Survival Analysis Competing Risk using Fine-Gray Subdistribution Model with the Maximum Partial Likelihood Estimation Method

A Ainurrochmah*, S W Purnami and Purhadi

Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

* Email: alifta.aaa@gmail.com

Abstract. Survival analysis is a statistical method which the variable of concern is the time until the event occurs. In survival analysis, there is a situation where an individual can experience more than one type of event and the occurrence of these events prevents the occurrence of other types of events. This situation is called competing risk. The situation causes the Kaplan Meier method, which is a method of estimating the survival function, cannot be used. The Cumulative Incidence Function method is proposed as a solution to solve competing risk events uses the probability of each type of event. The Cox regression model is also modified to allow for competing risk is called the Fine-Gray subdistribution model using the Maximum Partial Likelihood Estimation. This study examines the estimation of parameter Fine-Gray subdistribution model and applies it to melanoma case. Melanoma is a type skin cancer that can spread to other organs in the body. In the case of death, melanoma patients who experience death with other causes are called a competing risks and death with melanoma are considered as the event of interest. Study showed that factors that influence melanoma mortality are age, sex, tumor thickness, and cultured epidermal autografts. Patients death with melanoma was 1.400 times higher risk than death with other causes.

1. Introduction

Survival analysis is a statistical procedure, which is related to time and is often referred to as time to event. Survival analysis can be used to analyse the time of an individual survival time starting from the initial observation until an event occurs [1]. The response variable in the survival analysis is the initial time of observation until a certain event occurs, while the predictor variables are factors that might influence an event to occur.

In the survival analysis, there are several situations where the observation is not suitable for survival method which is commonly used time to event analysis. 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 risk [2]. The Cumulative Incidence Function (CIF) method is proposed as a solution to solve more than one event. The CIF method uses the probability of each event that occurs, and then is partitioned into probabilities for each type of event [3].

Survival analysis can also estimate the effect predictor variables on survival time. The analysis that is commonly used to determine the effect of predictor variables to the response variable is regression.
analysis. Hazard regression analysis can be used to estimate the effect of predictor variables on survival time. The hazard regression model that is often used is the Cox regression model. Fine-Gray (1999) modified the Cox proportional hazard regression model to allow for competing risk events [4]. Fine-Gray modelling is performed based on estimating the hazard subdistribution. The Fine-Gray subdistribution model can estimate the effect of predictor variables on survival time based on the cumulative incidence function at each event with competing risk, but in interpreting the Fine-Gray subdistribution model and cumulative incidence, more attention must be paid to determine the effect of the predictor variables [5]. If they do not fulfill the proportional hazard assumptions, we have to use basic regression model with stratified Cox [6] or extended Cox regression [7].

Some of the events which are competing risk for example in clinical oncology, cancer-related mortality may be of primary interest, but other causes of death can prevent its occurrence and deaths caused by reasons other than cancer are typical examples of competing risk. Melanoma is one of the most invasive and metastatic human cancers and accounts for the majority of skin cancer deaths despite comprising less than 5% of all cutaneous malignancies [8].

Based on the description above, this study will describe a different approach to survival analysis in the presence of competing risks with Fine-Gray subdistribution model and the application using Melanoma data.

2. Methodology

2.1 Survival Analysis

Survival analysis is a statistical method with the variable of concern is the time until an event occurs or it is called the survival time. Time used in the survival analysis can be in the form of days, weeks, months or even years obtained from the initial observation until the event occurs, while events can be positive events such as healing and negative events such as contracting a disease, recurrence of a disease and death [1].

In the survival analysis, there are two main functions, namely the survival function and the hazard function. The survival function is used to determine the probability of individual survival time from the initial observation time to time \( t \). In the survival function equation, if the survival time is denoted by \( T \) which is a non-negative random variable and has a chance density function, then the probability density function can be written with the following equation:

\[
f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}
\]

(1)

where the cumulative distribution function can be written as follows:

\[
F(t) = P(T \leq t) = \int_0^t f(t) \, dt
\]

(2)

The survival function is expressed in terms of the cumulative distribution function as follows:

\[
S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t)
\]

(3)

