Comparison of cox models in detecting factors affecting healing rate of dengue hemorrhagic fever

Miftahuddin1,*, I Salsabila1, A Gul2

1 Department of Statistics, Faculty of Mathematics and Sciences, Syiah Kuala University, Jln. Teuku Nyak Arief, Darussalam, Banda Aceh, Aceh, 23111, Indonesia
2Department of Statistics, Shaheed Benazir Bhutto Women University Peshawar, Pakistan

E-mail: miftah@unsyiah.ac.id

Abstract. Dengue Hemorrhagic Fever (DHF) is an epidemic of disease that usually becomes a benchmark of good or bad condition of the environment and health facilities of a region. When the disease is outbreak and does not get serious treatment, it causes death. One of the first steps that can be done to deal with this disease is to know the factors that affect the healing rate of DHF patients. This study aims to detect factors affecting the healing rate of DHF patients by using Cox regression analysis based on Cox model family, such as Cox Proportional Hazard (PH) model, extended Cox model with one and two heaviside function, and stratified Cox model and get best model for this study. The data used is secondary data consisting of 107 in-patients DHF period January-December 2016 at Regional General Hospital dr. Zainoel Abidin Banda Aceh, Indonesia. Based on Cox Proportional Hazard model, the variable that significantly affect the patient's healing rate is Age. Based on the Extended Cox model with one heaviside function, the variables that significantly affect the patient's healing rate are Age, Sex, Number of Platelets, Clinical Degrees III and Number of Leukocytes multiplied by the time function. Based on the Extended Cox model with two heaviside functions, the variables that significantly affect the patient's healing rate are Age, Sex, Number of Platelets, Clinical Degrees III and Number of Leucocytes multiplied by second time function. Based on the stratified Cox model, there are no variables that significantly affect the patient's healing rate. The best model based on Akaike Information Criterion value is the Stratified Cox model.

1. Introduction
Dengue Hemorrhagic Fever (DHF) is a disease caused by dengue virus that is transmitted by mosquito aedes aegypti. This disease often appears as an Extraordinary evidence with a relatively high number of morbidity and mortality. The number of dengue fever cases in Banda Aceh during 2016 were 152 (men 77 and women 75 cases). Number of DHF cases, according to Community Health Center, can be seen in Figure 1 below (Health Office of Banda Aceh, 2016). DHF is a disease that is often prevalent in the community and needs to get serious handling. For this purpose, identification of the factors that affect the healing rate of DHF patients are of crucial importance.

*Corresponding author: miftah@unsyiah.ac.id
Generally, the statistical method applied to analyze the relationship between dependent and independent variables is regression analysis. If the dependent variable is considered as the time until going out of remission, the analysis used is related to survival analysis. In epidemiological studies, one of the most commonly used regression analyzes for survival analysis is the Cox regression model. This model makes it possible to isolate variables that have small effect on survival. In addition, the model allows to estimate the risk or danger of death for an individual based on the prognostic variables. This study investigated the factors that influence the healing rate of DHF patients. Cox regression analysis technique is implemented, where comparisons are made among the Cox model family, such as Cox Proportional Hazard, extended Cox with one and two heaviside functions, and stratified Cox model to get best model. Based on these models, there are different results about the factors that affect the healing rate of DHF patients, with identification a set of factors.

2. Materials and Methods
This study used secondary data consisting of 107 Dengue Hemorrhagic Fever (DHF) patients in the period of January-December 2016 at Regional General Hospital dr. Zainoel Abidin Banda Aceh. The variables used are detailed in Table 1.

### Table 1. Research Variables.

| Symbol | Variable Name                      | Variable Type | Unit     |
|--------|------------------------------------|---------------|----------|
| T      | Duration of patient treated (survival time) | Numerical     | Days     |
| d      | Status 0 = censored, 1 = uncensored | Numerical     |          |
| X1     | Age                                | Numerical     | Years    |
| X2     | Sex 0 = female, 1 = male           | Categorical   |          |
| X3     | Hemoglobin level                   | Numerical     | gr/dl    |
| X4     | Hematocrit level                   | Numerical     | %        |
| X5     | Number of Platelets                | Numerical     | Thousand/mm$^3$ |
| X6     | Number of Leukocytes               | Numerical     | Thousand/mm$^3$ |
| X7     | Degree DHF 0 = none, 1 = Degree I  | Categorical   |          |
|        | 2 = Degree II, 3 = Degree III, 4 = Degree IV |               |          |

