College of Natural Sciences

Department of Statistics

Determinants of Time to First Recurrence of Women with Cervical Cancer Using Fine and Gray Model: A Case Study at Jimma University Medical Center.

By: Fayera Bayisa

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

September 26, 2022

Jimma, Ethiopia

Jimma University
College of Natural Sciences
Department of Statistics

Determinants of Time to First Recurrence of Women with Cervical Cancer Using Fine and Gray Model: A Case Study at Jimma University Medical Center

By: Fayera Bayisa

Advisor: Geremew Muleta (PhD Scholar)
Co-Advisor: Yasin Negash (MSc.)

September 26, 2022
Jimma, Ethiopia

Jimma University

School of Graduate Studies

Department of Statistics

Approval Sheet

As research advisors, we here by certify that we have read the research prepared by Fayera Bayisa Galata under our guidance, which is entitled **Determinants of Time to First Recurrence of Women with Cervical Cancer using Fine and Gray Model: A Case Study at Jimma University Medical Center**, in its final format it is consistent and acceptable. Hence we recommend that the research are accepted as it fulfills the university and department style requirements for the degree of Master of Science in Bio-statistics.

Geremew Muleta (PhD Scholar)		21/09/2022
Advisor	Signature	Date
Yasin Negash(MSc.)		23/09/2022
Co-advisor	Signature	Date

As the members of the board of examiners of M.Sc. thesis open defense examination, we certify 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.

Dr Habte Tadesse		21/09/2022
Name of External Examiner	Signature	Date
Dr. Tadele Akeba		23/09/2022
Name of Internal Examiner	Signature	Date
Abiyot Negash(Ass't Pro.)	For	23/09/2022
Name of Chairman	Signature	Date

Declaration

I declare that, this thesis is a result of my genuine work and all sources of materials used for writing it have been duly acknowledged. I have submitted this thesis to Jimma University in partial fulfillment for the Degree of Master of Science in Biostatistics. And also, I solemnly 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.

Fayera Bayisa Galata	.....	.....
Name	Signature	Date

Acknowledgement

First and for most, I would like to extend my unshared thanks to the almighty God for providing me the opportunity for what I have achieved and for his mercy. Next to this, I would like to express sincere appreciation to my Thesis advisor, Geremew Muleta (Ass't Prof. and PhD Scholar), for sharing his substantial experience to do this research in the expect way and giving wonderful personality through the time without any reservation time. I would like to express my heartfelt thanks to my co-advisor, Yasin Negash (M.Sc.), for his valuable comments and suggestion. My sincere thanks are due to the Jimma University Medical Center for giving me the data for study. My special appreciation and acknowledgment go to Assosa University for offering me the opportunity and financial support during my study at Jimma University.

Next, I want to express my gratitude to all of my brothers and sisters, but especially to Garoma and Urge Bayisa for standing by my side during difficult circumstances and providing moral support. Additionally, I would like to extend my sincere gratitude to all of my classmates for their invaluable assistance in all manners.

Abstract

Background: Cervical Cancer is a cancer arising from the cervix. The cancer recurrence is when cancer cells are detected following the initial treatment. When there are competing risks, techniques like traditional survival analysis that censor the competing event produce overestimate of the risks. The Fine-Gray model is favored in this case over other methods of survival analysis.

Objective: The aim of the study was to investigate the determinants of time to first recurrence of woman with cervical cancer.

Methods: A retrospective study was used to obtain data on women with cervical cancer that recorded in oncology department of Jimma Univeristy Medical Center. To reach the proposed objective, 280 women with cervical cancer were included in the study based on data taken from medical record card of patients enrolled starting from 1^{st} January 2017 to 31^{st} December 2021. Fine-Gray model were used to identifies which factor significantly affect time to first recurrence of cervical cancer by taking into account the occurrence of death as competing events.

Results: Of 280 cervical cancer patients, 60 (21.4%) experienced first recurrence, 64 (22.9%) died without recurrence, and 156 (55.7%) experienced censored. The International Federation of Gynecology and Obstetrics stage IV(SHR=3.71, 95%CI: 1.02-13.47, P=.046), smoker(SHR=3.34, 95%CI:1.43- 7.81, P=.0053), HIV positive(SHR=2.08, 95%CI:1.09- 3.94, P=.0058), age at diagnosis $\geq$ 50 years (SHR=0.29, 95%CI:0.09- 0.9) and oral contraceptives users(SHR=2.2, 95%CI:1.09- 4.47, P=.029) were independently associated with recurrence of cervical cancer.

Conclusion and recommendation: The International Federation of Gynecology and Obstetrics stage(IV), parity(Multipara and Grand multipara), HIV positive, smoker and oral contraceptive users increased the risk of recurrence. Age at diagnosis of $\geq$ 50 years and used chemo for two cycles, and three and more than three cycles decreases the risk of recurrence. To decrease the recurrence of cervical cancer, it is advised that policymakers, the ministry of health, and Jimma University Medical Center pay attention to individuals who are at a more advanced stage, smokers, HIV positive, and women who have several children.

Keywords: Cervical cancer, Competing risk, Cumulative incidence function and Recurrence

Table of Contents

List of Tables

List of Figures

List of Abbreviations

SHR	Sub-hazard Ratio
ACS	American Cancer Society
AIDS	Acquire Immune Deficiency Syndrome
CC	Cervical Cancer
CI	Confidence Interval
CIF	Cumulative Incidence Function
Cox-PH	Cox Proportional Hazards
DNA	Deoxyribonucleic acid
FIGO	International Federation of Gynecology and Obstetrics
HIV	Human Immune Virus
HPV	Human Papilloma Virus
HR	Hazard Ratio
JUMC	Jimma University Medical Center
KM	Kaplan-Meier
STI	Sexually Transmitted Infection
SSA	Sub-Saharan African
TASH	Tikur Anbessa Specialized Hospital
UK	United Kingdom
US	United States
WHO	World Health Organization

1 Introduction

1.1 Background of the study

Cervical Cancer(CC) is a cancer arising from the cervix(Raza, 2021). It is due to the abnormal growth of cells that have the ability to invade or spread to other parts of the body. It is caused primarily by a sexually transmitted infection with human papillomavirus (HPV), with which people become infected shortly after the onset of sexual activity(Nallbani & Agolli, 2022). The cervix is the lower part of the uterus, the place where a baby grows during pregnancy(Amabebe et al., 2022).

Worldwide, CC is the fourth most frequent cancer in women with an estimated 604,000 new cases and 342,000 deaths in 2020, about 90% of these occur in low- and middle-income countries(WHO, 2020). In developed countries, such as the United Kingdom(UK) and the United States (US), the incidence of CC has fallen dramatically since the 1960s, owing to the implementation of population-wide screening programmes, cytology-based, using HPV DNA testing more recently(Jedy-Agba et al., 2020).

In contrast, the incidence of CC in developing countries continues to rise due to the absence of an effective population-level screening programs, and a lack of knowledge about transmission routes, poor awareness about prevention, inequitable access to health services, poverty, and low socioeconomic status(Ginsburg et al., 2017). In Africa, where 267.9 million women aged 15 and older are at risk of having CC, 80,000 women are diagnosed with the disease, and slightly over 60,000 of them pass away every year(Masekwameng, 2020). In Ethiopia, there are nearly 26 million women who are over the age of 15 and believed to be at risk of getting HPV(Demissie et al., 2022). About 35.9 new cases of CC are diagnosed and 22.6 die from it, per 100,000 women annually(Asseffa, 2017).

According to different studies revealed, there are a variety of treatments for invasive CC patients(Monk et al., 2022). Those are surgery, radiation therapy, chemotherapy, or any combination of those. Despite these various available treatments, many patients experience recurrences after primary treatment(Li et al., 2022). A cancer recur when cancer cells are detected following the initial treatment at the place of origin or another part of the human body. Once the recurrence happens, the patients is left with very limited treatment options

and a risk of death for patients in general(Baiocchi et al., 2022). Shi et al. (2022) pointed out that CC survivors were at great risk of developing recurrence, and these patients were even more likely to die from their recurrence than from their initial cancers. Reducing the recurrence of CC requires an improved understanding of the actual recurrence rate, time to recurrence, and its related risk factors(Gennari et al., 2022).

After standard treatment, the recurrence rates of International Federation of Gynecology and Obstetrics (FIGO) reported that a 5 year recurrence rate of 28% and an overall mortality rate of 27.8% for females with CC(Okubo et al., 2021). According to some research, patients with advanced CC have a recurrence incidence of up to 70%(Chao et al., 2020). According to a study done in Ethiopia by Jaleta & Mokonnon (2018) at Tikur Anbessa Specialized Hospital(TASH), the 5-year recurrence rate of CC was 21.7 %. Numerous variables, which have prognostic relevance in the context of recurrence and thus alter patients' survival times, have been found to have prognostic significance in prior research(Origoni et al., 2022).

Survival data, failure time data or lifetime data are different names to describe data that deal with the time to an event(Wang et al., 2022). Survival analysis is a family of statistical techniques aimed at analyzing time-to-event data and/or assessing the relationship between a given exposure and the occurrence of an outcome after a follow-up period among a cohort of individuals(Abd ElHafeez et al., 2021). The specific difficulties relating to survival analysis arise largely from the fact that only some individuals have experienced the event and, subsequently, survival times will be unknown for a subset of the study group(Wreede et al., 2022). This phenomenon is called censoring. It may occur because the patient withdraws from a study, is lost to follow-up, or did not experience the event of interest before the end of the study(Xue et al., 2017). Standard survival models, such the Kaplan-Meier (KM) estimator, logrank test, and Cox Proportional Hazards (Cox-PH) regression, are frequently used when censoring is present, assuming that censoring is non-informative(Su et al., 2022).

In survival analysis, there are situations where the observation is not suitable for standard survival method which is commonly used time to event analysis(Austin & Fine, 2017). One such situation is when an individual can experience more than one type of event, and these events prevent other events from occurring. In general, this situation is called a competing risks(Ainurrochmah et al., 2021). A competing risk is, by definition, an event that either

hinders the observation of the event of interest or modifies the chance that this event occurs(Buzkova, 2021).

In the presence of competing risk, the effect of covariates on cause-specific hazard can be estimated with the cause-specific Cox-PH regression model(Schuster et al., 2020). This model assumes the same functional relationship between the cause-specific hazard function and covariates as the popular Cox-PH model for survival data without competing risks does for the relationship between the overall hazard and covariates(Wolbers et al., 2014). Even though the Cox-PH model is used to analyze the relationship between exposure and outcome, it have limitations(Noroozi et al., 2022). The major limitation of using cause-specific PH model in a competing risk setup is that during estimation of regression parameters under a specific cause it considers the individuals failing from causes other than cause of interest as censored observations(Mohammad et al., 2017). This difficulty led to the development of regression models that do not censor competing risks. Fine and Gray models were recommended for this issue(Scheike et al., 2022).

Fine-Gray regression model is based on an alternative failure rate summary measure, the sub-distribution hazard function(Bryson et al., 2021). The sub-distribution hazard for a specific cause is the instantaneous rate of experiencing that particular cause given the individual have not yet experienced failure from that cause(Chandra & Rehman, 2021). With the sub-distribution hazard, subjects who fail from competing risks cause remain in the risk set(Hsu et al., 2017). This is in contrast to the cause-specific hazards approach which censors such patients at the time of occurrence of the competing event. In Fine-Gray model, there is a direct link between the sub-distribution hazards and CIF(Ghosh et al., 2021).

1.2 Statement of the Problem

Despite the fact that HPV infection is the necessary cause in the etiology of CC recurrence, HPV infection alone is not a sufficient cause for the occurrence of cases(Śniadecki et al., 2022). Hence, there must be the factors that lead to the development of CC recurrence. To identify the factors, most of the epidemiologic research has been done in developed countries where CC recurrence declined significantly in the last three decades(Chao et al., 2020; Taarnhøj et al., 2018). But there is limited evidence about the extent of which of these factors

are prevalent in developing countries like ours.

Even though a variety of research has been done on factors that lead to a recurrence of CC using standard survival analysis, this approach in the presence of competing risks imposes additional challenges for clinical investigators(Zhang, 2017). The common survival models, such as the KM estimator, log rank test, and Cox-PH regression, make the assumption that censoring is ”non-informative,” which means that subject censoring times and event timings should be independent(Feakins et al., 2018). However, the assumption of non-informative censorship is invalid if the occurrence of one event eliminates the possibility of the occurrence of an event of interest(Donoghoe & Gebski, 2017). In presence of this event, conventional survival techniques overestimated the risks as compared to methods that take into account the competing risks because it censored the competing risks(Schmid & Berger, 2021).