The hazard function \( h(t) \) describes the failure rate of an individual that occurs in the time interval \( t \) to \( t + \Delta t \) the individual still survives until time \( t \). The hazard function can be written as follows:

\[
h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}
\]

(4)
2.2 Cumulative Incidence Function and Gray’s Method

Cumulative Incidence Function (CIF) describes the estimated marginal probability of survival of an individual at each event. The CIF method can separately estimate the failure rate for each event. The CIF value for the type \( c \) event at time \( t \) can be written as follows:

\[
\text{CIF}_c(t_{(f)}) = \sum_{j=1}^{f} I_c(t_{(f)}) = \sum_{j=1}^{f} S(t_{(f-j)}) \times h_c(t_{(f)})
\]  

(5)

Gray’s test is an extension of the Log Rank test which is used to evaluate the CIF similarity between two or more groups. The concept of Gray’s test is to compare hazard subdistribution functions. Gray’s test can be written as follows:

\[
\chi^2 = \frac{U^2}{\text{Var}(U)}
\]  

(6)

with,

\[
U = \sum_{\text{all } t_{(f)}} R_{(f)} \left( \frac{d_{1(f)} - d_{2(f)}}{R_{1(f)} - R_{1(f)} + R_{2(f)}} \right)
\]  

(7)

and

\[
\text{Var}(U) = \sum_{i=1}^{f} \left( \frac{d_1(t_{(f)}) + d_2(t_{(f)})}{m_1(t_{(f)}) + m_2(t_{(f)})} \right)
\]  

(8)

where:

\[
R_{(f)} = n_1(t_{(f)}) \frac{1 - F_1(t_{(f-j)})}{S_1(t_{(f-j)})}
\]

\[
m_c(t_{(f)}) = \frac{n_c(t_{(f)})}{S_c(t_{(f-j)})}
\]

\[
d_c(t_{(f)}) = \text{number of events of interest in the type } c \text{ event group at time } t
\]

The test criterion is that \( H_0 \) is rejected if \( \chi^2 > \chi^2_{\alpha,g-1} \) where \( g \) is many groups of event types or if \( p\text{-value} < \alpha \) [9].

2.3 Fine-Gray Subdistribution Model

Fine and Gray (1999) modified the Cox model to allow for competing risk by using a hazard subdistribution regression model by modeling the cumulative incidence function (CIF) curve with covariates. The CIF curve comes from the subdistribution function. This model is basically a development of the Cox model so that there is a proportional hazard assumption that must be tested [9]. Based on equation (1), which is the opportunity density function, it can be seen that the opportunity density function for event type \( c \) can be written as follows:

\[
f_c(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, C = c)}{\Delta t}
\]  

(9)

Fine-Gray subdistribution hazard function for event type \( c \) can be written as follows:

\[
h_c(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, C = c | T \geq t)}{\Delta t}
\]  

(10)

Then the CIF based Fine-Gray subdistribution model can be defined as follows:

\[
\hat{h}_c(t, X_i) = \hat{h}_c(t) \exp \left( \sum_{j=1}^{p} \hat{\beta}_j X_{ij} \right)
\]  

(11)
where:
\[ h_i(t, X_i) \] : estimate hazard subdistribution function event type \( c \)
\[ h_0(t) \] : estimate baseline hazard function event type \( c \)
\[ \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_p \] : regression parameter estimation event type \( c \)
\[ X_{i1}, X_{i2}, ..., X_{ip} \] : predictor variable object \( i \) event type \( c \)

The partial likelihood model for the Fine-Gray subdistribution model is as follows:

\[
L(\beta) = \prod_{t=1}^{n} \exp \left( \sum_{t \in R(t)} \frac{w_{id} \exp(\beta'X_i)}{ \hat{g}(t_i) \hat{g}(\min(t_i, t_l))} \right)
\]

A collection of risks \( R(t) \) defined by adding weights, which is based on non-censored objects. It can be concluded that collection of risks is formed based on the event of interest object with the following definition:

\[ R(t_i) = \{ i; T_i \geq t \text{ or } (T_i \leq t) \text{ and an individual in a competing risk} \} \]

\[ w_{id} = \text{the weight of the object } i \text{ in the event of interest, the weight can be defined as follows:} \]

\[ w_{id} = \frac{\hat{G}(t_i)}{\hat{G}(\min(t_i, t_l))} \]

The test criterion is that \( H_0 \) is rejected if \( T(D) \geq \chi_{m}^2 \) where \( m \) is covariates effect or if \( p-value < \alpha \), it means that the proportional hazard assumption is not fulfilled [13].
2.5 Parameter Significance Test

The parameter significance test was conducted to determine whether there is an effect of a predictor variable on the response variable. There are two tests in testing the significance of parameters, namely the test simultaneously using the likelihood ratio test and partially test using the Wald test.