2.1 Dengue Hemorrhagic Fever (DHF)
Dengue Fever is an acute febrile viral disease frequently presenting headaches, bone or joint and muscular pain, rashes and leukopenia as symptoms. Dengue Hemorrhagic Fever (DHF) is characterized by four major clinical manifestations: high fever, haemorrhagic phenomena, often with hepatomegaly and, in severe cases, signs of circulatory failure. Such patients may develop hypovolemic shock resulting from plasma leakage. This is called Dengue Shock Syndrome (DSS) and can be fatal [1]. According to [2], several tests are commonly used to determine whether antibodies to dengue virus are present in the body by counting the number of antibodies to dengue virus and complete blood count (hemoglobin, leucocytes, hematocrit and platelets).
Hemoglobin (Hb) is a complex protein containing iron and red in the erythrocytes. Normal Hb level for women is 11.4 to 15.1 g/dl while for males is 13.4 to 17.7 g/dl. Leukocytes or white blood cells are blood cells containing the nucleus. The number of leukocytes within normal limits for men and women is 4300-11300/mm$^3$. Platelets are the smallest part of cellular elements in the bone marrow and are important in clotting and hemostasis process. Under normal conditions, the number of Platelets for women and men is 150-400 thousand/mm$^3$. Hematocrit (Ht) is a number that indicates the percentage of solids in blood to blood fluid. For normal Ht levels for women 38-42% while for men 40-47% [3].

2.2 Survival Analysis

In survival analysis, is focused to time-related data, ranging from time origin or start point to the occurrence of a special event or end point [4]. The subjects are called censored when the follow up time ends before the event occurs. In such cases the actual duration of time to the event is not known or 'censored' by the study. The outcome variable for each subject is therefore composed of 'time' and the 'status' at the end. Mathematically, the status is 1 if the event takes place and 0 otherwise [5]. The causes of censored data include:
- a. Loss to follow up, occurs when objects move, dies or refuses to participate
- b. Drop out, occurs when the treatment is stopped for some reason.
- c. Termination of the study, occurs when the study period ends while the observed object has not reached a failure event [6].

2.3 Cox Proportional Hazard Model

Cox Proportional Hazard regression model was applied to find the correlation of predictor variables to survival time. The hazard function of different individuals is assumed to be proportional every time [7]. In general, the Cox Proportional Hazard regression model can be expressed as follows.

$$h(t; X) = h_0(t) \exp \left( \sum_{i=1}^{p} \beta_i X_i \right)$$

A key assumption of the Cox regression model is proportional hazards. The proportional hazards assumption means that the hazard ratio is constant over time, or that the hazard for an individual is proportional to the hazard for any other individual [8]. There are three methods commonly used to assess the PH assumption: (1) graphical, using, say, log–log survival curves; (2) using an extended Cox model; and (3) using a goodness-of-fit (GOF) test. There are two options to consider if the PH assumption is not satisfied for one or more of the predictors in the model. We suggest the application of a stratified Cox (SC) model, which stratifies on the predictor(s) not satisfying the PH assumption, while keeping in the model those predictors that satisfy the PH assumption. In addition the option, which involves using time-dependent variables is also considered [6].

2.4 Extended Cox Model

In the Cox regression model, there can be variables which involve t. Such variables are called time-dependent variables. A time dependent variable is defined as any variable whose value for a given subject may differ over time(t). If there are time-dependent variables in the model, the Cox regression model can be used but can no longer satisfy the proportional hazards assumption. Therefore, extended Cox regression model should be used instead [9]. The general model of this method is as follows.

$$h(t; X(t)) = h_0(t) \exp \left[ \sum_{i=1}^{p} \beta_i X_i + \sum_{j=1}^{p} \beta_j X_j(t) \right]$$

The formation of extended Cox model to overcome non-proportional hazard is done by adding time function, i.e. $g(t) = t$ on variable that does not satisfy the PH assumption. One Heaviside function is of the form $g(t)$, which takes on the value 1 if $t$ is greater than or equal to some specified value of $t$, called $t_0$, and takes on the value 0 if $t$ is less than $t$. An extended Cox model which contains a single Heaviside function is shown here, where $g(t) = 1$ if $t \geq t_0$ and 0 if $t < t_0$ then
\[ h(t, X(t)) = h_0(t) \exp \left[ \beta E + \delta E g(t) \right], \]
where for \( t \geq t_0 : g(t) = 1 \rightarrow E \times g(t) = E \)
so \( h(t, X(t)) = h_0(t) \exp \left[ (\beta + \delta)E \right], \) and \( \overline{RR} = \exp[\beta + \delta] \)
whereas for \( t < t_0 : g(t) = 0 \rightarrow E \times g(t) = 0 \) so \( h(t, X(t)) = h_0(t) \exp \left[ \beta E \right], \) and \( \overline{RR} = \exp[\beta] \).

