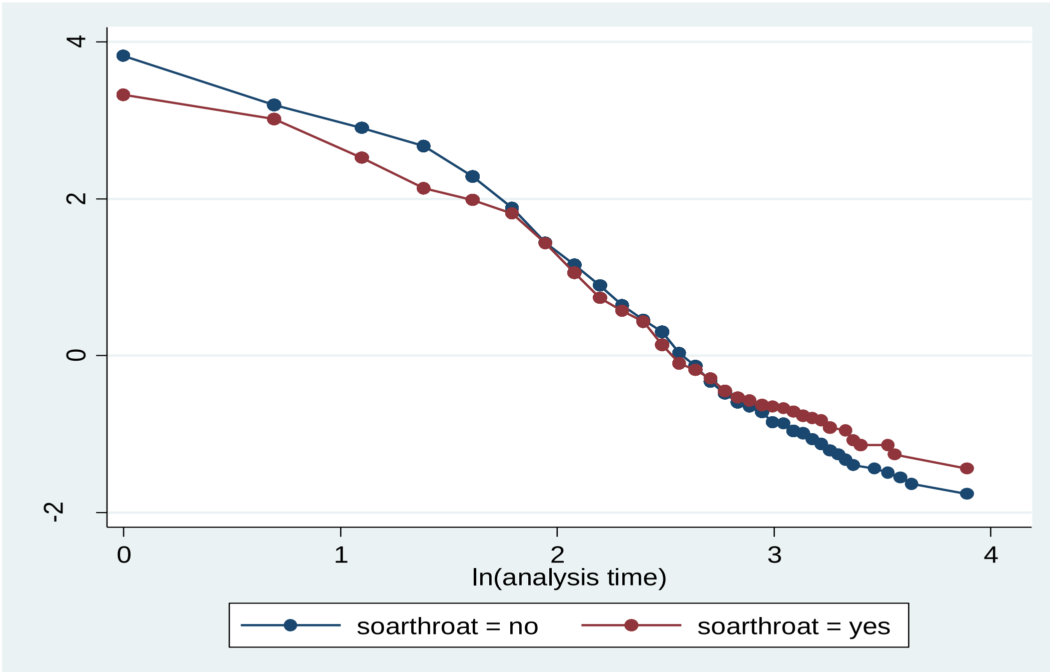
H) The plot of other pain out of covid-19 to check the validity of the PH assumption



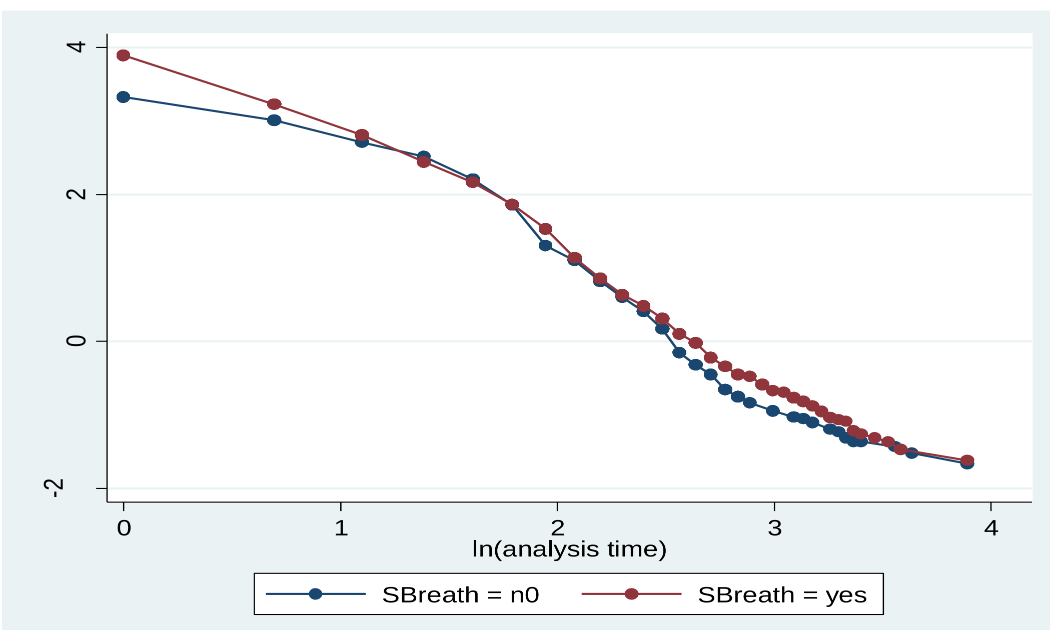
I) The plot of Generalized body weakens to check the validity of the PH assumption