There are studies that were conducted on the time to recurrence of women with CC. For instance, Li et al. (2022) performed a retrospective study using Cox-PH model for time to recurrence of women with CC. Since they considered the patients who experienced death as censoring, alternative methods are required that are specifically designed for analyzing competing risk data. The researcher used the Fine and Gray model to fill this gap. Unlike the standard survival approach, the Fine and Gray model considers death as a competing event rather than a censored event(Schellenberg et al., 2022). Generally, since the researcher did not yet find a study performed on determinants of time to the first recurrence of women with CC using the Fine and Gray model at the national level and the cases under study are found to be really a predominant issue, it happened to be a reason to conduct this study.

Thus, this study addressed the following research questions:

- What are the factors that significantly affect the time to first recurrence of CC patients?
- Is there difference between the cumulative incidence of groups of covariates?

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of the study is to assess the time to first recurrence of CC patients at Jimma University Medical Center by using Fine and Gray Model.

1.3.2 Specific Objectives

The specific objectives of the study are

- To identify the risk factors associated with the time to the first recurrence of CC patients.
- To determine whether there is a difference between the cumulative incidence of groups ofcovariates.

1.4 Significance of the Study

CC is one of the common gynecological malignancies with a high recurrence rate after initial treatment(Miccò et al., 2022). However, there is insufficient data regarding the prognostic variables that raise the risk of CC recurrence generally in Ethiopia. This study aims to identify prognostic factors which play a critical role in the time to first recurrence of women with CC. For academicians or statisticians, it will be direct to thoughts and genuine interest in the subject matter for further research, especially when competing events preclude the occurrence of an event of interest. Moreover, the study can be an input to policy makers, program managers, and health professionals to decide based on evidence about the recurrence of CC and serve as a base line data for further studies.

1.5 Limitations of the Study

Some of the limitations of the study are:-

- The study was conducted based on secondary data, which might have incomplete and biased information.
- As the data is gathered from patients' cards, the study has a limited number of variables considered as risk factors for the recurrence of women with CC.
- The JUMC recently started providing radiotherapy services for patients, so patients who received radiotherapy only or radiotherapy with other treatment during follow-up were not included in this study.

1.6 Organization of the Study

This study is presented in five chapters. The first chapter gives a general background of the study; a statement of the problem; an objective; its significance; and limitations of the study. Chapter 2 deals with the review of literature on the recurrence of CC in Ethiopia and the rest of the world, whereas chapter 3 specifies the data and methodology of the study, such as sources of data and variables to be included in the study with their coding and description. Methods of data analysis are also described in this chapter. Chapter 4 reports results from the statistical data analysis and provides discussions. Finally, the last chapter presents a conclusion and policy recommendations based on the findings of the study.

2 Literature Review

2.1 Overview of CC Disease

The cancer starts when cell in the body begin to grow out of control(Bozorgpour et al., 2021). Cells nearly any part of the body can become cancer and can spread to other area of the body(Soni & Soni, 2021). CC is a disease that results from failure of the mechanisms that regulate normal cell growth and cell death leading to uncontrollable proliferation of cervical cells(Zhu et al., 2021). This cancer can affect the deeper tissues of their cervix and often the lungs, liver, bladder, vagina, and rectum(Bhatla et al., 2021). It has a bad prognosis as it is frequently diagnosed in advanced stages of disease, while CC identified in the early stages has a good prognosis(Antunes & Cunha, 2013). It is curable disease if detected early and adequately treated(WHO, 2022).

Yet it remains one of the most common cancers cases and causes of cancer-related death in women across the globe(Khurshid et al., 2022). A study conducted by Arbyn et al. (2020) show that the annual number of new cases of CC has been projected to increase from 570,000 to 700,000 between 2018 and 2030 with the annual number of deaths projected to increase from 311,000 to 400,000. About more than 85% of those affected are young, women who live in the world's poorest countries(Mailhot Vega et al., 2019). CC is the second most commonly diagnosed cancer among Ethiopian women, killing an estimated 4700 women each year(Burrowes et al., 2022). As the government rolls out the countrys first national cancer control strategy, information on patient and provider experiences in receiving and providing CC screening, diagnosis, and treatment is critical.

2.2 Risk Factors of Cervical Cancer Recurrence

Age at diagnosis is independent risk factors of CC recurrenceLi et al. (2022). They conducted a long-term prospective cohort study on time to recurrence of CC using cox-PH regression model. Their findings revealed the risks of recurrence for age group $\geq$55 is 0.63 [HR= 0.63, 95% 0.45-0.89] times lower than that of less than 55 years. Furthermore, Gurmu (2018) conducted the study using retrospective study design at TASH to assessing survival time of women with CC shows that age at diagnoses has significance effect on risk of death of CC

patients. Finally, the study conducted by Oga et al. (2016) on the recurrence of cervical intraepithelial lesions after thermo-coagulation in HIV-positive and HIV-negative Nigerian women shows that women aged 30 years or older were much less likely to develop recurrence (HR=0.34,95%CI:0.13-0.93) as compared to those younger age.

Smoking is an important risk factor for the development of several squamous cell cancers, and smokers often present with more advanced tumor stages(Agarwal, 2021). A retrospective study conducted by Lu et al. (2006) shows that the odds of recurrence is 3.945 for smoker (HR, 3.945;95% CI: 1.545- 10.35; p=0.0044) as compared to non smoker. A retrospective study conducted by Jaleta & Mokonnon (2018) on recurrence of CC among patients under follows up since 2012 to 2015 in Ethiopia shows that the hazard of smoker is 1.711 times greater than that of non-smoker (HR=1.711, 95% CI:1.071-2.732, p=0.024). Finally, he concluded that women who smoke are about twice as likely as non-smokers to get CC.

HIV is the most common risk factors among women with CC(Stelzle et al., 2021). A study conducted by Béhanzin et al. (2022) on perceptions and knowledge about CC among women living with HIV in parakou using Cox-PH model. Their finds revealed that the patients with sero-positive HIV status were 2.8 times (HR=2.8; 95% CI:2.14-7.65) more likely to have recurrence of CC disease as compared to those with sero-negative HIV status. Addis (2010) conducted a retrospective study to identify the combating CC in Ethiopia. This finds showed that women living with HIV are more readily infected with certain types of HPV, more likely to develop precancerous lesions, and more vulnerable to rapid development of recurrence than HIV-negative women. Lodi et al. (2011) conducted a retrospective study on the factors associated with recurrence of cervical intraepithelial neoplasia after conization in HIV-infected and noninfected women. Their results shows that recurrence occurred in 75.6% of HIV-infected women versus 24.4% of non-HIV infected women (p = 0.001). In contrast to non-HIV, odds of recurrence were 4.29 (HR=4.29, 95% CI: 1.72-10.1) in women with HIV.

Tumors were staged according to guidelines set by the FIGO(Pecorelli et al., 2009). It has been categorized into four stages (I, II, III and IV) and reported as the most common risk factor for CC recurrence(Li et al., 2022). The study conducted in China reported that the risk of recurrence for subjects with clinical state II was 1.52 (95%CI 1.07-2.16) and clinical

stage III or IV was 1.84 (95%CI 1.17-2.90) compared to those with stage I(Li et al., 2022). In addition to this, Takehara et al. (2001) conducted the study on recurrence of invasive CC for more than 5 years after initial therapy. Their results shows that the probability of late recurrence in patients with stage I disease was significantly lower than that in stage II and stage III diseases (stage I compared with stage II, P= .038, stage I compared with stage III, P=.002).

There was much histology of cervical malignancy(Young, 2014). The predominant ones were squamous cell carcinoma and adenocarcinomas. The other histologies, such as small cell carcinoma, melanoma, and lymphoma, were included as different histology types. Ponce et al. (2020) conducted study the multi-center retrospective on risk factors for recurrence after robot-assisted radical hysterectomy for early-stage CC. Their result revealed that the risk of recurrence for adenocarcinoma is 2.51(1.03-6.07) times that of squamous cell carcinoma(Ponce et al., 2020). Another a retrospective study conducted by H. Li et al. (2016) revealed that the risk of recurrence is 2.25(HR = 2.25, 95% CI: 1.30-3.90)for non-squamous cell carcinoma as compared to squamous cell carcinoma.

Parity is the number of times the patients have given birth to a baby(Tidy & Payne, 2019). It was categorized into three groups, namely grand multipara (given birth more or equal to 5 times), multipara (given birth 2- 4 times), and primipara/ nullipara (for the patients gave birth once or never). Multiparity is believed to be a risk factor for CC, especially among human papilloma virus (HPV)-positive women(Eluf-Neto et al., 1994). A retrospective study conducted in Indonesia revealed that the women who have given birth more than 3 can increase the incidence of cancer by 3 times that of have 3 and below children(Teguh et al., 2021). Another retrospective study conducted by Sharma & Pattanshetty (2018) revealed that, the odds of recurrence was 4.55 times in women with parity ≥ 3 compared to women with parity 3. The study conducted on prognostic factors and relapse patterns in early-stage CC after brachytherapy and radical hysterectomy revealed that risk of recurrence is almost five times for those patients who had three and more children as compared to those had less than three children(relative risk, RR = 4.6)(Ye et al., 2022).

Family history of CC patients is also one factor that significantly predicts the survival of the patients. Women with a family history of CC, especially an affected mother or sister, have

a two-fold risk of developing CC, suggesting an inherited susceptibility(Shah et al., 2014). Bellinger et al. (2013) were conducted study on the role of family history of cancer on CC screening behavior in a population-based survey of women in the south-eastern united states founded those women who had no family history had long survival time than those who had family history. Some researchers suspect that some instances of this familial tendency are caused by an inherited condition that makes some women less able to fight off HPV infection than others(Sahoo et al., 2014).

Another most important risk factors of CC recurrence are treatment. The type of treatment applied individually or in combination depends on the stage of CC at time of diagnosis. For stage I CC, surgery such as conization or total/modified hysterectomy with internal radiation therapy is used. Within stage II CC, combinative radiation and chemotherapy following radical hysterectomy and removal of pelvic lymph nodes is often considered. For stage III CC combinative radiation and chemotherapy, followed by internal radiation therapy to shrink the tumor before full surgical hysterectomy and removal of pelvic lymph nodes, with follow-up chemotherapy often applied. Within stage IV CC chemotherapy and radiation therapy can be administered as palliative care to relive cancer symptoms, as well as for comfort. Other possible treatment options for stage IV CC however can include drastic surgical pelvic exenteration or clinical trials of targeted immunotherapies(Rydzewska et al., 2010; Brookfield et al., 2009).

A study was conducted on the time to recurrence of women with CC at TASH using a retrospective study design(ALTAYE, 2011). This study shows that the hazard rate for the time to recurrence of women with CC who took one, two, three, four, five, and six cycles of chemotherapy treatment was 0.083(HR = 0.083, 95%CI: 0.040-0.171, P=0.000), 0.507(HR = 0.507, 95% CI: 0.17-1.51, P=0.223), 0.382(HR = 0.382, 95% CI: 0.172-0.852, P=0.019), 0.214(HR=0.214, 95% CI=0.081-0.567,P=0.002), 0.432(HR=0.432,95% CI = 0.185-1.009, P=0.0052) and 0.801(HR=0.801, 95% CI=0.231-2.78,P= 0.727) as compared to patients who did not take the chemotherapy treatment, respectively.

Oral contraceptive (OC) pills are known to be a risk factor for CC recurrence. In an international collaborative epidemiological study of CC, the relative risk in current users increased with an increase in the duration of OC use. It has been reported that the use of OC for 5

years or more can double the risk of cancer(Cervical Cancer et al., 2007). And in a multi-center case-control study, among women who tested positive for HPV DNA, the risk of CC increased by 3 times for those who have used OC pills for 5 years or more(Muñoz et al., 2002). In addition, a recent systematic review and meta-analysis also suggested that OC pills use had a definite associated risk for developing CC especially for adenocarcinoma. Their study concluded that use of OC pills is an independent risk factor in causing CC(Asthana et al., 2020). The study conducted by Medeiros et al. (2005) compared the women without history of oral contraception, the risk in patients using oral contraception was increasing in accordance with duration of usage.