There is actually an equivalent way to write this model that uses two Heaviside functions in the same model. The alternative model is shown here. The two Heaviside functions are called \( g_1(t) \) and \( g_2(t) \). Alternative model with two Heaviside functions is shown here.

\[ g_1(t) = 1 \text{ if } t \geq t_0 \text{ and } 0 \text{ if } t < t_0 \text{ and } g_2(t) = 1 \text{ if } t < t_0 \text{ and } 0 \text{ if } t \geq t_0 \]

If \( t \geq t_0 : g_1(t) = 1, \) and \( g_2(t) = 0 \) then

\[ h(t, X) = h_0(t) \exp \left[ \delta_1(E \times 1) + \delta_2(E \times 0) \right] = h_0(t) \exp \left[ \delta_1 E \right], \text{ so that } \overline{RR} = \exp[\delta_1] \]

If \( t \geq t_0 : g_1(t) = 0, \) and \( g_2(t) = 1 \) then

\[ h(t, X) = h_0(t) \exp \left[ \delta_1(E \times 0) + \delta_2(E \times 1) \right] = h_0(t) \exp \left[ \delta_2 E \right], \text{ so that } \overline{RR} = \exp[\delta_2] \]

2.5 **Stratified Cox Model**

In comparison to the extended Cox model, a Stratified Cox models utilized. The method used is grouping of variables that do not satisfy the PH assumption into several strata. The general model of this method is as follows.

\[ h_g(t, X) = h_{g_k}(t) \exp \left[ \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \right] \quad g = 1, 2, \ldots, k^* \]

The stratified Cox regression model is a modification of the Cox regression model by the stratification of a covariate that does not satisfy the proportional hazards assumption. Covariates that are assumed to satisfy the proportional hazards assumption are included in the model, whereas the predictor being stratified is not included [10]. According to [11], a stratified Cox model can be used that allows separate baseline hazards but common coefficients between strata.

2.6 **Testing the Significance of Parameters**

The test used to test the significance of the \( \beta \) coefficient of the model may use partial test and simultaneous test.

2.6.1 **Partial Test**

Partial test is done to know the significance of parameters to the dependent variable. Testing the significance of parameters using Wald [12] with the following hypothesis. \( H_0 : \beta_j = 0 \) and \( H_1 : \beta_j \neq 0 \), with statistics test:

\[ w = \left( \frac{\beta_j}{SE(\hat{\beta}_j)} \right)^2 \]

The W test statistic also referred to as Wald's test statistic with \( SE(\hat{\beta}_j) \) is a standard estimate of parameter error. The rejection zone is if \( |W| > Z_{\alpha/2} or W^2 > \chi^2_{(\nu, \alpha)} \) with degrees of freedom \( \nu \).

2.6.2 **Simultaneous Test**

Tests simultaneously performed to examine the significance of the overall coefficient \( \beta \) with the following hypothesis. \( H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0 \) and \( H_1 : \) at least one \( \beta_j \neq 0; j = 1, 2, \ldots, p \)

Test statistics:

\[ G = -2 \left[ L(\hat{\beta}) - Lp(0) \right] \]

where, \( L(\hat{\beta}) \) = Partial log likelihood for models with predictor variables

\( Lp(0) \) = Partial log likelihood for models without predictor variables

The test statistic \( G \) follows the Chi-square distribution. The rejection criterion \( H_0 \) if the value of \( G > \chi^2_{(\nu, \alpha)} \) with \( \nu \) (degree of freedom) is degrees of freedom and \( \alpha \) is the error rate used.

2.7 **Model Selection**

Akaike Information Criterion (AIC) is used for selection of the optimal model. The AIC is a popular method for comparing the adequacy of multiple, possibly non-nested models. Current practice in cognitive psychology is to accept a single model on the basis of only the “raw” AIC values. It makes
difficult to un-ambiguously interpret the observed AIC differences in terms of a continuous measure such as probability [13]. The model with the smallest AIC value is the best model that can explain the prediction data with the real data. AIC is defined as:

$$AIC = 2k - \frac{2}{n} \sum_{i=1}^{n} u_i^2$$

Where $k$ is the number of parameters in the model including the intercept, $n$ is the number of observations (samples) and $u$ is the residual. The smallest standard value of the AIC depends on the number of parameters and the number of observations present in the model.

