ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 8, Nomor 1, Tahun 2019, Halaman 93 - 105
Online di: http://ejournal3.undip.ac.id/index.php/gaussian

PERBANDINGAN MODEL REGRESI COX *PROPORTIONAL HAZARD* MENGGUNAKAN METODE *BRESLOW* DAN *EFRON* (Studi Kasus: Penderita Stroke di RSUD Tugurejo Kota Semarang)

Eri Setiani[1], Sudarno[2], Rukun Santoso[3]
[1, 2, 3]Departemen Statistika FSM Universitas Diponegoro
Email: dsghani@gmail.com

ABSTRACT

Cox *proportional hazard* regression is a regression model that is often used in survival analysis. Survival analysis is phrase used to describe analysis of data in the form of times from a well-defined time origin until occurrence of some particular even or end-point. In analysis survival sometimes *ties* are found, namely there are two or more individual that have together event. This study aims to apply Cox model on *ties* event using two methods, *Breslow* and *Efron* and determine factors that affect survival of stroke patients in Tugurejo Hospital Semarang. Dependent variable in this study is length of stay, then independent variables are gender, age, type of stroke, history of hypertension, systolic blood pressure, diastolic blood pressure, blood sugar levels, and BMI. The two methods give different result, *Breslow* has four significant variables there are type of stroke, history of hypertension, systolic blood pressure, and diastolic blood pressure, while *Efron* contains five significant variables such as type of stroke, history of hypertension, systolic blood pressure, diastolic blood pressure and blood sugar levels. From the smallest AIC criteria obtained the best Cox *proportional hazard* regression model is *Efron* method.

Keywords: Stroke, Cox *Proportional Hazard Regression* model, *Breslow* method, *Efron* method.

1. PENDAHULUAN

Analisis ketahanan hidup atau analisis *survival* merupakan metode yang digunakan untuk menggambarkan analisis data dalam bentuk waktu dari waktu asal terdefinisi sampai kejadian tertentu terjadi (Collet, 2015). Terdapat beberapa metode yang digunakan untuk menganalisis data *survival*. Salah satunya yaitu metode regresi *survival* yang digunakan untuk mencari hubungan antara variabel bebas terhadap waktu *survival* sebagai variabel terikatnya. Model regresi yang sering digunakan adalah model semiparametrik, salah satunya yaitu regresi Cox *proportional hazard*. Hal ini dikarenakan pada model regresi Cox *proportional hazard* tidak memerlukan informasi tentang distribusi yang mendasari waktu ketahanan (Lee dan Wang, 2003). Dalam analisis *survival* terkadang ditemukan kejadian bersama atau sering disebut *ties*, yaitu keadaan ketika terdapat dua individu atau lebih yang mengalami event atau kejadian pada waktu yang bersamaan. Kejadian bersama tersebut dapat menimbulkan permasalahan pada estimasi parameter yang berhubungan dengan penentuan anggota dari himpunan risiko.

Penelitian ini menggunakan model Cox *proportional hazard* untuk mencari ada tidaknya pengaruh dari variabel bebas terhadap waktu *survival* dengan dua pendekatan dalam menyelesaikan kejadian bersama yaitu dengan pendekatan metode *Breslow* dan metode *Efron*.

Studi kasus pada penelitian ini adalah penderita stroke di RSUD Tugurejo Kota Semarang. Stroke merupakan salah satu penyakit kardiovaskuler, yaitu penyakit yang disebabkan gangguan fungsi jantung dan pembuluh darah.

Dari kriteria AIC terkecil didapatkan model terbaik regresi Cox *proportional hazard* menggunakan metode *Efron* dengan nilai AIC sebesar 453,157.

2. TINJAUAN PUSTAKA

2.1. Analisis *Survival*

Menurut Collet (2015), analisis *survival* merupakan metode yang digunakan untuk menggambarkan analisis data dalam bentuk waktu dari waktu asal terdefinisi sampai kejadian tertentu terjadi. Menurut Kleinbaum dan Klein (2005), waktu yang dimaksud adalah seperti jangka waktu tahun, bulan minggu, atau hari dari awal individu yang diamati sampai peristiwa yang diinginkan terjadi. Sedangkan peristiwa atau kejadian yang dimaksud seperti kematian, kejadian penyakit kambuh dari pengobatan, pemulihan atau pengalaman yang ditentukan oleh peneliti untuk kepentingan yang mungkin terjadi pada individu yang diteliti.