1. Likelihood Ratio Test

The likelihood ratio (LR) test denoted by $G^2$ is used to test the effect of the predictor variables in the regression model simultaneously.

$$G^2 = -2 \left( \max \ell(\omega) - \max \ell(\Omega) \right)$$

(16)

The $G^2$ test follows the chi-square distribution with degrees of freedom $p$ so that the testing criterion, namely $H_0$ is rejected if the value is $G^2 \geq \chi^2_{p, \alpha}$, or if $p$-value < $\alpha$, meaning that at least one predictor variable has a significant effect on survival time.

2. Wald Test

Wald test is used to test the effect of parameter $\beta_j$ individually

$$W^2 = \left( \frac{\hat{\beta}_j}{\text{SE}(\hat{\beta}_j)} \right)^2$$

(17)

The Wald test statistic follows the chi-square distribution with degrees of freedom 1. The test criterion is that $H_0$ is rejected if the value is $W^2 \geq \chi^2_{1, \alpha(1)}$ or if $p$-value < $\alpha$, meaning the $j$ predictor variable has a significant effect on survival time.

3. Result and Discussion

3.1 Estimated Fine-Gray Subdistribution Model Parameters

Estimated subdistribution Fine-Gray model parameters can be done using the maximum partial likelihood estimation. The sensor status in the survival data is defined as censored data and uncensored data so that the survival data can have a certain distribution such as the Logistic distribution with the opportunity density function as follows [3].

$$f(t) = \frac{\lambda \theta t^{\theta-1}}{1 + \lambda \theta t^\theta} \quad \text{where } \lambda, \theta > 0 \text{ and } t \geq 0$$

(18)

Then the likelihood function is obtained from multiplying the probability density function for the type $c$ event based on equations (9) and (10) as follows.

$$L(\beta) = \prod_{i=1}^{n} f_c(t_i, X_i)$$

$$= \prod_{i \in R(t_i)} \left[ h_c(t_i, X_i) \right] \prod_{i=1}^{n} \left[ S(t_i, X_i) \right]$$

$$= \prod_{i \in R(t_i)} \left[ h_c(t_i, X_i) \right] \exp \prod_{i=1}^{n} \left[ - \int_0^{t} h(t, X_i) dt \right]$$

$$= \prod_{i \in R(t_i)} \left[ h_{0c}(t_i) \exp (\beta' X_i) \right] \exp \prod_{i=1}^{n} \left[ -H_0(t_i) \exp (\beta' X_i) \right]$$

(19)
Based on equation (12), the partial likelihood function of the subdistribution Fine-Gray model can be known. Then form that, the individual opportunities for censored and uncensored data can be defined as follows.

\[
P(\text{censored}) = \frac{\exp(\beta' X_i)}{\sum_{i \in R(t)} w_i \exp(\beta' X_i)} \quad P(\text{uncensored}) = \frac{1}{\sum_{i \in R(t)} w_i \exp(\beta' X_i)}
\]

Therefore, the log-partial likelihood function as follows.

\[
\ell(\beta) = \log \left\{ \prod_{i=1}^n \left( \frac{\exp(\beta' X_i)}{\sum_{i \in R(t)} w_i \exp(\beta' X_i)} \right) \right\} = \sum_{i=1}^n \left\{ \log(\exp(\beta' X_i)) - \log \left( \sum_{i \in R(t)} w_i \exp(\beta' X_i) \right) \right\} = \sum_{i=1}^n \left[ \frac{\sum_{j=1}^p (\beta_j X_{ij})}{\sum_{j=1}^p \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)} \right] - \log \left( \sum_{i \in R(t)} w_i \exp(\beta' X_i) \right)
\]

After obtaining the log partial likelihood function in equation (21), here are the steps for estimating parameters \(\beta\) and the hypothesis test statistics on the Fine-Gray subdistribution model. To estimate parameters \(\beta\), the following steps are required.