3. Results and Discussion

Characteristics of patient data on Dengue Hemorrhagic Fever (DHF) can be seen from the descriptive statistics presented in tables 2 and 3 below.

| Variable                          | Minimum | Maximum | Mean  | Deviation Standard |
|-----------------------------------|---------|---------|-------|--------------------|
| Duration of patient treated (days)| 2       | 21      | 4663  | 2656               |
| Age (years)                       | 1       | 64      | 14.77 | 10.75              |
| Hemoglobin (Hb) (gr/dl)           | 6       | 19      | 11392 | 2338               |
| Hematocrit (%)                   | 20      | 57      | 36317 | 6535               |
| Platelet (thousand/mm$^3$)        | 12      | 344     | 622055| 51301              |
| Leukocyte (thousand/mm$^3$)       | 1       | 20      | 5.616 | 3982               |

Table 2. Descriptive statistics of DHF data.

| Variable                          | Category | Quantity | Percentage |
|-----------------------------------|----------|----------|------------|
| Sex                               | Female   | 35       | 33         |
|                                   | Male     | 72       | 67         |
| Clinical Degree                   | Degree I | 34       | 32         |
|                                   | Degree II| 63       | 59         |
|                                   | Degree III| 8       | 7          |
|                                   | Degree IV| 0        | 0          |
|                                   | None     | 2        | 2          |
| Status                            | Censored | 2        | 1.87       |
|                                   | Uncensored| 105     | 98.13      |

Table 3. Distribution of DHF patients by sex, clinical degree and status

Table 2, the average hospitalization of patients treated was 4 days with a standard deviation of 2 days. The minimum time of patients treated is 2 days and maximum is 21 days. The average patient age is 14 years old with the youngest of 1 year old and at most 64 years. Hemoglobin, hematocrit, platelets and leukocytes patients had an average of 11.3392 g/dl, 36317%, 62205 thousand/mm$^3$ and 5616 thousand/mm$^3$. The average platelet of patients was in abnormal condition, but the maximum value is still in normal condition. While the average leukosit still in normal condition, but its maximum value is in abnormal condition. Whereas table 3, the percentage of DBD male patients is (67%) and women is (33%). Based on Clinical Degree, DHF patients are more classified in DHF II, ie 59%, compared to Degree I, III, and no degree. Based on the status, DHF patients are more in the uncensored status (stated by the expert doctor repatriated because it has improved) that is 98.13% compared to censored status (run from the hospital and died) that is 1.87%.

3.1 Cox Proportional Hazard Model

The formation of Cox Proportional Hazard (PH) model was performed to find out the relationship between survival time and the variables suspected to affect survival time. The variables are age, sex, Hb level, hematocrit level, Number of Platelets, Number of Leukocytes, and dengue degrees. The results of Cox PH model formation can be seen from the following table 4. The Cox Proportional Hazard model equation obtained is as follows.
\[
\begin{align*}
    h(t, X) = h_0(t) \exp[-0.02445X_1 + 0.43895X_2 + 0.12534X_3 + 0.05925X_4 - 0.00472X_5 - 0.03806X_6 - 0.38981X_7(1) - 0.60638X_7(2) - 1.24453X_7(3)]
\end{align*}
\]

(4)

| Table 4. Summary of Cox PH model. |
|-------------------------------|-----------------|--------------------------|--------|
| **Coef** | **P** | **Likelihood ratio test** | **P-value** |
| Age \((X_1)\) | -0.02445 | 0.014 | 24.67 | 0.003 |
| Sex \((X_2)\) | 0.43895 | 0.063 |  |  |
| Hemoglobin level \((X_3)\) | 0.12534 | 0.129 |  |  |
| Hematocrit level \((X_4)\) | -0.03806 | 0.188 |  |  |
| Number of Platelets \((X_5)\) | -0.00472 | 0.056 |  |  |
| Number of Leukocytes \((X_6)\) | 0.05925 | 0.081 |  |  |
| Degree I \((X_7(1))\) | -0.38981 | 0.602 |  |  |
| Degree II \((X_7(2))\) | -0.60638 | 0.411 |  |  |
| Degree III \((X_7(3))\) | -1.24453 | 0.134 |  |  |

