College of Natural Sciences

Department of Statistics

Modeling Time to Recovery of Diabetic Patients from COVID-19: A Case Study at Eka Kotebe General Hospital COVID-19 isolation and treatment center, Addis Ababa Ethiopia

---

By: Kubie Diriba

A Thesis Submitted to the Department of Statistics, College of Natural Sciences, Jimma University as a Partial Fulfillment for the Requirements of Master of Sciences (MSc) Degree in Biostatistics

August, 2022

Jimma, Ethiopia

Jimma University

College of Natural Sciences

Department of Statistics

Modeling Time to Recovery of Diabetic Patients from COVID-19: A Case Study at Eka Kotebe General Hospital, Addis Ababa, Ethiopia

By: Kubie Diriba

Advisor: Tadele Akeba Diriba (PhD)

Co-Advisor: Jaleta Abdisa Fufa (MSc)

August, 2022

Jimma, Ethiopia

STATEMENT of AUTHORS

As author of this research study, I declare that the thesis is a result of my genuine work, support of my supervisors and help hands of other individuals. Thus, all those who had participated in the study and sources of the materials used for writing this thesis have been duly acknowledged. I have submitted this thesis to Jimma University as a partial fulfillment for the requirements of Degree of Master of Science in Biostatistics. The library directorate of Jimma University can deposit the copy of the thesis in the university library so that students and researchers can refer it. Moreover, I declare that I have not so far submitted this thesis to any other institution anywhere for that award of any academic degree, diploma or certificate and or to get prove of society's problems. Any brief quotations from this thesis are allowed without requiring special permission if an accurate acknowledgment and citation (after publication) of the source is made. In all other instances, however, permission must be obtained from the author.

Kubie Diriba Adugna

Date_____

Signature_____

Jimma, Ethiopia

DEPARTMENT OF STATISTICS, SCHOOL OF GRADUATE STUDIES JIMMA UNIVERSITY

As thesis research advisors, we here by certify that we have read the thesis prepared by Kubie Diriba under our guidance, which is entitled **Modeling Time to Recovery of Diabetic Patients from COVID-19: A Case Study at Eka Kotebe General Hospital, Addis Ababa, Ethiopia**, in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials, including tables and figures, are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready As members of the board of examiners of the M.Sc. thesis open defense examination, we certify

Tadele Akeba Diriba (PhD)		23/09/2022
Advisor	Signature	Date
Jaleta Abdisa Fufa (MSc)		23/09/2022
Co-advisor	Signature	Date

that we have read and evaluated the thesis and examined the candidate. Hence, we recommend that the thesis be accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

Abiy Disasa (MSc)	For	23/09/2022
Name of Internal-examiner	Signature	Date
Kasahun Takele (PhD)		26/09/2022
Name of External-examiner	Signature	Date
Yasin Negash (MSc)		26/09/2022
Name of Chairman	Signature	Date

DEDICATION

This thesis is dedicated to all of my family especially to my mother Kibnesh Bedhedhe and my aunt Aster Hundie for making me who I am today, for their support and for teaching me the value of education.

AKNOWLEDGMENT

I wish to express my sincere appreciation to my advisor, Tadele Akeba Diriba (PhD) who has the substance of a genius: for his convincingly guided, kind support, advice, and constructive comments; and for encouraging me to be professional and do the right thing in my thesis.

Likewise, I would like to pay my special regards to my co-advisor, Mr. Jaleta Abdisa Fufa (MSc), for his guidance, advice, and kind support from proposal construction to the final thesis. It is a great pleasure to express his commitment to share his experience and knowledge, and his smooth contacts is also appreciable.

I heartily thank Ambo University for helping me with the financial support for this study and Jimma University for all their facilitations, by giving me an advisor, co-advisor, and instructors.

I gratefully thank all of my parents, whose love, trust, encouragement, comfort, and support always surround me. Finally, I would like to acknowledge the Eka Kotebe General Hospital, COVID-19 Isolation and Treatment Center staff to undertake this study with their cooperation and permission in using the data with special thanks for Dr. Abebaw Bekele (MO) and Dr. Tadios Niguss (Mph) for their willingness to help me during the data collection period. Also I want to express a sincere acknowledgement to Ms. Tirhas Hagos, and Dr. Mikyas Teferi for those unflinching encouragement, advice and guidance starting from data collection.

ABSTRACT

Background: Corona virus 2019 (COVID-19) is a pandemic disease which is caused by SARS-Cov-2 and it emerged on December 31, 2019, in China. COVID-19 affects more patients with chronic diseases mainly diabetes. Those with diabetes were more likely to have serious complications and have delayed recovery time from the virus. One reason is that high blood sugar weakens the immune system and makes it less able to fight off infections. The main aim of this study was to model the time to recovery of diabetic patients from COVID-19 in Eka Kotebe General Hospital.

Methods: A retrospective cohort study design was conducted on diabetic patients with COVID-19 whose age 18 years and above from March 20, 2020, to April 30, 2022 G.C, at Eka Kotebe General Hospital (EKGH), Addis Ababa, Ethiopia. Kaplan-Meier estimation method and log-rank tests, were used to compare the survival experience of different groups and also Cox proportional hazard model was employed to identify the covariates that have a statistically significant effect on the recovery time of diabetic patients from COVID-19 with the help of R software's (R version 4.1.2) to analyze the data.

Results: Out of 481 diabetic patients with COVID-19, 306 (63.62%) were recovered from COVID-19, with a minimum and maximum recovery times of 5 and 59 days, respectively. The median recovery time was 16 days with 95% confidence interval (15, 17). The multivariate Cox regression model analysis showed that patients whose age groups 36-55 ($\hat{HR}$ = 0.585, 95% CI: 0.404, 0.848), age groups > 55 ($\hat{HR}$ = 0.663, 95% CI: 0.445, 0.989), female patients ($\hat{HR}$ = 1.382, 95% CI: 1.061, 1.8), rural resident (($\hat{HR}$ = 4.839, 95% CI: 2.82, 8.247), HIV ($\hat{HR}$ = 0.067, 95% CI: 0.024, 0.182), hypertension ($\hat{HR}$ = 0.38, 95% CI: 0.255, 0.567), symptom ($\hat{HR}$ = 0.514, 95% CI: 0.3, 0.881), asthma ($\hat{HR}$ = 0.491, 95% CI: 0.309, 0.78), stroke ($\hat{HR}$ = 0.508, 95% CI: 0.333, 0.776), TB ($\hat{HR}$ = 0.424, 95% CI: 0.277, 0.65), CLD ($\hat{HR}$ = 0.564, 95% CI: 0.356, 0.893), types of diabetes (type 2 diabetes ($\hat{HR}$ = 0.159, 95% CI: 0.107, 0.236) & Gestational diabetes ($\hat{HR}$ = 41.875, 95% CI: 9.279, 188.976)), and other co-factors ($\hat{HR}$ = 0.531, 95% CI: 0.398, 0.708) were statistically associated with time to recovery of diabetic patients from COVID-19.

Conclusion & Recommendation: Finally, the findings of this study implied that factors like, age, sex, residence, HIV, hypertension, asthma, TB, stroke, CLD, and other co-factors were a major factors related to time to recovery of diabetic patients from COVID-19. Based on study results, it is recommended that health professionals should be give more attention to diabetic patients with HIV, hypertension, TB, stroke, asthma, CLD, types of DM, and other co-factors to control COVID-19.

Key Words: Diabetes Mellitus, COVID-19, Kaplan-Meier estimator, Cox model, Recovery time

Contents

List of Tables

List of Figures

1 Introduction

1.1 Background of the Study

The coronavirus disease 2019 (COVID-19) pandemic is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which has resulted in millions of morbidities and mortalities worldwide [1–3]. COVID-19 has drawn a lot of attention and poses a healthcare threat globally. The devastating effects of COVID-19 pandemic were highly associated with chronic medical conditions, mainly diabetes mellitus (DM). DM is one of the most frequent comorbidities and has been a challenge for patients worldwide since the emergence of coronavirus disease. It includes different types, namely, type I and type II, which are the most prevalent subtypes, and the other is gestational. Type II DM usually occurs in middle-aged, older people and is independently associated with an increased risk of hospitalization, admission to intensive care, and death from COVID-19. Gestational diabetes occurs in some women during pregnancy and it usually goes away after pregnancy. People with gestational diabetes are at a higher risk of developing type II diabetes later in their life [4, 5]. DM is a chronic, life-long condition that affects both survival and the body's ability to use the energy found in food [6], and it occurs when the pancreas does not produce enough insulin or when the body does not use the insulin effectively [7].

The most common symptoms of DM with COVID-19 are like shortness of breath or difficult to breathing, cough, fever, running nose, abdominal pain, and headache, among others [8, 9]. Corona viruses are mainly transmitted from person to person through physical contact and respiratory droplets [10, 11]. When an infected person sneezes, talks, and coughs, droplets or tiny particles called aerosols carry the virus into the air, and this makes the transmission easy among the people [12, 13]. Individuals with diabetes create a viral disease, and it can be harder to treat due to vacillations in blood glucose levels and the nearness of diabetes complications. Recent studies have focused on patients with TII DM infected with COVID-19, and their findings result shows patients with TII DM hospitalized for COVID-19 are at an increased risk of mortality, longer hospital stays, and ICU admission [14].

Globally, 615,621,881 total confirmed cases of COVID-19, 594,740,021 total recoveries and 6,523,543 total deaths up to June 30, 2022 were reported to WHO [15]. The estimated global prevalence of diabetes in 2021 was 10.5% (537 million), a 16% (74 million) rise since 2019. According to the International Diabetes Federation (IDF), the number of diabetic patients will rise to 11.3% (643 million)

by 2030 and 12.2% (783 million) by 2045 [16]. The prevalence of diabetes among COVID-19 patients has been estimated up to 31% [17]. Diabetes mellitus is a well-known risk factor for worse clinical outcomes in patients with coronavirus disease and lethal outcomes such as intensive care unit (ICU) admission, invasive ventilation, and death independent of other comorbidities [18, 19]. Furthermore, the presence of typical complications of diabetes mellitus (cardiovascular disease, heart failure, hypertension, and chronic kidney disease) increases COVID-19 mortality and delays the recovery time from COVID-19 [18, 20]. It has been observed that 1.5% of in-hospital deaths related to COVID-19 occurred in patients with type I DM and 31.4% in patients with type II DM [21]. Gestational diabetes mellitus is one of the most frequent pregnancy complications, with a global prevalence of 13.4% in 2021.

According to International Diabetes Federation (IDF) predictions, the prevalence of diabetes is expected to rise by 156 percent in Africa, 16 percent in Europe, 35 percent in North America and the Caribbean, and 84 percent in Southeast Asia by 2045 [22]. The WHO analysis of 14 African countries, which provided information on COVID-19 and comorbidities, showed that the risk of complications or death from COVID-19 among people with diabetes increases with age and up to 18% (1 in 5)of COVID-19 deaths in the African region are among people with diabetes [23].

In Ethiopia, the first confirmed case of COVID-19 was announced on the 13th of March 2020, after two and a half months of the outbreak in China. Up to 30^{th} of June 2022, 486,831 confirmed total and 7,527 total deaths due to COVID-19, which have been reported to the WHO by the federal ministry of health of Ethiopia. According to a study conducted on COVID-19 Knowledge, Attitudes, and Prevention Practices Among People with Hypertension and Diabetes Mellitus in Ambo, Ethiopia, Concerning attitude, 79.2% strongly believed that DM and hypertension patients were more at risk of death because of COVID-19 [24].

A case-control study conducted on the assessment of hypertension and other factors associated with the severity of disease in COVID-19 pneumonia, in Eka Kotebe General Hospital, Addis Ababa, Ethiopia, factors like hypertension, diabetes, chronic cardiac, and lung diseases were the most frequent comorbidities, with 19.6%, 11.3%, 4.2%, and 4.2%, respectively [25].

From a methodological point of view, a study was conducted recently on the mortality of COVID-19 among diabetic patients in Addis Ababa, Ethiopia using a binary logistic regression model [26]. The statistical method employed by Migbar *et al* [26] does not incorporate the survival time of patients,

namely, the time to event of interest and censoring. Survival analysis is the most appropriate statistical method for data analysis where the outcome variable of interest is the time to the occurrence of an event. The methodological developments in survival analysis that have the most profound impact are the Kaplan-Meier method for estimating the survival function, the log-rank test for comparing the equality of two or more survival distributions, and the Cox proportional hazards (PH) model for examining the covariate effects on the hazard function. The Cox proportional hazard model is widely used for analyzing survival data in the presence of covariates or prognostic factors due to its simplicity and for not being based on any assumptions about the survival distribution [27–29].

For instance, a retrospective cohort study was conducted by Abrahim *et al* [30] on time to recovery and its predictors among adults hospitalized with COVID-19 by using Kaplan Meier for estimating the survival function and Cox proportional hazards for analyzing survival data in the presence of covariates that affect the time to recovery of patients from COVID-19. The findings showed that study participants without comorbidities recovered more quickly than those with at least one comorbidity. Similarly, a retrospective cohort study was conducted on 263 adult COVID-19 patients. Cox proportional hazard regression model was fitted to determine factors associated with recovery time. In this case, older age, presence of fever on admission, and comorbidity were found to have a statistically significant association with recovery time [31].

There has been considerable policy and programmatic progress over the past year to reduce and eradicate the COVID-19 disease worldwide. As the occurrences of infectious diseases like COVID-19 mostly affect people with underlying medical conditions, studying the survival time of patients and identifying factors affecting the event of interest is crucial. Therefore, the focus of this thesis is to model the time to recovery of diabetic patients from COVID-19 in the case of Eka Kotebe General Hospital. Eka Kotebe General Hospital is one of the first designated to manage COVID-19 cases in Ethiopia.

This thesis is presented in five chapters, including this introduction chapter up to the concluding chapter. Chapter 1 highlights major issues relating to diabetics with COVID-19 at a global level and in Ethiopia. The objectives, significance, and limitations of the study are also described in the introduction. Chapter 2 contains both theoretical and empirical literature reviews relating to factors associated with the DM/COVID-19 co-infected. Chapter 3 briefly describes the methodological issues of the study, and Chapter 4 gives the results and discussions of the study. Finally, the conclusions and recommendations of the study were presented in Chapter 5.

1.2 Statement of the problem

COVID-19 is a global public health problem causing high mortality worldwide, and the risk of dying due to COVID-19 disease was higher among patients with chronic diseases like diabetes [32, 33]. Diabetes is one of the leading causes of morbidity and it causes enormous health and financial burdens worldwide [34]. The available evidence demonstrates that diabetes predisposes people to developing infectious diseases, and patients with diabetes are at greater risk of infection-related mortality [35, 36]. Wang *et al* [37] and Guan *et al* [38] reported that patients with diabetes accounted for 10.1% and 7.4% of COVID-19 patients, respectively. Recent publications showed that 20-30% of nonsurviving COVID-19 patients had underlying diabetes [39, 40]. This evidence indicates that COVID-19 patients with diabetes might be at a higher risk of mortality.

The previous study was conducted to determine prognostic covariates in diabetics with COVID-19 by using binary logistic regression [26]. However, binary logistic regression is not capable of considering the survival time of the patients in the hospital since it does not account for the censoring of observations and does not hold for time-to-event data. Many studies haven't been done on the recovery time of diabetic patients from COVID-19. Some of the previous studies in this area considered the prevalence of diabetes in COVID-19, to explore the prevalence of diabetes in the COVID-19 pandemic [41, 42]. This study, therefore, has tried to fill these methodological and knowledge gaps in understanding the status of diabetes with COVID-19 patients by identifying the risk factors for the recovery time of diabetic patients from COVID-19. Additionally, there was no research done on the time to recovery of diabetic patients from COVID-19 at Eka Kotebe General Hospital. But this study was conducted on time to recovery of diabetic patients from COVID-19 at Eka Kotebe General Hospital to fulfill this geometrical gap.

Generally, this study has attempted to answer the following basic research questions:

- What is the estimated median recovery time & recovery rate of diabetic patients from COVID-19?
- Which type of diabetes Mellitus is a faster recovery time from COVID-19?
- Which factors affect the time to recovery of diabetic patients from COVID-19?

1.3 Objective of the Study

1.3.1 General objective

The general objective of the study is to model the time to recovery of diabetic patients from COVID-19 in case of Eka Kotebe General Hospital.

1.3.2 Specific objective

The specific objectives are:

- To estimate the median recovery time and recovery rate of diabetic patients from COVID-19;
- To identify which types of diabetes Mellitus have faster recovery time from COVID-19; and
- To identify the factor which affects the time to recovery of diabetic patients from COVID-19

1.4 Significance of the study

This study identified and discussed the factors that influence the recovery time of diabetic patients from COVID-19 at Eka Kotebe General Hospital. This will help for health professional to focus on the risk factors for COVID-19 contracted patients with diabetic cases in their care follows. Solved disease risk factors after treatment have their own benefit for stakeholders and families to save themselves financially and from the disease. And it provides information to the government and non-government and also to other concerned bodies in setting policies, strategies, and further investigation for increasing the recovery of diabetic patients from COVID-19. The findings can also be used as a key input for any future interventions aimed at increasing the recovery rate of diabetic patients from COVID-19. Even in the future, infectious diseases like these pandemic disease can cause serious damage to other chronic diseases, and this research will be useful for people to beware of them. Additionally, this study would be used as a reference for other researchers to further analyze by using other an appropriate methods or models for approving or modifying this result.

1.5 Limitation of the study

This study also has its own limitations related to retrospective study design and the use of secondary data, which might have incomplete information. Consequently, essential variables that couldn't be recorded in patients' medical records and had the potential to affect the recovery time of diabetic

patients from COVID-19 infection were excluded. Due to this, the number of participants is minimized, and some necessary information may be lost. Moreover, there was a lack of literature in our country related to the subject under study.

2 Literature review

2.1 Overview of Diabetes with COVID-19

Coronavirus-2019 (COVID-19) is an infection that is caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and was first reported last December in Wuhan, China. The World Health Organization (WHO) declared it a global pandemic on March 11, 2020 [43, 44]. It infected more than 5 million people in the United States, and the death rate continued to rise, reaching more than 219,000 at the time of publication. According to early research, roughly 25 percent of people admitted to hospitals with severe COVID-19 infections had diabetes. DM is a collection of metabolic illnesses defined by a persistently high blood sugar level. Because they aggravate inflammation and modify immune system responses, coronavirus infections have been shown to significantly affect the management of diabetes mellitus. This causes problems with glycaemic control. Patients with diabetes mellitus are more likely to develop cardiorespiratory failure from SARS-CoV-2 infection than patients without diabetes mellitus [45].