1. Obtain the first partial derivative of the log-partial likelihood function based on equation (21) as follows.

\[
\frac{\partial \ell(\beta)}{\partial \beta_p} = \sum_{i=1}^n \left[ \frac{\sum_{j=1}^p (\beta_j X_{ij})}{\sum_{j=1}^p \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)} \right] - \frac{\sum_{i \in R(t)} w_i X_{ip} \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)}{\sum_{i \in R(t)} w_i \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)}
\]

So that it forms a vector \(g(\beta)\) which contains the first derivative of the log-partial likelihood function of size \(p \times 1\).

2. Obtain the second partial derivative of the log-partial likelihood function against \(\beta\) based on equation (22).

\[
\frac{\partial^2 \ell(\beta)}{\partial \beta_k \partial \beta_k'} = -\sum_{i=1}^n \left[ \frac{\sum_{i \in R(t)} w_i X_{ik} X_{ki} \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)}{\sum_{i \in R(t)} w_i \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)} \right] + \sum_{i=1}^n \left[ \frac{\sum_{i \in R(t)} w_i X_{ki} \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right) \left( \sum_{i \in R(t)} w_i X_{ik} \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right) \right)^2}{\sum_{i \in R(t)} w_i \exp \left( \sum_{j=1}^p (\beta_j X_{ij}) \right)^2} \right]
\]
where \( k, k' = 1, 2, \ldots, p \)

So that it can form a Hessian matrix \( H(\hat{\beta}) \) which contains the second derivative of the log-partial likelihood function of size \( p \times p \).

3. Then perform the Newton-Raphson iteration using the following equation.

\[
\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} - H^{-1}(\hat{\beta}^{(t)}) g(\hat{\beta}^{(t)})
\]

where \( H^{-1}(\hat{\beta}^{(t)}) \) is the inverse of the Hessian matrix \( H(\hat{\beta}^{(t)}) \).

4. Newton Raphson's iteration will stop if the value of \( \| \hat{\beta}^{(t+1)} - \hat{\beta}^{(t)} \| \leq \epsilon \), where \( \epsilon \) is a very small positive number.

After getting the parameter estimate \( \hat{\beta} \), then a Fine-Gray subdistribution model can be formed based on equation (11). Hypothesis testing and statistical tests simultaneously on the parameters of the Fine-Gray subdistribution model are carried out using the maximum likelihood ratio test (MLRT) method. This method involves two log-likelihood functions, namely \( \ell(\Omega) \) which is the log-likelihood function of the parameter set under the model population by involving the predictor variable and \( \ell(\omega) \) which is the log-likelihood function of the parameter set under \( H_0 \) model without involving the existing predictor variables, with steps as follows.

1. Obtain the log-likelihood function from the set of parameters under population \( \Omega = \{\beta, \lambda, \theta\} \) model by involving predictor variables based on the likelihood function in equation (12).

\[
\ell(\hat{\Omega}) = \log \left\{ \prod_{i=1}^{n} \frac{\exp(\hat{\beta}'X_i)}{\sum_{l \in R(t_i)} w_{il} \exp(\hat{\beta}'X_i)} \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ \log(\exp(\hat{\beta}'X_i)) - \log \left\{ \sum_{l \in R(t_i)} w_{il} \exp(\hat{\beta}'X_i) \right\} \right\} \]

\[
= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{k} (\hat{\beta}_j X_{ij}) - \log \left\{ \sum_{l \in R(t_i)} w_{il} \exp \left( \sum_{j=1}^{k} (\hat{\beta}_j X_{ij}) \right) \right\} \right\}
\]

2. Obtain the log-likelihood function from the set of parameters under the \( H_0 \) \( \omega = \{h_0(t)\} \) model without involving the existing predictor variables based on the likelihood function in equation (12)

\[
\ell(\hat{\omega}) = \log \left\{ \prod_{i=1}^{n} \frac{1}{\sum_{l \in R(t_i)} w_{il}} \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ \log(1) - \log \left\{ \sum_{l \in R(t_i)} w_{il} \right\} \right\} \]

\[
= -\sum_{i=1}^{n} \left\{ \log \left\{ \sum_{l \in R(t_i)} w_{il} \right\} \right\}
\]
3. Obtain the likelihood ratio test statistics $G^2$ according to equation (16) as follows.