Menurut Lee dan Wang (2003), data waktu pada analisis *survival* digambarkan atau ditandai oleh tiga fungsi, yaitu fungsi ketahanan hidup (*survival function*), fungsi kepadatan peluang (*density function*), dan fungsi kegagalan (*hazard function*).

2.1.1. Fungsi Ketahanan Hidup (*Survival Function*)

Menurut Lee dan Wang (2003), variabel random yang menyatakan waktu ketahanan hidup sebuah objek disimbolkan dengan T dan fungsi ketahanan hidupnya dinotasikan dengan S(t) yang menunjukan probabilitas suatu individu bertahan hidup lebih dari waktu t, dengan t > 0. Maka S(t) dapat didefinisikan sebagai berikut:

$$\begin{aligned} S(t) &= P\,(\text{individu bertahan hidup lebih dari waktu t}) \\ &= P\,(T > t) \\ &= 1 - P\,(\text{individu gagal sebelum waktu t}) \\ &= 1 - P\,(T \leq t). \end{aligned} \quad (1)$$

2.1.2. Fungsi Kepadatan Peluang (*Density Function*)

Fungsi ini dinotasikan dengan f(t) dan didefinisikan sebagai probabilitas suatu individu gagal pada interval waktu t sampai t + Δt atau peluang kegagalan dalam interval per satuan waktu. Fungsi kepadatan peluang dinyatakan dengan

$$\begin{aligned} f(t) &= \lim_{\Delta t \to 0} \left[\frac{\text{P (objek gagal pada interval } (t,\ t+\Delta t)}{\Delta t} \right] \\ &= \lim_{\Delta t \to 0} \left[\frac{\text{P}\,(t < T < t + \Delta t)}{\Delta t} \right] \end{aligned} \quad (2)$$

2.1.3. Fungsi Kegagalan (*Hazard Function*)

Fungsi ini dinotasikan dengan h(t) dan didefinisikan sebagai tingkat kegagalan bersyarat yaitu probabilitas suatu individu gagal bertahan dalam interval waktu yang sangat pendek dari t sampai t +Δt, jika diketahui bahwa individu tersebut telah bertahan hingga waktu t. Fungsi kegagalan dari waktu tahan hidup T didefinisikan sebagai berikut:

$$h(t) = \frac{\lim\limits_{\Delta t \to 0} \text{P}\,(t < T < t + \Delta t \mid \text{T} > t)}{\Delta t} \quad (3)$$

2.2. Data Tersensor

Penyensoran terjadi apabila kejadian memiliki informasi tentang waktu *survival* individu, tetapi tidak tahu secara pasti waktu *survival*nya, maka data tersebut termasuk data tersensor (Kleinbaum dan Klein, 2005).

Pada umumnya ada tiga penyebab terjadinya data tersensor, yaitu:

1. *Study ends no event* yaitu subjek tidak mengalami kejadian sebelum penelitian berakhir.
2. *Lost to follow up* yaitu subjek menghilang selama masa pengamatan.
3. *Withdraws* yaitu subjek ditarik dari penelitian sebelum pengamatan berakhir karena kematian atau alasan lain.

Menurut Lee dan Wang (2003) ada tiga tipe penyensoran yang sering digunakan dalam eksperimen uji hidup, yaitu sebagai berikut:

1. Sensor Tipe I

   Sensor tipe I adalah tipe penyensoran ketika semua *n* objek masuk pada waktu yang sama dan percobaan akan dilakukan selama waktu T yang telah ditentukan dan akan berakhir setelah mencapai waktu T.
2. Sensor Tipe II

   Sensor tipe II adalah tipe penyensoran ketika semua objek yang diteliti (n) masuk pada waktu yang bersamaan dan pengujian akan dihentikan sampai *r* dari *n* unit mengalami kematian.
3. Sensor Tipe III

   Sensor tipe III adalah tipe penyensoran ketika objek yang diteliti masuk ke dalam percobaan pada waktu yang berlainan selama periode waktu tertentu.