A retrospective study conducted in Ethiopia on the predictors of advanced stage and prolonged time to diagnosis of CC by Begoihn et al. (2019) shows the adjusted Hazard Ratio of CC for rural residence was 1.23 (CI: 1.11-1.36) as compared to those live in urban residence. The study conducted on the risk factors associated with CC by Panjaliya et al. (2015) revealed that maximum number of the patients belonged to the rural areas (81.6%) and 18.4% belonged to urban areas. They reports that the incidence of CC is higher among the patients living in the rural areas.

2.3 Overview of Competing Risk Models

The effect of competing risks was first acknowledged by d'Alembert and Bernoulli in the 1760s in relation to the effects of inoculation on short- and long-term mortality from smallpox (Messerli et al., 2013). In the presence of competing risk events, different models were proposed to take into account the relationship between the effect of predictors and the outcome of interest(Fine & Gray, 1999a).

Prentice et al. (1978) proposed the use of standard survival models like Cox-PH regression on the cause-specific hazard. In the cause-specific hazard model the effect of the investigated covariates on the competing event(s) is ignored, so there is no direct connection between the regression coefficients and the incidence of events. Fine & Gray (1999a) introduced a regression approach focusing on the so called sub-distribution hazard. In the Fine and Gray model the regression coefficients are monotonously linked to the CIF and the occurrence of competing events has an influence on the coefficients. The modified standard survival models can be fit to estimate the influence of the investigated covariates on the sub-distribution hazard.

In the last three decades, different articles in the presence of competing risks have been published. Saeedi et al. (2020) was conducted the study to determine the significant prognostic factors for the recurrence of pediatric Acute lymphoblastic leukemia by considering the first recurrence in children with Acute lymphoblastic leukemia to be the event of interest and non-relapse mortality to be the competing event. In addition to this Brandstorp-Boesen et al. (2016) conducted the study on the risk of recurrence in laryngeal cancer by taking recurrence as event of interest and death due to any cause as competing event by using competing risk model.

In addition to investigation of the factors that affect the recurrence of CC, we need to compute and compare the CIF estimate for groups of covariates. CIF of an event is often of interest in medical research and is frequently presented in medical articles. The KM method has been a widely used tool for estimating survival function and CIF(Kim, 2007). However, if there is more than one type of event (or failure), and if these events are dependent, KM estimates are biased. This bias arises because the KM method assumes that all events are independent, and thus, censors events other than the event of interest. To overcome the lim-

itation of the KM estimator, Prentice et al. (1978) introduced the CIF. The CIF approach partitions the probability of all events to their constituent probabilities such that at any point in time, the probability of all events is the sum of the probabilities of interest and those of the competing risks (Pintilie, 2007).

After estimating the CIF of an event, it is often of interest to determine whether there is a difference in the CIF among different groups. In standard survival analysis, this is done using the log-rank test with the KM method. The log-rank statistic is a non-parametric test based on the difference between the observed and expected numbers of events summed over all time points(Gaubatz et al., 2019). In the presence of competing risks, the CIF was compared by the gray's test rather than the log-rank test(Gray, 1988). Generally this study is conducted to assess the most risk factors of the CC recurrence by taking into account the patients those who died during follow-up period as competing risk.

3 Data and Methodology

3.1 Description of the study area

The study was conducted at Jimma University Medical Center(JUMC). JUMC is one of the oldest public hospitals in Ethiopia. It was established in 1930 E.C by Italian invaders for service of their soldiers. After the withdrawal of the colonial occupants, it has been governed by the name of Ras Desta Damtew Hospital and later Jimma Hospital during Dergue regime and currently it called JUMC. This time the hospital provides services for more than 2 million patients with 800 bedded. The hospital is located in Jimma city and, Jimma is the largest city in South-western of Oromiya Region at a distance of 355.2 Km from Addis Ababa, the capital city of Ethiopia. It has latitude and longitude of 7040'N 36050'E. Jimma has relatively cool tropical monsoon climate. The temperatures are in comfortable range, with the daily mean staying between $200°C$ and $250°C$ year-round.

The oncology department of the JUMC is providing chemotherapy, radiation therapy, complain therapy and other supportive and palliative cares. It is the main center for cancer registry, early detection, prevention, standard treatment and palliative care in Jimma and it is the only the second cancer center in the Ethiopia.

3.2 Study Design

A retrospective cohort study design was carried out to retrieve relevant information.

3.3 Study Population and Period

The CC patients were the source of the population for the study. The data was collected from the medical chart and patient's registration card in the Oncology department at JUMC starting from 1^{st} January 2017 to 31^{st} December 2021.

3.4 Inclusion and exclusion criteria

Inclusion criteria: All CC patients diagnosed and treated at JUMC starting from 1^{st} January 2017 to 31^{st} December 2021.

Exclusion criteria:- Patients with insufficient information about one of the vital variables either in the registration book or in the card were not eligible.

3.5 Data Collection Procedure

The data set used for this study was collected from patients' individual cards. All the data had been carefully reviewed from the registration log book and patients' registration cards. For the data collection, health professional and the experienced data collectors under the supervision of the researcher were contributed.

3.6 Data Structure for Time to First Recurrence of CC

The table below illustrates the standard way of representing competing risks data for determinants of time to first recurrence of CC as an event of interest and death due to any cause of CC patients as a competing risk event.

Table 3.1: Data Structure for the Time to First Recurrence of CC

Id	Time	Code	Event status	Treatment taken	Family History
1	5	1	Recurred	Chemotherapy	Yes
2	10	0	Censored	Combination	No
3	20	2	Death	Chemotherapy	yes
4	22	2	Death	Combination	yes
5	20	1	Recurred	Surgery	No

Where:

Id:- Patients

Time:- Time in months at which the event of interests, competing events or censoring events occurs from the day at patients registered at hospital.

Events status:- 1 event of interest, 2 competing risk event and 0 censoring.

Survival times in this data set are the actual time at first recurrence in months and death due to any causes of CC patients. Censoring is caused by refer to other hospital or end of the study. Treatment and the family history are covariates.

3.7 Study Variables

3.7.1 Response Variable

The response variable in this study is the time in months to the first recurrence of CC starting from the day the CC patients registered at the hospital.

3.7.2 Independent Variables

The predictor variables that are thought to influence the recurrence of CC in women are listed below:-

1. Smoking habit(Non-smoker, Smoker)
2. Age($\leq$34, 35-49, $\geq$50)
3. Family history(No, Yes)
4. HIV status(Negative, Positive)
5. Stage(I, II, III, IV)
6. Treatment taken (Chemotherapy, Surgery, Combination of two)
7. Parity(Nullipara, Multipara, Grand multipara)
8. Use oral contraceptive(No, Yes)
9. Histology type(Squamous cell carcinoma, Adenocarcinoma, other)
10. Place of residence(Rural, Urban)
11. Cycle of chemotherapy(No chemo, One cycle of chemo, Second cycles of chemo, Third and more cycles of chemo)

3.8 Statistical Methods

3.8.1 Survival Data Analysis

Survival Function:- is the probability that the survival time of a randomly selected subject is greater than or equal to some specified time. Thus, it gives the probability that an individual

surviving beyond a specified time. The distribution of survival time is characterized by survivorship, probability density function and hazard function.

Let T be a random variable associated with the survival times and 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 survivor function, S(t), is given by

$$S(t) = P(T > t) = 1 - CIF(t). \tag{1}$$

where CIF(t) is cumulative distribution function

Cumulative Incidence Function(CIF) is represents the probability that a subject selected at random have a survival time less than or equal to some stated value t, given by:-

$$CIF(t) = P(T \leq t) = \int_0^t f(u)du, t \geq 0. \tag{2}$$

The probability density function, f(t), is given by

$$f(t) = \frac{d}{dt}CIF(t) = \frac{-d}{dt}S(t). \tag{3}$$

The hazard function is the instantaneous probability of having an event at time t given that one has survived up to time t(Kleinbaum et al., 2013). It is given by

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{4}$$

The **cumulative hazard function** is defined as the sum of the hazard function going from duration 0 to t. It is given as:-

$$\Lambda(t) = \int_0^t \lambda(t)dt = -lnS(t). \tag{5}$$

Where λ(t) and S(t) are hazard and survival function at time t, respectively.

The relationship between survivor function and hazard function are:-

$$S(t) = e^{-\Lambda(t)} \tag{6}$$

3.9 Competing Risk Analysis of Survival Data

In this classic analysis, there is a favorite event and all other events are censored(Mansournia et al., 2022). The assumption of this method is that of non-informative censoring which is based on the idea that censored patients are more likely to experience the event as follow-up patients. However, this assumption has not been confirmed in the presence of numerous competing risks(Teixeira et al., 2013). Competing risks encountered in studies where the subjects under study are at risk of multiple failure causes(Noroozi et al., 2022). For example, in a follow-up study of recurrence of patients, patients may die due to any case before occurrence of recurrence, and death is said to be a competing risk(Schellenberg et al., 2022).

3.9.1 CIF

The primary interest in describing competing risks data is often to estimate the absolute risk of the occurrence of an event of interest up to a follow-up time point t. This risk is formalized by the CIF which is defined for each event type separately and increases with time t(Andersen et al., 2012). Let T and C denote the failure and censoring times, respectively. For data with k causes of failure, the pair ($\mathbf{X}$, Y, δ) is observed, where Y=min(T,C), $\mathbf{X}$= covariates and δ= 0,1..., k is an indicator with values 0 for censoring and other values that designate specific failure causes. Then CIF for k^{th} of event can be written as(Andersen et al., 2012)

$$CIF_k(t) = \int_0^t S(t)\lambda_k(t)dt. \tag{7}$$

where $S(t) = e^{\sum_{i=1}^{k} -\Lambda_k(t)}$ is the survival function at time t and is determined by the event of interest and competing events. $\lambda_k(t)$ is hazard of k^{th} failure case, and $\Lambda_k(t)$ is cumulative hazard of k^{th} failure case at time t. In our case k=1,2. The **cuminc()** function shipped with the **cmprsk** package can estimate the CIFs for different causes of failure(Zhang, 2017).

3.9.2 Gray test

In addition to estimating the CIF of an event, it is often of interest to determine whether there is a difference in the CIF among different groups(Austin et al., 2021). In standard survival analysis, this is done using the log-rank test to compare curves generated with the KM method(Mondal et al., 2021). However, when competing risks are present, the CIF of an

event is not defined solely by its corresponding cause-specific hazard(Zhang, 2017). Instead of Log-rank, Gray's investigated this issue and proposed a class of tests for comparing the CIF curves of a particular type of failure among different groups in the presence of competing risks(Ainurrochmah et al., 2021). Grays test is a g-sample test that was introduced by (Gray, 1988). This test is performed under the null hypothesis that there is no difference in CIF between the g-groups versus the alternative hypothesis that at least one of the CIF curves differs. The concept of Grays test can be written as follows:

$$X^2 = \frac{U^2}{Var(U)} \tag{8}$$

with,

$$U = \sum_{allt_r} R_1 t_r \frac{d_1(r)}{R_1(r)} - \frac{d_1(r)+d_2(r)}{R_1(r)+R_2(r)} \tag{9}$$

and

$$Var(U) = \sum_{i=1}^{r} \frac{d_1 t(r)+d_2 t(r)}{R_1 t(r)+R_2 t(r)} \tag{10}$$

$$R_1 = n_1(t_r)\frac{1-CIF_1(t_{(r-1)})}{S_1(t_{(r-1)})}$$

$$m_c(t_r) = \frac{n_c(t_r)}{S_1(t_{(r-1)})}$$

$d_c(t_r) =$ number of events of interest in the type c event group at time t. The test criterion is that H_0 is rejected if $X^2 > X^2_{\alpha,g-1}$ where g is many groups of event types or if P-value $<$ α(Z. Zhang et al., 2017).

3.9.3 Fine-Gray Model

Fine and Gray is the modified Cox-PH model to allow for the presence of competing risks(Schuster et al., 2020). The technical modification consists of keeping the competing risks observations in the risk set with a diminishing weight(Noroozi et al., 2022). The risk set is the set of individuals /subjects under investigation and vulnerable to the event. In this way the Fine and Gray not censored the subjects who experienced the competing risks(Donoghoe & Gebski, 2017). It directly models the covariate effect on CIF and reports sub-distribution hazard ratio (SHR). However, just like standard survival analysis and cause-specific hazard model approach, the subdistribution hazard can be modelled in a proportional hazards frame-

work(Rossello & González-Del-Hoyo, 2022).