More than 1.2 million deaths and 48.1 million cases of SARS-CoV-2 infections have been reported globally as of the 5th of November 2020 [46]. Diabetes has been identified as a risk factor for a poor prognosis among people with COVID-19, along with other concurrent medical disorders (such as underlying CVD, respiratory illnesses, hypertension, and obesity) [47–49]. According to a number of systematic reviews and meta-analyses on the subject, people with diabetes have a two-to three-fold higher risk of dying from COVID-19 than people without diabetes [50, 51].

According to the World Health Organization (WHO) 2014 country profile, about 30% of total deaths in Ethiopia were associated with non-communicable diseases (NCDs), of which cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes are the leading causes of morbidity and mortality. Similarly, the report revealed disproportionate age-specific death rates with a significant rise in deaths from non-communicable diseases between the ages of 30 and 70 years [52]. Globally, the overall case fatality rates of COVID-19 vary between countries. For instance, 4.1% in China, 4.6% in Spain, 8.3% in Italy, 2.73% in Egypt, and 1.6% in Ethiopia [53]. The fatality rate of patients with COVID-19 was highest in people aged 80, ranging from 13% to 16.7%, followed by 7.2%–8.9% among those aged 70–79 years [54]. However, patients with NCDs are more likely to have severe disease and subsequent mortality. The COVID-19 pandemic has had widespread health impacts, revealing the

particular vulnerability of those with underlying conditions [55]. From the previous research in China, a total of 258 patients, 63 of whom were diabetic, used the Cox PH model. Diabetes is one prognostic factor for increased COVID-19 severity and a higher risk of death. The median time length of the patient from admission to discharge in the hospital was 12 days within a 95% confidence interval (7, 15) with a P value of 0.022 in China [56].

2.2 Literature related to the variables used in the study

Different studies conducted on the relationship between diabetic mellitus and COVID-19, determinants of diabetes with COVID-19 with different design and statistical method of analysis. A study conducted by Nafakhi *et al* [57] showed that out of 67 diabetic patients and 125 non-diabetic patients, 58% and 70% of diabetic and non-diabetic patients were completely recovered from COVID-19, respectively. Among the factors they used, age, sex, hypertension, and clinical symptoms were statistically significant. Additionally, Belice *et al* [58] study results show diabetic men were at a higher risk of COVID-19 than women patients.

Antonio *et al* [59] investigated cardiovascular risk management issues in people with diabetes in the COVID-19 era, and their findings revealed that diabetes, along with other classic cardiovascular (CV) risk factors such as hypertension, obesity, and smoking, was associated with COVID-19. Other meta-analysis results showed that the most prevalent cardiovascular metabolic comorbidities with COVID-19 were hypertension, cardio-cerebrovascular disease, and followed by diabetes [60]. Furthermore, the results of a cross-sectional study on COVID-19 co-infection with type 2 diabetes, hypertension, and obesity showed the risk of mortality was higher in patients with diabetes and hypertension [61]. Evidence from meta-analysis suggested that hypertension was independently associated with a significantly increased risk of critical COVID-19 and in-hospital mortality from COVID-19 [62].

Shibru *et al* [63] studied the effect of COVID-19 on poor treatment control among ambulatory diabetes and/or hypertension patients in Northwest Ethiopia from December 2020 to February 2021 using a multivariable binary logistic regression mixed model to identify the determinants of poor treatment control. The study covered 836 patients (410 with diabetes and 526 with hypertension). The man diagnosis was diabetes in 410 (49%) patients. The median age of the study participants was 52 years, with an Inter Quartile Range (IQR) of 18 (43-61). According to their descriptive results, nearly two-thirds of the study participants, 543 (65%), were urban, which implies that urban diabetes and/or hypertension patients are more affected by the COVID-19 pandemic disease.

A retrospective cohort study was conducted on a total of 42 COVID-19 patients with type 2 diabetes mellitus (TII DM) with ages ranging from 41 to 66 years. The result of this study showed that stroke is one of the risk factors for diabetes with COVID-19 diseases [64]. Furthermore, a study conducted by Nannoni et al [65] results showed that factors like hypertension, diabetes, strokes, coronary artery disease, and older age were statistically associated with COVID-19.

Similarly, a retrospective cohort study was conducted on severe COVID-19 in people with TI DM and TII DM in Sweden. The study covers 44,639 and 411,976 adult patients with TI DM and TII DM diabetes alive on Jan 1, 2020, using Cox regression analyses. According to Cox regression analysis, variables like age, sex, stroke, asthma, and hypertension were statistically significant at a 5% level of statistical significance [66].

Migbar *et al* [26] studied on determinants of mortality among COVID-19 patients with diabetes mellitus in Addis Ababa, Ethiopia, they used a binary logistic regression model. The study covered a total of 340 COVID-19 patients (114 with diabetes and 226 non-diabetes). In these findings factors, like age, severity of COVID-19 disease, obesity, hypertension, anemia at presentation, RVI and AKI after hospital admission were a statistically association with diabetic patients in COVID-19 infection.

A cross-sectional study conducted on diabetes with COVID-19 psychosocial consequences of the COVID-19 pandemic in people with diabetes by distributing online questionnaires to 2430 adult-age (>18 years old) members of two user panels at Steno Diabetes Center Copenhagen and the Danish Diabetes Association, respectively, by using multivariate logistic regression. In this study, sociodemographic and health status items like age, gender, and type of diabetes were included, and the result showed women more often than men experienced worries about being overly affected by COVID-19 and types of diabetes, but variable age has no association with COVID-19 co-infection in DM patients [67]. Luzi *et al* [68] conducted on telemedicine and urban diabetes during the COVID-19 pandemic in Milano, Italy. And the result of their findings showed that rural residents are more likely to be older and to have health conditions including cardiovascular disease, chronic lung disease (CLD), diabetes, and obesity that exacerbate the effects of COVID-19 [69].

A cross-sectional study was conducted on the psychological impact of COVID-19 on diabetes mellitus patients in Cape Coast, Ghana.The results showed older people are at greater risk of developing diabetes with COVID-19 [70]. Similarly, in the study of Elemam *et al* [71] the result showed COVID-19 patients with diabetes who belong to an older age group are more affected by COVID-19. Moreover,

retrospective cohort study results showed that factors like age and gender were statistically associated with increased risk of adverse outcomes in diabetic patients with COVID-19. In another study, the results showed that aged patients had a slower recovery rate from COVID-19 infection as compared to young respondents [72]. A preliminary South African data shows that having HIV or active tuberculosis (TB) increases a person's risk of dying from COVID-19. However, the effect is small compared with other known risk factors such as old age and diabetes [73].

A case study conducted on the recovery of an elderly woman with diabetes and asthma from COVID-19 infection shows that a 61-year-old woman was hospitalized due to dyspnea and low blood O_2 saturation and was later diagnosed with COVID-19, as confirmed by PCR. The patient also had a history of asthma and diabetes. However, all these COVID-19 complications were successfully managed, and she was discharged with a good clinical condition after the improvement of respiratory complications. Finally, they conclude that despite having an advanced age and underlying diseases (diabetes and asthma), the recovery of this elderly woman shows that even such patients can defeat COVID-19 as long as the disease is not progressed to advanced phases and appropriate therapeutic measures are taken [74].

Similarly, a case-control study [26] conducted on 114 cases (diabetic patients) and 226 controls (non-diabetic patients) that died due to COVID-19 disease and variables like age, Sex, Obesity, Hypertension, Symptom, and RVI were statistically significant. Additionally, a retrospective cross-sectional result showed patients with diabetes were more likely to have a longer length of stay (LOS) in hospital, with 14.4 (SD $\pm$ 9.6) days, compared to patients without diabetes, 9.8 (SD $\pm$ 17.1) days. In this study male patients were more likely to be admitted to hospitals with COVID-19 than females, and older COVID-19 patients with diabetes were less likely to survive compared to younger [75].

A retrospective study conducted on the clinical characteristics of 382 confirmed COVID-19 patients, consisting of 108 with and 274 without preexisting TII DM, from January 8 to March 7, 2020, in Tianyou Hospital in Wuhan, China, was collected and analyzed. In this study, univariate and multivariate Cox regression models were performed to identify specific clinical factors associated with the mortality of COVID-19 patients with TII DM. The results showed that older age, elevated glucose level, increased serum amyloid A (SAA), diabetes treatment with only oral diabetes medication, and oral medication plus insulin were independent prognostic factors [76].

A meta-analysis study was conducted on a total of 1558 patients with coronavirus 2019 (COVID-19). According to their results, factors like hypertension, diabetes, chronic lung disease (CLD), cardiovascular disease, and cerebrovascular disease were independent risk factors associated with COVID-19 patients [77]. Additionally, a retrospective cohort study was conducted on a total of 42 COVID-19 patients with type 2 diabetes mellitus (TII DM) who presented with acute ischemic stroke (AIS). In this study, variables like age, sex, smoking and other cofactors like IHD and dyslipidemia were independent risk factors associated with diabetes in COVID-19 patients [64].

2.3 Survival Analysis

Survival analysis is the analysis of time-to-event data. Such data describe the length of time from a time origin to an endpoint of interest [78]. It can be used solely for investigations of mortality and morbidity based on vital registration statistics. The most punctual arithmetical analysis of human survival processes can be traced back to the 17th century when the English statistician John Graunt published the first life table Glass *et al* [79]. For assessing conditional survival functions, non-parametric estimators can be favored to parametric and semi-parametric estimators due to loose presumptions that empower estimation [80]. The Kaplan-Meier estimator, which appears to have been first proposed by Bohmer, may be a non-parametric estimator [81]. It is utilized to assess the survival conveyance work from censored information. However, it was misplaced by subsequent analysts and was not investigated until the immunocompetent could become forceful and profoundly harmful in patients with an obstructed safe framework, for example, those with hematological malignancies, immunosuppression, ineffectively controlled diabetes, press overburden, and critical injury.

Several methods can be used for survival analysis. These include the Kaplan-Meier method, which is an estimator of survival probabilities [82] and the Cox Proportional Hazard Model (CPHM). [83]. These two methods are considered among the methods that contributed significantly to the development of survival analysis. Most of the studies were conducted to model the survival time and to predict the mortality risk for COVID-19. Guillermo-Salinas Escudero *et al* [84] applied survival analysis to study the effect of COVID-19 in Mexican. The factors they used in their study included age, sex, comorbidities, hospitalization, and admission to the intensive care unit. They applied the Kaplan-Meier and Cox regression models to assess the proportion of survival time and the random-effects model to find the hazard ratio of prognostic variables. Their results show that men and older people had higher mortality rates than women and young people, respectively. Monira Mollazehi *et al* [85] modeled

survival time to recover from COVID-19. Additionally, a Tolossa *et al* [31] study was conducted on time to recovery from COVID-19 and its predictors among patients admitted to the treatment center of Welloga University referral hospital, and a Cox proportional hazard regression model was used to determine factors associated with recovery time. PH modeling is the most frequently used type of survival analysis modeling in many research areas, having been applied to topics such as smoking relapse and employee turnover, Sterensund [86], and in medical areas for the identification of important covariates that have a significant impact on the response of the interested variables.

Another study was conducted by Ketema [87] at the Armed Forces General Teaching Hospital (AFGTH) located in Addis Ababa, Ethiopia. The Kaplan-Meier method and log-rank test were used to compare the survival experiences of different categories of patients. The proportional hazards Cox regression model was employed to identify predictors of mortality. A total of 734 patients with ART were included in the study. Predictors like CD4 count, ambulatory, bedridden, & co-infection were significant at 5%.

Generally, the previous study was conducted to determine prognostic covariates in diabetics with COVID-19 by using binary logistic regression [26]. However, binary logistic regression is not capable of considering the survival time of the patients in the hospital since it does not account for the censoring of observations and does not hold for time-to-event data. By considering this gap using survival analysis is an appropriate Model which is incorporate the time to event data and capable of censoring.

3 Methodology

3.1 Description of the study area

The data set for this study was obtained from Eka Kotebe General Hospital (EKGH). EKGH is located in the Yeka sub-city, Woreda 12, around Karalo. The center was the first hospital designated to manage positive COVID-19 cases in Ethiopia. It has a capacity of 600 beds, with 16 beds dedicated to intensive care services. The hospital has given treatment to 6706 COVID-19 patients until this study was conducted. Among the total patients in the hospital, 5428 recovered. Over 130 nurses, 90 general practitioners, two internists, one anesthesiologist, two emergency physicians, one pulmonology and critical care sub-specialist, two obstetricians and gynecologists, two surgeons, two psychiatrists, two radiologists, and two pediatricians were involved in the care of those patients.

3.2 Study Design and Target Population

This study is a retrospective cohort study of diabetic patients with COVID-19 who were treated at Eka Kotebe General Hospital. The data was extracted from the patient's chart. The target population of this study was all diabetic patients with COVID-19, whose age is 18 years and above, who were admitted to Eka Kotebe General Hospital from March 20, 2020, to April 30, 2022, G.C.

3.3 Data Collection Procedure

The secondary data was collected from the patient's individual charts for investigation of the time to recovery of diabetic patients from COVID-19. A checklist of variables was prepared and used to collect relevant data by reviewing patients' cards. One health professional and two experienced data collectors under the supervision of the researcher contributed to the data collection. The overall activity was controlled by the principal investigator. The data was coded and analyzed by using the statistical packages for social science (SPSS version 26) and R software (version 4.1.2), respectively.

3.4 Study period

The study period for this study was from March 20, 2020, to April 30, 2022, G.C. Entry of the data was considered from the day that patients started the treatment after admission up to the day they were discharged from the hospital.

3.5 Inclusion and exclusion criteria

All 18-year-old and above diabetic patients positive for COVID-19 with full information were included in the study, whereas diabetic patients negative for COVID-19 and with insufficient information regarding study variables on the registration book or in the card were not included in the study.

3.6 Variable description

3.6.1 Dependent variable

The response variable for this study is time to recovery of diabetic patients from COVID-19, which is measured in days. The status variable is coded as 0 for censored and 1 for the event (recovery).

Time to recovery:- means the time (in day) when diabetic patients were diagnosed positive for COVID-19 to the time (in day) when the diabetic patients are diagnosed with two consecutive negative results of the virus by the rRT-PCR tests in 24 hours.

Censored:- can occur if diabetic patients with COVID-19 have not recovered from COVID-19 up to the end of the study; death due to different causes, even death due to COVID-19; and if they are transferred to other hospitals.

3.6.2 Independent variable

The candidate covariates used in the research and the codes of categorical variables are described in the following (Table3.1).

Table 3.1: Description of covariates together with their values/codes

Covariate	Categories
Age	0 = 18-35, 1 = 36-55 , 2 = >55
Sex	0 = Male, 1= Female
Residence	0 = Urban, 1 = Rural
HIV	0 = Negative, 1 = Positive
Hypertension	0 = No, 1 = Yes
CVD	0 = No, 1 = Yes
Symptoms	0 = No, 1 = Yes
Asthma	0 = No, 1 = Yes
RVI	0 = No, 1 = Yes
Stroke	0 = No, 1 = Yes
TB	0 = No, 1 = Yes
CLD	0 = No, 1 = Yes
Obesity	0 = No, 1 = Yes
Types of diabetes	0 = Type I, 1 = Type II, 2 = Gestational
Other co-factor	0 = No, 1 = Yes

HIV = human immunodeficiency virus, CVD = Cardiovascular disease, RVI = Respiratory viral infection, TB = Tuberculosis, CLD = Chronic lung disease, Other co-factor = Patients with additional comorbidity (like ischemic heart disease (IHD), heart failure (HF), chronic kidney disease (CKD), ...), and according to [88, 89] studies, age of patients were categorized into (18-35, 36-55, and >55).

3.7 Method for survival data analysis

Survival analysis is a statistical method for data analysis where the outcome variable of interest is the time to the occurrence of an event [90]. One of the most important differences between the outcome variables modeled via linear and logistic regression analyses and the time variable in the survival data is the fact that we may only observe the survival time partially. The variable time actually records two different things. For those subjects who experienced the event, it is the outcome variable of interest, the actual survival time. However, for subjects who were alive at the end of the study, for patients who were lost to follow-up, patients withdrawing from the study, and competing events (e.g., death due to some cause other than the cause of interest), time indicates the length of follow-up (which is a partial or incomplete observation of survival time). These incomplete observations are referred to as being censored. Most survival analyses consider a key analytical problem called censoring. In

essence, censorship occurs when we have some information about an individual's survival time but we do not know the exact survival time (Collett [91]). Survival time is said to be right-censored when it is recorded from its beginning to a defined time before its end time. This type of censoring is commonly recognized in survival analysis and was also considered in this study. If an individual develops an event of interest prior to the start of the study, the survival time is said to be left-censored; this is not common in survival studies. Survival time is said to be interval censored when it is only known that the event of interest occurs within an interval of time but the exact time of its occurrence is not known.

Survival data is generally described and modeled in terms of two related probabilities, namely survival and hazard (Hosmer and Lemeshow) [92]. The survivor function (also known as the survival probability) S(t) is the likelihood that a randomly selected subject will live longer than or equal to a given time. Hence, it gives the probability that an individual survives beyond a specified time. In addition, the survival time distribution is characterized by the survivorship function, probability density function, and hazard function.

Let T be a random variable associated with the survival times, t be the specified value of the random variable T and f(t) be the underlying probability density function of the survival time T. The cumulative distribution function $F(t) = Pr(T < t)$, which represents the probability that a subject selected at random will have a survival time less than some stated value t, is given as:

$$F(t) = P(T < t) = \int_0^t f(u)d_u, \;\; t \geq 0 \tag{1}$$

Using the above function; the survival function S (t) can be given as: $S(t)$ = Prob(individual(T) survives longer than time t). This is the probability of surviving at least as long as t, which is one minus the integral of f(t) up to time t, that is:

$$S(t) = p(T \geq t) = 1 - F(t) \tag{2}$$

From equations (2) and (3) the relationship between f (t) and S (t) can be derived as:

$$f(t) = \frac{d}{d_t}F(t) = \frac{d}{d_t}(1 - S(t)) = \frac{-d}{d_t}S(t) \tag{3}$$

Theoretically, the survivor function can be graphed as a smooth curve as t ranges from 0 to infinity [0,∞]. Survival function are non-increasing or decreasing function, at time t = 0, S (t) = S (0) =1; that

is, at the start of the study, since no one has experienced the event yet, the probability of surviving past time 0 is one, and at time $t \to \infty$, $S(t) = S(\infty) \to 0$; that is, theoretically, if the study period increased without limit, eventually nobody would survive, so the survivor curve must eventually converge to zero.