$$G^2 = -2 \left( \frac{\ell(\hat{\theta})}{\ell(\Omega)} \right)$$

$$= 2 \left[ \sum_{i=1}^{n} \left[ \sum_{j=1}^{p} (\hat{\beta}_j X_{ij}) - \log \left( \sum_{i \in R(t)} w_i \exp \left( \sum_{j=1}^{p} (\hat{\beta}_j X_{ij}) \right) \right) \right] - \left[ -\sum_{i=1}^{n} \left( \log \left( \sum_{i \in R(t)} w_i \right) \right) \right] \right]$$

$$= 2 \left[ \sum_{i=1}^{n} \left( \sum_{j=1}^{p} (\hat{\beta}_j X_{ij}) - \log \left( \sum_{i \in R(t)} w_i \exp \left( \sum_{j=1}^{p} (\hat{\beta}_j X_{ij}) \right) \right) + \log \left( \sum_{i \in R(t)} w_i \right) \right) \right]$$

(27)

The $G^2$ test follows the chi-square distribution with degrees of freedom $p$ so that the testing criterion, namely $H_0$ is rejected if the value is $G^2 \geq \chi^2_{p, \alpha}$ or if $p$-value < $\alpha$. If there are predictor variables that have a significant effect on survival time together, then it can be done using partial hypothesis testing. Hypothesis testing of parameter $\hat{\beta}_j$ is partially carried out to determine the significance of the partial test parameters, with steps as follows:

1. Obtain the parameter estimate of the $j$ predictor variable $\hat{\beta}_j$ using the Newton-Raphson iteration.

2. Obtain the standard error estimation of the parameter of the $j$ predictor variable. The standard error estimation of the parameter of the $j$ predictor variable is obtained from the second partial derivative of the log-partial likelihood function which is the diagonal element of the Hessian matrix $H(\beta)$.

3. Get the Wald test statistics according to equation (17).

The Wald test statistic follows the chi-square distribution with degrees of freedom 1. The test criterion is that $H_0$ is rejected if the value is $W^2 \geq \chi^2_{0, \alpha}$ or if $p$-value < $\alpha$.

3.2 Application on Melanoma Data

Data were used is data on patients with Melanoma (cancer of skin) had a radical surgery performed at the University Hospital of Odense, Denmark. All patients were followed which time 134 were still alive while 71 had died (of out whom 57 had died from cancer and 14 from other causes). The dataset can be seen in software $R$. The research variables are shown in Table 1.

| Variables                      | Describes                                      |
|-------------------------------|------------------------------------------------|
| Status                        | 0 = Cencored Data                              |
|                               | 1 = Death with Melanoma (event of interest)    |
|                               | 2 = Death with other causes (competing risk)   |
| Age                           | Age at operation (years)                       |
| Sex                           | 0 = Female ; 1 = Male                          |
| Tumor Thickness               | Tumor thickness (1/100 mm)                     |
| Cultured Epidermal Autografts | 0 = Not present Epicel ; 1 = Present Epicel    |

Therefore, modeling was carried out using a Fine-Gray subdistribution to determine death with melanoma of competing risks case, namely death with other causes and estimate the CIF of death with melanoma and death with other causes. The cumulative curves of the categorical predictor variables such as sex and cultured epidermal autografts (epicel) can be seen in Figure 1.
In Figure 1(a), the risk of death based on sex can be seen, namely male patients have higher risk of death with melanoma and other causes than female. Figure 1(b) shown that patients who did not have cultured epidermal autografts have higher risk of death with melanoma than death with other causes and patients who had cultured epidermal autografts have higher risk of death with melanoma than death with other causes. The difference between the cumulative incidence curve can be determined with certainty using the Gray’s test statistic based on equation (6) are shown in Table 2.