After obtaining parameter estimation for Cox PH model formation, all parameters must be tested for significance. The significance test can be done with two stages, i.e. simultaneous test using Likelihood Ratio Test, and partial test using Wald Test. Based on Table 4, with a significance level of 5% \((\alpha = 0.05)\), it can be done simultaneous significance test so it is known that there is at least one independent variable that affects the model. It is concluded pursuant to p-value \((0.003)\) which is smaller than \(\alpha (0.05)\). To determine which variables are influential, then it can be done partial test. Based on table 4, the p-values for the partial test \((p)\) of all independent variables are different. The P-value for the Age\((X_1)\) is smaller than \(\alpha\), whereas the p-values for the Sex\((X_2)\), Hb level\((X_3)\), Hematocrit level\((X_4)\), Number of Platelets\((X_5)\), Number of Leukocytes\((X_6)\), Degree I \((X_7(1))\), II \((X_7(2))\) and III\((X_7(3))\) is greater than \(\alpha\). Therefore, it can be concluded that the Age variable has significant effect on the model, while the Hb level, Hematocrit level, Number of Platelets, Number of Leukocytes, Degree I, II and III have no significant effect on the model. The next step is to test the PH assumption to find out whether the PH assumption for the model has been met. The PH assumption can be tested using Goodness-of-Fit (GoF) test. The following table presents the results using GoF test.

| Table 5. The Result of PH assumption test using Goodness-of-Fit test |
|-----------------|-----------------|-----------------|--------|
| **Rho** | **Chisq** | **P** |
| Age\((X_1)\) | 0.0946 | 0.913 | 0.33929 |
| Sex\((X_2)\) | 0.1097 | 1.493 | 0.22176 |
| Hemoglobin level\((X_3)\) | -0.0417 | 0.269 | 0.60429 |
| Hematocrit level\((X_4)\) | -0.0815 | 0.95 | 0.32983 |
| Number of Platelets\((X_5)\) | 0.112 | 1.866 | 0.17191 |
| Number of Leukocytes\((X_6)\) | 0.0221 | 7.588 | 0.00588 |
| Degree I \((X_7(1))\) | 0.0932 | 0.954 | 0.32858 |
| Degree II \((X_7(2))\) | -0.0495 | 0.263 | 0.60830 |
| Degree III \((X_7(3))\) | -0.0829 | 0.786 | 0.37533 |
| **Global** | NA | 12.201 | 0.20220 |

Based on table 5, with \(H_0\) the data satisfy the PH assumption and the 5% significance level, the decision to reject \(H_0\) for the Number of Leukocytes\((X_6)\), and receive \(H_0\) for the Age\((X_1)\), Sex\((X_2)\), Hb level\((X_3)\), Hematocrit level\((X_4)\), Number of Platelets\((X_5)\), Degree I \((X_7(1))\), II \((X_7(2))\) and III \((X_7(3))\). This means that the assumption of PH is satisfy for the Age, Sex, Hb level, Hematocrit level, Number of Platelets, Degree I, II and III variables and that are not satisfy for the Number of Leukocytes variable. Because there are variables that do not satisfy the PH assumption, another alternative is needed for this data modeling in order to avoid non-proportional hazard conditions.

3.2 Extended Cox Model

The formation of extended Cox model to overcome non-proportional hazard is done by adding time function, i.e. \(g(t) = t\) on variable that does not satisfy the PH assumption. Note
that the Heaviside function is the form \( g(t) \), which takes the value of 1 if \( t \) is greater than or equal to some certain \( t \) value, called \( t_0 \), and takes the value 0 if \( t \) is less than \( t_0 \).

For this case, suppose we want to assess the assumption of PH using the Heaviside function (HF) to produce a constant hazard ratio for less than 10 days, and a constant hazard ratio of 10 days or more of follow-up. There are two possible ways to model the involvement of one HF depending on whether the experiment uses one or two Heaviside functions.

Model 1 (with one Heaviside function):

\[
h(t, X(t)) = h_0(t) \exp \left[ \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_7(1) + \beta_7 X_7(2) + \beta_8 X_7(3) + \delta_1(X_6 x g(t)) \right]
\]

where \( g(t) = 1 \) if \( 0 \leq t < 10 \) days and 0 if \( t \geq 10 \) days.

Model 2 (with two Heaviside functions):

\[
h(t, X(t)) = h_0(t) \exp \left[ \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_7(1) + \beta_7 X_7(2) + \beta_8 X_7(3) + \delta_1(X_6 x g_1(t)) + \delta_2(X_6 x g_2(t)) \right]
\]

where \( g_1(t) = 1 \) if \( 0 \leq t < 10 \) days and 0 if \( t \geq 10 \) days, \( g_2(t) = 1 \) if \( t \geq 10 \) days and 0 if \( 0 \leq t < 10 \) days.