The sub-distribution hazard denotes the instantaneous risk of the event of interest in subjects that have not (yet) experienced the event of interest. In our example, this means that the risk set consists of both individuals that have not (yet) developed recurrence and individuals that died before the onset of recurrence. Because of there is direct relation between the covariates and the CIF in Fine-Gray model, the subdistribution hazard model is considered the right model in presence of competing events(F. Zhang et al., 2019). The sub-distribution hazard function for event type k can be expressed as:

$$\lambda_k^*(t) = \lim_{\Delta t \to 0} \frac{p(t \leq T_k < t + \Delta t | T_k \geq t \cup T_{k'} \leq t, k \neq k')}{\Delta t} \tag{11}$$

Fine & Gray (1999b) proposed a semi-parametric proportional regression model for the sub-distribution hazard function

$$\lambda_k^*(t) = \lambda_{0k}^*(t) exp(X\beta_k) \tag{12}$$

Where $\lambda_{0k}^*(t)$ is the baseline sub-distribution hazard for the cause of k and e^{β_k} is the relative risk probability of k^{th} cause associated with the given **X** covariates. The Fine-Gray model can be fit using **FGR()** function shipped with **riskRegression** package. This function calls another function **crr()** from the **cmprsk** package(Zhang, 2017).

3.10 Method of Parametric Estimation

3.10.1 Likelihood Ratio Test

Estimation of parameters in the Fine-Gray model uses the partial likelihood approach similar to the standard Cox model since a proportional hazard assumption is imposed on the sub-distribution hazards(Kuk & Varadhan, 2013). However, in this model, the parameters are estimated by incorporating weights in the partial likelihood. The weight partial likelihood for the Fine-Gray model is given as(Kuk & Varadhan, 2013)

$$L(\beta) = \prod_{j=1}^{r} \frac{exp(X\beta_j)}{\sum_{i \varepsilon R^*(t_j)} w_{ji} exp(X\beta_i)} \tag{13}$$

The product is taken over all r time points $t_1 < t_2 < .... < t_r$, where r is the total number of primary events $(\sum_{i=1}^{n} I(\varepsilon_i = 1))$. The risk set, $R^*(t_j)$ is a set of individuals who are still at risk for the primary event at time t (i.e., those who did not experience the primary event and are not censored by time t)(Fine & Gray, 1999b). This is an unusual risk set in that it also includes those who have experienced one of the competing events (note that these individuals will never experience the primary event), but it facilitates the mathematical development necessary for direct estimation of the CIF. The weight, w_{ji},is defined as

$$w_{ji} = \frac{\hat{G}(t_j)}{\hat{G}(min(t_j, t_i))} \tag{14}$$

where $t_i = min(T_i; C_i)$ for i such that $\varepsilon_i \neq 1$ and t_j is the time of the j^{th} primary event. $\hat{G}$ is the KM estimate of the survivor function of the censoring distribution $(G(t) = P(C \geq t))$. The weight is 1 for the individuals who did not have any type of event by time t_j and less than 1 for those who had a competing event before t_j. As a result, individuals who experience a competing event at time t_i do not participate fully in the partial likelihood; the further the time point t_k is from the time of the competing event t_i, the smaller the weight. When there is only one event of interest, the weights are all equal to 1, and the risk set contains only those at risk at the specified time point (Pintilie, 2006).

3.11 Model Diagnostics

The main assumption when modeling survival data is the proportionality of hazards. When the Fine-Gray model is used, the hazards of the CIF must be proportional whereas, in the Cox proportional hazard model, it is the cause-specific hazards that need to be proportional(Katsahian et al., 2006). The proportionality assumption is the most common in competing risk regression model, which considers the sub-distribution with covariates X is a constant shift on the complementary log-log scale from a baseline sub-distribution function. If the curves do not cross with each other then we say that the model does not violate the assumption of proportionality(Kuk & Varadhan, 2013).

3.11.1 Proportionality of the sub-hazards of the CIF

The proportional hazard assumption is used to measure whether the relationship between cumulative hazards is constant over time(Kleinbaum & Klein, 2012). To investigate the proportionality assumption for the competing risks regression, log (-log (1-F)) can be plotted against log (time) where F is the CIF for the event of interest (Pintilie, 2006).

3.12 Ethical Consideration

The research ethics review board of Jimma University provided an ethical clearance for the study. The data have been be collected after written permeation was obtained from oncology department of JUMC and department of statistics write an official co-operation letter to the Hospital for the permeation. The study was conducted without informed consent since retrospective study design would be applied. Confidentiality of any information related to the patients and their clinical history would be maintained by keeping both the hard-copy and soft-copy of every collected data in a locked cabinet and password secured computer. Only the researcher would access to the de-identified data that has been kept in a secure place. All data would be coded with numbers and without personal identifiers. Since this study would be secondary data analysis, researcher did not have direct contact with the participants. The study was noninvasive and without any harm to patients.

Statistical Software Used

The statistical software used was:-

- R version 4.1.2 used for data analysis.

4 Results and Discussion

4.1 Descriptive Analysis

The researcher reviewed a total of 304 CC patients from medical records who registered at JUMC from 2017 to 2021 GC for inclusion in this study. Among these, 280 records qualified for inclusion, while 24 were excluded. Of 280 patients, 60 (21.4%) of them experienced first recurrence, 64(22.9%) died without evidence of recurrence, and 156 (55.7%) experiences censored. The minimum of time to the first recurrence was 9 months and a maximum of time to the first recurrence was 52 months after registering at JUMC. The median time to the first recurrence of CC is 19 months.

About half of patients 141(50.4%) took chemotherapy from this 45(16.1%) of them were experienced of recurrence while 66(23.6%) and 30(10.7%) of them were censored and died respectively. From the total of the patients 80(28.6%) have taken surgery of which 11(3.9%) were experienced recurrence while 47(16.8%) and 22(7.9%) of them were censored and died, respectively. Among 59(21.1%)patients treated in the hospital with a combination of surgery and chemotherapy, 4(1.4%) of them experienced recurrence and 12(4.3%)of them were died. Regards to smoking status, non-smoker incorporates 245(87.5%) of the total patients where, 152(54.3%), 30(10.7%)and 63(22.5%) of them were experience of censored, recurrence and death respectively. About 202(72.1%) patients were from rural communities, 100(35.7%), 51(18.2%) and 51(18.2%) were experience of censored, recurrence and death,respectively.

This study included 186 (66.4%) participants, who had not taken oral contraceptives, 129(46.1%), 23(8.2%) and 34(12.1%) of them were experience of censored, recurrence and death, respectively. Similar to this, 224 (80.0%) of all patients were those without a family history of CC disease, among them 35(48.2%), 30(10.7%) and 59(21.1%) were experience of censored, recurrence and death, respectively. From the total of patients, 240(85.7%) were HIV negative during the follow-up period, among this, 153(54.6%), 31(11.0%) and 56(20.0%) of them were experienced censoring, recurrence and death, respectively.

Table 4.1: Descriptive Statistics of Variables in the Study.

		Event			
Variables	Category	Censored	Recurrence	death	Total
Smoking status	Non-smoker	152(54.3%)	30(10.7%)	63(22.5%)	245(87.5%)
	Smoker	4(1.4%)	30(10.7%)	1(0.4%)	35(12.5%)
Age	≤ 34	30(10.7%)	22(7.9%)	16(5.7%)	68(24.3%)
	$35-49$	87(31.1%)	34(12.1%)	41(14.7)	162(57.9%)
	≥ 50	39(13.9%)	4(1.4%)	7(2.5%)	50(17.8%)
Family History	No	135(48.2%)	30(10.7%)	59(21.1%)	224(80.0%)
	Yes	21(7.5%)	30(10.7%)	5(1.8%)	56(20.0%)
HIV status	No	153(54.6%)	31(11.0%)	56(20.0%)	240(85.7%)
	Yes	3(1.1%)	29(10.4%)	8(2.9%)	40(14.3%)
FIFO Stage	I	41(14.6%)	3(1.0%)	0(0.0%)	44(15.7)
	II	69(24.6%)	17(6.1%)	26(9.3%)	112(40.0%)
	III	39(13.9%)	21(7.5%)	28(10.0%)	88(31.4%)
	IV	7(2.5%)	19(6.8%)	10(3.6%)	36(12.9%)
Histology	Squamous cell carcinoma	125(44.6%)	31(11.1%)	40(14.3%)	196(70.0%)
	Adenocarcinoma	10(6.8%)	13(4.6%)	12(4.3)	44(15.7%)
	other	12(4.3%)	16(5.7%))	12(4.3%)	40(14.3%)
Treatment	Chemotherapy	66(23.6%)	45(16.1%)	30(10.7%)	141(50.4%)
	Surgery	47(16.8%)	11(3.9%)	22(7.9%)	80(28.6%)
	Combination of two	43(15.4%)	4(1.4%))	12(4.3%)	53(26.0%)

Variables	Category	Event			
		Censored	Recurrence	death	Total
Parity	Nullipara	59(21.1%)	5(1.8%)	7(2.5%)	71(25.4%)
	Multipara	56(20.0%)	25(8.9%)	34(12.1)	115(41.0%)
	Grand multipara	41(14.6%)	30(10.7%))	23(8.3%)	94(33.6%)
Oral contraceptive	No	129(46.1%)	23(8.2%)	34(12.1%)	186(66.4%)
	Yes	27(9.6%)	37(13.2%)	30(10.1%)	94(33.6%)
Place of residence	Rural	100(35.7%)	51(18.2%)	51(18.2%)	202(72.1%)
	Urban	56(20.0%)	9(3.2%)	13(4.7%)	78(28.9%)
Cycles of chemotherapy	No chemo cycle	54(19.3%)	43(15.4%)	28(10.0%)	125(44.7%)
	First cycle	19(6.8%)	7(2.5%)	8(2.9%)	34(12.1%)
	Second cycles	42(15.0%)	7(2.5%)	20(7.1%)	69(24.6%)
	Third cycles and above	41(14.6%)	3(1.0%)	8(2.9%)	52(18.6%)

Source: JUMC, Ethiopia, from 1st January 2017 to 31st December 2021.

4.2 Non-parametric estimate for CIF

4.2.1 The CIF Estimate of Time-to-First Recurrence of CC Patients

The time to the first recurrence of women with CC patients was used as the event of interest, and death was considered as competing risk. The CIF of time to first recurrence at 10-, 20-, 30-, 40-, 50-, and 60-months was 0.01, 0.14, 0.27, 0.33, 0.33, and 0.35, respectively.

4.2.2 The CIF Curves for Different Groups of Covariates

From the results of this study, the CIF curve is higher for those patients who had used oral contraceptives as compared to those who had not used oral contraceptives(p-value$<$ 0.0011). It is higher for those whose histology type was adenocarcinoma as compared to those who had squamous cell carcinoma(p-value$<$ 0.001). It was higher for those patients who had a family history of CC when compared to those whose family had no CC (p-value = 0.001). According to the study's findings, the CIF curve is lower for patients aged 35 to 49 years and 50 and older during the follow-up period than for those aged less than or equal to 34 years (p-value = 0.004). Regarding HIV status, patients who are HIV positive have a higher CIF than patients who are HIV negative(p-value = 0.000). The CIF curve was higher for smokers as compared with non-smokers(p-value = 0.001)(Appendix A.1).

4.2.3 The Comparison of CIF Curves for Different Groups of Covariates

To test for equality of the CIF curves for categorical predictor variables the Gray's test was performed. From the Gray's test the researcher have observed that the CIF curves are statistically significant different for all the groups of predictors($p \leq 0.05$) except for place of residence variable($p=0.074$)(Table 4.2).

Table 4.2: Gray Test of the Difference Between Pairs of CIF for Time to First Recurrence of CC

Variables	df	p-value
Age	2	0.004
Family history	1	0.0001
Smoking status	1	$<$0.0001
HIV status	1	0.000
Treatment taken	2	$<$0.0001
FIGO Stage	3	0.0001
Histology type	2	$<$0.0001
Oral contraceptive	1	$<$ 0.0001
Parity	2	0.0015
Place of residence	1	0.074
Cycles of chemotherapy	3	$<$ 0.0001

Source: JUMC, Ethiopia; from 1st January 2017 to 31st December 2021.