### Table 2. Gray’s Test

| Predictor Variables | $\chi^2$ | $p$-value |
|---------------------|---------|-----------|
| Sex                 |         |           |
| Male                | 5.814   | 0.015     |
| Female              | 0.854   | 0.355     |
| Cultured Epidermal  |         |           |
| Autografts          |         |           |
| Not Present Epicel  | 7.524   | 0.016     |
| Present Epicel      | 2.629   | 0.355     |

Based on Table 2, it can be seen that both patient, male patients and patients who did not cultured epidermal autografts have a $p$-value $< \alpha$, therefore there is a significant difference between the cumulative incidence curve of patients who experienced death with melanoma and death with other causes evaluated by male sex and did not have cultured epidermal autografts. Then modeling using the Fine-Gray subdistribution model and the following results are obtained.

Fine-Gray subdistribution model for patients who experienced death with melanoma:

$$
\hat{h}_M(t, X_1) = \hat{h}_M(t) \exp (0.669 \text{male} + 0.013 \text{age} + 0.126 \text{thick} - 0.947 \text{epicel})
$$

Fine-Gray subdistribution model for patients who experienced death with other causes:

$$
\hat{h}_{OC}(t, X_1) = \hat{h}_{OC}(t) \exp (0.121 \text{male} + 0.054 \text{age} + 0.029 \text{thick} - 0.683 \text{epicel})
$$

Based on Fine-Gray subdistribution model for patients death with melanoma and death with other causes, the likelihood ratio test is used to test the effect of the predictor variables in the regression model simultaneously based on equation (16) are shown in Table 3.
Table 3. Likelihood Ratio Test

| Models                          | $G^2$  |
|--------------------------------|--------|
| Death with Melanoma            | 31.600 |
| Death with Other Causes        | 13.000 |

In Table 3, it can be seen that Fine-Gray subdistribution model for patients death with melanoma and death with other causes have a likelihood ratio test value $\geq \chi^2_{0.05;4} = 0.711$, therefore there is a significant influence of predictor variables on survival time. Then a partial test is using Wald test based on equation (17) are shown in Table 4.

Table 4. Wald Test

| Models                          | Predictor Variables                          | $p$-value   |
|--------------------------------|---------------------------------------------|-------------|
| Death with Melanoma            | Sex (male)                                  | $1.5 \times 10^{-2}$ * |
|                                | Age                                         | $1.9 \times 10^{-1}$ * |
|                                | Tumor thickness                              | $3.6 \times 10^{-5}$ * |
|                                | Cultured epidermal autografts (did epicel)  | $2.7 \times 10^{-3}$ * |
| Death with Other Causes        | Sex (male)                                  | 0.830       |
|                                | Age                                         | $1 \times 10^{-4}$ * |
|                                | Tumor thickness                              | 0.720       |
|                                | Cultured epidermal autografts (did epicel)  | 0.300       |

Based on Table 4, it can be seen that the predictor variables that affect the risk of death with melanoma patients are sex, age, tumor thickness, and cultured epidermal autografts, while the predictor variable that affects the risk of death with other causes patients is age. There is a basic assumption that a subdistribution has a covariate value is constant, it can be seen based on schoenfeld residual plots as follows (See Figure 2):

Figure 2. Schoenfeld Residual Based on Predictor Variables
Schoenfeld residuals against failure time for each covariate, if the proportional assumption is true, the residual should have a constant mean across time. In Figure 2, it appears that the schoenfeld residuals based predictor variables constant residuals across time, indicating potential proportional assumptions is fulfilled and based on hazard ratio in models. Then tested the proportional hazard assumption with goodness of fit can be seen in Table 5.

**Table 5. Goodness of fit for proportional hazard assumption**

| Predictor Variables | p-value |
|---------------------|---------|
| Tumor thickness x t  | 0.330   |
| Tumor thickness x t² | 0.800   |
| Age x t             | 0.880   |
| Age x t²            | 0.900   |

Based on Table 5, it can be seen that the predictor variables for tumor thickness and age (continuous predictor variables) did not have a significant effect on the time function, so the proportional hazard assumption has been fulfilled. Then, based on Fine-Gray subdistribution model for patients who experienced death with melanoma and other causes, the hazard ratio of patient mortality can be found as follows.