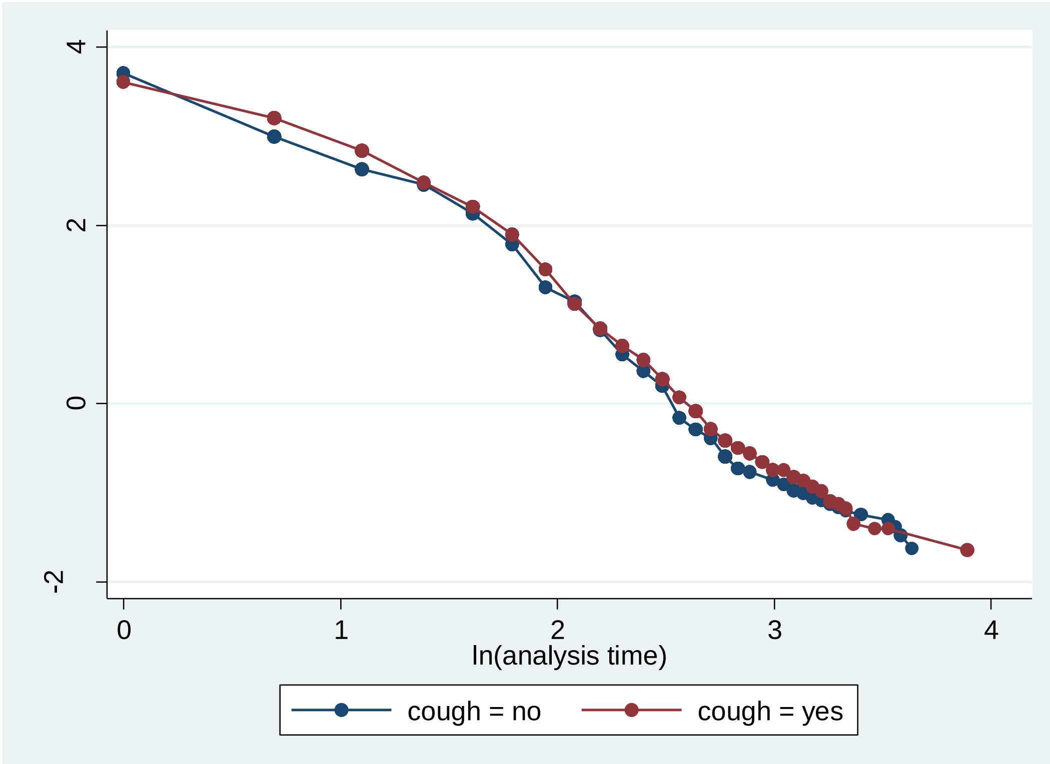
J) The plot of head ache to check the validity of the PH assumption



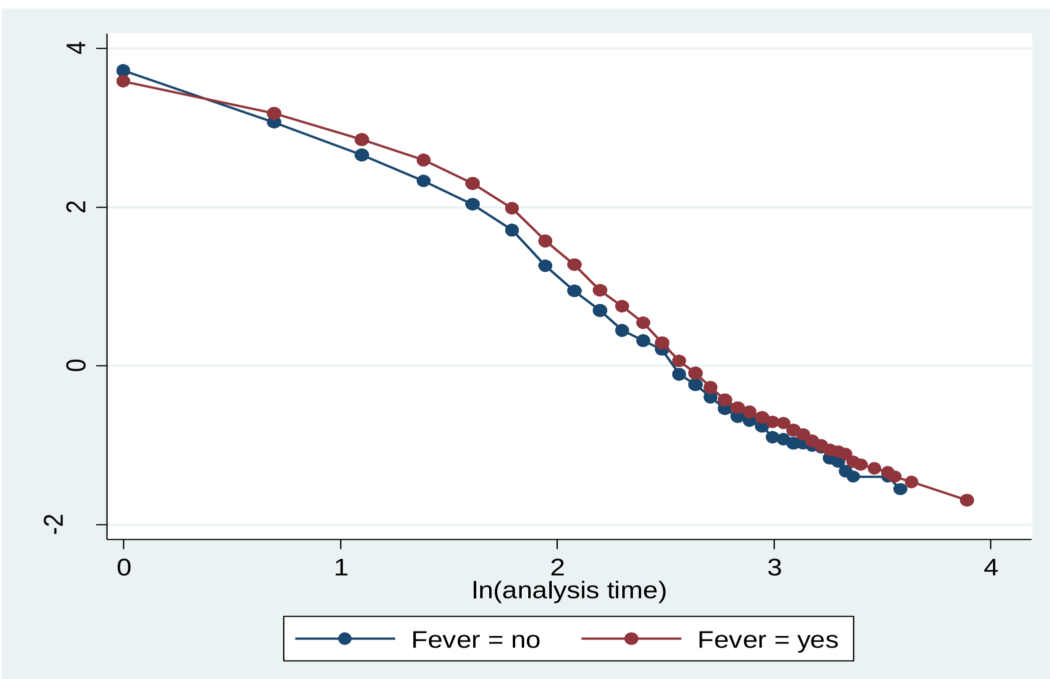
K) The plot of soarthroat to check the validity of the PH assumption