4.3 Model Diagnostics

Checking the Assumption of Proportional sub-distribution hazard

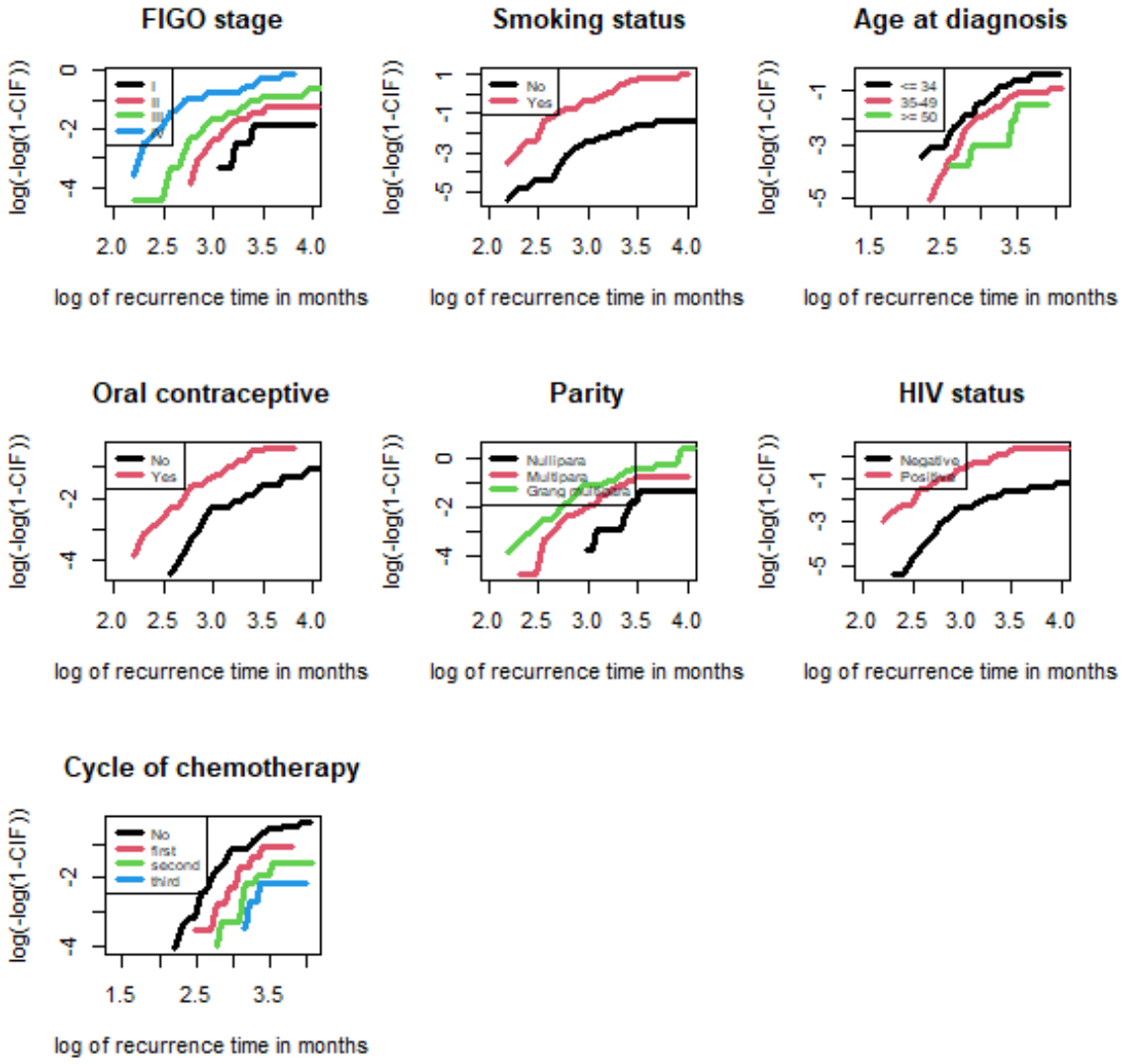

Figure 4.1: The Plot of Proportionality of Sub-Distribution Hazard

The log-minus-log of CIF with the log of time to first recurrence was used to test the proportional assumption for a Fine-Gray model. A parallel curve is visible on the plot. The researcher concluded that since the curves do not cross, the model does not deviate from the proportionality condition. Simply, the proportionate sub-distribution hazard assumption is valid(Appendix B.1 for other covariates)

4.4 Fine and Gray Model

Uni-variable and multi-variable analysis were applied. In uni-variable analysis, the model was fitted one covariate at a time to determine the variables that could be included in the multi-variable analysis. Covariateswith a p-value less than or equal to 25% in the uni-variable analysis were considered for multi-variable analysis(Assemie et al., 2018). For multi-variable analysis, variables with a p-value of less than 5% were selected as significant covariates.

4.4.1 Uni-Variable Fine and Gray Model Analysis

The prognostic factors considered in the study were age at diagnosis, family history, HIV status, FIGO stage, smoking status,treatment taken, oral contraceptives, parity, area of residence, cycles of chemotherapy, and histology type of CC patients. Outputs from uni-variable analysis Table 4.3 showed that all covariates were statistically significant at a 25% level of significance.

Table 4.3: Uni-Variable Fine and Gray Model Analysis for Recurrence of CC

Variables	Categories	SHR[95%]	P-Value
Smoking status	Non smoker	-	-
	Smoker	12.7[7.81- 20.7]	0.0001
Age	$\leq$34	-	-
	35-49	0.6[0.35-1.02]	0.061
	$\geq$ 50	0.212[0.07-0.61]	0.0039
Treatment taken	Chemotherapy	-	-
	Surgery	0.39[0.2- 0.76]	0.0053
	Combination	0.18[0.07 -0.5]	0.0009
FIFO Stage	I	-	-
	II	2.37[0.71 -7.93]	0.16
	III	3.89[1.19-12.78]	0.025
	IV	10.46[3.14- 34.8]	0.00013
Family history	No	-	-
	Yes	5.13 [3.11-8.46]	0.0000

Source: JUMC, Ethiopia; from 1st January 2017 to 31st December 2021.

Variables	Categories	SHR[95%]	P-Value
HIV status	Negative	-	-
	Positive	8.71[5.3-14.3]	0.000
Oral contraceptive	Not used	-	-
	Used	3.82[2.28-6.39]	< 0.0001
Parity	Nullipara	-	
	Multipara	3.41[1.33-8.73]	0.01
	Grand multipara	5.39[2.14-3.58]	0.00035
Type of histology	Squamous	-	-
	Adenocarcinoma	2.0[1.07-3.84]	0.03
	Others	3.02[1.65-5.51]	0.00033
Place of residence	Rural	-	-
	Urban	0.4[0.21-0.85]	0.016
Cycles of chemotherapy	No chemo cycle	-	-
	First cycle	0.54[0.24-1.18]	0.12
	Second cycle	0.25[0.11-0.55]	0.0005
	Third cycle and above	0.14[0.04-0.44]	0.00075

Source: JUMC, Ethiopia; from 1st January 2017 to 31st December 2021.

4.4.2 The Multi-Variable Fine-Gray Model Analysis

After fitting a uni-variable Fine and Gray model, all the predictor variables were found to be significant predictors for the recurrence of CC at a 25% level of significance. This means that all covariates were fitted to the multi-variable Fine and Gray model. As a result, a multi-variable Fine and Gray model revealed that FIGO stage, age at diagnosis, HIV status, smoking habits, oral contraceptives, chemotherapy cycles, and parity were risk factors for the recurrence of CC patients. The estimated parameters for the sub-distribution hazard model are presented in Table 4.4 below and are interpreted as following after controlling for other prognostic factors and accounting for competing risk.

For patients who smoke, the sub-hazard ratio of recurrence is 3.34 [aSHR= 3.34:95%: CI: 1.43-7.81] and the p-value is 0.0058. This demonstrates that smokers have a sub-hazard of CC recurrence that is three times higher than that of non-smokers. The sub-hazard ratio of

recurrence for patients diagnosed with CC at age 50 and older is 0.29 (aSHR= 0.29:95%: CI:0.09-0.9). This suggests that the sub-hazard of recurrence is 0.29 times lower for individuals diagnosed at 50 years and older than for those diagnosed at 34 years and younger (aSHR= 0.29:95%: CI:0.09-0.9).

Taking into account the parity of women with CC, the sub-hazard ratio of recurrence for multi-para women is 2.37 with [95% CI:1.07-5.25] and p-value 0.034, and grand multi-para is 3.34 with [95% CI: 1.61-6.94] and p-value 0.0012. This indicates that multi-para and grand multi-para women were increasing risk of recurrence as compared to nulli-para women. The sub-hazard of recurrence for multi-para women is 2.37 times that of nulli-para women. Similarly, the sub-hazard of recurrence for grand multi-para women is 3.34 times that of nulli-para.

The sub-hazard ratio of CC recurrence is 2.08[95%CI: 1.09-3.94] for HIV positive women among CC patients and the p-value is 0.0053. This shows that the sub-hazard of recurrence for HIV positive women is 2.08 times that of HIV negative women. In addition to this, patients who had FIGO stage-IV increased the sub-hazard of recurrence by 3.71 (aSHR= 3.71, 95% CI: 1.02-13.47) times as compared to patients who had FIGO stage-I.

Regarding oral contraceptive use in women CC, the sub-hazard ratio of time to the first recurrence for the patients who had used oral contraceptives is 2.2[95%CI: 1.09-4.47] and the p-value is 0.029. This shows that the sub-hazard of time to the first recurrence for the patients who had used oral contraceptives is 2.2 times that for the patients who had not used oral contraceptives for a long period of time.

Finally, observing for women using chemotherapy, the sub-hazard ratio of time to recurrence for patients who had used chemotherapy for two cycles is 0.27(aSHR=0.27 :95%: CI:0.11-0.66). This implies that the sub-hazard of recurrence for patients who have used chemotherapy for two cycles is 0.27 times lower than that of patients who have not used chemotherapy. In similar ways, patients who had used three or more cycles of chemotherapy decreased the sub-hazard of recurrence by 0.25 times as compared to those who had not used cycles of chemotherapy(aSHR=0.25 :95%: CI:0.09-0.71) and p-value is 0.0094.

Table 4.4: Multi-Variable Fine and Gray Model Analysis for Time to First Recurrence of CC

Variables	Category	aSHR [95% CI]	Sig.
Smoking status	Non-smoker	-	-
	Smoker	3.34[1.43-7.81]	0.0053
Age	≤ 34	-	-
	35-49	0.63[0.32-1.24]	0.18
	≥ 50	0.29[0.09- 0.9]	0.032
Family History	No	-	-
	Yes	1.59[0.79-3.2]	0.19
FIFO Stage	I	-	-
	II	1.93[0.6-6.2]	0.27
	III	2.1[0.63-7.1]	0.23
	IV	3.71[1.02-13.47]	0.046
HIV status	No	-	-
	Yes	2.08[1.09-3.94]	0.0058
Histology	Squamous cell carcinoma	-	
	Adenocarcinoma	1.73[0.93-3.2]	0.081
	other	0.81[0.4-2.07]	0.47
Treatment taken	Chemotherapy	-	-
	Surgery	0.61[0.23-1.33]	0.21
	Combination of two	0.31[0.09-1.03]	0.057
Oral contraceptives	No	-	-
	Yes	2.2[1.09-4.47]	0.029
Number of children	Nullipara	-	-
	Multipara	2.37[1.07-5.25]	0.034
	Grand multipara	3.34[1.61-6.94]	0.0012

Variables	Category	aSHR [95% CI]	Sig.
Place of residence	Rural	-	-
	Urban	0.83[0.37-1.85]	0.64
Cycles of chemotherapy	No chemo cycle	-	-
	First cycle	0.7[0.31-1.56]	0.38
	Second cycle	0.27[0.11-0.66]	0.0039
	Third cycle and above	0.25[0.09, 0.71]	0.0094

Source: JUMC, Ethiopia, from 1st January 2017 to 31st December 2021.

4.5 Discussion

Recurrence of CC is one of the most important and significant discussions in Gynecologic Oncology especially in patients with locally advanced stage. Therefore, it's crucial to identify risk factors that could make the patient vulnerable to recurrence. To our knowledge, some research has been done on the prognosis of CC recurrence, however the majority of the survival analyses have only investigated at one end point(Tewari et al., 2022). However, in many clinical trials, the number of endpoints is not limited to one; they therefore have a competing risks(Dutz & Löck, 2019). In that case, statistical methods taking into account the competing risks are needed; otherwise, the results obtained might be incorrect(Hou et al., 2019).