2.3. Regresi Cox *Proportional Hazard*

Menurut Kleinbaum dan Klein (2005), regresi Cox *proportional hazard* adalah model matematika yang popular digunakan untuk menganalisis data ketahanan hidup. Melalui model Cox *proportional hazard* dapat diketahui pengaruh dari variabel bebas terhadap variabel terikat yaitu waktu *survival* suatu objek terhadap suatu peristiwa tertentu. Model regresi Cox *proportional hazard* sebagai berikut:

$$h_i(t) = exp\left(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}\right) h_0(t) = exp\ (\boldsymbol{\beta}'\boldsymbol{x}_i)\ \ h_0(t) \qquad (4)$$

dengan:

$h_i(t)$: fungsi kegagalan individu ke-i

$h_0(t)$: fungsi kegagalan dasar

$\boldsymbol{\beta}'$: vektor koefisien dari variabel bebas $x_1, x_2, \ldots, x_p$.

$\boldsymbol{x}_i$: vektor nilai variabel bebas $x_1, x_2, \ldots, x_p$ individu ke-i, dengan i = 1,2,…,n.

2.4. Estimasi Parameter Tanpa Kejadian Bersama

Parameter regresi pada persamaan Cox *proportional hazard* dapat diketahui dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Koefisien regresi Cox *proportional hazard* ditaksir terlebih dahulu sebelum menaksir fungsi kegagalan dasar. Fungsi likelihood dari model regresi Cox *proportional hazard* adalah sebagai berikut:

$$\mathrm{L}\ (\beta) = \prod_{i=1}^{n} \frac{\exp(\sum_{j=1}^{p} \beta_j\, x_{ji})}{\sum_{l \epsilon R(t_i)} \exp(\sum_{j=1}^{p} \beta_j\, x_{jl})} \qquad (5)$$

dimana

β_j : koefisien regresi

x_{ji} : variabel dari individu yang gagal pada waktu ke t_i

$R(t_i)$: himpunan individu yang bertahan pada waktu ke t_i

x_{jl} : variabel individu yang masih hidup dan merupakan elemen dari $R(t_i)$.

Fungsi log likelihood yang bersesuaian dapat dituliskan

$$\log \mathrm{L}(\beta) = \sum_{i=1}^{n} \sum_{j=1}^{p} \beta_j\, x_{ji} - \sum_{i=1}^{n} \log\left[\sum_{l \epsilon R(t_i)} \exp(\sum_{j=1}^{p} \beta_j\, x_{jl})\right] \qquad (6)$$

Pendugaan parameter β_j pada model regresi Cox *proportional hazard* dapat diperoleh dengan memaksimumkan fungsi log likelihood. Karena persamaan (7) bersifat *close form* maka penyelesaian dilakukan menggunakan salah satu metode numerik, yaitu metode iterasi Newton-Raphson.

Misalkan $\boldsymbol{u}(\boldsymbol{\beta})$ sebagai vektor $p \times 1$ turunan pertama dari fungsi log likelihood terhadap $\boldsymbol{\beta_j}$.

$$\frac{\partial logL(\beta)}{\partial \beta_j} = \sum_{i=1}^{n}\sum_{j=1}^{p} x_{ji} - \sum_{i=1}^{n} \frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})} = \sum_{i=1}^{n}\left\{\sum_{j=1}^{p} X_{ji} - \left[\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} X_{jl}\exp(\sum_{j=1}^{p}\beta_j X_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j X_{jl})}\right]\right\}$$

sehingga,

$$\boldsymbol{u(\beta)} = \sum_{i=1}^{n}\left\{\sum_{j=1}^{p} x_{ji} - \left[\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}\right]\right\} = 0 \qquad (7)$$

$$\boldsymbol{u(\beta)}_{px1} = \begin{pmatrix} \frac{\partial logL(\beta)}{\partial \beta_1} \\ \vdots \\ \frac{\partial logL(\beta)}{\partial \beta_p} \end{pmatrix}.$$

Langkah selanjutnya adalah dimisalkan $\boldsymbol{I(\beta)} = -\left\{\frac{\partial^2 logL(\beta)}{\partial \beta_j \partial \beta_s}\right\}$, dengan $\boldsymbol{I(\beta)}$ adalah matrik yang berukuran $p \times p$ dari turunan kedua negatif dari fungsi log likelihood.