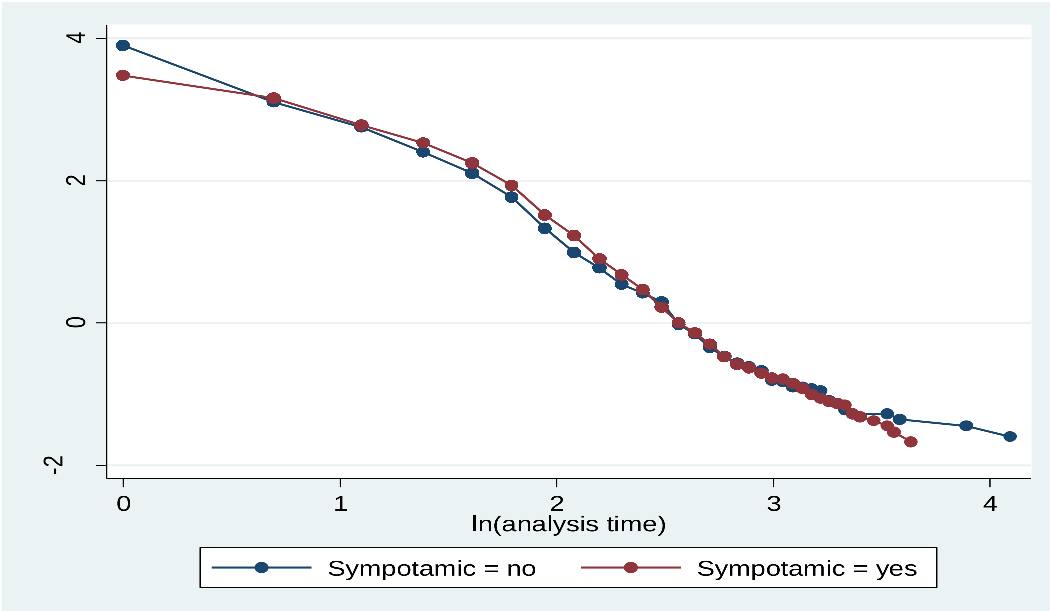
L) The plot of shortness of breath to check the validity of the PH assumption



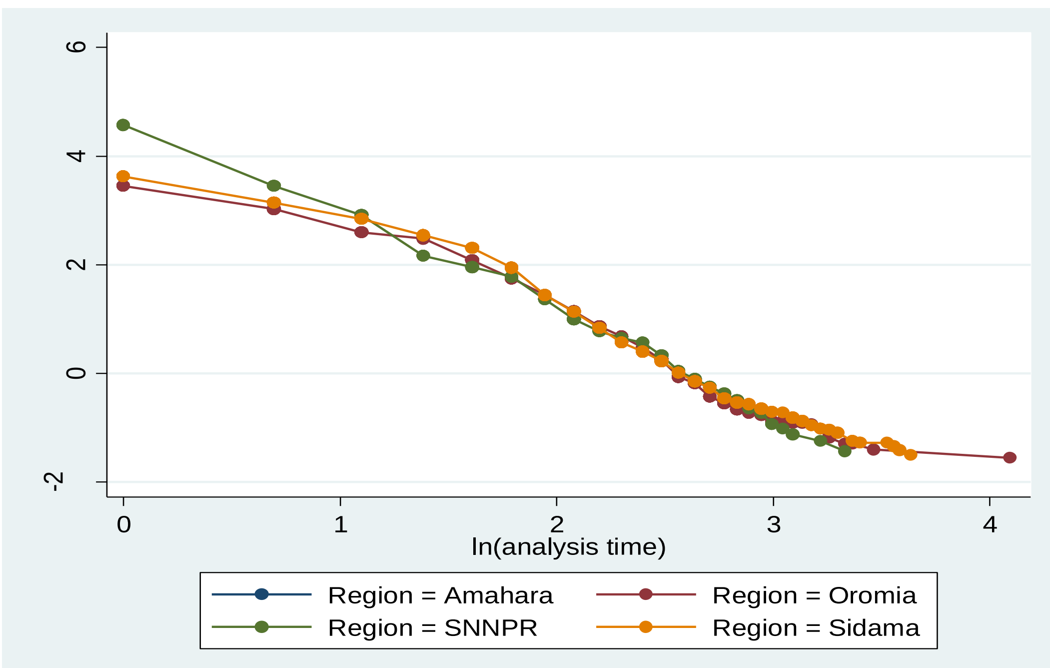
M) The plot of cough to check the validity of the PH assumption