In this study, the Fine and Gray model is used on the CC recurrence datasets obtained from JUMC. From the total of 280 CC patients participated in this study, 60 (21.4%) of them experienced recurrence. Sasidharan et al. (2020) reported that the 5 years recurrence rate was 23% in India. This suggests that, in comparison to our study(21.4%), the rate of recurrence in India was higher. The median recurrence time of the CC patients is 19 months.

The result of the study revealed that the age factor is found to be a major predictor of CC recurrence. According to the study, the likelihood of CC recurrence is higher among younger patients than in others. This is consistent with an earlier research conducted in China(Li et al., 2022). According to their research, the risk of recurrence was 0.63 times lower for patients over the age of 55 compared to the patients under 55 (HR=0.63, 95%CI: 0.45-0.89

and P value=0.009). This may be due to the fact that younger patients are more likely to smoke, have higher blood hormone levels, engage in more sexual activity, and express more survivin, all of which enhance the chance of CC recurrence.

In the study, the cycles of chemotherapy that patients took was significant in their uni-variate as well as in multi-variate of the Fine and Gray regression models. This finding is supported by the studies conducted in Ethiopia by (ALTAYE, 2011). Similarly, the HIV status variable was found to be a significant predictor factor in CC recurrence. According to the study, HIV-infected female patients were more likely to have CC recurrence(aSHR = 2.08 and p = 0.0058). This could be related to the HIV affects the bodys immune system by specifically targeting CD4 cells, which are helps to protect themselves from the disease. This is also supported by the research undertaken by Lodi et al. (2011). They reported that the odds of recurrence was 4.17 (95% CI: 1.72-10.10) in women with HIV as compared to those without HIV.

According to the study's findings, oral contraceptives was a major risk factor for CC recurrence. The sub-hazard ratio of recurrence for the patient who had used the oral contraceptive was 1.91, which shows that being an oral contraceptive user increased the sub-hazard of recurrence by 91% as compared to not being an oral contraceptive user. This is in line with the study conducted by Muñoz et al. (2002) in order to analyze the survival analysis of patients with CC using cox regression. This is due to its association with elevated levels of the female hormone, estrogen, and changing the susceptibility of cervical cells to persistent infection with high-risk HPV types.

The findings of this study suggested that smoking behavior was a significant risk factor for the recurrence of CC. The sub-hazard of recurrence for patients who had a smoking habit is 3.7 times that of patients who had no smoking habit. This is in line with the findings in other studies like Jaleta & Mokonnon (2018). The odds of recurrence was 3.95 for smokers (HR, 3.95;95% CI: 1.55-10.35; p=0.0044) as compared to non-smokers. These substances damage the DNA of cervix cells and make the immune system less effective in fighting HPV infection. Lastly, it accelerates the onset of CC recurrence.

Accordingly, the results of this study suggest that the FIGO stage was significantly associated with the time to the first recurrence of CC. The sub-hazard ratio of recurrence for

patients who had stage four was 3.71[1.02-13.47]. This shows that patients whose cancer was stage IV increased the sub-hazard of recurrence by 3.71[1.02-13.47] times as compared to stage I. This is consistent with the study conducted by Li et al. (2022) in order to analyze the risk factors for recurrence of CC. According to their study, the hazard rate of recurrence for stage IV cancer was 1.84 (95%:1.17-2.9) times that of stage I. This may be due to stage IV occurring by invasion or spread in any body organ and it is difficult to treat it at all.

Considering the parity of women with CC, the results of this study suggest that parity was significantly associated with the time to first recurrence of CC. The sub-hazard of recurrence for multipara woman is 2.37 times that of the nullipara woman. In similar ways, the sub-hazard of recurrence for grand multipara women was 3.34[95% CI: 1.61-6.94] and p-value 0.0012. The sub-hazard of recurrence for grand multipara is 3.34 times that of nulli-para women. This study's findings are supported by the study of Sharma & Pattanshetty (2018). This is related causes new dynamics of immature metaplastic epithelium which can increase the risk of cell transformation and trauma to the cervix to facilitate HPV infection. For decades, high parity has been suspected of being associated with an increased risk of CC. Confounding with sexual behaviour, specifically with age at first sexual intercourse, however, thought to account for the apparent adverse effect of multiparity.

5 Conclusion and Recommendation

5.1 Conclusion

The conclusions drawn by this investigation are as follows.

From the result of Fine-Gray model age at diagnosis, smoking behavior, HIV status, parity, FIGO stage, cycles of chemotherapy and oral contraceptives were found to be statistically significant factors for recurrence of women with CC. Of all this significant covariates, smoking status(smoker), HIV status (positive), parity(Multipara and Grand multipara), stage(IV), and oral contraceptives (users) were significantly increases the risk of recurrence of women with CC. While, age at diagnosis($\geq$ 50 year) and cycles of chemotherapy(two, and three and more than three cycles of chemotherapy) were significantly decreases the risk of recurrence of women with CC.

From the result of Gray's, except for the place of residence covariate, the CIF of all groups of covariates were statistically significantly different.

5.2 Recommendations

Based on the study finding the following recommendations are forwarded:-

- It is better if patients minimize the use of tobacco and oral contraceptive substances to protect themselves from a recurrence of CC.
- The time to first recurrence of women with CC risk is high for younger women and women with higher FIGO stage, so it is better to give special care to them.
- The physicians are expected to record additional information of the patients history such as physical activities, age at marriage, age at sexual intercourse, age at giving birth and etc., because these are the expected risk factors from many literatures.

References

Abd ElHafeez, S., DArrigo, G., Leonardis, D., Fusaro, M., Tripepi, G., & Roumeliotis, S. (2021). Methods to analyze time-to-event data: the cox regression analysis. *Oxidative Medicine and Cellular Longevity*, *2021*.

Addis, T. (2010). Combating cervical cancer in ethiopia addressing the screening and treatment gap: the single-visit approach. *Pathfind Int*, 1–2.

Agarwal, S. K. (2021). Smoking and its relationship with cancer. *Journal of Cancer Research Reviews & Reports. SRC/JCRR-141. DOI: https://doi. org/10.47363/JCRR/2021 (3)*, *136*, 26–36.

Ainurrochmah, A., Purnami, S., et al. (2021). Survival analysis competing risk using fine-gray subdistribution model with the maximum partial likelihood estimation method. In *Iop conference series: Materials science and engineering* (Vol. 1115, p. 012014).

ALTAYE, A. B. (2011). Statistical modeling for the recurrence of cervical cancer: A case study at tikur anbessa referral hospital, addis ababa, ethiopia.

Amabebe, E., Ogidi, H., & Anumba, D. O. (2022). Matrix metalloproteinase-induced cervical extracellular matrix remodelling in pregnancy and cervical cancer. *Reproduction and Fertility*, *3*(3), R177–R191.

Andersen, P. K., Geskus, R. B., De Witte, T., & Putter, H. (2012). Competing risks in epidemiology: possibilities and pitfalls. *International journal of epidemiology*, *41*(3), 861–870.

Antunes, D., & Cunha, T. M. (2013). Recurrent cervical cancer: how can radiology be helpfull. *OMICS journal of radiology*, *2*, 2–6.

Arbyn, M., Weiderpass, E., Bruni, L., Sanjosé, S. de, Saraiya, M., Ferlay, J., et al. (2020). Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *The Lancet Global Health*, *8*(2), e191–e203.

Asseffa, N. A. (2017). Cervical cancer: Ethiopias outlook. *J Gynecol Womens Health*, *5*(2), 555660.

Assemie, M. A., Muchie, K. F., & Ayele, T. A. (2018). Incidence and predictors of loss to follow up among hiv-infected adults at pawi general hospital, northwest ethiopia: competing risk regression model. *BMC research notes*, *11*(1), 1–6.

Asthana, S., Busa, V., & Labani, S. (2020). Oral contraceptives use and risk of cervical cancera systematic review & meta-analysis. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *247*, 163–175.

Austin, P. C., & Fine, J. P. (2017). Practical recommendations for reporting f ine-g ray model analyses for competing risk data. *Statistics in medicine*, *36*(27), 4391–4400.

Austin, P. C., Steyerberg, E. W., & Putter, H. (2021). Fine-gray subdistribution hazard models to simultaneously estimate the absolute risk of different event types: cumulative total failure probability may exceed 1. *Statistics in Medicine*, *40*(19), 4200–4212.

Baiocchi, G., Ribeiro, R., Dos Reis, R., Falcao, D. F., Lopes, A., Costa, R. L. R., et al. (2022). Open versus minimally invasive radical hysterectomy in cervical cancer: The circol group study. *Annals of Surgical Oncology*, *29*(2), 1151–1160.

Begoihn, M., Mathewos, A., Aynalem, A., Wondemagegnehu, T., Moelle, U., Gizaw, M., et al. (2019). Cervical cancer in ethiopia–predictors of advanced stage and prolonged time to diagnosis. *Infectious agents and cancer*, *14*(1), 1–7.

Béhanzin, L., Brun, L. V. C., Salanon, E., Pognon, M.-C. A. O. B., Goma-Matsétsé, E., Agonnoudé, M., et al. (2022). Perceptions and knowledge about cervical cancer among women living with hiv in parakou in 2017. *Open Journal of Epidemiology*, *12*(1), 31–49.

Bellinger, J. D., Brandt, H. M., Hardin, J. W., Bynum, S. A., Sharpe, P. A., & Jackson, D. (2013). The role of family history of cancer on cervical cancer screening behavior in a population-based survey of women in the southeastern united states. *Women's Health Issues*, *23*(4), e197–e204.

Bhatla, N., Aoki, D., Sharma, D. N., & Sankaranarayanan, R. (2021). Cancer of the cervix uteri: 2021 update. *International Journal of Gynecology & Obstetrics*, *155*, 28–44.

Bozorgpour, A., Azad, R., Showkatian, E., & Sulaiman, A. (2021). Multi-scale regional

attention deeplab3+: Multiple myeloma plasma cells segmentation in microscopic images. *arXiv preprint arXiv:2105.06238*.

Brandstorp-Boesen, J., Sørum Falk, R., Folkvard Evensen, J., Boysen, M., & Brøndbo, K. (2016). Risk of recurrence in laryngeal cancer. *PLoS One*, *11*(10), e0164068.

Brookfield, K. F., Cheung, M. C., Lucci, J., Fleming, L. E., & Koniaris, L. G. (2009). Disparities in survival among women with invasive cervical cancer: a problem of access to care. *Cancer*, *115*(1), 166–178.

Bryson, A., Buraimo, B., Farnell, A., & Simmons, R. (2021). Time to go? head coach quits and dismissals in professional football. *De Economist*, *169*(1), 81–105.

Burrowes, S., Holcombe, S. J., Leshargie, C. T., Hernandez, A., Ho, A., Galivan, M., et al. (2022). Perceptions of cervical cancer care among ethiopian women and their providers: a qualitative study. *Reproductive Health*, *19*(1), 1–18.

Buzkova, P. (2021). Competing risk of mortality in association studies of non-fatal events. *Plos one*, *16*(8), e0255313.

Cervical Cancer, I. C. of Epidemiological Studies of, et al. (2007). Cervical cancer and hormonal contraceptives: collaborative reanalysis of individual data for 16 573 women with cervical cancer and 35 509 women without cervical cancer from 24 epidemiological studies. *The Lancet*, *370*(9599), 1609–1621.

Chandra, N., & Rehman, H. (2021). Applications of lifetime modeling with competing risks in biomedical sciences. In *Deep learning in biomedical and health informatics* (pp. 137–155). CRC Press.

Chao, X., Fan, J., Song, X., You, Y., Wu, H., Wu, M., et al. (2020). Diagnostic strategies for recurrent cervical cancer: a cohort study. *Frontiers in oncology*, *10*, 591253.

Demissie, B. W., Azeze, G. A., Asseffa, N. A., Lake, E. A., Besha, B. B., Gelaw, K. A., et al. (2022). Communities perceptions towards cervical cancer and its screening in wolaita zone, southern ethiopia: A qualitative study. *PloS one*, *17*(1), e0262142.

Donoghoe, M. W., & Gebski, V. (2017). The importance of censoring in competing risks analysis of the subdistribution hazard. *BMC medical research methodology*, *17*(1), 1–11.

Dutz, A., & Löck, S. (2019). Competing risks in survival data analysis. *Radiotherapy and Oncology*, *130*, 185–189.

Eluf-Neto, J., Booth, M., Muñoz, N., Bosch, F., Meijer, C., & Walboomers, J. (1994). Human papillomavirus and invasive cervical cancer in brazil. *British journal of cancer*, *69*(1), 114–119.