If we choose to use model 1, then the estimation of the extended Cox model parameter is given by the following table 6. The equation of the extended Cox model with one heaviside function is obtained as follows.

\[
h(t, X(t)) = h_0(t) \exp \left[ -0.02034 X_1 + 0.38485 X_2 + 0.10662 X_3 - 0.02945 X_4 - 0.00436 X_5 - 0.17576 X_7(1) - 0.40253 X_7(2) - 0.95231 X_7(3) + 0.05917 (X_6 x g(t)) \right]
\]

### Table 6. Summary of Extended Cox model with one Heaviside function.

| Variable                  | Coef   | P       | Likelihood ratio test | P-value |
|---------------------------|--------|---------|-----------------------|---------|
| Age(X(1))                 | -0.02034 | 0.0036 | 18.63                 | 0.03    |
| Sex(X(2))                 | 0.38485 | 0.0256 |                       |         |
| Hemoglobin level(X(3))    | 0.10662 | 0.1150 |                       |         |
| Hematocrit level(X(4))    | -0.02945 | 0.1735 |                       |         |
| Number of Platelets(X(5)) | -0.00436 | 0.0172 |                       |         |
| Degree I (X(7)(1))        | -0.17576 | 0.5476 |                       |         |
| Degree II (X(7)(2))       | -0.40253 | 0.1194 |                       |         |
| Degree III (X(7)(3))      | -0.95231 | 0.0383 |                       |         |
| Number of Leukocytes(X(6)) x g_1(t) | 0.05917 | 0.0364 |                       |         |

Based on table 6, with a significance level of 5% (\( \alpha = 0.05 \)), p-value (0.03) smaller than \( \alpha (0.05) \) can conclude that there is at least one independent variable that significantly affects the model. To determine which variables are influential, then it can be done partial test. Based on Table 6, the p-values for the Age(X(1)), Sex(X(2)), Number of Platelets(X(5)), Degree I (X(7)(1)) and Number of Leukocytes(X(6)) x g_1(t) smaller than \( \alpha \), while p-value for Hb level(X(3)), Hematocrit level(X(4)), Degree II (X(7)(2)) greater than \( \alpha \). Therefore, it can be concluded that the variables Age, Sex, Number of Platelets, Degree III and Number of Leukocytes x time function have a significant effect on the model, while the Hb level, Hematocrit level, Degree I and II have no significant effect on the model. Then to find out whether the condition of non-proportional hazard has been resolved, it can be tested the PH assumption using GoF test. The following table presents the test results using GoF test.

Based on table 7, with Null Hypothesis (H_0) the data satisfy the PH assumption and the 5% significance level (\( \alpha = 0.05 \)), the decision to receive H_0 for the Age(X(1)), Sex(X(3)), Hb level(X(3)), Hematocrit level(X(4)), Number of Platelets(X(5)), Degree I (X(7)(1)), Degree II (X(7)(2)), Degree III (X(7)(3)) and Number of Leukocytes(X(6)) x g_1(t). This means that the PH assumption is satisfied for all variables, i.e. Age, Sex, Hb level, Hematocrit level, Number of Platelets, Degree I, II, III, Number of Leukocytes x time function. Since all variables satisfy the PH assumption, the non-proportional hazard condition is resolved.
The equation of the extended Cox model with two heaviside function is obtained as follows.

\[ h(t, X(t)) = h_0(t) \exp[-0.01860X_1 + 0.39466X_2 + 0.10956X_3 - 0.03092X_4 - 0.00406X_5 - 0.17615X_6(1) - 0.40552X_7(2) - 0.97773X_8(3) + 0.05658(X_6 \times g_1(t)) - 0.25119(X_6 \times g_2(t))] \]  

Table 8, with a significance level of 5% (\( \alpha = 0.05 \)), p-value (0.03) smaller than \( \alpha = 0.05 \) can conclude that there is at least one independent variable that significantly affects the model. To determine which variables are influential, then it can be done partial test. Based on table 8, the p-values for the Age(X_1), Sex(X_2), Number of Platelets(X_3), Degree III (X_3(3)) and Number of Leukocytes(X_6) \( \times g_2(t) \) smaller than \( \alpha = 0.05 \), while p-value for the Hb level(X_1), Hematocrit level(X_4), Degree I (X_3(1)), II (X_3(2)) and Number of Leukocytes(X_6) \( \times g_1(t) \) greater than \( \alpha = 0.05 \). Therefore, it can be concluded that the Age, Sex, Number of Platelets, Degree III and Number of Leukocytes \( \times g_2(t) \) have significant effect on the model, while Hb level, Hematocrit level, Degree I, II and Number of Leukocytes \( \times g_1(t) \) have no significant effect on the model. Furthermore, the following table presents the PH assumption test results using Goodness-of-Fit test.