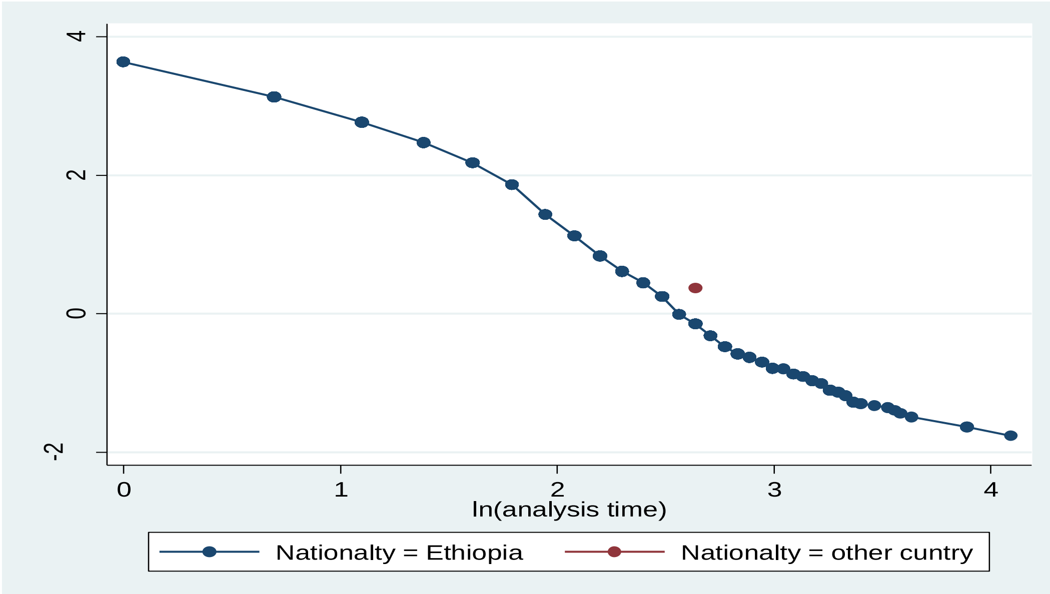
N) The plot of Fever to check the validity of the PH assumption



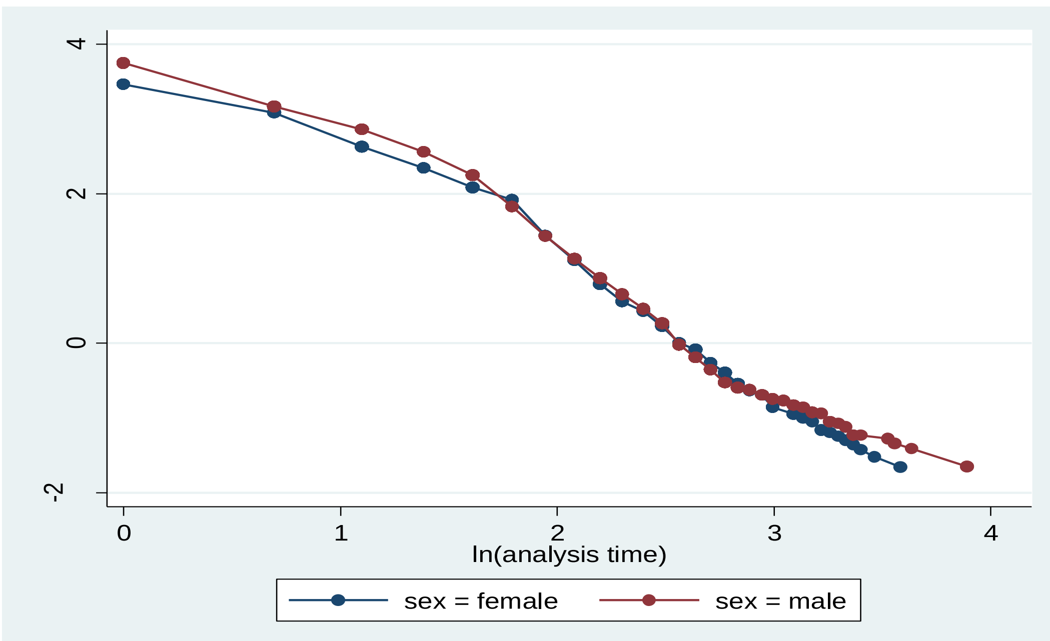
O) The plot of Sympotomic to check the validity of the PH assumption