The hazard function is widely used to express the risk or hazard of experiencing an event (recovery) at some time t and is obtained from the probability that an individual will experience the event at time t, conditional on whether he or she has survived (censoring) to that time. That is, the function represents the instantaneous failure rate for an individual surviving until time t. The hazard function is a measure of the probability of failure during a very small interval, assuming that the individual has survived at the beginning of the interval. It is defined as:

$$h(t) = \frac{\lim_{dt \to 0} P(t < T \leq t + dt / T \geq t)}{dt}, t > 0$$

$$= \frac{1}{P(T \geq t)} \lim_{\Delta t \to 0} \frac{P(t \leq T < +\Delta t)}{\Delta t} \tag{4}$$

By applying the theory of conditional probability and the relationship in equation (4), the hazard function can be expressed in terms of the underlying probability density function and the survivor function becomes

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

The corresponding cumulative hazard function H(t) is defined as;

$$H(T) = \int_0^t h(u) d_u = -\ln S(t). \tag{5}$$

Then, S(t) = $\exp$(-H(t)) and f(t) = h(t)S(t)

3.7.1 Non-parametric methods for survival analysis

The estimation of the survival distribution provides estimates of descriptive statistics such as the median survival time. These methods are said to be Kaplan-Meier estimator of the survival function since they require no assumptions about the distribution of survival time. Non-parametric methods are often very easy and simple to understand as compared to parametric methods. Furthermore, non-parametric analyses are more widely used in situations where there is doubt about the exact form of distribution.

A commonly reported summary statistic in survival studies is the median survival time. This is the time beyond which 50% of the individuals in the population under study are expected to survive and is given by that value t(50) which is such that $S(50) = 0.5$. When there is no possible realistic estimated survival time that makes the survival function exactly equal to 0.5, the estimated median survival time, t(50) defined to be the smallest observed survival time for which the value of the estimated survival function is less than 0.5 [29].

Mathematically,

$$\hat{t}(50) = min\left\{t_i / S(t_i) < 0.5\right\}, \tag{6}$$

where t_i is the observed survival time for i^{th} individual, i = 1, 2, ..., n. t_j is j^{th} ordered recovery time, j = 1,2,...,r.

Kaplan-Meier (KM) is the standard non-parametric estimator of the survival function , also known as the product-limit method. It incorporates information from all observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. The Kaplan-Meier estimator is used to estimate the survival time (time of censoring) of a patient and construct survival curves to compare the survival experience of a patient across different categorical variables [92].

The survival probability of developing the disease at any particular time is estimated as:

$$\hat{S}(t) = \prod_{t_i<t} \left(\tfrac{n_i - d_i}{n_i}\right) = \prod_{t_i<t} \left(1 - \hat{h_j}\right). \tag{7}$$

Where $\hat{S}(t)$ denotes the estimated time of developing diabetes with covid-19 at t_i , and n_i is the number of patients at risk and d_i is the number of individual who experienced to diabetes with covid-19, at time t_i.

The Kaplan-Meier estimator survival curve gives the estimate of survivor function among different strata or groups of covariates to make comparisons. Separate graphs of the estimates of the Kaplan-Meier survivor functions are constructed for different categorical covariates. In general, the pattern that one survivorship function lies below another means the group defined by the lower curve has better recovery time than the group defined by the upper curve.

The log rank test is a non-parametric test for comparing two or more independent survival curves. This

test is most powerful in detecting a higher cured proportion in one group than the other group [93]. Mathematically,

$$LR = \frac{[\sum_{i=1}^{m}(O_i - \hat{E}_{1i})]^2}{\sum_{i=1}^{m}\hat{V}_{Ei}} \sim X_1^2 \tag{8}$$

3.8 Regression Models for Survival Data

In most medical studies which give rise to survival data, supplementary information referred to as covariates or independent variables needs to be collected on each individual, so that the relationship between the survival experience of individuals and various explanatory variables has to be investigated.

In the analysis of survival data, interest centers on the risk of hazard of failure at any time after the time origin of the study. As a consequence, the hazard function is modeled directly in the survival analysis. There are two broad reasons for modeling survival data. One objective of the modeling process is to determine which combinations of potential explanatory variables affect the form of the hazard function. Another reason for modeling the hazard function is to obtain an estimate of the hazard function itself for an individual from a set of explanatory variables (Klein and Moeschberger [93]). One of the most popular types of regression models used in survival analysis is the Cox proportional hazard model (Cox, [94]).

3.8.1 The Cox Proportional Hazards Regression Model

The Cox Proportional Hazard (PH) Model is a multiple regression method and is used to evaluate the effect of multiple covariates on survival. Cox (1972) 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 specified as a function of time and the covariates:

$$h(t, X, \beta) = h_0(t) \exp(\beta' X),$$

or, equivalently:

$$h(t, X, \beta) = h_0(t) exp(\beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} = exp(\beta' x)) \quad i = 1, 2, ..., n, \tag{9}$$

where $h_0(t)$ denotes baseline hazard function, which is the hazard function for an individual for whom all the variables included in the model are zero, $X = (x_1, ..., x_p)$ is the values of the vector of explanatory variables for a particular individual, $\beta = (\beta_1, \beta_2, ..., \beta_p)$ is a vector of regression coefficients, and

$e^{\beta' x}$ characterizes how the hazard function changes as a function of subject covariates, and n is total number of observations in the study. If all of the covariates are zero the above equation (9) becomes $h_i(t) = h_0(t)$. For this reason the term $h_0(t)$ is called the baseline hazard function.

Assumptions of the Cox proportional hazards model are; i) the baseline hazard function $h_0(t)$ depends on t, but not on covariates, $x_1, x_2, ..., x_p$, ii) the hazard ratio $\exp\beta$, depends on the covariates $x_1, x_2, ..., x_p$ not on time (t), and iii) the covariates X_i are time-independent.

The survival function for Cox proportional hazard model is:

$$S(t, X) = S_0^{exp(\sum_{i=1}^{p} \beta_i X_i)}.$$

And the estimated survival function:

$$\hat{S}(t, X) = \left(\hat{S}_0(t)\right)^{exp(\sum_{i=1}^{p} \hat{\beta}_i X_i)},$$

where $\hat{S}_0(t)$, $\hat{\beta}_i$ are estimated baseline survival and coefficient of covariates respectively.

The measure of effect is called hazard ratio. The hazard ratio of two subjects with covariates x and x^* is given by

$$\hat{HR} = \frac{h_0(t)exp(\hat{\beta}'x)}{h_0(t)exp(\hat{\beta}'x^*)} = exp \sum \hat{\beta}'(x - x^*)$$

This hazard ratio is time independent, which is why this is called the proportional hazard model. The parameter of the Cox proportional hazard model refers to the hazard ratio of one group in comparison to the other groups for categorical covariates and change in hazard ratio with a unit change of the covariate for the continuous variable when other covariates are fixed.

So the change in hazard ratio for the continuous covariaties are given by:

$$\frac{h(t/x_k + 1)}{h(x_k)} = \frac{exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_k(x_k + 1) + ... + \beta_p x_p)}{exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + ... + \beta_p x_p)} = e_k^{\beta} \tag{10}$$

Which represents change (equivalently, exp(β_k)*100% percentage change) hazard function with unit change in covariate provided that other covariates remains fixed.

Interpreting outputs from the Cox model involves examining the coefficients for each explanatory variable. The negative regression coefficient for an explanatory variable indicates that the hazard is

lower and thus the prognosis worse. Conversely, a positive regression coefficient implies a better prognosis for patients with higher values of that variable when time to event is recovery and conversely for death.

3.9 Estimation of Parameters in proportional hazard model

The regression coefficients in the proportional hazards Cox model, which are the unknown parameters in the model, can be estimated using the method of maximum likelihood. In Cox proportional hazards model we can estimate the vector of parameters β without having any assumptions about the baseline hazard $h_0(t)$. Suppose the survival data based on n independent observations are denoted by the triplet $(t_i, \delta_i, x_i), i = 1, 2, ..., n$.

Where

t_i is the survival time for i^{th} individual

δ_i an indicator of censoring for the i^{th} individual given by 0 for censored and 1 for event/recovery.

x_i a vector of covariates for individual $i(x_{i1}, x_{i1}, ...x_{ip})$.

The full maximum likelihood is defined as the following.

$$L(\beta) = \prod_{i=1}^{n} h(t - i, x_i, \beta)^{\delta_i} S(t_i, x_i, \beta), \tag{11}$$

where $h(t_i, x_i, \beta_i) = h_0(t_i)e^{\beta' x_i}$ is the hazard function for individual i and $S(t_i, x_i, \beta) = (S_0(t_i)) \exp(\beta' x_i)$ is the survival function for individual i.

Then, the full maximum likelihood becomes;

$$L(\beta) = \prod_{i=1}^{n} (h_0(t_i) \exp(\beta' x_i))^{\delta_i} (S_0(t_i) \exp(\beta' x_i)) \tag{12}$$

Full maximum likelihood requires that we maximize (12) with respect to the unknown parameter of interest, β, and unspecified baseline hazard and survival functions. This indicates that unless we explicitly specify the baseline hazard, $h_0(t)$, we cannot obtain the maximum likelihood estimators for the full likelihood. But, Cox (1972) proposed using an expression called "partial likelihood function" that depends only on the parameter of interest.

3.9.1 Partial likelihood estimate for Cox PH model

Kalbfleisch and prentice derive a likelihood involving only β and X (not h_0) based on the marginal distribution of the ranks of the observed failure times (in the absence of censoring). Cox (1972), derived the same likelihood and generalized it for censoring using the idea of a partial likelihood. Instead of constructing a full likelihood, we consider the probability that an individual experiences an event at time t_i given that an event occurred at that time.

Let $\mathrm{R_i}$ denote the set of individuals at risk at time just prior to $t_{(i)}$. Assume that for the present case there is only one failure at time t_i, i.e, no ties. The probability that individual i with covariates x_i is the one who experience the event at time $t_{(i)} = \mathrm{P}\,(\,$ individual i has experiences an event at time $t_{(i)} \mid$ one event at time $t_{(i)})$.

$$\frac{\lambda\left(t, x_i\right)}{\sum_{jut_h)} \lambda\left(t, x_j\right)} \tag{13}$$

And under the proportional hazards assumption on using equation (13), the ratio

$$\frac{\lambda_o(t) \exp\left(\beta' x_i\right)}{\sum_{jaa_{i=1}} \lambda_o(t) \exp\left(\beta' x_j\right)} \tag{14}$$

shows the contribution to the partial likelihood at each death time $t_{(i)}$ by the individuals with covariate x_i in the risk set $\mathrm{R_{t(i)}}$. Where $\mathrm{R_{t(i)}}$ is the overall subjects in the risk set at time $t_{(i)}$. By eliminating the baseline hazards function, in the numerator and denominator, equation (14) becomes

$$\frac{\exp\left(\beta' x_i\right)}{\sum \exp\left(\beta' x_j\right)} \tag{15}$$

Thus, the partial likelihood is the product over all failure time $t_{(i)}$ for $i = 1, 2, \ldots, m$ of the conditional probability (15) to give the partial likelihood.

$$L_p(\beta) = \prod_{i=1}^{m} \frac{\exp\left(\beta' x_i\right)}{\sum_{j=R_{i(1)}} \exp\left(\beta' x_j\right)} \tag{16}$$

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

$$L_p(\beta) = \sum_{i=1}^{\mathrm{m}} \left[\beta' x_i - \ln\left(\sum_{ja\mathbb{R}_{i(j)}} \exp\left(\beta' x_j\right)\right)\right] \tag{17}$$

We obtain the maximum partial likelihood estimator by differentiating the right-hand side of (17) with respect to the component of β, setting the derivative equal to zero, and solving for the unknown parameters. The partial likelihood derived above is valid when there are no ties in the data set. But in most real situations, tied survival times are more likely to occur. In addition to the possibility of more than one event at a time, there might also be more than one censored observation at a time of event. To handle this real-world fact, partial likelihood algorithms have been adopted to handle ties. There are three approaches commonly used to estimate regression parameters when there are ties. These are Breslow (1974), Efron (1977) and Cox (1972) approximations (Collett, 2003). The most popular and easy approach is Breslow's approximation. In many applied settings there will be little or no practical difference among the estimators obtained from the three approximations. Because of this, and since the Breslow approximation is more commonly available, otherwise, analysis presented in this study was based on it. This approximation is proposed by Breslow and Peto to modify the partial likelihood and has the form;

$$L_B(\beta) = \prod_{i=1}^{m} \frac{\exp\left(\beta' X_i\right)}{\left[\sum_{leR_{l_{(i)}}} \exp(\beta' X_j)\right]^{d_i}}, \tag{18}$$

where d_i is the number of experienced an event occurred at time t_i and X_i the sum of covariates over d_i subjects at time t_i.

Then, the partial log likelihood of (18) is given as

$$l_B(\beta) = \sum_{i=1}^{m} \left[\beta' X_i - d_i \ln \sum_{leR_{((i)} } \exp\left(\beta' X_j\right)\right] \tag{19}$$

Breslow maximum partial likelihood estimator, adjusted for tied observation is obtained, by differentiating equation (19) with respect to the components of β and setting the derivative equal to zero and solving for the unknown parameters.

3.10 Model Development and Adequacy

In any applied setting, performing a proportional hazard regression analysis of survival data requires a number of critical decisions. It is likely that we will have data on more covariates than we can reasonably expect to include in the model, so we must decide on a method to select a subset of the total number of covariates. When selecting a subset of the covariates, we must consider such issues as clinical importance and statistical significance (Hosmer and Lemeshow, [95]).

When the number of variables is relatively large, it can be computationally expensive to fit all possible models. In this situation, automatic routines for variable selection that are available in many software packages might seem an attractive prospect. These routines are based on forward selection, backward elimination or the combination of the two known as the step wise procedure. The model selection strategy depends to some extent on the purpose of the study. In a situation where the aim is to identify variables upon which the hazard function depends, instead of using the automatic variable selection procedures, the following procedure is recommended.

1. The first step is fitting a univariable model for each of the explanatory variables and identifying the variables that are significant at some level from 20% to 25% is recommended in [96], 2. The variables that appear to be important in step 1 are then fitted together in a multivariable model. In the presence of certain variables, others may cease to be important. Consequently, backward elimination is used to omit non-significant variables from the model. Once a variable has been dropped, the effect of omitting each of the remaining variables should be examined in turn.

3.10.1 Proportional hazard assumption checking

The proportional hazards assumption is so important to cox-regression that we often include it in the name (the cox-proportional hazards model). What it essentially means is that the ratio of hazard functions for two individuals with different regression covariates does not vary with time. The most popular graphical technique for evaluating the PH assumption involves comparing estimated $-\ln(-\ln)$ survival curves over different (combinations) of categories of variables and scaled Schoenfeld residual being investigated.

The tests of the proportional hazards assumption for each covariate were done by correlating the corresponding set of scaled Schoenfeld residuals with a suitable transformation of time, with the default being based on the KM estimate of the survival function. If the plot of scaled Schoenfeld residuals versus the logarithm of time is a random, smooth, straight line about zero, the proportional hazards assumption will be satisfied. Otherwise, a plot of scaled Schoenfeld residuals for a given covariate may reveal a violation of the proportional hazards assumption [97]. Furthermore, if we plot the estimated log(-log(survival)) versus survival time for two or more groups, we would see parallel curves, indicating that the cox proportional assumption is met.

3.10.2 Model Diagnostic

The Cox-Snell residual is given by Cox and Snell. The Cox-Snell residual for the i^{th} individual with observed survival time t_i is defined as

$$r_{Ci} = \exp(\hat{\beta}' x_i)\hat{H}_0(ti) = \hat{H}_i(ti) = -\log \hat{S}_i(ti), \tag{20}$$

where $\hat{H}_0(ti)$ and $\hat{S}_i(ti)$ are an estimate of the baseline cumulative hazard function and survivor functions, respectively, for the i^{th} individual at the censored survival time, which was derived by Kalbfleisch and Prentice. Then the modified Cox-Snell residual is given by

$$r'_{Ci} = 1 - \delta_i + r_C$$

Plotting $-\log \hat{S}(T_i)$ vs t should yield a straight line and Plotting $\log\left[-\log\left(\hat{S}(T_i)\right)\right]$ vs $\log(\mathrm{t})$ should yield a straight line through the origin with slope $= 1$.

3.11 Ethical consideration

The data was obtained from Eka Kotebe General Hospital, and ethical clearance for the study was provided by the research ethics review board of Jimma University. An official letter of co-operation was written to Eka Kotebe General Hospital by the department of statistics. Qualified data collectors were carefully recruited and trained before the start of the data collection phase. As the study was conducted through the review of medical records, any personal information regarding the study subject was replaced by a number, and patient evidence was kept confidential without disclosing it to others. The recorded data is not accessed by a third person except the principal investigator and is kept confidentially.

4 Result and Discussion

4.1 Descriptive statistics

This study included 481 diabetic patients with COVID-19 whose age was 18 years and above and were placed under treatments that followed between March 20, 2,020, and April 30, 2022, in Eka Kotebe General Hospital. Descriptive statistics of baseline covariate for diabetic patients with COVID-19 were illustrated in (Table 4.1).