$$\left\{\frac{\partial^2 logL(\beta)}{\partial \beta_j \partial \beta_s}\right\} = \left\{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p} X_{ji} - \sum_{i=1}^{n}\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}}{\partial \beta_s}\right\}$$

Misalkan u = $\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})$

$u' = \sum_{l\epsilon R(t_i)}\sum_{j=1}^{p}\sum_{s=1}^{p} x_{jl}\, x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})$

$v = \sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})$

$v' = \sum_{l\epsilon R(t_i)}\sum_{s=1}^{p} x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})$

$$\frac{u'v - v'u}{v^2} = \frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p}\sum_{s=1}^{p} x_{jl}x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl}) - \sum_{l\epsilon R(t_i)}\sum_{s=1}^{p} x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}$$

$$= \sum_{i=1}^{n}\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p}\sum_{s=1}^{p} x_{jl}\, x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})} - \sum_{i=1}^{n}\left[\frac{\sum_{l\epsilon R(t_i)}\sum_{s=1}^{p} x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}\right]\left[\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}\right]$$

$$\left\{\frac{\partial^2 logL(\beta)}{\partial \beta_j \partial \beta_s}\right\} = -\left\{\sum_{i=1}^{n}\left\{\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p}\sum_{s=1}^{p} x_{jl}\, x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})} - \left[\frac{\sum_{l\epsilon R(t_i)}\sum_{s=1}^{p} x_{sl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}\right]\left[\frac{\sum_{l\epsilon R(t_i)}\sum_{j=1}^{p} x_{jl}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}{\sum_{l\epsilon R(t_i)}\exp(\sum_{j=1}^{p}\beta_j x_{jl})}\right]\right\}\right\} \qquad (8)$$

Matrik yang diharapkan adalah

$$\boldsymbol{I(\beta)}_{p\times p} = -\left\{\frac{\partial^2 logL(\beta)}{\partial \beta_j \partial \beta_s}\right\}$$

$$\boldsymbol{I(\beta)}_{p\times p} = -\begin{pmatrix} \frac{\partial^2 logL(\beta)}{\partial \beta_1 \partial \beta_1} & \vdots & \frac{\partial^2 logL(\beta)}{\partial \beta_1 \partial \beta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 logL(\beta)}{\partial \beta_p \partial \beta_1} & \cdots & \frac{\partial^2 logL(\beta)}{\partial \beta_p \partial \beta_p} \end{pmatrix}$$

Untuk menyelesaikan taksiran parameternya digunakan metode iterasi Newton Raphson dengan statistik $\widehat{\boldsymbol{\beta}}_{\boldsymbol{0}}$ yaitu suatu nilai awal yang ditentukan, metode iterasinya adalah sebagai berikut:

$$\widehat{\boldsymbol{\beta}}_{S+1} = \widehat{\boldsymbol{\beta}}_{S} + \boldsymbol{I}^{-1}(\widehat{\boldsymbol{\beta}}_{S})\,\boldsymbol{u}(\widehat{\boldsymbol{\beta}}_{S})$$

iterasi berhenti apabila $\|\hat{\boldsymbol{\beta}}_{S+1}-\hat{\boldsymbol{\beta}}_{S}\| < \varepsilon$, untuk suatu ε yang kecil. Dalam penelitian ini diambil $\varepsilon = 10^{-5}$

dengan $\|\hat{\boldsymbol{\beta}}_{S+1}-\hat{\boldsymbol{\beta}}_{S}\| = \sqrt{(\beta_{S+1,1}-\beta_{S,1})^2 + (\beta_{S+1,2}-\beta_{S,2})^2 + \cdots + (\beta_{S+1,p}-\beta_{S,p})^2}$.

Menurut Collet (2015) setelah didapat estimasi parameter β, fungsi kegagalan individu dapat ditaksir jika nilai dari $\hat{h}_0(t)$ telah diketahui dengan

$\hat{h}_0(t) = 1 - \xi_i$

untuk $d_i = 1, \xi_i = \left(1 - \frac{\exp(\hat{\boldsymbol{\beta}}'\boldsymbol{x}_i)}{\sum_{l\epsilon R(t_i)}\exp(\hat{\boldsymbol{\beta}}'\boldsymbol{x}_l)}\right)^{\exp(-\hat{\boldsymbol{\beta}}'\boldsymbol{x}_i)}$

untuk $d_i > 1, \xi_i = \exp\left(\frac{-d_i}{\sum_{l\epsilon R(t_i)}\exp(\hat{\boldsymbol{\beta}}'\boldsymbol{x}_l)}\right)$

dengan:

ξ_i : fungsi kegagalan individu ke-i pada waktu ke t dengan i=1,2,…,n.