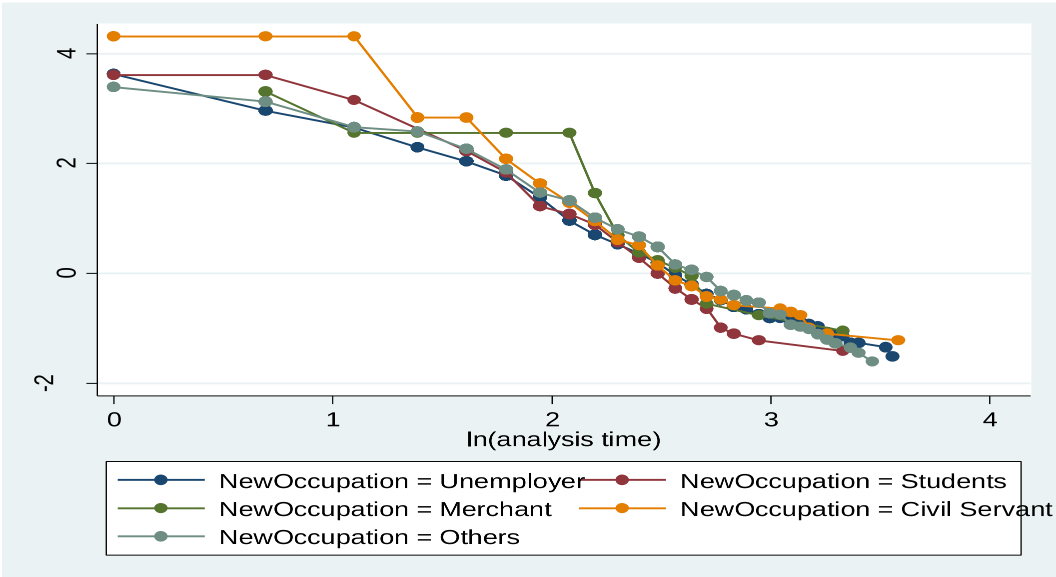
P) The plot of Region to check the validity of the PH assumption



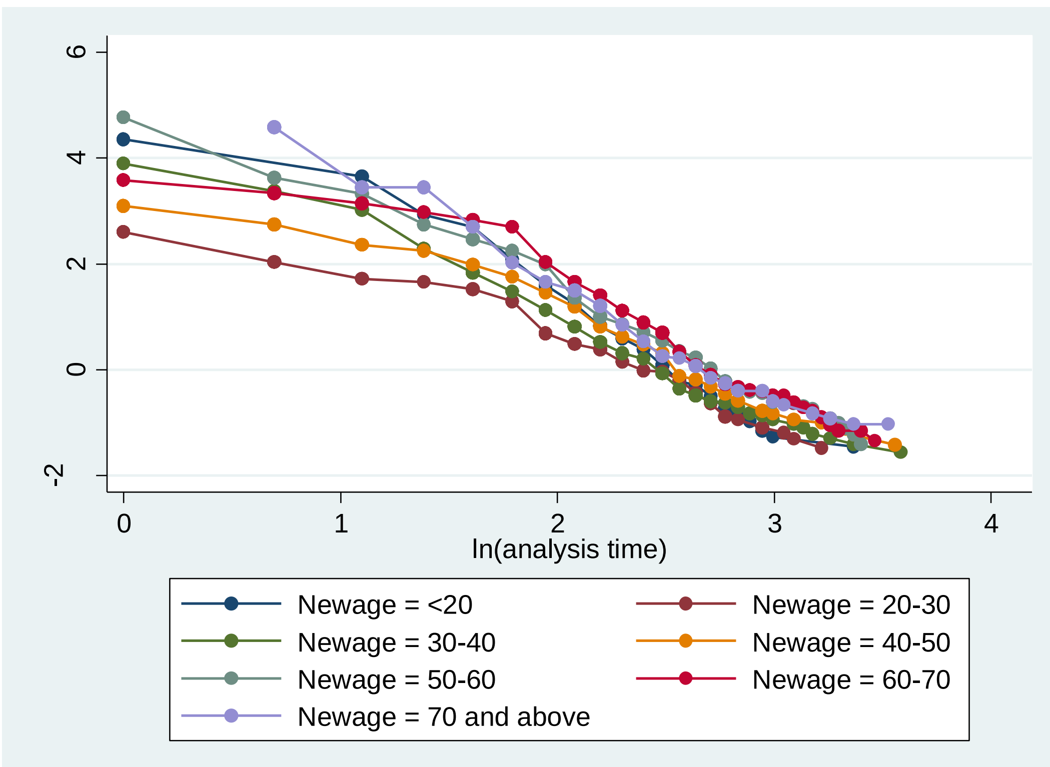
Q) The plot of national to check the validity of the PH assumption