Among 481 patients eligible for the study, 292 (60.71%) were males and the remaining 189 (39.29%) were females. Female patients were faster to recover, with a median recovery time of 15 days within 95% C.I (14, 16) than male patients, with a median recovery time of 17 days within (15, 18) 95% confidence interval. Of the total 306 (63.62%) recoveries and 175 (36.38%) censored observations. A large proportion of these patients, 433 (90.02%) were HIV negative and 269 (62.12%) of them were recovered, whereas 164 (37.88%) of them were censored, and 48(9.98%) were HIV positive, of which 37 (77.08%) were recovered with a median recovery time of 31 days within 95% C.I (28,33). Patients whose age were above 55 years took a longer period of time to recover from COVID-19, with a median recovery time of 19 days within 95% C.I (18, 21) than the reference group (18-35 age categories). Most of the patients 343, or 71.31% were from urban areas, and a large number of patients 254, or 74.05% recovered in this group in comparison with rural areas. Of the total 481 patients, 231(48.03%) have hypertension and 76(32.9%) of them were recovered from COVID-19 with a median recovery time of 26 days within 95% C.I (23, 28) when we made a descriptive comparison with the non-hypertension group of the total patients. Regarding cardiovascular disease, 331 (68.82%) had not, whereas 150 (31.18%) had cardiovascular disease. Among the cardiovascular patients, 61(40.67%) recovered from COVID-19, with a median recovery time of 24 days within 95% C.I (21, 26).

In addition, of the total of diabetics with COVID-19 patients, 165 (34.3%) had asthma, and of the total, 46 (27.88%) recovered, with a median recovery time of 28 days within 95% C.I (26, 32). Among diabetic patients with COVID-19, 151 (31.39%) had a respiratory viral infection (RVI) and 67(44.37%) of them were recovered, with a median recovery time of 22 days within 95% C.I (22, 26). Of the total diabetic patients with COVID-19, 116(24.12%) had a stroke, 178(37.01%) were TB, and 96(19.96%) were CLD. Of the total diabetic patients with COVID-19, 106(22.04%) had obesity and 33(31.13%) of them were recovered, with a median recovery time of 29 days within 95% C.I (26, 33). Most of the

patients were exposed to TII DM, which is 289 (60.08%) of the patients included in this study, and of them, 224(77.51%) were recovered with a median recovery time of 18 days within 95% C.I (17, 18). Considering other co-factor diseases, 299(62.16%) patients were those who had no other co-factors at their baseline and the rest (182, or 37.84%) had been co-infected by other diseases at their baseline. The minimum and maximum recovery times observed in the data were 5 and 59 days, respectively (see Table 5.1 under Appendix I. A).

Table 4.1: Descriptive Statistics for Categorical Variables included in the analysis.

		Survival status		
Covariate	Categories	Event/Recovery (%)	Censored(%)	Total(%)
Age	18-35	71(83.53%)	14(16.47%)	85(17.67%)
	36-55	113(79.02%)	30(20.98%)	143(29.73%)
	>55	122(42.22%)	131(51.78%)	253(52.6%)
Sex	Male	179(61.3%)	113(38.7%)	292(60.71%)
	Female	127(67.2%)	62(32.8%)	189(39.29%)
Residence	Urban	254(74.05%)	89(25.95%)	343(71.31%)
	Rular	52(37.68%)	86(62.32%)	138(28.69%)
HIV	Negative	269(62.12%)	164 (37.88%)	433(90.02%)
	Positive	37(77.08%)	11(22.92%)	48(9.98%)
Hypertension	No	230(92%)	20(8%)	250(51.98%)
	Yes	76(32.9%)	155(67.1%)	231(48.02%)
CVD	No	245(74.02%)	86(25.98%)	331(68.82%)
	Yes	61(40.67%)	89(59.33%)	150(31.18%)
Asthma	No	260 (82.28%)	56 (17.72%)	316 (65.7%)
	Yes	46(27.88%)	119(72.12%)	165(34.3%)

RVI	No	239(72.42%)	91(27.58%)	330(68.61%)
	Yes	67(44.37%)	84(55.63%)	151(31.39%)
Stroke	No	263(72.05%)	102(27.95%)	365(75.88%)
	Yes	43(37.07%)	73(62.93%)	116(24.12%)
TB	No	254(83.83%)	49(16.17%)	303(62.99%)
	Yes	52(29.21%)	126(70.79%)	178(37.01%)
CLD	No	263 (68.31%)	122(31.69%)	385(80.04%)
	Yes	43(44.79%)	53(55.21%)	96(19.96%)
Obesity	No	273(72.8%)	102(27.2%)	375(77.96%)
	Yes	33(31.13%)	73(68.87%)	106(22.04%)
TDM	TypeI	62(56.36%)	48(43.64%)	110(22.87%)
	TypeII	224(77.51%)	65(22.49%)	289(60.08%)
	Gestational	20(24.39%)	62(75.61%)	82(17.05%)
Other co-factor	No	208(69.57%)	91(30.43%)	299(62.16%)
	Yes	98(53.85%)	84(46.15%)	182(37.84%)

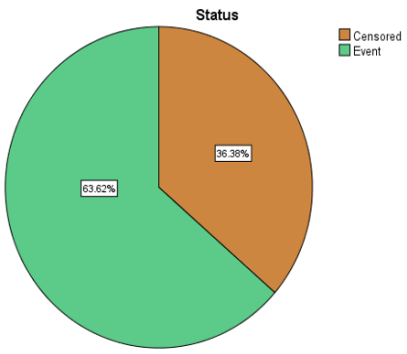

Figure 1: Treatment outcome of DM plus Covid-19 patients admitted to EKGH treatment center, Addis Ababa Ethiopia from March 20, 2020, to April 30, 2022.

The above figure1 shows a total of four hundred eighty-one diabetic patients with COVID-19 were considered in this study. Among these patients, 306 (63.62%) were recovered from COVID-19 pandemic disease, and the remaining 175 (36.38%) were censored during the study period.

4.2 Non-parametric Survival Analysis

The Kaplan Meier (KM) estimator of survivorship function (survival probability) of diabetic patients from COVID-19 is analyzed in Table 5.1 (Appendix I. A). The output shows that the survival estimation of diabetic patients with COVID-19 on day five (5^{th}) for two individual patients is 0.996. The corresponding standard error of survival estimation of patients on day five was 0.003 with a 95% C.I (0.989, 1.000). Also, the survival estimation of the diabetic patients from COVID-19 on day six for four diabetic patients was 0.987, and its standard error of survival estimation on day six was 0.005, within 95% C. I (0.976, 0.997). Although the survival estimation of patients with DM plus COVID-19 on fifty-nine days was 0.008 with a standard error of 0.008, within 95% C. I (0.001, 0.052), it is much lower than the survival time in the beginning, but the time is still increasing.

Median recovery time, or 50% recovery time, is the recovery time where the probability of recovery time is equal to 0.5. Hence, the overall median recovery time was 16 days within a 95% C.I (15, 17). It is explained in the following (Table 4.2):

Table 4.2: Median survival time of DM plus COVID-19 patents

Total n. of patients	Recovery	Median(50(%))	95% CI	
			Upper CI	Lower CI
481	306	16	15	17

Additionally, the following (figure 2) shows the estimate of the overall Kaplan-Meier median survivor function and the horizontal line shows the overall 50% median recovery time of diabetic patients from COVID-19, which appears at 16 days within a (15, 17) 95% confidence interval.

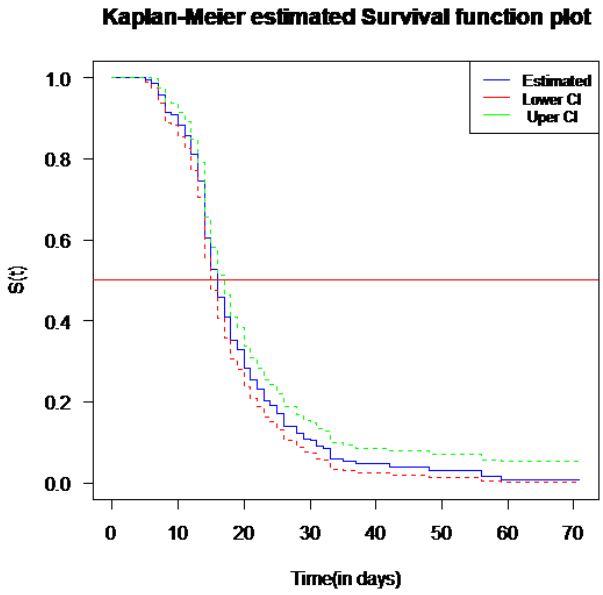

Figure 2: A plot of overall estimated survivor function at median 16 days with 95% CI (15, 17) of diabetic patients with COVID-19 disease.

4.2.1 Survival Time-to-Recovery for Different Groups of Covariates

From these results displayed in figure 3, the survival time to recovery of diabetic patients from COVID-19, whose age group of 36-55, has a prolonged time to recovery from COVID-19 compared to the 18-35 age group. Also, the age group above 55 has a more prolonged time to recovery compared to the 18-35 age group. The log-rank test (Table 4.3) also illustrated that, there were significant differences among age categories with respect to survival probability ($p < 0.001$). Female patients have a better recovery time from COVID-19 compared to the male patients.

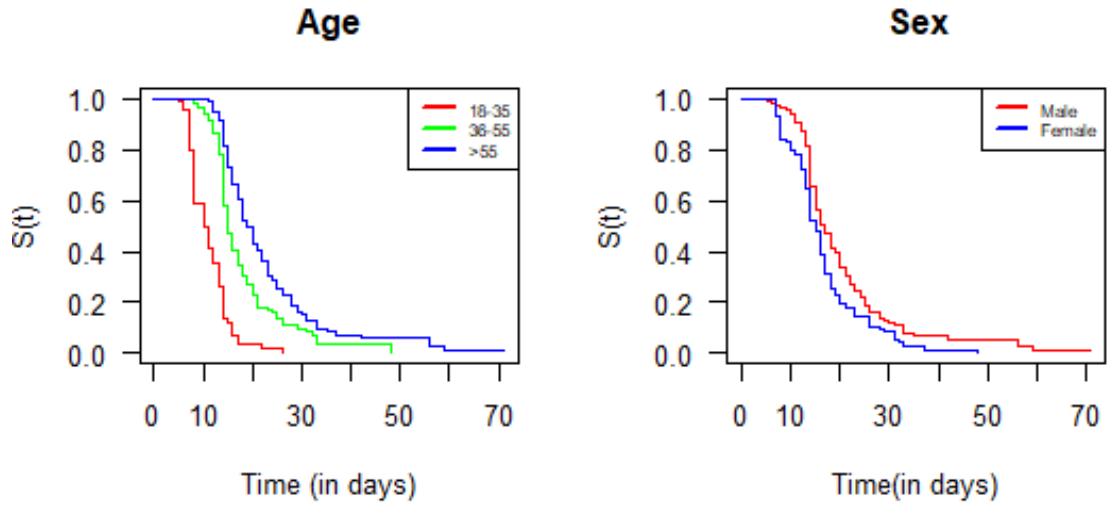

Figure 3: Kaplan-Meier survival estimates by Age and Sex

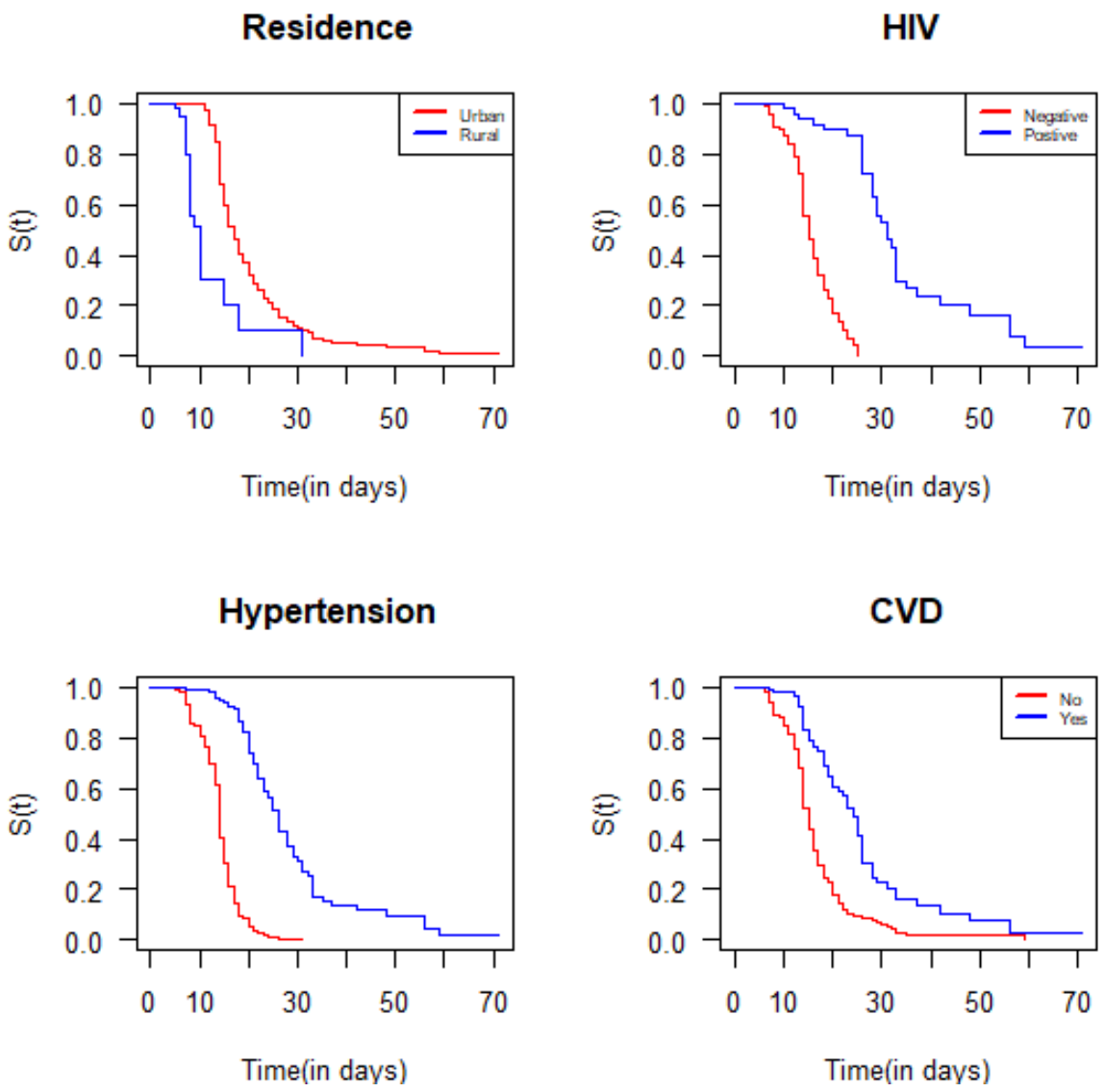

Figure 4: Kaplan-Meier survival estimates by Residence, HIV, Hypertension, and CVD

KM curve in the above Figure 4 indicates that diabetic patients who live in rural areas have a faster time of recovering from COVID-19 than the patients who live in urban areas. The log rank test also revealed that residence of patients had significant association with time to recovery of diabetic patients from COVID-19 ($p < 0.001$) at 5% level of significance. Diabetic patients with HIV positivity have a prolonged time to recover from COVID-19 compared to the patients with HIV negative. The log rank test in Table 4.3 also showed HIV status was significant association with time to recovery of diabetic patients from COVID-19 ($p < 0.001$). Diabetic patients with high blood pressure (hypertension) also have a prolonged time to recover from COVID-19 than diabetic patients with normal high blood pressure (no hypertension). Additionally, diabetic patients who have cardiovascular disease have a prolonged time to recover from COVID-19 compared to the reference group (no CVD).

KM curve in the below Figure 8 in (Appendix I C.), indicates that there is an effect on time-to-recovery due to Type II diabetic patients with COVID-19. Patients with type II have a longer time duration to recover from COVID-19 than the type I diabetes mellitus. However, patients with gestational have a shorter recovery time compared to the type I diabetes mellitus. The result of the log rank test is also revealed the difference is significant ($p < 0.001$) at 5% level of significance. Moreover, diabetic patients who had asthma, RVI, symptoms, stroke, TB, CLD, obesity, and other co-factors had a prolonged time to recover from COVID-19 than patients who had no asthma, RVI, symptoms, stroke, TB, CLD, obesity, and other co-factors. The log-rank in Table 4.3 also shows that difference was significant ($p < 0.001$) at a 5% level of significance.

In general, to explore whether there are significant differences among different groups of categorical predictors, we employ the log-rank statistical test. The null hypothesis test stated that there is no difference between the survival curves, whereas the alternative hypothesis test is stated that there is a difference between survival curves.

Table 4.3: Log rank test for each covariates of diabetics with COVID-19 patents in Eka Kotebe General Hospital.

Covariates	Chi-square	DF	P-value
Age	183	2	< 0.001
Sex	12.7	1	0.0004
Residence	187	1	< 0.001
HIV	107	1	< 0.001
Hypertension	212	1	< 0.001
CVD	50.5	1	< 0.001
Symptom	113	1	< 0.001
Asthma	140	1	< 0.001
RVI	42.5	1	< 0.001
Stroke	55.2	1	< 0.001
TB	118	1	< 0.001
CLD	104	1	< 0.001
Obesity	86.6	1	< 0.001
Types of DM	367	2	< 0.001
Other co-factor	95.6	1	< 0.001

DF = degree freedom

The above Table 4.3 shows that the log-rank test results that there is a statistically significant difference in the cumulative incidence of recovery time for all categorical covariates (age, sex, residence, HIV, hypertension, CVD, symptoms, asthma, RVI, stroke, TB, CLD, obesity, and types of DM, P-value < 0.001). These log-rank test results show that rejecting the null hypothesis since the groups have different distribution curves as shown on the Kaplan-Meier survival plots above (figure 3 and figure 4).

4.3 Results of the Cox proportional hazards model

In order to study the relationship between survival time and covariates, a regression modeling approach to survival analysis using the Cox proportional hazards model can be employed for estimating the regression coefficients, making interpretation based on the hazard function, conducting statistical tests, constructing confidence intervals, checking the adequacy of model and its development precede interpretation of results obtained from the fitted model.

4.3.1 Univariate Analysis

The univariate Cox proportional hazard model analysis is an appropriate method that is used to show potentially important variables before directly included in the multivariate model. The relationship between each covariates and survival probability of diabetic patients with COVID-19 are presented in Table 4.4. For each variable, there is a univariate Cox proportional hazards model analysis that contains a single independent variable, which is used to know the significance of each variable with survival time. The variables that are significant in the univariate analysis in relation to time to the occurrence of an event (or recovery) due to diabetics with COVID-19 were selected at a 25% modest level of significance [96]. Accordingly, the univariate Cox proportional hazards regression models are fitted for every covariate shown in Table 4.4. All covariates (age, sex, residence, HIV, hypertension, CVD, symptoms, asthma, RVI, stroke, TB, CLD, obesity, types of DM, and other co-factors) are extracted to be included at a 25% level of significance and they are candidate variables for multivariate analysis.