\[
HR = \frac{\hat{h}_M(t, X_i)}{\hat{h}_{OC}(t, X_i)}
\]

\[
= \frac{\hat{h}_{0i}(t)\exp(0.669\text{male} + 0.013\text{age} + 0.126\text{thick} - 0.947\text{epicel})}{\hat{h}_{0i}(t)\exp(0.121\text{male} + 0.054\text{age} + 0.029\text{thick} - 0.683\text{epicel})}
\]

\[
= \frac{\hat{h}_{0i}(t)\exp(0.669 + 0.013 + 0.126 - 0.947)}{\hat{h}_{0i}(t)\exp(0.121 + 0.054 + 0.029 - 0.683)}
\]

\[= 1.400\]

Based on equation (30), it can be seen that the patients death with melanoma was 1.400 times higher risk than death with other causes.

**4. Conclusion**

Based on the analysis described above, then it can be concluded that, the estimation the Fine-Gray subdistribution model can be written as follows.

\[
\hat{h}_i(t, X_i) = \hat{h}_{0i}(t)\exp\left(\sum_{j=1}^{p} \beta_j X_{ij}\right)
\]

Fine-Gray subdistribution model for patients who experienced death with melanoma as follows.

\[
\hat{h}_M(t, X_i) = \hat{h}_{0i}(t)\exp(0.669\text{male} + 0.013\text{age} + 0.126\text{thick} - 0.947\text{epicel})
\]

Fine-Gray subdistribution model for patients who experienced death with other causes as follows.

\[
\hat{h}_{OC}(t, X_i) = \hat{h}_{0i}(t)\exp(0.121\text{male} + 0.054\text{age} + 0.029\text{thick} - 0.683\text{epicel})
\]

Predictor variables that affect the risk of patients who experienced death with melanoma are sex, age, tumor thickness, and cultured epidermal autografts, while the predictor variable that affects the risk of patients who experienced death with other causes is age. The patients death with melanoma was 1.400 times higher risk than death with other causes.
References

[1] Kleinbaum D G and Klein M 2012 *Survival Analysis A Self Learing Text Third Edition* New York: Springer

[2] Gichangi A and Vach W 2005 The Analysis of Competing Risks Data: A Guided Tour *Stat Med* **132** (4) 1–41

[3] Pintilie M 2011 An Introduction to Competing Risks Analysis *Revista Espanola de Cardiología (English Edition)* **64** (7) 599–605

[4] Fine J P and Gray R J 1999 A Proportional Hazards Model for the Subdistribution of a Competing Risk *Journal of the American Statistical Association* **94** (446) 496–509

[5] Austin P, Latouche A and Fine J 2019 A Review of the use Time-Varying Covariate in the Fine–Gray Subdistribution Hazard Competing Risk Regression Model *Statistics in Medicine* **39** (8) 103-113

[6] Purnami S W, Inayati K D, Sari N W W, Chosuvivatwong V and Sriplung H 2016 Survival Analysis of Cervical Cancer using Stratified Cox Regression *AIP Conference Proceedings* **1723** (1) 030018

[7] Edina A I, Purnami S W, Purjadi and Sukur E 2020 Estimation of Baseline Hazard Function on Cox Model using Maximum Penalized Likelihood *International Journal of Advanced Science and Technology* **29** (7) 3253-3263

[8] Ward W H and Farma J M 2017 *Cutaneous Melanoma: Etiology and Therapy* Brisbane: Codon Publications

[9] Zhang Z, Geskus R B and Kattan M 2017 Nomogram Survival Analysis in the Presence of Competing Risks *Annals of Translational Medicine* **5** (20) 1–9

[10] Pintilie M 2006 *Competing Risks A Practical Perspective* John London: John Wiley and Sons

[11] Collett D 2003 *Modelling Survival Data in Medical Research (Second Edition)* London: Chapman and Hall

[12] Zhou B, Fine J and Laird G 2013 Goodness of Fit Test for Proportional Subdistribution Hazards Model *Statistics in Medicine* **32** (22) 3804–3811.

[13] Husain H, Thamrin S A, Tahir S, Mukhlisin A and Apriani M M 2018 The Application of Extended Cox Proportional Hazard Method for Estimating Survival Time of Breast Cancer *Journal of Physics: Conference Series* **979** 012087 1-8