R) The plot of Sex to check the validity of the PH assumption



S) The plot of occupation to check the validity of the PH assumption



T) The plot of age to check the validity of the PH assumption

Appendix 3 All Parametric Models

Exponential regression

. streg i.severity i.Hvisit i.TravelingHx i.comorbidty i.Co_morbidity i.Other_Pain i.HeadAche i.Newage, dist(exponential)

failure _d: Status == 1

analysis time _t: Time

Iteration 0: log likelihood = -853.07839

Iteration 1: log likelihood = -829.37205

Iteration 2: log likelihood = -828.77295

Iteration 3: log likelihood = -828.77244

Iteration 4: log likelihood = -828.77244

Exponential regression -- log relative-hazard form

No. of subjects = 810 Number of obs = 810

No. of failures = 637

Time at risk = 8818

LR chi2(19) = 48.61

Log likelihood = -828.77244 Prob > chi2 = 0.0002

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
severity |						
mild |	1.254536	.1505961	1.89	0.059	.9915254	1.587312
moderet |	1.171189	.1422395	1.30	0.193	.9231016	1.485951
sever |	.9538607	.1494051	-0.30	0.763	.7017131	1.296613
crtical |	.8494319	.1319952	-1.05	0.294	.6264096	1.151857
Assymptotic |	1.211676	.269793	0.86	0.389	.7831715	1.874631

Hvisit						
yes	1.274475	.1495734	2.07	0.039	1.01259	1.604091
TravelingHx						
yes	.7091488	.0946362	-2.58	0.010	.545939	.9211505
comorbidty						
yes	.7622502	.0999189	-2.07	0.038	.5895472	.9855452
Co_morbidity						
one comorbodity	1.199265	.1633605	1.33	0.182	.9182623	1.566258
two co morbodity	1.078195	.1917617	0.42	0.672	.7608647	1.527873
morethan two	1.303953	.4258434	0.81	0.416	.6875077	2.473127
Other_Pain						
yes	.8153776	.0742265	-2.24	0.025	.6821362	.974645
HeadAche						
yes	.8786044	.0785141	-1.45	0.148	.7374423	1.046788
Newage						
20-30	1.175284	.185823	1.02	0.307	.8621052	1.602232
30-40	1.052453	.1541989	0.35	0.727	.7897503	1.402541
40-50	.9025514	.1435239	-0.64	0.519	.6608665	1.232623
50-60	.8352352	.133711	-1.12	0.261	.6102977	1.143078
60-70	.7673424	.1209615	-1.68	0.093	.563389	1.045129
70 and above	.7388687	.13005	-1.72	0.086	.5232935	1.043252
_cons	.0866523	.0133998	-15.82	0.000	.0639959	.1173299

Gompertz regression

. streg i.severity i.Hvisit i.TravelingHx i.comorbidty i.Co_morbidity i.Other_Pain i.HeadAche i.Newage, dist(gompertz)

failure _d: Status == 1

analysis time _t: Time

Fitting constant-only model:

Iteration 0: log likelihood = -853.07576

Iteration 1: log likelihood = -780.42818

Iteration 2: log likelihood = -776.43075

Iteration 3: log likelihood = -776.41004

Iteration 4: log likelihood = -776.41004

Fitting full model:

Iteration 0: log likelihood = -776.41004

Iteration 1: log likelihood = -719.412

Iteration 2: log likelihood = -715.98128

Iteration 3: log likelihood = -715.97043

Iteration 4: log likelihood = -715.97043

Gompertz regression -- log relative-hazard form

No. of subjects = 810 Number of obs = 810

No. of failures = 637

Time at risk = 8818

LR chi2(19) = 120.88

Log likelihood = -715.97043 Prob > chi2 = 0.0000

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]	
severity						
mild	1.32801	.1621629	2.32	0.020	1.045349	1.687102
moderet	1.144569	.1410952	1.10	0.273	.8989	1.45738
sever	.8942612	.1434084	-0.70	0.486	.6530723	1.224525
crtical	.8251433	.1291736	-1.23	0.220	.6071226	1.121456
Assymptotic	1.323111	.2957425	1.25	0.210	.8537593	2.050489
Hvisit						
yes	1.618548	.1939306	4.02	0.000	1.279784	2.046984
TravelingHx						
yes	.5285635	.0729925	-4.62	0.000	.4032271	.6928586
comorbidty						
yes	.5532681	.0712357	-4.60	0.000	.4298726	.7120846
Co_morbidity						
one comorbodity	1.600322	.2106481	3.57	0.000	1.236417	2.071333
two co morbodity	1.468947	.2569582	2.20	0.028	1.042577	2.069684
morethan two	1.879481	.6133942	1.93	0.053	.9913722	3.563191
Other_Pain						
yes	.5892468	.0568388	-5.48	0.000	.4877422	.7118757
HeadAche						
yes	.8215838	.0744657	-2.17	0.030	.6878626	.9813006
Newage						
20-30	1.248515	.197708	1.40	0.161	.9153814	1.702885