Table 4.4: Results of the univariable proportional hazards Cox regression model of diabetics plus COVID-19 patients.

Covariates		DF	$\hat{\beta}$	SE($\hat{\beta}$)	Z	Pr(>|z|)	$\hat{HR}$	CI ($\hat{HR}$)
Age	36-55	2	-1.351	0.158	-8.562	0.000	0.259	(0.190, 0.353)
	>55		-1.863	0.159	-11.693	0.000	0.155	(0.114 0.212)
Sex	Female	1	0.394	0.117	3.362	0.001	1.482	(1.178, 1.865)
Residence	Rural	1	2.059	0.174	11.84	0.000	7.839	(5.575, 11.02)
HIV	Positive	1	-3.166	0.428	-7.405	0.000	0.042	(0.018, 0.098)
Hypertension	Yes	1	-2.024	0.161	-12.6	0.000	0.132	(0.096, 0.181)
CVD	Yes	1	-0.950	0.146	-6.509	0.000	0.387	(0.290, 0.515)
Symptom	Yes	1	-1.608	0.173	-9.302	0.000	0.2	(0.143, 0.281)
Asthma	Yes	1	-1.876	0.183	-10.24	0.000	0.153	(0.107, 0.219)
RVI	Yes	1	-0.866	0.144	-6.011	0.000	0.421	(0.317, 0.558)
Stroke	Yes	1	-1.182	0.176	-6.706	0.000	0.307	(0.217, 0.433)
TB	Yes	1	-1.629	0.169	-9.619	0.000	0.196	(0.141, 0.273)
CLD	Yes	1	-1.560	0.1740	-8.964	0.000	0.210	(0.149, 0.296)
Obesity	Yes	1	-1.597	0.196	-8.162	0.000	0.203	(0.138, 0.297)
TDM	Type II	2	-1.893	0.163	-11.643	0.000	0.151	(0.11, 0.207)
	Gestational		3.11	0.404	7.693	0.000	22.41	(10.148, 49.486)
Other co-factor	Yes	1	-1.157	0.131	-8.861	0.000	0.314	(0.243, 0.406)

Source: Eka Kotebe General Hospital, Addis Ababa, Ethiopia; from March 20, 2020 to April 30, 2022. DF= Degree freedom, $\hat{\beta}$= Parameter Estimate, SE= standard error, $\hat{HR}$= estimated hazard ratio, CI= confidence interval

4.3.2 Multivariate Analysis

Multivariable survival analysis instead of univariate analysis considers the possibility that weakly associated variables could become important predictors of the outcome when taken together. All potential variables that were supposed to have a significant impact (p-value < 0.25) on the survival time to recovery of diabetic patients from COVID-19 in univariate analysis were included in the multivariate cases. Accordingly, the relationship between covariates and survival time to recovery of diabetic patients from COVID-19 is presented in Table 4.5 below. The result of the multivariate analysis of the cox-PH model in (Table 4.5) indicates that variables like age, sex, residence, HIV, hypertension, asthma, stroke, symptoms, TB, CLD, types of DM, and other co-factors were significantly associated

with time to recovery of diabetic patients from COVID-19 at a 5% level of significance.

The results reveal that after accounting for other confounders in the data, time to recovery takes a longer time in the age groups of 36-55 and older than 55. Also, diabetic patients whose age is older than 55 years take a prolonged time to recover from COVID-19 with the estimated hazard ratio ($\hat{HR}$ = 0.663, 95% C.I: 0.445, 0.989). This indicates that patients whose age is older than 55 years take a prolonged time to recover from COVID-19 disease compared with patients whose age group is 18-35. Sex is seen to be significantly associated with the recovery time of diabetic patients from COVID-19. Female diabetic patients recover from COVID-19 about 1.382 times higher than the males patients. That is, females recover faster than males ($\hat{HR}$ = 1.382, 95% CI: 1.061, 1.8).

Patients with HIV positive are less likely about 0.067 to recover from COVID-19 disease compared to those who had HIV negative (holding other covariates constant). In other words, diabetic patients with HIV had a 93.3% lower rate of recovery from COVID-19 infection compared to those without the disease. Likewise, the estimated hazard ratio for hypertension is 0.38 with a 95% confidence interval (0.255, 0.567). This indicates that patients with hypertension at baseline are less likely about 0.38 to recover from COVID-19 compared to patients with normal blood pressure or no hypertension (keeping other covariates constant). In other words, diabetic patients with hypertension had a 62% lower rate of recovery from COVID-19 infection compared to those without hypertension. Patients symptoms are less likely about 0.514 to recover from COVID-19.

The estimated hazard ratio for asthma is 0.491 with a 95% confidence interval (0.309, 0.78). This indicates that the time for recovery is prolonged for patients with the presence of asthma compared to those who have no asthma. Patients with asthma are less likely about (0.491) to recover from the COVID-19 disease compared to those who have no asthma. Also, the estimated hazard ratios for patients with stroke and TB were ($\hat{HR}$ = 0.508, 95% C.I: 0.333, 0.776) and ($\hat{HR}$ = 0.424, 95% C.I: 0.277, 0.65) respectively. It shows patients with comorbidities like stroke and TB are less likely about 0.508 and 0.424 to recover from COVID-19 compared to patients who have no stroke and TB.

The estimated hazard ratio for chronic lung disease (CLD) is 0.564 with a 95% confidence interval (0.356, 0.893). This indicates that the time for recovery is prolonged for patients with the presence of chronic lung disease compared to those who have no CLD. Patients with CLD are less likely about (0.564) to recover from COVID-19 disease compared to those who have no chronic lung disease (CLD).

Considering the types of diabetes after adjusting for other confounding variables, the hazard ratio of patients with type II and gestational diabetes was 0.159 and 41.875 ($\hat{HR}$ = 0.159, 95% C.I: 0.107, 0.236, $\hat{HR}$ = 41.875, 95% C.I: 9.279, 188.976) respectively; this indicates that patients who have gestational diabetes have a faster time of recovery from COVID-19 disease compared to those who have type I diabetes. However, the recovery time of patients who had type II diabetes was reduced by 84.1% when compared with patients who had type I diabetes. Additionally, the estimated hazard ratio of patients with other co-factors is 0.531 with a (0.398, 0.708) 95% confidence interval. This indicates that the time for recovery is prolonged for patients with the presence of other co-factors compared to those who have no other co-factors, which shows patients with other co-factors are less likely about 0.531 to recover from COVID-19 disease compared to those who have no other co-factors (holding other covariates constant).

4.4 Assessment of Model Adequacy

The adequacy of the model needs to be assessed after the model has been fitted to the observed survival data. At this point, we have a preliminary fitted model and the next step is to assess the adequacy of the fitted model. This should be done in order to evaluate how well the fitted regression describes the data set. In Cox-proportional hazard survival regression analysis assessment of model adequacy, the study must test the assumption of proportional hazards and make the model diagnostic.

4.4.1 Checking for Proportional Hazard Assumption

As it was described in chapter three, graphical diagnoses of the Log-cumulative hazard plot against time t and scaled Schoenfeld residuals were applied to assess the proportional hazard assumption of covariates that are significant in the multivariate analysis. The PH assumption holds if the plot of the log-cumulative hazard plot against time is parallel and the distribution of residuals over time is random, and the LOWESS smoothing line should be a straight line around zero.

Figure 5 below shows that the $\log(-\log(survival)$ against the time plot for age, sex, residence, HIV, hypertension, and symptoms is parallel and also the log (-log (survival) against time for all the remaining categorical variables is parallel (see Figure 9 under Appendix II A.). Therefore, the proportional hazard assumption holds for all categorical variables.

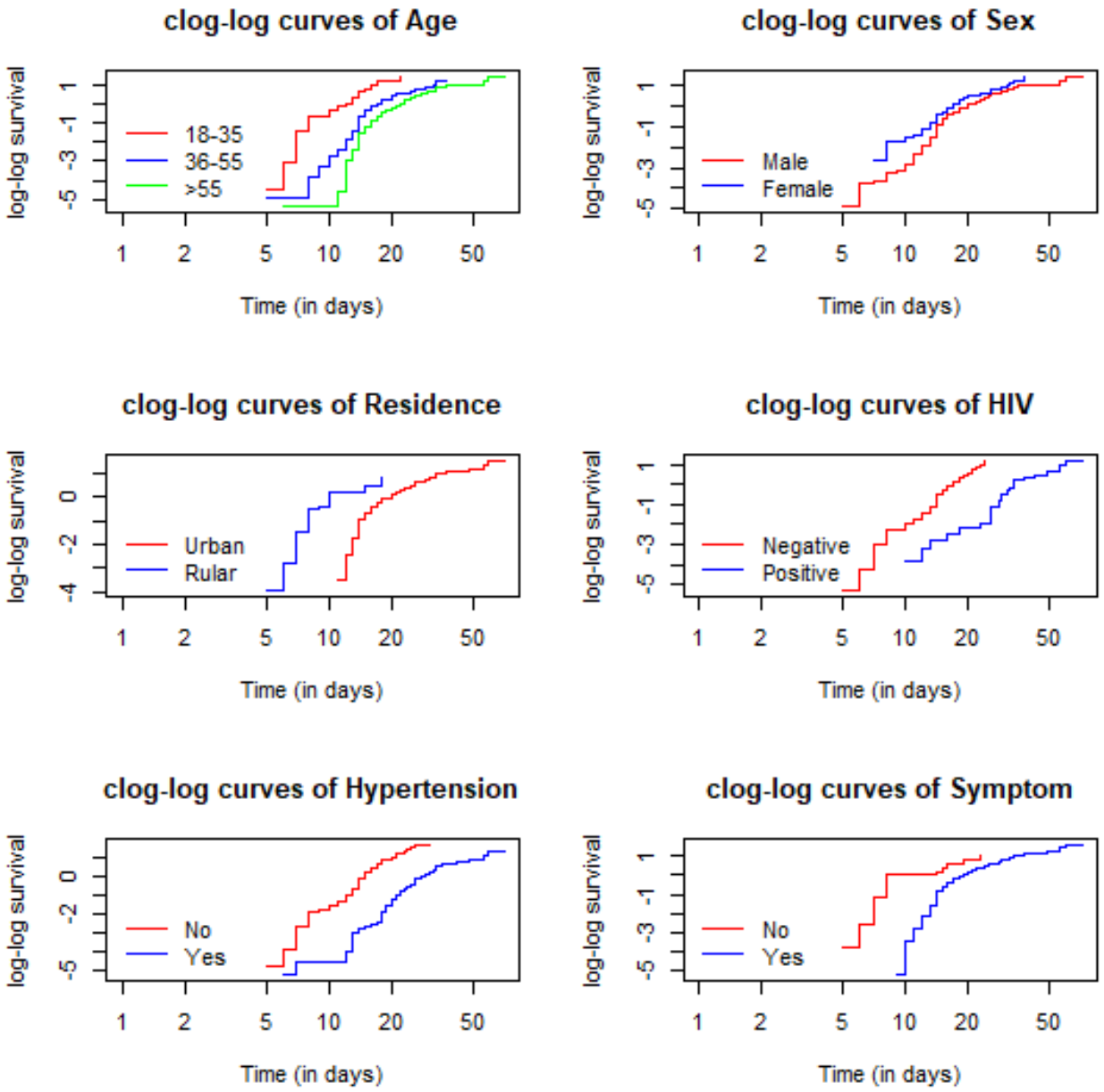

Figure 5: Log(-log(survival)) plot against time (Age, Sex, Residence, HIV, Hypertension, & Symptom)

The PH assumption checking with a graphical method based on the Schoenfeld residual has been described in Figure 6 and is included in the model. Systematic departures from a horizontal line are indicative of non-proportional hazards since PH assumes that estimates $\beta_1, \beta_2, ..., \beta_p$ do not vary much over time. Also, the graphs for some of the categorical variables displayed in Figure 5 using Kaplan appeared to be parallel, implying that the proportional-hazards assumption of categorical variables has not been violated. Plotting the scaled Schoenfeld residuals of each covariate against log time will be used to check whether the assumption of proportional hazards is violated or not. Clearly, a close look at Figure 6 indicates that the residuals are random and that the LOESS curve is smooth and

horizontal with zero slopes. This also suggests that the plots support the proportionality assumption to hold, and the scaled Schoenfeld residual plots for the remaining variables are available in Appendix II B.

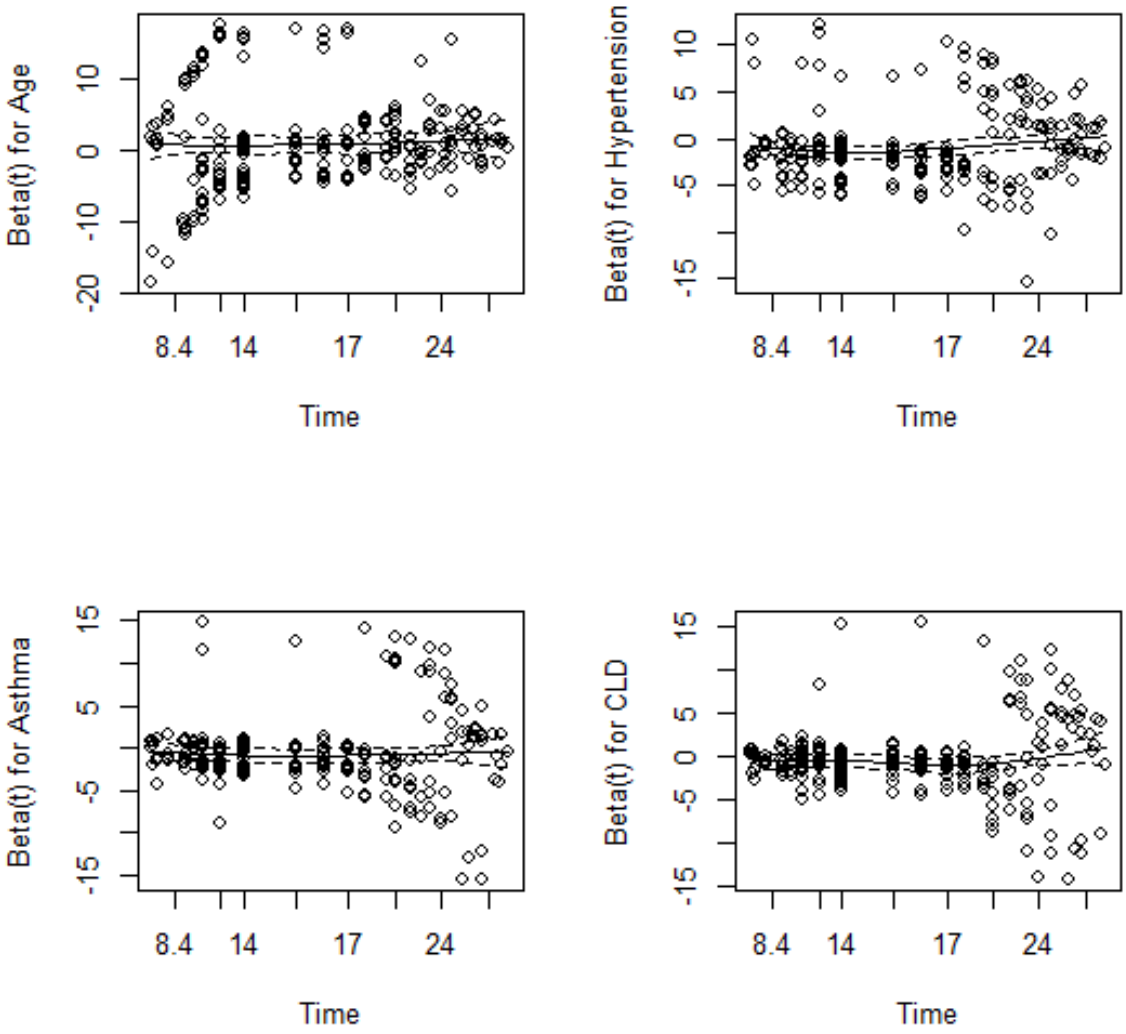

Figure 6: Scaled Schoenfeld Residual Plots for Age, Hypertension, Asthma,and CLD

4.4.2 Diagnostics for the Cox proportional hazards model

Having identified the final preliminary model, the next step and most important in statistical analysis is to diagnose the fit of the model. After a model has been fitted to an observed set of survival data the adequacy of the fitted model needs to be assessed. The use of diagnostic procedures for model checking is an essential part of the model in process.

A plot of the Cox-Snell residuals against the cumulative hazard of the residuals is presented in Figure 7. The hazard function follows 45-degree line very closely, except for very large values of the time. It

is very common for models with censored data to have some wiggling at large values of time, and it is not something that should cause much concern. Overall, the final model fits the data.

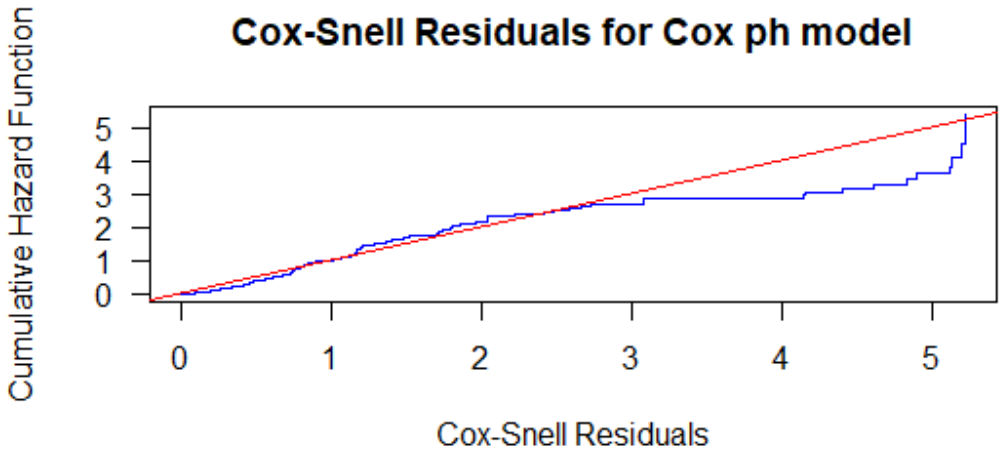

Figure 7: Cox-Snell residuals obtained from fitting Cox PH model to the diabetics with COVID-19.