**Table 9. The Result of PH assumption test using Goodness-of-Fit test.**

|                | Rho  | Chisq | P   |
|----------------|------|-------|-----|
| Age(X_1)      | 0.1224 | 0.7461 | 0.388 |
| Sex(X_2)      | 0.0998 | 0.6935 | 0.405 |
| Hb level(X_1) | -0.0259 | 0.0738 | 0.786 |
| Hematocrit level(X_4) | -0.0687 | 0.4272 | 0.513 |
| Number of Platelets(X_3) | 0.1087 | 0.9569 | 0.328 |
| Degree I (X_3(1)) | -0.1352 | 0.8673 | 0.352 |
| Degree II (X_3(2)) | -0.0505 | 0.0623 | 0.803 |
| Degree III (X_3(3)) | -0.1085 | 1.1865 | 0.276 |
| Number of Leukocytes(X_6) \( \times g_1(t) \) | -0.1344 | 2.2268 | 0.136 |
| Number of Leukocytes(X_6) \( \times g_2(t) \) | -0.1969 | 0.7281 | 0.393 |
| Global        | NA   | 5.6424 | 0.844 |
Table 9, with Null Hypothesis (H₀) the data satisfy the PH assumption and the 5% significance level (α = 0.05), the decision to receive H₀ for the Age(X₁), Sex(X₂), Hb level(X₃), Hematocrit level(X₄), Number of Platelets(X₅), Degree I (X₇(1)), II (X₇(2)), III (X₇(3)), Number of Leukocytes(X₆) x g₁(t) and Number of Leukocytes(X₆) x g₂(t). This means that the PH assumption is satisfied for all variables, i.e. Age, Sex, Hb level, Hematocrit level, Number of Platelets, Degree I, II, III, Number of Leukocytes x time function (both of g₁(t) and g₂(t)). Since all variables satisfy the PH assumption, the non-proportional hazard condition is resolved.

3.3 Stratified Cox Model
How to resolve non-proportional hazard using the time-dependent variable method or extended Cox model is done. In comparison of these results, we can create a stratified Cox model. Below is the estimation of model parameters using stratified Cox model.

**Table 10. Summary of Stratified Cox Model.**

| Coef   | P     | Likelihood ratio test | P-value |
|--------|-------|-----------------------|---------|
| Age (X₁) | -0.02640 | 0.33 | 7.54 | 0.5 |
| Sex (X₂) | 0.52183 | 0.21 | |
| Hemoglobin level (X₃) | 0.11796 | 0.47 | |
| Hematocrit level (X₄) | -0.01067 | 0.11 | |
| Number of Platelets (X₅) | -0.00588 | 0.90 | |
| Degree I (X₇(1)) | -0.34255 | 0.82 | |
| Degree II (X₇(2)) | -0.27620 | 0.86 | |
| Degree III (X₇(3)) | -1.05038 | 0.55 | |

Based on table 10, with a significance level of 5% (α = 0.05), we can test the significance of the parameters simultaneously. Based on p-value (0.5) greater than α (0.05) it can be concluded that all independent variables have no significant effect on the model so no partial test is required. The equation of the stratified Cox model is obtained as follows.

\[ h_{d}(t, X) = h_{0}(t) \exp[-0.02640X_{1} + 0.52183X_{2} + 0.11796X_{3} - 0.01067X_{4} - 0.00588X_{5} - 0.34255X_{7(1)} - 0.27620X_{7(2)} - 1.05038X_{7(3)}] \]

Based on table 11, with Null Hypothesis (H₀) the data satisfy the PH assumption and the 5% significance level (α = 0.05), the decision to receive H₀ for the Age(X₁), Sex(X₂), Hb level(X₃), Hematocrit level(X₄), Number of Platelets(X₅), Degree I (X₇(1)), II (X₇(2)) and III (X₇(3)). This means that the PH assumption is satisfied for all variables, i.e Age, Sex, Hb level, Hematocrit level, Number of Platelets, Degree I, II and III. Because all variables satisfy the PH assumption, the non-proportional hazard condition is resolved. The following table presents the PH assumption test results using GoF.