30-40	1.011706	.1489374	0.08	0.937	.7581328	1.350091
40-50	.789883	.1267552	-1.47	0.142	.5767234	1.081827
50-60	.6839279	.1106698	-2.35	0.019	.4980503	.939177
60-70	.674155	.1071804	-2.48	0.013	.4936643	.9206357
70 and above	.6156684	.1104878	-2.70	0.007	.4331021	.8751923
_cons	.0558407	.0087579	-18.40	0.000	.041063	.0759365
/gamma	.0693458	.0041101	16.87	0.000	.0612902	.0774015

Loglogistic regression

. streg i.severity i.Hvisit i.TravelingHx i.comorbidty i.Co_morbidity i.Other_Pain i.HeadAche i.Newage, dist(loglogistic)

failure _d: Status == 1

analysis time _t: Time

Fitting constant-only model:

Iteration 0: log likelihood = -1070.1699

Iteration 1: log likelihood = -768.95254

Iteration 2: log likelihood = -723.75035

Iteration 3: log likelihood = -694.05425

Iteration 4: log likelihood = -693.80335

Iteration 5: log likelihood = -693.8032

Iteration 6: log likelihood = -693.8032

Fitting full model:

Iteration 0: log likelihood = -693.8032

Iteration 1: log likelihood = -660.01456

Iteration 2: log likelihood = -646.30972

Iteration 3: log likelihood = -646.29921

Iteration 4: log likelihood = -646.29921

Loglogistic regression -- accelerated failure-time form

No. of subjects = 810 Number of obs = 810

No. of failures = 637

Time at risk = 8818

LR chi2(19) = 95.01

Log likelihood = -646.29921 Prob > chi2 = 0.0000

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
severity						
mild	-.1679623	.064323	-2.61	0.009	-.294033	-.0418917
moderet	-.084011	.0655214	-1.28	0.200	-.2124306	.0444085
sever	.0530474	.0817384	0.65	0.516	-.107157	.2132517
crtical	.1080086	.082869	1.30	0.192	-.0544116	.2704289

Assymptotic	-.1394776	.1287071	-1.08	0.279	-.3917389	.1127837
Hvisit						
yes	-.2057168	.0608995	-3.38	0.001	-.3250776	-.086356
TravelingHx						
yes	.3295015	.0695028	4.74	0.000	.1932785	.4657246
comorbidty						
yes	.2472546	.0733398	3.37	0.001	.1035111	.390998
Co_morbidity						
one comorbodity	-.198475	.0767289	-2.59	0.010	-.348861	-.0480891
two co morbodity	-.1536604	.0954193	-1.61	0.107	-.3406789	.0333581
morethan two	-.2283374	.1640474	-1.39	0.164	-.5498644	.0931896
Other_Pain						
yes	.1316638	.049798	2.64	0.008	.0340616	.229266
HeadAche						
yes	.1232863	.0485566	2.54	0.011	.0281172	.2184555
Newage						
20-30	-.1930068	.0877908	-2.20	0.028	-.3650736	-.02094
30-40	-.1149158	.0788831	-1.46	0.145	-.2695239	.0396923
40-50	.0062271	.086761	0.07	0.943	-.1638214	.1762756
50-60	.1216003	.0852446	1.43	0.154	-.0454761	.2886766
60-70	.1550882	.0824998	1.88	0.060	-.0066085	.3167848
70 and above	.1026961	.0916819	1.12	0.263	-.0769971	.2823893
_cons	2.347631	.0815457	28.79	0.000	2.187804	2.507457

/ln_gam | -1.105638 .033422 -33.08 0.000 -1.171144 -1.040132

------------------+--

gamma | .3309996 .0110627 .3100121 .353408

Lognormal regression

. streg i.severity i.Hvisit i.TravelingHx i.comorbidty i.Co_morbidity i.Other_Pain i.HeadAche i.Newage, dist(lognormal)

failure _d: Status == 1

analysis time _t: Time

Fitting constant-only model:

Iteration 0: log likelihood = -871.41234

Iteration 1: log likelihood = -783.08257

Iteration 2: log likelihood = -741.42804

Iteration 3: log likelihood = -741.00468

Iteration 4: log likelihood = -741.00434

Iteration 5: log likelihood = -741.00434

Fitting full model:

Iteration 0: log likelihood = -741.00434

Iteration 1: log likelihood = -705.44461

Iteration 2: log likelihood = -691.84773

Iteration 3: log likelihood = -691.81058

Iteration 4: log likelihood = -691.81058

Lognormal regression -- accelerated failure-time form

No. of subjects = 810 Number of obs = 810

					No. of failures =	637
					Time at risk =	8818
					LR chi2(19) =	98.39
Log likelihood =	-691.81058				Prob > chi2 =	0.0000