4.5 Discussion

The research was targeted at a positive event, which is the recovery time of diabetic patients from COVID-19. COVID-19 disease remains one of the leading causes of death seeking public health attention. Several studies conducted in the COVID-19 era also reported a strong association between DM and COVID-19 mortality. Diabetes Mellitus has been identified as a risk factor for adverse outcomes from various types of infections, including those caused by respiratory viruses [98]. This thesis revealed that the overall median recovery time of diabetic patients from COVID-19 for the patients in Eka Kotebe General Hospital is 16 days, within a 95% confidence interval (15, 17). This median recovery time was higher than a study in China 12 days within 95% confidence interval (7, 15) [56]. The possible reasons for the observed discrepancy between the studies might be variations in study setting and time. Patients were previously given better care and treatment than they are now.

The main purpose of this study was to model the time to recovery of diabetic patients from COVID-19, which was obtained from Eka Kotebe General Hospital. A Cox proportional hazard analysis was applied to this data. The covariates that were included in this study were age, sex, residence, HIV, hypertension, CVD, symptoms, stroke, asthma, RVI, TB, CLD, obesity, types of DM, and other co-factors, and the outcome variable of interest was the recovery time of diabetic patients from COVID-19 which, measured in days.

The findings of this study revealed that patients with age groups of 36-55 and above 55 years old significantly associated with the time to recovery of diabetic patients from COVID-19. This indicates that the time for recovery time from COVID-19 is prolonged for patients in the age group of 35-36 and above 55 years old. The current study is also consistent with other findings by Nafakhi *et al* [57], Abdene *et al* [99] and World health organization [100].

The results of this study indicated that male diabetic patients were at a higher risk of COVID-19 than female diabetic patients. The current study is consistent with other findings of Belice et al [58]. Furthermore, the current study results showed that the majority of diabetic patients with COVID-19 (71.31%) were from urban areas, indicating that urban participants were more affected by COVID-19 than rural participants. The current study is agreed with the study conducted by Shibru *et al* [63].

The results of this study indicated that status of HIV was a significant predictive factor for the time to recovery of diabetic patients from COVID-19. Patients who lived with HIV positive had a longer recovery time than patients who lived with HIV negative. The current study agreed with study conducted by Abdene *et al* [99] and Nordling [73]. Similarly, this study revealed that TB is significantly association with the recovery time of diabetic patients from COVID-19. This study consistent with Abdene *et al* [99] and Nordling [73] findings.

The results of this study indicated that hypertension was a significant predictive factor for the time to recovery of diabetic patients from COVID-19. Patients with high blood pressure (hypertension) had a longer recovery time than patients with no hypertension. The current study agreed with a study conducted by Migbar *et al*, Shadnoush *et al*, and DU *et al* [26, 61, 62]. The results of this study suggested that symptom was significant predictive factor for time to recovery of diabetic patients from COVID-19. Patients who had no symptoms have a better recovery time than patients with symptoms. Studies of Migbar *et al* [26] have similar conclusion to this finding.

Asthma is another important significant factor in the time to recovery of diabetic patients from COVID-19 disease. Diabetic patients with asthma have a prolonged recovery time from COVID-19 compared to those who have no asthma. This finding is in line with studies conducted in Rawshani *et al* [66].

The results of the study revealed that having a stroke has a significant association with the time to recovery of diabetic patients from COVID-19. Diabetic patients who had a stroke had a prolonged recovery time from COVID-19 than patients with no stroke. The current study is consistent with

Rawshani *et al* and Nannoni *et al* [65,66]. The results of this study indicated that CLD was a significant predictive factor for the time to recovery of diabetic patients from COVID-19. This study is in line with a study conducted by Wang *et al* [77].

Furthermore, the study findings showed that types of diabetes are another important significant factor for time to recovery of diabetic patients from COVID-19. Those patients with type II diabetes were less likely to recover from COVID-19 than that of patients with type I diabetes. This result is supported by study of Sisman *et al* [101] and Sonmez *et al* [14]. Similarly, this study result showed that diabetic patients with other co-factors were also significantly associated with the time to recovery of diabetic patients from COVID-19. Patients who had no other co-factors have a better recovery time than patients with other cofactors. The current study consistent with Al *et al* [64] findings.

5 Conclusion and Recommendation

5.1 Conclusion

This study used the survival time of diabetic patients with COVID-19 dataset of patients who started their treatment from March 20, 2020, to April 30, 2022, G.C with the aim of modeling the time-to-recovery of diabetic patients from COVID-19 in EKGH. Out of the total 481, about 306 (63.62%) diabetic patients recovered from COVID-19, and the remaining 175 (36.38%) were censored during the study period. The overall median recovery time of diabetic patients from COVID-19 was 16 days within 95% confidence intervals (15, 17).

The results of the Kaplan-Meier and log-rank showed that patients who have HIV, hypertension, asthma, stroke, TB, CLD, and other cofactors have a prolonged time to recovery from COVID-19 compared to patients with no HIV, hypertension, asthma, stroke, TB, CLD, and other co-factors. Similarly, the results of the Kaplan-Meier survival estimate showed that gestational diabetes mellitus has a shorter time to recovery from COVID-19 than patients with other types of diabetes mellitus.

The result of multivariate Cox proportional hazards regression model showed that factors like age, sex, residence, HIV, hypertension (HTN), symptoms, asthma, stroke, TB, CLD, types of DM, and other co-factors were the major factors that affect time to recovery of diabetic patients from COVID-19 at a 5% level of significance. Cox proportional hazard assumption was checked by using Schoenfeld residual and graphical ($\log(-\log$(survival probability)) versus survival time) and fitted by Cox-Snell.

5.2 Recommendation

Based on the results of this study, the following recommendations were made:

- Health professionals should be given more attention to diabetic patients with HIV, hypertension, TB, stroke, asthma, and CLD to improve the recovery time of those patients with COVID-19. Further, considering the delay in recovery, those diabetic patients with other cofactor (like IHD, HF, CKD...) should be given priority in shielding them from contracting the SARS-CoV-2.

- Every diabetic patient should be able to manage themselves from COVID-19 attacks by adhering to WHO principles, especially the elderly.

- Further studies should be conducted to identify other factors that are not identified in this study.

- Health professionals should record carefully the patients' data like socio-demographic characteristics and clinical variables as it is important for further investigation of factors affecting the recovery time of diabetic patients from COVID-19.

References

[1] Na ZHU, Dingyu ZHANG, Wenling WANG, Xingwang LI, Bo YANG, Jingdong SONG, Xiang ZHAO, Baoying HUANG, Weifeng SHI, Roujian LU *et al.* : A novel coronavirus from patients with pneumonia in china, 2019. *New England journal of medicine*, 2020.

[2] Nadia JEBRIL : World health organization declared a pandemic public health menace: a systematic review of the coronavirus disease 2019 "covid-19". *Available at SSRN 3566298*, 2020.

[3] Chidiebere V UGWUEZE, Basil Chukwuma EZEOKPO, Bede I NNOLIM, Emmanuel A AGIM, Nnamdi C ANIKPO et Kenechukwu E ONYEKACHI : Covid-19 and diabetes mellitus: the link and clinical implications. *Dubai Diabetes and Endocrinology Journal*, 26(2):69–77, 2020.

[4] Gian Franco Del PRETE, Corrado BETTERLE, Domenico PADOVAN, Giuseppe ERLE, Antonella TOFFOLO et Giorgio BERSAHI : Incidence and significance of islet-cell autoantibodies in different types of diabetes mellitus. *Diabetes*, 26(10):909–915, 1977.

[5] Eliana M WENDLAND, Maria Regina TORLONI, Maicon FALAVIGNA, Janet TRUJILLO, Maria Alice DODE, Maria Amélia CAMPOS, Bruce B DUNCAN et Maria Inês SCHMIDT : Gestational diabetes and pregnancy outcomes-a systematic review of the world health organization (who) and the international association of diabetes in pregnancy study groups (iadpsg) diagnostic criteria. *BMC pregnancy and childbirth*, 12(1):1–13, 2012.

[6] Carl J LAVIE, Richard V MILANI et Hector O VENTURA : Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *Journal of the American college of cardiology*, 53(21):1925–1932, 2009.

[7] World Health ORGANIZATION *et al.* : Global report on diabetes: executive summary. Rapport technique, World Health Organization, 2016.

[8] Yousef ALIMOHAMADI, Mojtaba SEPANDI, Maryam TAGHDIR et Hadiseh HOSAMIRUDSARI : Determine the most common clinical symptoms in covid-19 patients: a systematic review and meta-analysis. *Journal of preventive medicine and hygiene*, 61(3):E304, 2020.

[9] Dante S HARBUWONO, Dwi OTL HANDAYANI, Endang S WAHYUNINGSIH, Novita SUPRAPTOWATI, Farid KURNIAWAN, Syahidatul WAFA, Melly KRISTANTI, Nico I PANTORO, Robert SINTO, Heri KURNIAWAN *et al.* : Impact of diabetes mellitus on covid-19 clinical symptoms and mortality: Jakarta's covid-19 epidemiological registry. *Primary Care Diabetes*, 16(1):65–68, 2022.

[10] Michael KLOMPAS, Meghan A BAKER et Chanu RHEE : Airborne transmission of sars-cov-2: theoretical considerations and available evidence. *Jama*, 324(5):441–442, 2020.

[11] Shahram YAZDANI, Majid HEYDARI, Zeynab FOROUGHI et Hadi JABALI : Factors affecting covid-19 transmission and modelling of close contact tracing strategies. *Iranian Journal of Public Health*, 50(10):2121, 2021.

[12] Lidia MORAWSKA, Julian W TANG, William BAHNFLETH, Philomena M BLUYSSEN, Atze BOERSTRA, Giorgio BUONANNO, Junji CAO, Stephanie DANCER, Andres FLOTO, Francesco FRANCHIMON *et al.* : How can airborne transmission of covid-19 indoors be minimised? *Environment international*, 142:105832, 2020.

[13] Smriti MALLAPATY : Why does the coronavirus spread so easily between people? *Nature*, 579(7798):183–184, 2020.

[14] Alper SONMEZ, Ibrahim DEMIRCI, Cem HAYMANA, Ilker TASCI, Selcuk DAGDELEN, Serpil SALMAN, Naim ATA, Ibrahim SAHIN, Rifat EMRAL, Erman CAKAL *et al.* : Clinical characteristics and outcomes of covid-19 in patients with type 2 diabetes in turkey: A nationwide study (turcovidia). *Journal of diabetes*, 13(7):585–595, 2021.

[15] Ensheng DONG, Hongru DU et Lauren GARDNER : An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534, 2020.

[16] Hong SUN, Pouya SAEEDI, Suvi KARURANGA, Moritz PINKEPANK, Katherine OGURTSOVA, Bruce B DUNCAN, Caroline STEIN, Abdul BASIT, Juliana CN CHAN, Jean Claude MBANYA *et al.* : Idf diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes research and clinical practice*, 183:109119, 2022.

[17] Sian A BRADLEY, Maciej BANACH, Negman ALVARADO, Ivica SMOKOVSKI et Sonu MM BHASKAR : Prevalence and impact of diabetes in hospitalized covid-19 patients: A systematic review and meta-analysis. *Journal of diabetes*, 14(2):144–157, 2022.

[18] Qiao SHI, Xiaoyi ZHANG, Fang JIANG, Xuanzhe ZHANG, Ning HU, Chibu BIMU, Jiarui FENG, Su YAN, Yongjun GUAN, Dongxue XU *et al.* : Clinical characteristics and risk factors for mortality of covid-19 patients with diabetes in wuhan, china: a two-center, retrospective study. *Diabetes care*, 43(7):1382–1391, 2020.

[19] Marcus W BUTLER, Aoife O'REILLY, Eleanor M DUNICAN, Patrick MALLON, Eoin R FEENEY, Michael P KEANE et Cormac MCCARTHY : Prevalence of comorbid asthma in covid-19 patients. *Journal of Allergy and Clinical Immunology*, 146(2):334–335, 2020.

[20] Jing YANG, YA ZHENG, Xi GOU, Ke PU, Zhaofeng CHEN, Qinghong GUO, Rui JI, Haojia WANG, Yuping WANG et Yongning ZHOU : Prevalence of comorbidities and its effects in patients infected with sars-cov-2: a systematic review and meta-analysis. *International journal of infectious diseases*, 94:91–95, 2020.

[21] Naomi HOLMAN, Peter KNIGHTON, Partha KAR, Jackie O'KEEFE, Matt CURLEY, Andy WEAVER, Emma BARRON, Chirag BAKHAI, Kamlesh KHUNTI, Nicholas J WAREHAM *et al.* : Risk factors for covid-19-related mortality in people with type 1 and type 2 diabetes in england: a population-based cohort study. *The lancet Diabetes & endocrinology*, 8(10):823–833, 2020.

[22] Ben BEPOUKA, Ossam ODIO, Donat MANGALA, Nadine MAYASI, Madone MANDINA, Murielle LONGOKOLO, Jean Robert MAKULO, Marcel MBULA, Jean Marie KAYEMBE et Hippolyte SITUAKIBANZA : Diabetes mellitus is associated with higher covid-19 mortality rates in sub-saharan africa: A systematic review and meta-analysis. *Cureus*, 14(7), 2022.

[23] Andrew Peter KYAZZE, Felix BONGOMIN, Sandra NINSIIMA, Gloria NATTABI, Winnie NABAKKA, Rebecca KUKUNDA, Henry ODANGA, Phillip SSEKAMATTE, Joseph Baruch BALUKU, Davis KIBIRIGE *et al.* : Optimizing diabetes mellitus care to improve covid-19 outcomes in resource-limited settings in africa. *Therapeutic advances in infectious disease*, 8:20499361211009380, 2021.

[24] Getu Melesie TAYE, Lemma BOSE, Tamirat Bekele BERESSA, Gosaye Mekonnen TEFERA, Biruk MOSISA, Hunduma DINSA, Adamu BIRHANU et Gurmu UMETA :

Covid-19 knowledge, attitudes, and prevention practices among people with hypertension and diabetes mellitus attending public health facilities in ambo, ethiopia. *Infection and Drug Resistance*, 13:4203, 2020.

[25] Andargew Yohannes ASHAMO, Abebaw BEKELE, Adane PETROSE, Tsegaye GEBREYES, Eyob Kebede ETISSA, Amsalu BEKELE, Deborah HAISCH, Neil W SCHLUGER, Hanan YUSUF, Tewodros HAILE *et al.* : Assessment of hypertension and other factors associated with the severity of disease in covid-19 pneumonia, addis ababa, ethiopia: A case-control study. *PloS one*, 17(8):e0273012, 2022.

[26] Migbar Mekonnen SIBHAT, Melsew Tsegaw GETNET, Wuletaw Zewde CHANE, Kassie Tiruneh GEBEYEHU, Asaminew Sane HABTAMU, Taye Ashine MEZGEBU, Hailu Beyene ASMARE, Melkie Mengistie AMBAW et Edmialem Mesfin GETAHUN : Determinants of mortality among covid-19 patients with diabetes mellitus in addis ababa, ethiopia, 2022: An unmatched case-control study. *medRxiv*, 2022.

[27] David G KLEINBAUM, Mitchel KLEIN *et al.* : *Survival analysis: a self-learning text*, volume 3. Springer, 2012.

[28] Jerald F LAWLESS : *Statistical models and methods for lifetime data*. John Wiley & Sons, 2011.

[29] David COLLETT : *Modelling survival data in medical research*. CRC press, 2015.

[30] Saro Abdella ABRAHIM, Masresha TESSEMA, Atkure DEFAR, Alemayehu HUSSEN, Eshetu EJETA, Getachew DEMOZ, Addisu Birhanu TEREDA, Enatenesh DILLNESSA, Altaye FELEKE, Misiker AMARE *et al.* : Time to recovery and its predictors among adults hospitalized with covid-19: A prospective cohort study in ethiopia. *PloS one*, 15(12):e0244269, 2020.

[31] Tadesse TOLOSSA, Bizuneh WAKUMA, Dejene SEYOUM GEBRE, Emiru MERDASSA ATOMSSA, Motuma GETACHEW, Getahun FETENSA, Diriba AYALA et Ebisa TURI : 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. *Plos one*, 16(6):e0252389, 2021.

[32] Addis Adera GEBRU, Tadesse BIRHANU, Eshetu WENDIMU, Agumas Fentahun AYALEW, Selamawit MULAT, Hussen Zakir ABASIMEL, Ali KAZEMI, Bosenu Abera TADESSE, Beniam Adera GEBRU, Berhanu Senbeta DERIBA *et al.* : Global burden of covid-19: Situational analyis and review. *Human antibodies*, 29(2):139–148, 2021.

[33] Abdene Weya KASO, Gebi AGERO, Zewdu HURISSA, Taha KASO, Helen Ali EWUNE, Habtamu Endashaw HARERU et Alemayehu HAILU : Survival analysis of covid-19 patients in ethiopia: A hospital-based study. *Plos one*, 17(5):e0268280, 2022.

[34] Jonathan PEARSON-STUTTARD, Samkeliso BLUNDELL, Tess HARRIS, Derek G COOK et Julia CRITCHLEY : Diabetes and infection: assessing the association with glycaemic control in population-based studies. *The lancet Diabetes & endocrinology*, 4(2):148–158, 2016.

[35] Baiju R SHAH et Janet E HUX : Quantifying the risk of infectious diseases for people with diabetes. *Diabetes care*, 26(2):510–513, 2003.

[36] Alain G BERTONI, Sharon SAYDAH et Frederick L BRANCATI : Diabetes and the risk of infection-related mortality in the us. *Diabetes care*, 24(6):1044–1049, 2001.

[37] Dawei WANG, Bo HU, Chang HU, Fangfang ZHU, Xing LIU, Jing ZHANG, Binbin WANG, Hui XIANG, Zhenshun CHENG, Yong XIONG *et al.* : Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in wuhan, china. *Jama*, 323(11):1061–1069, 2020.