$\boldsymbol{x}_i$: vektor dari variabel individu ke-i yang gagal pada waktu ke t dengan i=1,2,…,n.

$\boldsymbol{x}_l$: vektor variabel individu yang masih hidup dan merupakan elemen dari $R(t_i)$

$R(t_i)$: himpunan individu yang bertahan pada waktu t_i dengan i=1,2,…,n

d_i : banyak kasus ties pada waktu t_i

Taksiran fungsi ketahanan hidup dasar dapat dihitung dengan:

$\hat{S}_0(t) = \prod_{i=1}^{n} \xi_i$, n=1,2,…,r.

r merupakan individu yang teramati.

2.5. Estimasi Parameter Pada Kejadian Bersama

Menurut Collet (2015) kejadian bersama (*ties*) merupakan keadaan ketika terdapat dua individu atau lebih yang mengalami kejadian gagal pada waktu yang bersamaan atau memiliki nilai waktu tahan hidup yang sama. Kejadian bersama tersebut dapat menimbulkan permasalahan pada estimasi parameter yang berhubungan dengan penentuan anggota dari himpunan risiko. Ada beberapa cara pendekatan pada kejadian bersama, dua diantaranya yaitu dengan pendekatan metode *Breslow* dan metode *Efron*.

2.5.1. Metode *Breslow*

Metode *Breslow* mengasumsikan bahwa ukuran dari himpunan risiko pada kejadian bersama adalah sama. Berikut ini adalah fungsi likelihood untuk metode *Breslow* menurut Collet (2015) :

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\boldsymbol{\beta}' s_i)}{\{\sum_{l\epsilon R(t_i)}\exp(\boldsymbol{\beta}' x_l)\}^{d_i}} \tag{9}$$

dimana:

s_i : jumlah nilai variabel x_j dari semua individu yang mati pada saat t_i

x_l : variabel dari individu yang masih bertahan dan merupakan elemen dari $R(t_i)$.

d_i : banyaknya kejadian bersama pada waktu t_i

$R(t_i)$: himpunan individu yang berisiko pada waktu t_i yang terdiri dari individu yang bertahan pada waktu t_i

2.5.2. Metode *Efron*

Pada pendekatan metode *Efron*, himpunan risikonya diselesaikan dengan pengurangan terhadap rata-rata dari nilai fungsi hazard dari variabel ke-j. Berikut ini adalah fungsi likelihood untuk metode *Efron* menurut Collet (2015) :

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\boldsymbol{\beta}' s_i)}{\prod_{j=1}^{di}\left[\sum_{l\in R(t_i)} \exp(\boldsymbol{\beta}'\, x_l) - \frac{j-1}{d_i}\sum_{l\in D(t_i)} \exp(\boldsymbol{\beta}'\, x_l)\right]} \quad (10)$$

dimana:

s_i : jumlah nilai variabel x_j dari semua individu yang mati pada saat t_i

$R(t_i)$: himpunan individu yang berisiko pada waktu t_i yang terdiri dari individu yang bertahan pada waktu t_i

x_l : variabel dari individu yang masih bertahan dan merupakan elemen dari $R(t_i)$.

2.6. Asumsi Cox *Proportional Hazard*

Dalam menentukan asumsi *proportional hazard* yaitu melalui pendekatan *Goodness of Fit* menggunakan *schoenfeld residual*. langkah-langkah untuk pengujian asumsi *proportional hazard* menggunakan *Goodness of Fit* sebagai berikut:

1. Membangun model Cox *proportional hazard* dan *schoenfeld residual* untuk masing-masing individu pada setiap variabel bebas.
2. Membuat variabel yang menyatakan peringkat dari waktu survival.
3. Menguji korelasi antara variabel *schoenfeld residual* dengan waktu survival pada langkah ke-2.