Feakins, B. G., McFadden, E. C., Farmer, A. J., & Stevens, R. J. (2018). Standard and competing risk analysis of the effect of albuminuria on cardiovascular and cancer mortality in patients with type 2 diabetes mellitus. *Diagnostic and prognostic research*, *2*(1), 1–9.

Fine, P., & Gray, R. (1999a). J am stat assoc.

Fine, P., & Gray, R. J. (1999b). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, *94*(446), 496–509.

Gaubatz, M. E., Bukatko, A. R., Simpson, M. C., Polednik, K. M., Boakye, E. A., Varvares, M. A., et al. (2019). Racial and socioeconomic disparities associated with 90-day mortality among patients with head and neck cancer in the united states. *Oral oncology*, *89*, 95–101.

Gennari, P., Tchaikovski, S., Mészáros, J., Gerken, M., Klinkhammer-Schalke, M., Toth, G., et al. (2022). Protective effect of pre-operative conization in patients undergoing surgical treatment for early-stage cervical cancer. *Gynecologic Oncology*.

Ghosh, S., Samanta, G., & Mubayi, A. (2021). Comparison of regression approaches for analyzing survival data in the presence of competing risks. *Letters in Biomathematics*, *8*(1), 29–47.

Ginsburg, O., Bray, F., Coleman, M. P., Vanderpuye, V., Eniu, A., Kotha, S. R., et al. (2017). The global burden of womens cancers: a grand challenge in global health. *The Lancet*, *389*(10071), 847–860.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of statistics*, 1141–1154.

Gurmu, S. E. (2018). Assessing survival time of women with cervical cancer using various parametric frailty models: a case study at tikur anbessa specialized hospital, addis ababa, ethiopia. *Annals of Data Science*, *5*(4), 513–527.

Hou, Y., Guo, S., Lyu, J., Lu, Z., Yang, Z., Liu, D., et al. (2019). Prognostic factors in asian and white american patients with cervical cancer, considering competing risks. *Current Oncology*, *26*(3), 277–285.

Hsu, J. Y., Roy, J. A., Xie, D., Yang, W., Shou, H., Anderson, A. H., et al. (2017). Statistical methods for cohort studies of ckd: survival analysis in the setting of competing risks. *Clinical Journal of the American Society of Nephrology*, *12*(7), 1181–1189.

Jaleta, T. B., & Mokonnon, T. M. (2018). Survival analysis for the recurrence of cervical cancer among patients under follows up since 2012 to 2015. *Survival*, *56*.

Jedy-Agba, E., Joko, W. Y., Liu, B., Buziba, N. G., Borok, M., Korir, A., et al. (2020). Trends in cervical cancer incidence in sub-saharan africa. *British journal of cancer*, *123*(1), 148–154.

Katsahian, S., Resche-Rigon, M., Chevret, S., & Porcher, R. (2006). Analysing multicentre competing risks data with a mixed proportional hazards model for the subdistribution. *Statistics in Medicine*, *25*(24), 4267–4278.

Khurshid, Z., Abdullahi, K. O., Parveen, K., & Khan, S. (2022). Factors contributing to cervical cancer among women: A systematic review and meta-analysis: Cervical cancer among women. *Pakistan BioMedical Journal*.

Kim, H. T. (2007). Cumulative incidence in competing risks data and competing risks regression analysis. *Clinical cancer research*, *13*(2), 559–565.

Kleinbaum, D. G., & Klein, M. (2012). Evaluating the proportional hazards assumption. In *Survival analysis* (pp. 161–200). Springer.

Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Rosenberg, E. S. (2013). *Applied regression analysis and other multivariable methods*. Cengage Learning.

Kuk, D., & Varadhan, R. (2013). Model selection in competing risks regression. *Statistics in medicine*, *32*(18), 3077–3088.

Li, Liu, G., Luo, J., Yan, S., Ye, P., Wang, J., et al. (2022). Cervical cancer recurrence, prognosis, and related risk factors for patients with cervical cancer: A long-term prospective cohort study.

Li, H., Ma, J., & Zhou, H. (2016). An analysis on postoperative recurrence and prognostic factors of patients with cervical carcinoma. *International Journal of Clinical & Experimental Medicine*, *9*(7), 14542–14547.

Lodi, C. T., Michelin, M. A., Lima, M. I., Teixeira, N. C., Adad, S. J., Murta, E. F., et al. (2011). Factors associated with recurrence of cervical intraepithelial neoplasia after conization in hiv-infected and noninfected women. *Archives of gynecology and obstetrics*, *284*(1), 191–197.

Lu, C.-H., Liu, F.-S., Kuo, C.-J., Chang, C.-C., & Ho, E. S.-C. (2006). Prediction of persistence or recurrence after conization for cervical intraepithelial neoplasia iii. *Obstetrics & Gynecology*, *107*(4), 830–835.

Mailhot Vega, R. B., Balogun, O. D., Ishaq, O. F., Bray, F., Ginsburg, O., & Formenti, S. C. (2019). Estimating child mortality associated with maternal mortality from breast and cervical cancer. *Cancer*, *125*(1), 109–117.

Mansournia, M. A., Nazemipour, M., & Etminan, M. (2022). A practical guide to handling competing events in etiologic time-to-event studies. *Global Epidemiology*, *4*, 100080.

Masekwameng, M. J. (2020). *Profile cervical cancer in patients seen in the gynaecology clinic at mankweng hospital, capricorn district, limpopo province*. Unpublished doctoral dissertation.

Medeiros, L. R., Ethur, A. B. d. M., Hilgert, J. B., Zanini, R. R., Berwanger, O., Bozzetti, M. C., et al. (2005). Vertical transmission of the human papillomavirus: a systematic quantitative review. *Cadernos de saude publica*, *21*, 1006–1015.

Messerli, F. H., Bangalore, S., Torp-Pedersen, C., Staessen, J. A., & Kostis, J. B. (2013). Cardiovascular drugs and cancer: of competing risk, smallpox, bernoulli, and d'alembert. *European heart journal*, *34*(15), 1095–1098.

Miccò, M., Lupinelli, M., Mangialardi, M., Gui, B., & Manfredi, R. (2022). Patterns of recurrent disease in cervical cancer. *Journal of Personalized Medicine*, *12*(5), 755.

Mohammad, K. A., Fatima-Tuz-Zahura, M., & Bari, W. (2017). Fine and gray competing

risk regression model to study the cause-specific under-five child mortality in bangladesh. *BMC international health and human rights*, *17*(1), 1–8.

Mondal, S., Bhattacharya, N., Das, T., Ghosh, Z., & Khatua, S. (2021). Implication of statistical methods on patient data: An approach for cancer survivability prediction. In *Intelligent healthcare* (pp. 57–80). Springer.

Monk, B. J., Enomoto, T., Kast, W. M., McCormack, M., Tan, D. S., Wu, X., et al. (2022). Integration of immunotherapy into treatment of cervical cancer: Recent data and ongoing trials. *Cancer Treatment Reviews*, 102385.

Muñoz, N., Franceschi, S., Bosetti, C., Moreno, V., Herrero, R., Smith, J. S., et al. (2002). Role of parity and human papillomavirus in cervical cancer: the iarc multicentric case-control study. *The Lancet*, *359*(9312), 1093–1101.

Nallbani, G., & Agolli, L. (2022). Informing young girls about the risk of hpv infection in developing cervical cancer and preventing. *European Journal of Medicine and Natural Sciences*, *5*(1), 9–14.

Noroozi, M., Khalkhali, H. R., Bahadori, R., Omidi, T., Ghazizadeh, F., Hejazi, S., et al. (2022). The survival of childhood acute lymphoblastic leukemia and its related factors using competing risks model: A retrospective study from 2011 to 2019 in northwestern iran. *Middle East Journal of Cancer*, *13*(3), 531–542.

Oga, E. A., Brown, J. P., Brown, C., Dareng, E., Adekanmbi, V., Odutola, M., et al. (2016). Recurrence of cervical intraepithelial lesions after thermo-coagulation in hiv-positive and hiv-negative nigerian women. *BMC women's health*, *16*(1), 1–8.

Okubo, M., Itonaga, T., Saito, T., Shiraishi, S., Yunaiyama, D., Mikami, R., et al. (2021). Predicting factors for primary cervical cancer recurrence after definitive radiation therapy. *BJR— Open*, *3*, 20210050.

Origoni, M., Cantatore, F., Candotti, G., & Candiani, M. (2022). Prognostic significance of neutrophil/lymphocytes ratio (nlr) in predicting recurrence of cervical dysplasia. *BioMed Research International*, *2022*.

Panjaliya, R., Dogra, V., Gupta, S., et al. (2015). Study of the risk factors associated with cervical cancer. *Biomedical and Pharmacology Journal*, *3*(1), 179–182.

Pecorelli, S., Zigliani, L., & Odicino, F. (2009). Revised figo staging for carcinoma of the cervix. *International Journal of Gynecology & Obstetrics*, *105*(2), 107–108.

Pintilie, M. (2006). *Competing risks: a practical perspective*. John Wiley & Sons.

Pintilie, M. (2007). Analysing and interpreting competing risk data. *Statistics in medicine*, *26*(6), 1360–1367.

Ponce, J., Fernandez-Gonzalez, S., Gil-Moreno, A., Coronado, P. J., Rosa, J. De la, Nabais, H., et al. (2020). Risk factors for recurrence after robot-assisted radical hysterectomy for early-stage cervical cancer: A multicenter retrospective study. *Cancers*, *12*(11), 3387.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., & Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 541–554.

Raza, T. (2021). Molecular biology of human papillomavirus, cervical carcinoma and its management. *Canadian Journal of Clinical Nutrition*.

Rossello, X., & González-Del-Hoyo, M. (2022). Survival analyses in cardiovascular research, part ii: statistical methods in challenging situations. *Revista Española de Cardiología (English Edition)*, *75*(1), 77–85.

Rydzewska, L., Tierney, J., Vale, C., & Symonds, P. (2010). Chemoradiotherapy for cervical cancer meta-analysis collaboration (cccmac). reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: individual patient data meta-analysis. *Cochrane Database Syst Rev*, *1*, CD008285.

Saeedi, A., Baghestani, A., Maboudi, A. K., & Farhangi, H. (2020). Determining the significant prognostic factors for the recurrence of pediatric acute lymphoblastic leukemia using a competing risks approach. *Iranian Journal of Medical Sciences*, *45*(4), 304.

Sahoo, C. K., Satyanarayana, K., Nayak, P. K., Sahoo, K. R., & Champati, M. (2014). An overview on cervical cancer.

Sasidharan, A., Mahantshetty, U. M., Gurram, L., Chopra, S., Engineer, R., Maheshwari, A., et al. (2020). Patterns of first relapse and outcome in patients with locally advanced cervical cancer after radiochemotherapy: a single institutional experience. *Indian Journal of Gynecologic Oncology*, *18*(1), 1–6.

Scheike, T. H., Martinussen, T., & Ozenne, B. (2022). Efficient estimation in the fine and gray model. *Journal of the American Statistical Association*, 1–9.

Schellenberg, A. E., Moravan, V., & Christian, F. (2022). A competing risk analysis of colorectal cancer recurrence after curative surgery. *BMC gastroenterology*, *22*(1), 1–9.

Schmid, M., & Berger, M. (2021). Competing risks analysis for discrete time-to-event data. *Wiley Interdisciplinary Reviews: Computational Statistics*, *13*(5), e1529.

Schuster, N. A., Hoogendijk, E. O., Kok, A. A., Twisk, J. W., & Heymans, M. W. (2020). Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. *Journal of clinical epidemiology*, *122*, 42–48.

Shah, R., Rosso, K., & Nathanson, S. D. (2014). Pathogenesis, prevention, diagnosis and treatment of breast cancer. *World journal of clinical oncology*, *5*(3), 283.

Sharma, P., & Pattanshetty, S. M. (2018). A study on risk factors of cervical cancer among patients attending a tertiary care hospital: A case-control study. *Clinical Epidemiology and Global Health*, *6*(2), 83–87.

Shi, Y., Chen, J., Shi, B., & Liu, A. (2022). Cost-effectiveness analysis of pembrolizumab for treatment of us patients with persistent, recurrent, or metastatic cervical cancer. *Gynecologic Oncology*, *164*(2), 379–385.