[38] Wei-jie GUAN, Zheng-yi NI, Yu HU, Wen-hua LIANG, Chun-quan OU, Jian-xing HE, Lei LIU, Hong SHAN, Chun-liang LEI, David SC HUI *et al.* : Clinical characteristics of 2019 novel coronavirus infection in china. *MedRxiv*, 2020.

[39] Tao CHEN, DI WU, Huilong CHEN, Weiming YAN, Danlei YANG, Guang CHEN, Ke MA, Dong XU, Haijing YU, Hongwu WANG *et al.* : Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *bmj*, 368, 2020.

[40] Fei ZHOU, Ting YU, Ronghui DU, Guohui FAN, Ying LIU, Zhibo LIU, Jie XIANG, Yeming WANG, Bin SONG, Xiaoying GU *et al.* : Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study. *The lancet*, 395(10229):1054–1062, 2020.

[41] Awadhesh Kumar SINGH, Ritesh GUPTA, Amerta GHOSH et Anoop MISRA : Diabetes in covid-19: Prevalence, pathophysiology, prognosis and practical considerations. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(4):303–310, 2020.

[42] Ahmed H ABDELHAFIZ, Demelza EMMERTON et Alan J SINCLAIR : Diabetes in covid-19 pandemic-prevalence, patient characteristics and adverse outcomes. *International Journal of Clinical Practice*, 75(7):e14112, 2021.

[43] Divya VINAYACHANDRAN et Saravanakarthikeyan BALASUBRAMANIAN : Salivary diagnostics in covid-19: Future research implications. *Journal of Dental Sciences*, 15(3):364, 2020.

[44] Syed Ghulam Sarwar SHAH et Alexandra FARROW : A commentary on "world health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19)". *International journal of surgery (London, England)*, 76:128, 2020.

[45] Soo LIM, Jae Hyun BAE, Hyuk-Sang KWON et Michael A NAUCK : Covid-19 and diabetes mellitus: from pathophysiology to clinical management. *Nature Reviews Endocrinology*, 17(1):11–30, 2021.

[46] Meg MILLER : 2019 novel coronavirus covid-19 (2019-ncov) data repository: Johns hopkins university center for systems science and engineering. *Bulletin-Association of Canadian Map Libraries and Archives (ACMLA)*, (164):47–51, 2020.

[47] Xintao LI, Bo GUAN, Tong SU, Wei LIU, Mengyao CHEN, Khalid Bin WALEED, Xumin GUAN, Tse GARY et Zhenyan ZHU : Impact of cardiovascular disease and cardiac injury on in-hospital mortality in patients with covid-19: a systematic review and meta-analysis. *Heart*, 106(15):1142–1147, 2020.

[48] Diana C SANCHEZ-RAMIREZ et Denise MACKEY : Underlying respiratory diseases, specifically copd, and smoking are associated with severe covid-19 outcomes: a systematic review and meta-analysis. *Respiratory medicine*, 171:106096, 2020.

[49] Mária FÖLDI, Nelli FARKAS, Szabolcs KISS, Noémi ZÁDORI, Szilárd VÁNCSA, Lajos SZAKÓ, Fanni DEMBROVSZKY, Margit SOLYMÁR, Eszter BARTALIS, Zsolt SZAKÁCS *et al.* : Obesity is a risk factor for developing critical condition in covid-19 patients: a systematic review and meta-analysis. *Obesity Reviews*, 21(10):e13095, 2020.

[50] Kunal NANDY, Abhijeet SALUNKE, Subodh Kumar PATHAK, Apurva PANDEY, Chinmay DOCTOR, Ketul PUJ, Mohit SHARMA, Abhishek JAIN et Vikas WARIKOO : Coronavirus disease (covid-19): A systematic review and meta-analysis to evaluate the impact of various comorbidities on serious events. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(5):1017–1025, 2020.

[51] Ian HUANG, Michael Anthonius LIM et Raymond PRANATA : Diabetes mellitus is associated with increased mortality and severity of disease in covid-19 pneumonia–a systematic review, meta-analysis, and meta-regression. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(4):395–403, 2020.

[52] World Health ORGANIZATION *et al.* : Noncommunicable diseases country profiles 2014. 2014.

[53] World Health ORGANIZATION *et al.* : Estimating mortality from covid-19: scientific brief, 4 august 2020. Rapport technique, World Health Organization, 2020.

[54] G MARINONI, HV LAND et T JENSEN : The impact of covid-19 on higher education around the world: A global survey report. iau, 2020.

[55] K GREEN, S GRAZIADIO, P TURNER, T FANSHAWE et J ALLEN : The centre for evidence-based medicine develops, promotes and disseminates better evidence for healthcare, 2020.

[56] Yan ZHANG, Yanhui CUI, Minxue SHEN, Jianchu ZHANG, Ben LIU, Minhui DAI, Lingli CHEN, Duoduo HAN, Yifei FAN, Yanjun ZENG *et al.* : Association of diabetes mellitus with disease severity and prognosis in covid-19: a retrospective cohort study. *Diabetes research and clinical practice*, 165:108227, 2020.

[57] Hussein NAFAKHI, Mohammed ALAREEDH, Karrar AL-BUTHABHAK, Foaad SHAGHEE, Ahmed NAFAKHI et Samet KASIM : Predictors of adverse in-hospital outcome and recovery in patients with diabetes mellitus and covid-19 pneumonia in iraq. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 15(1):33–38, 2021.

[58] Tahir BELICE et Ismail DEMIR : The gender differences as a risk factor in diabetic patients with covid-19. *Iranian Journal of Microbiology*, 12(6):625, 2020.

[59] Antonio CERIELLO, Eberhard STANDL, Doina CATRINOIU, Baruch ITZHAK, Nebojsa M LALIC, Dario RAHELIC, Oliver SCHNELL, Jan ŠKRHA et Paul VALENSI : Issues of cardiovascular risk management in people with diabetes in the covid-19 era. *Diabetes Care*, 43(7):1427–1432, 2020.

[60] Akhtar HUSSAIN, Bishwajit BHOWMIK et Nayla Cristina do VALE MOREIRA : Covid-19 and diabetes: Knowledge in progress. *Diabetes research and clinical practice*, 162:108142, 2020.

[61] Mahdi SHADNOUSH, Soghra RABIZADEH, Alireza ESTEGHAMATI, Manouchehr NAKHJAVANI, Nasrin Baiat PARIDARI, Mostafa KHOSHABI, Armin RAJAB et Fatemeh GHAEMI : Covid-19 infection mortality risk in iranian patients with type 2 diabetes, hypertension and obesity. *Eastern Mediterranean Health Journal*, 28(3):221–224, 2022.

[62] Yanbin DU, Nan ZHOU, Wenting ZHA et Yuan LV : Hypertension is a clinically important risk factor for critical illness and mortality in covid-19: A meta-analysis. *Nutrition, Metabolism and Cardiovascular Diseases*, 31(3):745–755, 2021.

[63] Tadesse Awoke AYELE, Habtewold SHIBRU, Malede MEQUANENT SISAY, Tesfahun MELESE, Melkitu FENTIE, Telake AZALE, Tariku BELACHEW, Kegnie SHITU et Tesfa Sewunet ALAMNEH : The effect of covid-19 on poor treatment control among ambulatory hypertensive and/or diabetic patients in northwest ethiopia. *PloS one*, 17(5):e0266421, 2022.

[64] Hayder M AL-KURAISHY, Ali I AL-GAREEB, M ALBLIHED, Natália CRUZ-MARTINS et Gaber El-Saber BATIHA : Covid-19 and risk of acute ischemic stroke and acute lung injury in patients with type ii diabetes mellitus: the anti-inflammatory role of metformin. *Frontiers in Medicine*, page 110, 2021.

[65] Stefania NANNONI, Rosa de GROOT, Steven BELL et Hugh S MARKUS : Stroke in covid-19: a systematic review and meta-analysis. *International Journal of Stroke*, 16(2):137–149, 2021.

[66] Aidin RAWSHANI, Elin Allansson KJÖLHEDE, Araz RAWSHANI, Naveed SATTAR, Katarina EEG-OLOFSSON, Martin ADIELS, Johnny LUDVIGSSON, Marcus LINDH, Magnus GISSLÉN, Eva HAGBERG *et al.* : Severe covid-19 in people with type 1 and

type 2 diabetes in sweden: A nationwide retrospective cohort study. *The Lancet Regional Health-Europe*, 4:100105, 2021.

[67] LE JOENSEN, KP MADSEN, L HOLM *et al.* : Research: educational and psychological aspects diabetes and covid-19 pandemic in people with diabetes in denmark-what characterizes people with high levels of covid-19-releated worries. *Diabet. Med*, pages 1–9, 2020.

[68] Livio LUZI, Michele CARRUBA, Roberta CRIALESI, Stefano DA EMPOLI, Regina DAGANI, Elisabetta LOVATI, Antonio NICOLUCCI, Cesare C BERRA, Elisa CIPPONERI, Ketty VACCARO *et al.* : Telemedicine and urban diabetes during covid-19 pandemic in milano, italy during lock-down: epidemiological and sociodemographic picture. *Acta Diabetologica*, 58(7):919–927, 2021.

[69] J BRADFORD, E COE, K ENOMOTO et M WHITE : Covid-19 and rural communities: Protecting rural lives and health. *McKinsey & Company [Internet]*, 2020.

[70] Richard Kobina Dadzie EPHRAIM, Evans DUAH, Charles NKANSAH, Samuel AMOAH, Emmanuel FOSU, Justice AFRIFA, Felix BOTCHWAY, Perditer OKYERE, Samuel ESSIEN-BAIDOO, Kofi MENSAH *et al.* : Psychological impact of covid-19 on diabetes mellitus patients in cape coast, ghana: a cross-sectional study. *The Pan African Medical Journal*, 40, 2021.

[71] Noha M ELEMAM, Haifa HANNAWI, Issa AL SALMI, Kashif Bin NAEEM, Fahdah ALOKAILY et Suad HANNAWI : Diabetes mellitus as a comorbidity in covid-19 infection in the united arab emirates. *Saudi Medical Journal*, 42(2):170, 2021.

[72] Abdene Weya KASO, Habtamu Endashaw HARERU, Taha KASO et Gebi AGERO : Time to recovery from covid-19 and its associated factors among patients hospitalized to the treatment center in south central ethiopia. *Environmental Challenges*, 6:100428, 2022.

[73] L NORDLING : Hiv and tb increase death risk from covid-19, study finds—but not by much. *Science*, 382(25):2411, 2020.

[74] Abas GHAYSOURI, Maryam SAFARI, Elham MOHAMMADYARI, Firooz BALAVANDI, Ali NAZARI, Jalil FEIZI, Mohamad KARIMIAN, Elham BASTANI, Ebrahim SALIMI et Hamed TAVAN : The recovery of an elderly woman with asthma and diabetes from covid-19 infection: a case report. *Infectious disorders drug targets*, 2021.

[75] Alamin ALKUNDI, Ibrahim MAHMOUD, Abdelmajid MUSA, Saima NAVEED et Mohammed ALSHAWWAF : Clinical characteristics and outcomes of covid-19 hospitalized patients with diabetes in the united kingdom: a retrospective single centre study. *Diabetes Research and Clinical Practice*, 165:108263, 2020.

[76] Yuanyuan FU, Ling HU, Hong-Wei REN, Yi ZUO, Shaoqiu CHEN, Qiu-Shi ZHANG, Chen SHAO, Yao MA, Lin WU, Jun-Jie HAO *et al.* : Prognostic factors for covid-19 hospitalized patients with preexisting type 2 diabetes. *International journal of endocrinology*, 2022, 2022.

[77] Bolin WANG, Ruobao LI, Zhong LU et Yan HUANG : Does comorbidity increase the risk of patients with covid-19: evidence from meta-analysis. *Aging (albany NY)*, 12(7):6049, 2020.

[78] Stephen P JENKINS : Survival analysis. *Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK*, 42:54–56, 2005.

[79] DV GLASS : Graunt's life table. *Journal of the Institute of Actuaries*, 76(1):60–64, 1950.

[80] Scott E GRAVES : Congress, litigants, and judicial review of bureaucracy: A competing risks model of administrative case terminations. *Available at SSRN 913840*, 2006.

[81] James N MCNAIR, Anusha SUNKARA et Daniel FROBISH : How to analyse seed germination data using statistical time-to-event analysis: non-parametric and semi-parametric methods. *Seed Science Research*, 22(2):77–95, 2012.

[82] Steven J STAFFA et David ZURAKOWSKI : Competing risks analysis of time-to-event data for cardiovascular surgeons. *The Journal of thoracic and cardiovascular surgery*, 159(6):2459–2466, 2020.

[83] David R COX : Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

[84] Guillermo SALINAS-ESCUDERO, María Fernanda CARRILLO-VEGA, Víctor GRANADOS-GARCÍA, Silvia MARTÍNEZ-VALVERDE, Filiberto TOLEDANO-TOLEDANO et Juan GARDUÑO-ESPINOSA : A survival analysis of covid-19 in the mexican population. *BMC public health*, 20(1):1–8, 2020.

[85] Monira MOLLAZEHI, Mohammad MOLLAZEHI et Abdel-Salam G ABDEL-SALAM : Modeling survival time to recovery from covid-19: a case study on singapore. 2020.

[86] Adinew Arficho HANDISO, Yasin NEGASH et Getachew Tekle MEKISO : Modeling time to death of hiv infected patients on antiretroviral therapy in case of hossana queen elleni mohammad memorial hospital, south ethiopia. *International Journal of Public Health*, 5(1):071–082, 2019.

[87] Ketema KEBEBEW et Eshetu WENCHEKO : Survival analysis of hiv-infected patients under antiretroviral treatment at the armed forces general teaching hospital, addis ababa, ethiopia. *Ethiopian Journal of Health Development*, 26(3):186–192, 2012.

[88] Kai LIU, Ying CHEN, Ruzheng LIN et Kunyuan HAN : Clinical features of covid-19 in elderly patients: A comparison with young and middle-aged patients. *Journal of Infection*, 80(6):e14–e18, 2020.

[89] Amer S ALALI, Abdulaziz O ALSHEHRI, Ahmed ASSIRI, Shahd KHAN, Munirah A ALKATHIRI, Omar A ALMOHAMMED, Waleed BADOGHAISH, Saeed M ALQAHTANI, Musaad A ALSHAMMARI, Mohamed MOHANY *et al.* : Demographics, comorbidities, and outcomes among young and middle-aged covid-19 patients in saudi arabia. *Saudi Pharmaceutical Journal*, 29(8):833–842, 2021.

[90] David G KLEINBAUM et Mitchel KLEIN : *Survival analysis*. Springer, 2004.

[91] Takeshi EMURA et Yi-Hau CHEN : *Analysis of survival data with dependent censoring: Copula-Based Approaches*. Springer, 2018.

[92] Taane G CLARK, Michael J BRADBURN, Sharon B LOVE et Douglas G ALTMAN : Survival analysis part i: basic concepts and first analyses. *British journal of cancer*, 89(2):232–238, 2003.

[93] John P KLEIN et Melvin L MOESCHBERGER : *Survival analysis: techniques for censored and truncated data*, volume 1230. Springer, 2003.

[94] Dhananjay KUMAR et Bengt KLEFSJÖ : Proportional hazards model: a review. *Reliability Engineering & System Safety*, 44(2):177–188, 1994.

[95] Susanne MAY et David W HOSMER : Hosmer and lemeshow type goodness-of-fit statistics for the cox proportional hazards model. *Handbook of statistics*, 23:383–394, 2003.

[96] DW HOSMER : lemeshow, s.[1999]: applied survival analysis: regression modeling of time to event data. *JohnWiley and sons, new York*.

[97] David SCHOENFELD : Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–241, 1982.

[98] Giuseppe PUGLIESE, Martina VITALE, Veronica RESI et Emanuela ORSI : Is diabetes mellitus a risk factor for coronavirus disease 19 (covid-19)? *Acta diabetologica*, 57(11):1275–1285, 2020.

[99] Abdene Weya KASO, Habtamu Endashaw HARERU, Taha KASO et Gebi AGERO : Factors associated with poor treatment outcome among hospitalized covid-19 patients in south central, ethiopia. *BioMed Research International*, 2022, 2022.

[100] Sex AGE : Existing conditions of covid-19 cases and deaths. *WorldoMeter. URL: https://www. worldometers. info/coronavirus/coronavirus-age-sex-demographics/[accessed 2020-02-29]*, 2020.

[101] Pinar SISMAN, Irmak POLAT, Ensar AYDEMIR, Remzi KARSI, Ozen Oz GUL, Soner CANDER, Canan ERSOY et Erdinc ERTURK : How the covid-19 outbreak affected patients with diabetes mellitus? *International Journal of Diabetes in Developing Countries*, 42(1):53–61, 2022.

Appendix I

A The below table shows the Kaplan-Meier estimation of survival function of diabetic plus COVID-19.