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
severity							
mild	-.2107485	.0751895	-2.80	0.005	-.3581171	-.0633798
moderet	-.1547826	.0762194	-2.03	0.042	-.30417	-.0053953
sever	.0053591	.0957188	0.06	0.955	-.1822463	.1929645
crtical	.0796046	.0946909	0.84	0.401	-.1059861	.2651953
Assymptotic	-.2691216	.1402187	-1.92	0.055	-.5439453	.005702
Hvisit							
yes	-.2078416	.0724118	-2.87	0.004	-.3497662	-.065917
TravelingHx							
yes	.3847487	.0832028	4.62	0.000	.2216742	.5478232
comorbidty							
yes	.2538321	.0831807	3.05	0.002	.0908009	.4168634
Co_morbidity							
one comorbodity	-.2358298	.0869596	-2.71	0.007	-.4062675	-.065392
two co morbodity	-.1517392	.1095977	-1.38	0.166	-.3665469	.0630684
morethan two	-.161223	.2054602	-0.78	0.433	-.5639175	.2414716

Other_Pain						
yes	.1765298	.0576895	3.06	0.002	.0634605	.2895991
HeadAche						
yes	.1475292	.0560302	2.63	0.008	.037712	.2573463
Newage						
20-30	-.3062852	.1017555	-3.01	0.003	-.5057223	-.1068481
30-40	-.1280038	.0946349	-1.35	0.176	-.3134848	.0574771
40-50	-.0587057	.1020658	-0.58	0.565	-.258751	.1413395
50-60	.1151392	.1020178	1.13	0.259	-.0848121	.3150904
60-70	.1311497	.0995992	1.32	0.188	-.0640612	.3263606
70 and above	.137409	.108862	1.26	0.207	-.0759567	.3507747
_cons	2.343812	.0977606	23.98	0.000	2.152205	2.535419
/ln_sig	-.4241944	.0277201	-15.30	0.000	-.4785249	-.3698639
sigma	.6542967	.0181372			.6196968	.6908283

Weibull regression -- accelerated failure-time form

. . streg i.severity i.Hvisit i.TravelingHx i.comorbidty i.Co_morbidity i.Other_Pain i.HeadAche i.Newage, dist(weibull)

Fitting constant-only model:

Iteration 3: log likelihood = -694.06302

Fitting full model:

Iteration 4: log likelihood = -632.34112

Weibull regression -- log relative-hazard form

LR chi2(19) = 123.44

Log likelihood = -632.34112 Prob > chi2 = 0.0000

Table:4.7 Parameter estimates, standard errors and the hazard ratios in the final Weibull regression model (at Hawass Referral Hospital, during 2013-2014)

	Coffi.	Std. Err.	z	P>|z|	Haz. Ratio	95% CI for HR
severity						
mild	.3264644	.1689908	2.68	0.007	1.386059	1.091444 1.760199
moderet	.1339984	.140327	1.09	0.275	1.143391	.898934 1.454326
sever	-.0723883	.1471774	-0.46	0.647	.9301696	.6821495 1.268366
crtical	-.2463373	.1226847	-1.57	0.117	.7816585	.5746684 1.063205
Assymptotic	.3054374	.3036688	1.37	0.172	1.357219	.8753847 2.104266
Hvisit						
yes	.5082829	.2014198	4.20	0.000	1.662434	1.311032 2.108025
TravelingHx						
yes	-.640271	.0728463	-4.63	0.000	.5271496	.402075 .6911314
comorbidty						
yes	-.5681494	.0732555	-4.39	0.000	.566573	.4397431 .7299829
Co_morbidity						
one comorbodity	.4268915	.2035902	3.21	0.001	1.532486	1.181177 1.988284
two co morbodity	.3273931	.2450887	1.85	0.064	1.387347	.9813224 1.961365
morethan two	.6436124	.62844	1.95	0.051	1.903344	.9964841 3.635501
Other_Pain						
yes	-.4675198	.0579452	-5.06	0.000	.6265543	.5226821 .751069
HeadAche						
yes	-.2346949	.0717008	-2.59	0.010	.7908121	.6620597 .9446033

Newage						
20-30	.2592005	.2055328	1.63	0.102	1.295894	.9496555 1.768368
30-40	.0522361	.1551102	0.35	0.723	1.053624	.7895425 1.406035
40-50	-.220926	.1280067	-1.38	0.166	.801776	.5863479 1.096354
50-60	-.3494256	.1134637	-2.17	0.030	.705093	.5143643 .9665446
60-70	-.4207173	.1046997	-2.64	0.008	.6565757	.4803408 .8974705
70 and above	-.3756402	.1222438	-2.11	0.035	.6868494	.4845815 .9735456
_cons	-4.947399	.0015296	-22.97	0.000	.0071019	.0046563 .0108319
/ln_p |	.6930414	.0291055	23.81	0.000	.6930414	.6359956 .7500871
p |	1.999788	.0582049			1.999788	1.888902 2.117184
1/p |	.5000529	.0145543			.5000529	.4723254 .5294081