**Table 11. The Result of PH assumption test using Goodness-of-Fit test.**

| Rho   | Chisq | P   |
|-------|-------|-----|
| Age(X₁) | -0.02736 | 0.08708 | 0.768 |
| Sex(X₂) | -0.00219 | 0.00055 | 0.981 |
| Hemoglobin level(X₃) | 0.02153 | 0.06023 | 0.806 |
| Hematocrit level(X₄) | 0.06403 | 0.37673 | 0.539 |
| Number of Platelets(X₅) | -0.01598 | 0.04203 | 0.838 |
| Degree I (X₇(1)) | -0.03221 | 0.05858 | 0.309 |
| Degree II (X₇(2)) | -0.02026 | 0.02335 | 0.879 |
| Degree III (X₇(3)) | -0.02374 | 0.04298 | 0.836 |
| Global | NA | 0.72754 | 0.999 |

After getting some Cox models, then the best Cox model is selected by comparing the AIC values for each model. Here is a comparison table of each model.
Table 12. The comparison of Cox PH, Extended Cox, and Stratified Cox Model

| Model                      | AIC       |
|----------------------------|-----------|
| Cox PH                     | 777.162   |
| Extended Cox with one Heaviside | 822.6516 |
| Extended Cox with two Heaviside | 823.6601 |
| Stratified Cox             | 121.0142  |

Based on table 12, the AIC value for the stratified Cox model (121.0142) is lower than the Cox PH model (777.162), extended Cox model with one Heaviside function (822.6516), and two Heaviside function (823.6601). According to AIC criteria, the model with the lowest AIC value is the best model that can explain the data. Thus, the Stratified Cox model is the best model based on the results of this study.

4. Conclusions

Based on Cox Proportional Hazard model, the variable that significantly affect the DHF patient's healing rate is Age. Whereas the Extended Cox model with one Heaviside function, the variables that significantly affect the DHF patient's healing rate are Age, Sex, Number of Platelets, Clinical Degrees III and Number of Leukocytes multiplied by the time function.

For the Extended Cox model with two Heaviside functions, the variables that significantly affect the DHF patient's healing rate are Age, Sex, Number of Platelets, Clinical Degrees III and Number of Leucocytes multiplied by second time function. Whereas based on the stratified Cox model, there are no variables that significantly affect the DHF patient's healing rate. By using AIC criteria, the best model that can explain the data of the DHF patients is stratified Cox model because it has minimum AIC value compared to Cox PH, extended Cox model with one and two Heaviside function.

5. Acknowledges

The authors would like to thank the Department of Statistics, Faculty of Mathematics and Sciences, and the Institute of Research and Public Services of Syiah Kuala University. We also would like to thank the anonymous reviewers for their valuable suggestions.

6. References

[1] World Health Organization 1997 Dengue Haemorrhagic Fever: Diagnosis, Treatment, Prevention and Control (WHO Library Cataloguing in Publication Data).
[2] Chuansumrit, A. Tagnararatchakit K, Hongeng S 2003 Dengue Hemorrhagic Fever in Children with Underlying Hematologic-oncologic diseases. (Thai J Hematol Transf Med).
[3] Hadinegoro S.R. and Sattari, H.I 2004 Tatalaksana Demam Berdarah Dengue di Indonesia (Departemen Kesehatan RI dan Direktorat Jenderal Pemberantas ten Penyakit Menular dan Penyehatan Lingkungan, Jakarta).
[4] Collet, D 1994 Modeling Survival Data in Medical Research (New York)
[5] Chongsuvivatwong, Virasakdi 2012 Analysis of Epidemiological Data Using R and Epicalc. Epidemiology Unit (Prince of Songkla University).
[6] Kleinbaum, D.G. and Klein, M 2005 Survival Analysis a Self-Learning Text (Springer, New York)
[7] Hutahaean, L.P., M.A. Mukidand T. Wuryandari 2014 Jurnal Gaussian 3(2): 173 – 181.
[8] Therneau, T. M. and Grambsch, P. M 2000 Modelling Survival Data: Extending the Cox Model (Springer, New York).
[9] Fisher, L. D. ve Lin D. Y 1999 Annual Review of Public Health 20, 145 – 157.
[10] Klein, John P. and Moeschberger, M. L. Survival Analysis Techniques for Censored and Truncated Data, (Springer, New York, 1997).
[11] Everitt, Brian. and Rabe-Hesketh, Sophia 2001 Analyzing Medical Data Using S-PLUS. (Springer, New York).
[12] Hosmer, D.W. and Lemeshow, S 2000 Applied Survival Analysis Regression Modeling of Time to Event Data (John Wiley and Sons, Inc, New York).
[13] Wagenmakers, Eric-Jan. and Farrel, Simon 2004 Psychonomic Bulletin & Review, 11 (1), 192-196.