Śniadecki, M., Poniewierza, P., Jaworek, P., Szymańczyk, A., Andersson, G., Stasiak, M., et al. (2022). Thousands of womens lives depend on the improvement of polands cervical cancer screening and prevention education as well as better networking strategies amongst cervical cancer facilities. *Diagnostics*, *12*(8), 1807.

Soni, V. D., & Soni, A. N. (2021). Cervical cancer diagnosis using convolution neural

network with conditional random field. In *2021 third international conference on inventive research in computing applications (icirca)* (pp. 1749–1754).

Stelzle, D., Tanaka, L. F., Lee, K. K., Khalil, A. I., Baussano, I., Shah, A. S., et al. (2021). Estimates of the global burden of cervical cancer associated with hiv. *The Lancet Global Health*, *9*(2), e161–e169.

Su, P.-F., Lin, C.-C. K., Hung, J.-Y., & Lee, J.-S. (2022). The proper use and reporting of survival analysis and cox regression. *World Neurosurgery*, *161*, 303–309.

Taarnhøj, G. A., Christensen, I. J., Lajer, H., Fuglsang, K., Jeppesen, M. M., Kahr, H. S., et al. (2018). Risk of recurrence, prognosis, and follow-up for danish women with cervical cancer in 2005-2013: A national cohort study. *Cancer*, *124*(5), 943–951.

Takehara, K., Shigemasa, K., Sawasaki, T., Naito, H., & Fujii, T. (2001). Recurrence of invasive cervical carcinoma more than 5 years after initial therapy. *Obstetrics & Gynecology*, *98*(4), 680–684.

Teguh, S., Hariadi, H., Armanza, F., & Sabaruddin, H. (2021). Survival rate analysis of early stage cervical cancer patients after radical hysterectomy in ulin banjarmasin general hospital, indonesia. *Berkala Kedokteran*, *17*(1), 39–46.

Teixeira, L., Rodrigues, A., Carvalho, M. J., Cabrita, A., & Mendonça, D. (2013). Modelling competing risks in nephrology research: an example in peritoneal dialysis. *BMC nephrology*, *14*(1), 1–8.

Tewari, K. S., Monk, B. J., Vergote, I., Miller, A., Melo, A. C. de, Kim, H.-S., et al. (2022). Survival with cemiplimab in recurrent cervical cancer. *New England Journal of Medicine*, *386*(6), 544–555.

Tidy, C., & Payne, J. (2019). *Gravidity and parity definitions: implications in risk assessment [blog post].*

Wang, Y., Flowers, C. R., Li, Z., & Huang, X. (2022). Condis: A conditional survival distribution-based method for censored data imputation overcoming the hurdle in machine learning-based survival analysis. *Journal of Biomedical Informatics*, 104117.

WHO. (2020). Global strategy to accelerate the elimination of cervical cancer as a public health problem.

WHO. (2022). *Accelerating the elimination of cervical cancer as a public health problem: Towards achieving 90–70–90 targets by 2030* (Tech. Rep.). World Health Organization. Regional Office for South-East Asia.

Wolbers, M., Koller, M. T., Stel, V. S., Schaer, B., Jager, K. J., Leffondre, K., et al. (2014). Competing risks analyses: objectives and approaches. *European heart journal*, *35*(42), 2936–2941.

Wreede, L. C. de, Schetelig, J., & Putter, H. (2022). Analysis of survival outcomes in haematopoietic cell transplant studies: Pitfalls and solutions. *Bone Marrow Transplantation*, 1–7.

Xue, X., Agalliu, I., Kim, M. Y., Wang, T., Lin, J., Ghavamian, R., et al. (2017). New methods for estimating follow-up rates in cohort studies. *BMC medical research methodology*, *17*(1), 1–10.

Ye, Z., Zhang, Y., Liang, Y., Lang, J., Zhang, X., Zang, G., et al. (2022). Cervical cancer metastasis and recurrence risk prediction based on deep convolutional neural network. *Current Bioinformatics*, *17*(2), 164–173.

Young, R. (2014). Who classification of tumours of female reproductive organs. *Kurman RJ Carcangiu ML Herrington CS Young RH Monodermal teratomas and somatic-type tumours arising from a dermoid cyst*, 63–66.

Zhang. (2017). Survival analysis in the presence of competing risks. *Annals of translational medicine*, *5*(3).

Zhang, F., Peng, H., & Zhou, Y. (2019). Fine–gray proportional subdistribution hazards model for competing risks data under length-biased sampling. *Statistics and Its Interface*, *12*(1), 107–122.

Zhang, Z., Geskus, R. B., Kattan, M. W., Zhang, H., & Liu, T. (2017). Nomogram for survival analysis in the presence of competing risks. *Annals of translational medicine*, *5*(20).

Zhu, Y., Xu, L., Ma, W., & Chen, Z. (2021). Research on radiosensitivity of the protein kinase b signaling pathway in cervical cancer. *Computational and Mathematical Methods in Medicine*, *2021*.

A CIF estimate curves of first recurrence for each covarites categories

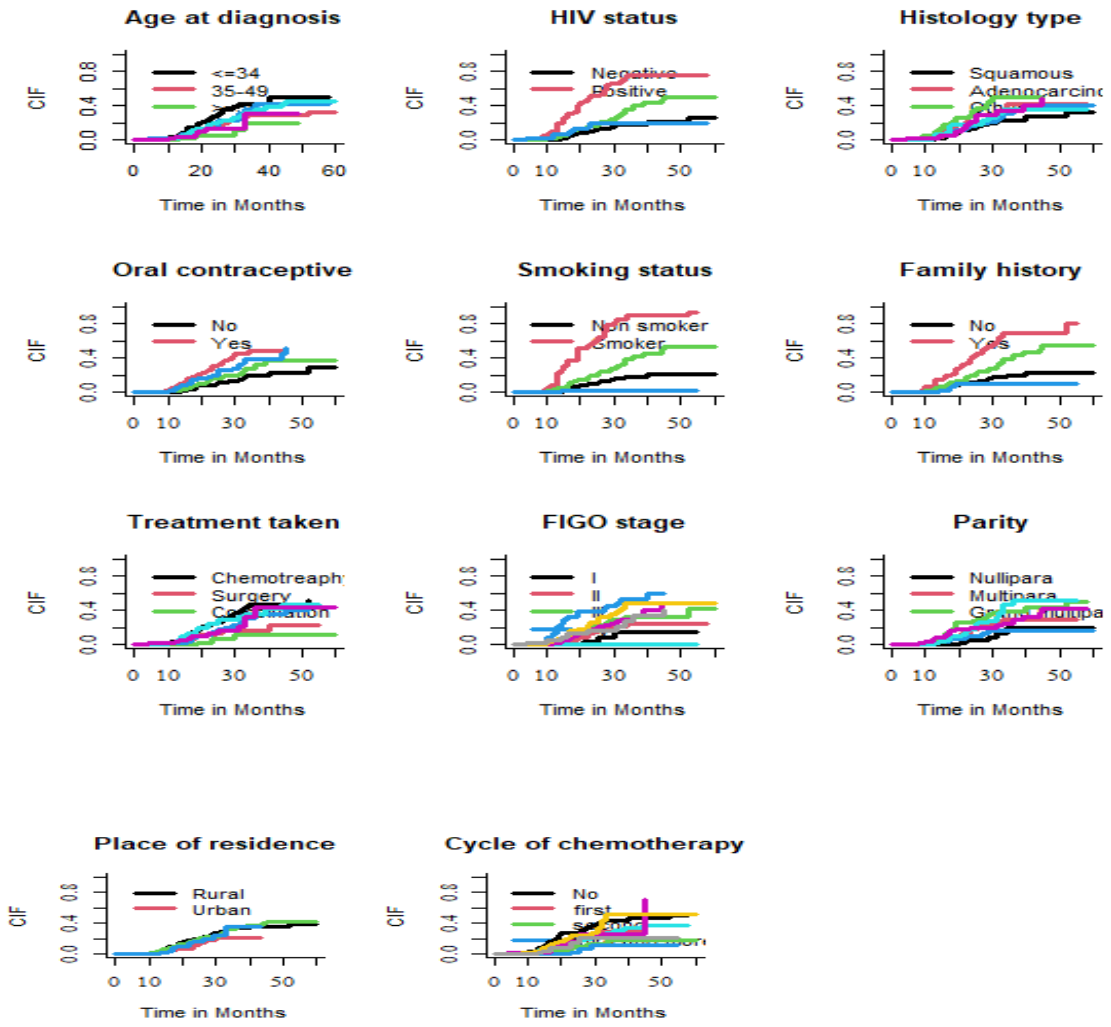

Figure A.1: CIF of time to first recurrence for the group of each covariates

B Proportionality Assumption Checking

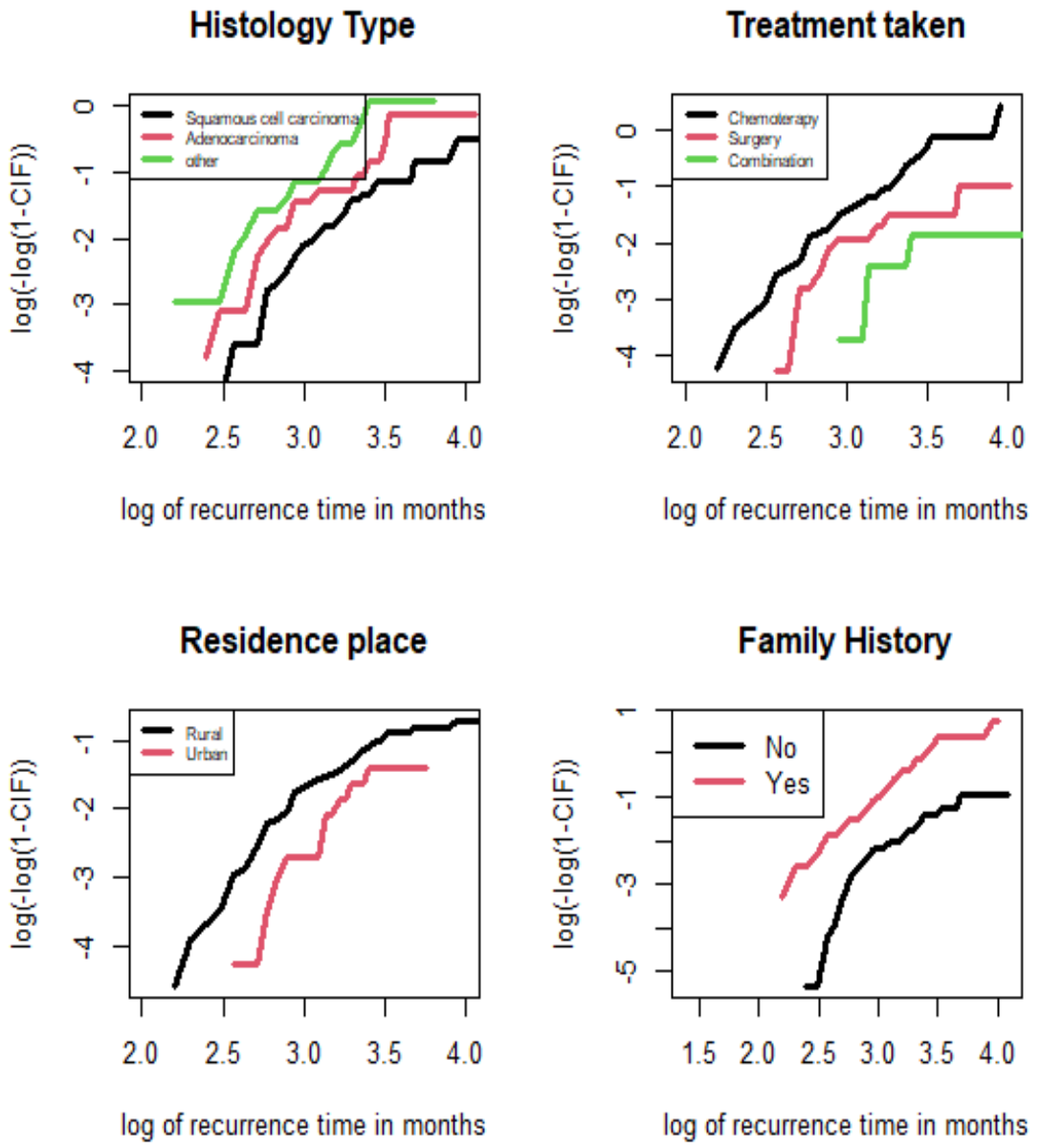

Figure B.1: Sub-distribution hazard assumption checking