Table 5.1: Estimates of survival function of DM plus COVID

Time	No. at risk(n_i)	Event/recovery	survival	SE of S(t)	95% CI	
			S(t)	time(S(t))	Lower	Upper
5	450	2	0.996	0.003	0.989	1.000
6	445	4	0.987	0.005	0.976	0.997
7	429	13	0.957	0.010	0.938	0.976
8	400	18	0.914	0.014	0.887	0.941
9	373	2	0.909	0.014	0.882	0.937
10	366	10	0.884	0.016	0.85384	0.9151
11	344	10	0.858	0.017	0.825	0.893
12	319	18	0.810	0.020	0.772	0.849
13	293	23	0.746	0.022	0.704	0.791
14	257	49	0.604	0.026	0.556	0.656
15	194	25	0.526	0.027	0.477	0.581
16	161	21	0.458	0.027	0.408	0.514
17	139	15	0.408	0.027	0.359	0.465
18	119	16	0.353	0.027	0.305	0.409
19	101	7	0.329	0.026	0.281	0.385
20	93	13	0.283	0.026	0.237	0.338

21	78	8	0.254	0.025	0.209	0.308
22	67	6	0.231	0.024	0.188	0.284
23	58	7	0.203	0.024	0.165	0.255
24	50	3	0.191	0.023	0.151	0.242
25	45	5	0.170	0.023	0.131	0.220
26	40	7	0.140	0.021	0.104	0.188
28	31	4	0.123	0.020	0.088	0.169
29	27	3	0.108	0.020	0.076	0.154
30	24	1	0.104	0.019	0.072	0.149
31	22	3	0.090	0.018	0.060	0.134
32	17	1	0.085	0.018	0.056	0.128
33	16	5	0.058	0.016	0.034	0.099
35	10	1	0.052	0.015	0.030	0.092
37	9	1	0.047	0.015	0.025	0.086
42	7	1	0.040	0.014	0.020	0.079
48	5	1	0.032	0.013	0.014	0.072
56	4	2	0.016	0.010	0.005	0.057
59	2	1	0.008	0.008	0.001	0.052

B. The following table shows that the analysis of median recovery of diabetic patients from COVID-19.

Table 5.2: Diabetes mellitus with COVID-19 patients baseline covariates of median recovery, percentage and frequencies

Covariate	Categories	Recovery (%)	Median(%)	Total	95% CI
Age	18-35	71(83.53%)	11	85(17.67%)	(11,12)
	36-55	113(79.02%)	15	143(29.73%)	(15,16)
	>55	122(42.22%)	19	253(52.6%)	(18,21)
Sex	Male	179(61.3%)	17	292(60.71%)	(15,18)
	Female	127(67.2%)	15	189(39.29%)	(14,16)
Residence	Urban	254(74.05%)	17	343(71.31%)	(16,18)
	Rular	52(37.68%)	10	138(28.69%)	(8,10)
HIV	Negative	269(62.13%)	15	433(90.02%)	(15,16)
	Positive	37(77.08%)	31	48(9.98%)	(28,33)
Hypertension	No	230(92%)	14	250(51.98%)	(14,14)
	Yes	76(32.9%)	26	231(48.03%)	(23,28)
CVD	No	245(74.02%)	15	331(68.82%)	(14,15)
	Yes	61(40.67%)	24	150(31.19%)	(21,26)
Symptom	No	43(37.39%)	8	115(23.91%)	(8,14)
	Yes	263(71.85%)	17	366(76.09%)	(16,18)
Asthma	No	260 (87.28%)	14	316 (65.7%)	(14,15)
	Yes	46(27.88%)	28	165(34.3%)	(26,32)

RVI	No	239(72.42%)	15	330(68.61%)	(14,16)
	Yes	67(44.37%)	22	151(31.39%)	(20,26)
Stroke	No	263(72.06%)	15	365(75.88%)	(14,16)
	Yes	43(37.07%)	26	116(24.12%)	(23,31)
TB	No	254(83.83%)	14	303(62.99%)	(14,15)
	Yes	52(29.21%)	26	178(37.01%)	(25,31)
CLD	No	263 (68.31%)	14	385(80.04%)	(14,15)
	Yes	43(44.79%)	28	96(19.96%)	(25,31)
Obesity	No	273(72.8%)	14	375(77.96%)	(14,15)
	Yes	33(31.13%)	29	106(22.04%)	(26,33)
TDM	TypeI	62(56.36%)	12	110(26.85%)	(11,12)
	TypeII	224(77.51%)	18	289(60.08%)	(17,18)
	Gestational	20(24.39%)	7	82(17.05%)	(7,_)
Other co-factor	No	208(69.57%)	14	299(62.16%)	(14,15)
	Yes	98(53.85%)	22	182(37.84%)	(20,25)

C. The following figure (8) shows the plot of the estimate of Kaplan-Meier survivor function of diabetics with COVID-19 patients by categories of covariates

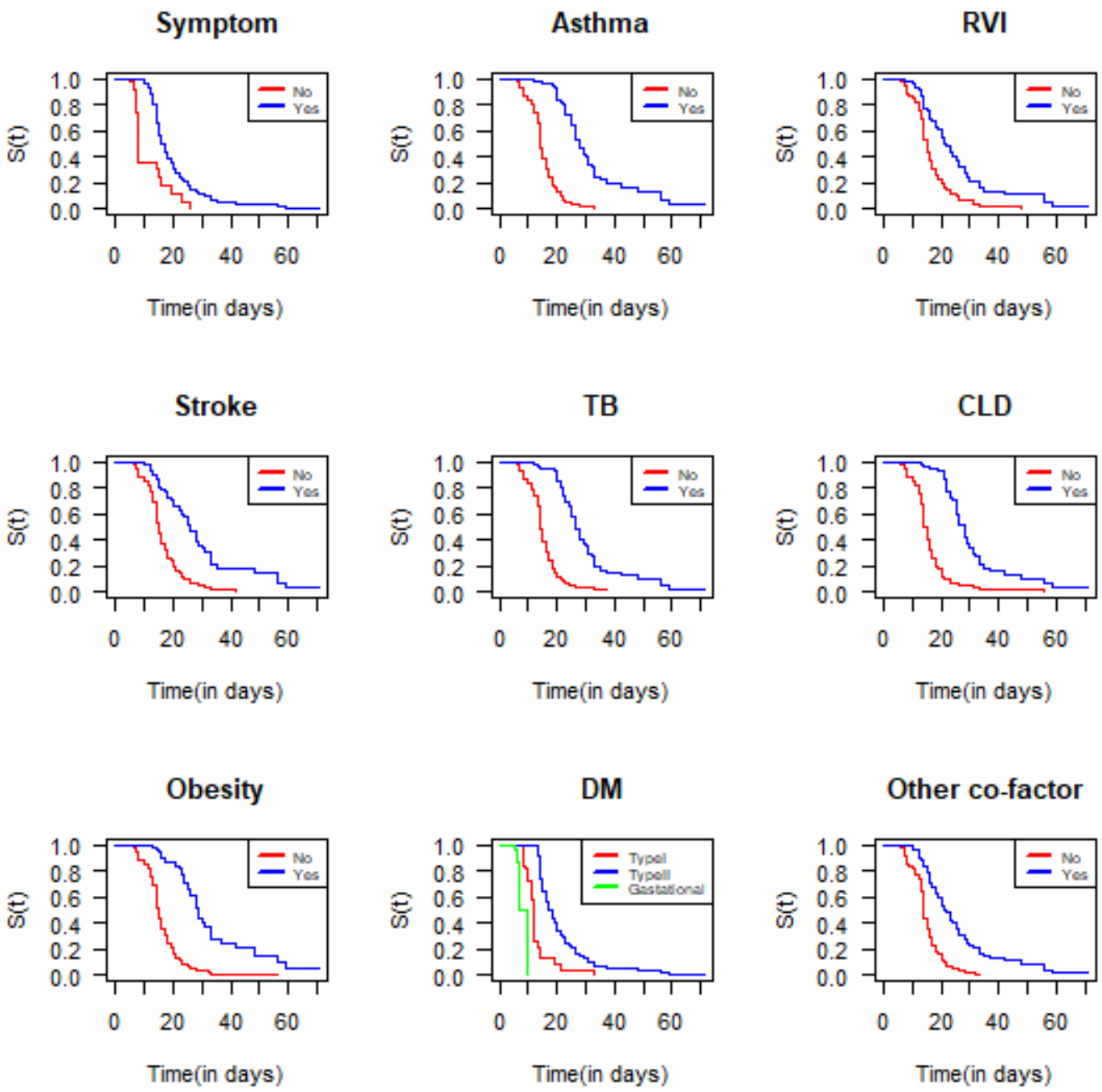

Figure 8: Kaplan-meier survival estimates by Symptom, Asthma, RVI, Stroke, TB, CLD, Obesity, Types of DM, and other co-factors

Appendix II

A. The following figure shows $\log(-\log(survival))$ versus a time is parallel for categorical variables Asthma, Stroke, TB, CLD, Types of DM & other co-factor are used to check proportional hazard assumption.

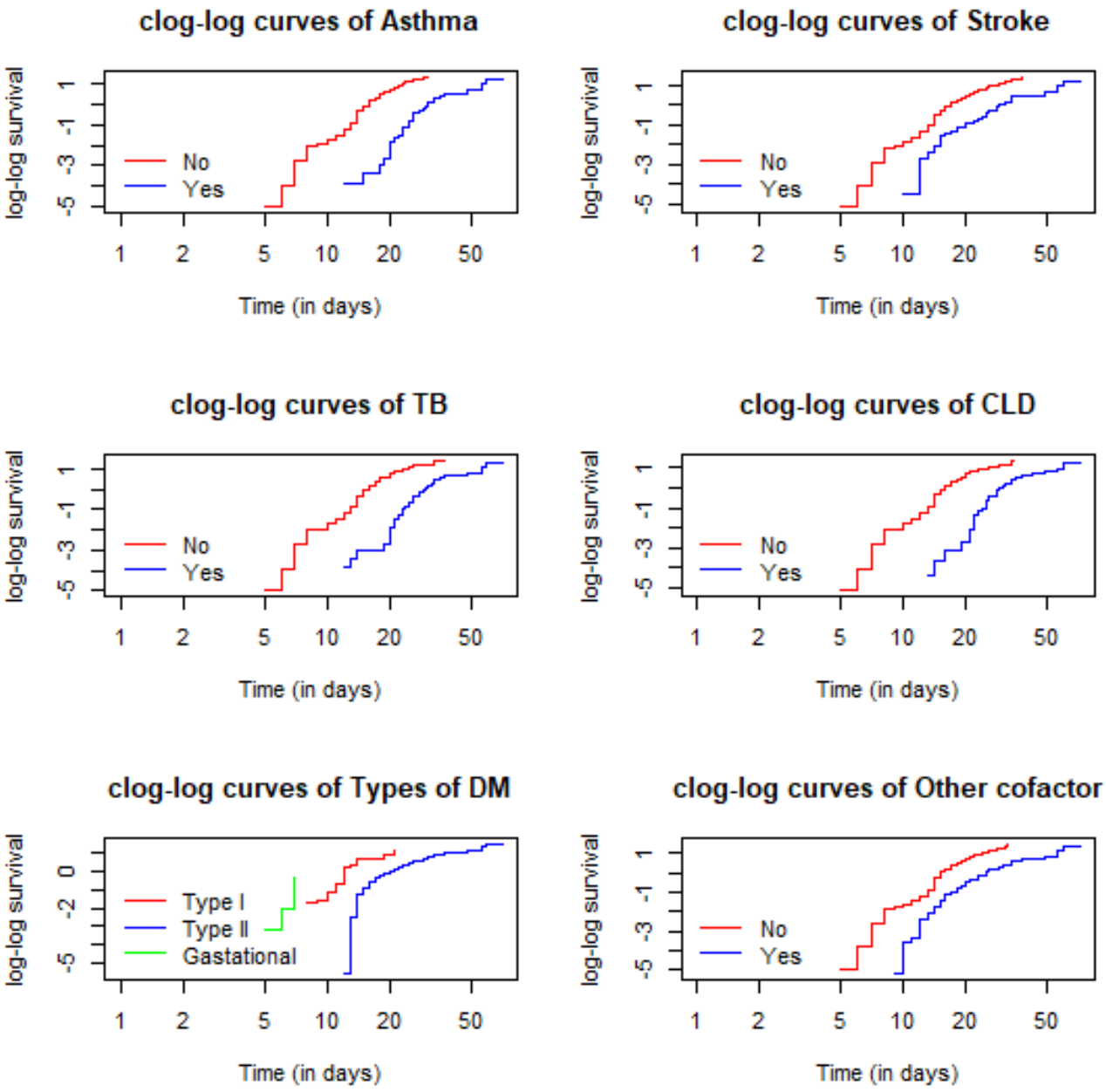

Figure 9: Log(-log(survival)) plot against time (Asthma, Stroke, TB, CLD, Types of DM & other co-factor)

B. Plot of scaled Schoenfeld residual for test of proportionality assumption.

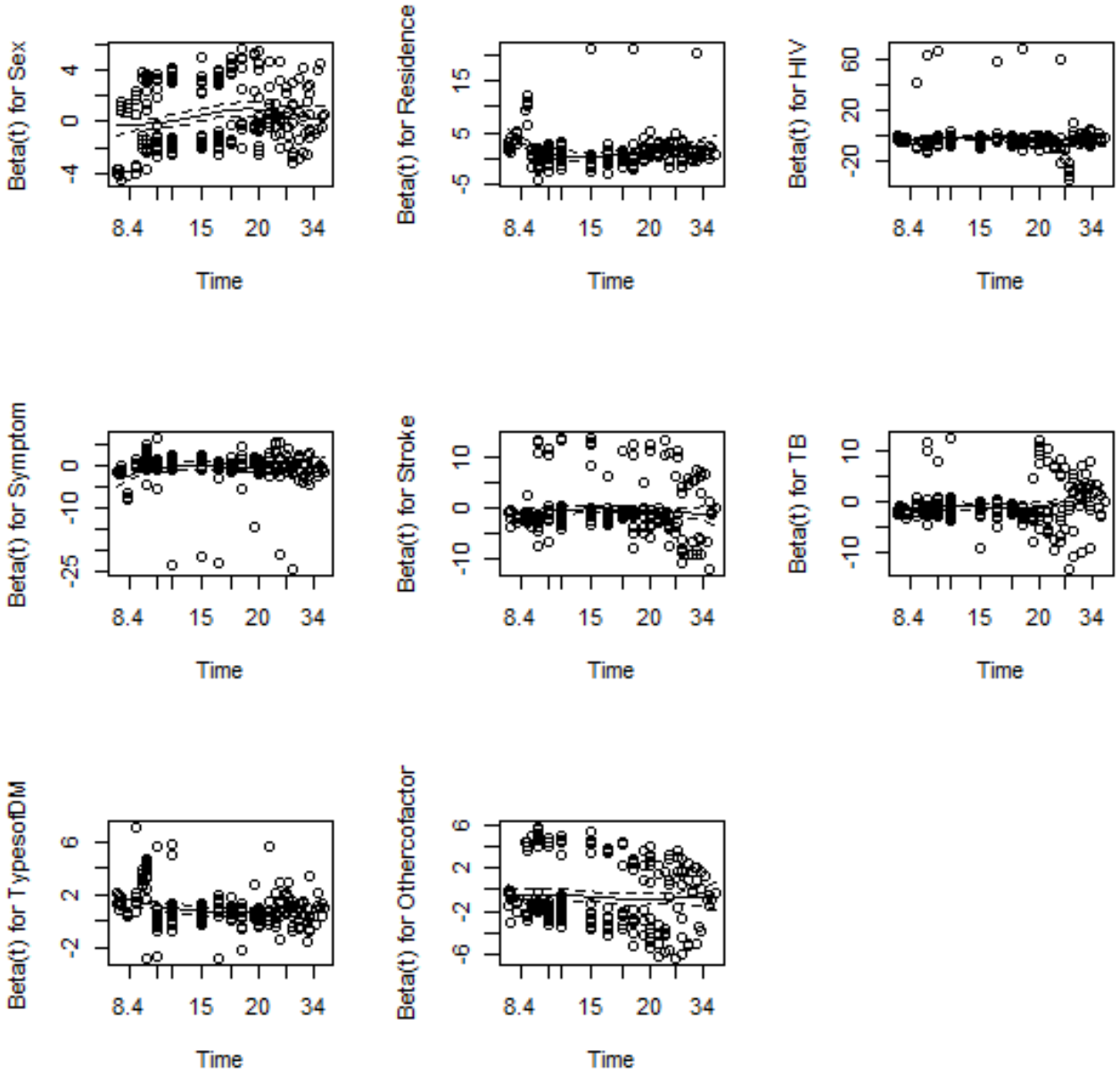

Figure 10: Scaled Schoenfeld Residual Plots for Sex, residence, HIV, symptom,stroke, TB, types of DM for, and other co-factors

Table 4.5: Results of the multivariate proportional hazards Cox regression model of DM plus Covid-19

Co-variate	Categories	Coef ($\hat{\beta}$)	SE (Coef($\hat{\beta}$))	P-value	Hazard ratio ($\hat{HR}$)	95% CI ($\hat{HR}$)
Age	18-35 (Ref.)	-	-	-	-	-
	36-55	-0.536	0.189	0.005	0.585	(0.404, 0.848)
	>55	-0.410	0.204	0.044	0.663	(0.445, 0.989)
Sex	Male (Ref.)	-	-	-	-	-
	Female	0.324	0.135	0.017	1.382	(1.061, 1.8)
Residence	Urban (Ref.)	-	-	-	-	-
	Rural	1.577	0.272	< 0.001	4.839	(2.82, 8.247)
HIV	Negative (Ref.)	-	-	-	-	-
	Positive	-2.709	0.517	< 0.001	0.067	(0.024, 0.182)
Hypertension	No. (Ref.)	-	-	-	-	-
	Yes	-0.967	0.204	< 0.001	0.38	(0.255, 0.567)
CVD	No. (Ref.)	-	-	-	-	-
	Yes	-0.211	0.168	0.211	0.81	(0.582, 1.127)
Symptom	No. (Ref.)	-	-	-	-	-
	Yes	-0.666	0.275	0.015	0.514	(0.3, 0.881)
Asthma	No. (Ref.)	-	-	-	-	-
	Yes	-0.712	0.236	0.003	0.491	(0.309, 0.78)
RVI	No. (Ref.)	-	-	-	-	-
	Yes	0.269	0.174	0.122	1.308	(0.931, 1.84)
Stroke	No. (Ref.)	-	-	-	-	-
	Yes	-0.677	0.216	0.002	0.508	(0.333, 0.776)
TB	No.(Ref.)	-	-	-	-	-
	Yes	-0.858	0.218	< 0.001	0.424	(0.277, 0.65)
CLD	No. (Ref.)	-	-	-	-	-
	Yes	-0.573	0.234	0.015	0.564	(0.356, 0.893)
Obesity	No. (Ref.)	-	-	-	-	-
	Yes	0.151	0.239	0.526	1.163	(0.729, 1.858)
TDM	Type I (Ref.)	-	-	-	-	-
	Type II	-1.841	0.202	< 0.001	0.159	(0.107, 0.236)
	Gestational	3.735	0.769	< 0.001	41.875	(9.279, 188.976)
Other co-factor	No. (Ref.)	-	-	-	-	-
	Yes	-0.634	0.147	< 0.001	0.531	(0.398, 0.708)

Source: Eka Kotebe General Hospital, Addis Ababa, Ethiopia; from March 20, 2020 to April 30, 2022. SE = Standard error, P-value < 0.05 = statistically significant, $\hat{HR}$ = estimated hazard ratio, Ref = Reference, CL = Confidence interval, and Coef($\hat{\beta}$) = Coefficient of estimated parameter β