   Hipotesis:

   H_0 : $\rho = 0$ (asumsi proportional hazard terpenuhi) Vs

   H_1 : $\rho \neq 0$ (asumsi proportional hazard tidak terpenuhi).

   Statistik Uji:

$$r_{hitung} = \frac{\sum_i^n (R_{ji} - \bar{R}_{ji})\,(RT_i - \overline{RT}_i)}{\sqrt{\sum_i^n (R_{ji} - \bar{R}_{ji})^2}\,\sqrt{\sum_i^n (RT_i - \overline{RT}_i)^2}} \quad (11)$$

   dimana:

   r_{hitung} : koefisien korelasi

   R_{ji} : *schoenfeld residual* untuk masing-masing variabel bebas ke-j dengan j=1,2,..,p untuk individu ke=i dengan i=1,2,…,n.

   RT_i : waktu *survival* individu ke-i

   Kriteria Penolakan:

   H_0 ditolak jika | r_{hitung} | > $r_{n-2,\frac{\alpha}{2}}$ atau *p-value* < α

2.7. Pengujian Parameter

Menurut Hosmer *et al.* (2008) untuk menguji signifikansi parameter dapat dilakukan menggunakan uji rasio likelihood dan uji *wald*. Pengujian signifikansi parameter bertujuan untuk memeriksa apakah variabel bebas memiliki pengaruh dalam model.

a. Pengujian Secara Serentak (Uji Rasio Likelihood)

   Hipotesis:

   H_0 : $\beta_1 = \beta_2 = \ldots = \beta_p = 0$

   H_1 : minimal ada satu $\beta_j \neq 0$, dengan j = 1,2,…,p

   Statistik Uji:

$$G = -2\,[\ln L(0) - \ln L(\hat{\beta})] \quad (12)$$

   Kriteria Penolakan:

   H_0 ditolak jika $G \geq x^2_{(\alpha:db=p)}$ atau p-value < α

b. Pengujian Secara Parsial (Uji *Wald*)

   Hipotesis:

   H_0 : $\beta_j = 0$

$H_1 : \beta_j \neq 0$, dengan $j = 1,2,\ldots,p$

Statistik Uji:

$$Z^2 = \left[\frac{\hat{\beta}_j}{SE\hat{\beta}_j}\right]^2 \tag{13}$$

dengan $SE\hat{\beta}_j = \sqrt{var\ \hat{\beta}_j}$

Kriteria Penolakan:

H_0 ditolak jika $Z^2 > x^2_{(1;\alpha)}$ atau p-value < α

2.8. Model Terbaik

Model terbaik adalah model yang memiliki nilai *Akaike's Information Criterion* (AIC) terkecil. Rumus menghitung nilai *Akaike's Information Criterion* (AIC) adalah sebagai berikut:

$$AIC = -2log\hat{L} + 2p \tag{14}$$

dengan:

$\hat{L}$: fungsi likelihood

p : banyaknya parameter β

2.9. Rasio Kegagalan

Menurut Kleinbaum dan Klein (2005), rasio kegagalan adalah kegagalan untuk satu kelompok individu dibagi dengan kegagalan untuk kelompok individu yang berbeda. Rasio kegagalan dapat dinyatakan ke dalam bentuk sebagai berikut:

$$\widehat{HR} = \frac{\widehat{h_0}(t)\, e^{\sum_{j=1}^{p}\hat{\beta}_j X_{ij}^*}}{\widehat{h_0}(t)\, e^{\sum_{j=1}^{p}\hat{\beta}_j X_{ij}}} = e^{\sum_{j=1}^{p}\hat{\beta}_j (X_{ij}^* - X_{ij})} \tag{15}$$

2.10. Normalisasi *Z-Score*

Menurut Bhandare dan Jain (2014) dalam metode normalisasi *Z-Score* nilai dari masing-masing variabel dinormalisasi berdasarkan rata-rata dan standar deviasi dari variabel. Persamaan untuk normalisasi Z-Score sebagai berikut:

$$Z = \frac{x_i - \mu}{\sigma} \tag{16}$$

dimana:

Z : nilai masing-masing variabel yang dinormalisasi

x_i : nilai masing-masing variabel awal

μ : rata-rata dari variabel

σ : standar deviasi dari variabel

2.11. Penyakit Stroke

Stroke merupakan salah satu jenis Penyakit Tidak Menular (PTM). Dalam bahasa medis stroke disebut CVA (*celebro-vascular accident*). Merujuk pada istilah medis, stroke didefinisikan sebagai gangguan saraf permanen akibat terganggunya peredaran darah ke otak yang terjadi sekitar 24 jam atau lebih (Lingga, 2013). Secara garis besar, faktor risiko stroke dibagi menjadi dua, yaitu faktor tidak terkendali atau faktor yang bersifat menetap dan faktor yang dapat dikendalikan atau faktor tidak tetap. Faktor tidak terkendali antara lain faktor genetik (ras), usia, gender, serta riwayat penyakit yang dialami oleh orang tua atau saudara sekandung, sedangkan faktor yang dapat dikendalikan terdiri atas gaya hidup tidak sehat yang memicu terjadinya penyakit-penyakit tertentu yang mendorong serangan otak seperti kebiasaan merokok, mengkonsumsi minuman beralkohol dan malas berolahraga.

3. METODE PENELITIAN

3.1. Jenis dan Sumber Data

Jenis data yang digunakan pada penelitian ini adalah data sekunder Data sekunder yang dimaksud dalam penelitian ini adalah data rekam medis pasien stroke di Rumah Sakit Umum Daerah Tugurejo Kota Semarang periode Januari 2018 sampai Desember 2018.

3.2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel terikat dan variabel bebas. Variabel terikat adalah lama waktu survival atau tahan hidup. Sedangkan variabel bebas antara lain jenis kemalin (X_1), usia (X_2), jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), tekanan darah diastolik (X_6), kadar gula darah sewaktu (X_7) dan indeks massa tubuh (X_8).

3.3. Analisis Data

Langkah-langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan analisis deskriptif
2. Melakukan standardisasi pada variabel bebas kontinyu menggunakan normalisasi *Z-Score.*
3. Pemodelan regresi Cox *proportional hazard* dengan metode *Breslow* dan *Efron.*
   a. Membuat model awal regresi Cox *proportional hazard.*
   b. Melakukan uji asumsi Cox *proportional hazard* menggunakan uji *Goodness of Fit*.
   c. Melakukan uji signifikansi parameter yang terdiri dari uji secara serentak dan uji secara parsial.
   d. Membentuk model akhir regresi Cox *proportional hazard.*
4. Membandingkan model regresi Cox *proportional hazard* dengan metode *Breslow* dan *Efron* dengan melihat nilai AIC masing-masing model. Model terbaik adalah model yang memiliki nilai AIC yang lebih kecil.
5. Melakukan interpretasi model terbaik regresi Cox *propotional hazard* yang telah terbentuk.

4. HASIL DAN PEMBAHASAN

4.1 Analisis Deskriptif

Jumlah ukuran sampel pada penelitian ini adalah 288 dengan 51 sampel atau 18% sampel merupakan data teramati atau data tidak tersensor dan sisanya yaitu 237 sampel atau 82% merupakan data tersensor.

4.2. Pemodelan Regresi Cox *Proportional Hazard* dengan Metode *Breslow*

4.2.1. Pemodelan Awal Regresi Cox *Proportional Hazard*

Model regresi Cox *proportional hazard* digunakan untuk mengetahui pengaruh variabel bebas terhadap variabel terikat. Pada penelitian ini, variabel bebas yang digunakan yaitu jenis kelamin (X_1), usia (X_2), jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), tekanan darah diastolik (X_6), kadar gula darah sewaktu (X_7) dan indeks massa tubuh (X_8). Model awal regersi Cox *proportional hazard* dengan metode *Breslow* sebagai berikut:

$$h(t,X) = h_0(t)\exp(-0{,}08414X_1 + 0{,}12040X_2 + 1{,}96797X_3 - 1{,}38160\,X_4 + 0{,}63952X_5 - 0{,}44002X_6 + 0{,}21005X_7 + 0{,}10943X_8)$$

4.2.2. Pengujian Asumsi *Proportional Hazard*

Pengujian asumsi *proportional hazard* dalam penelitian ini dilakukan secara formal dengan uji *Goodness of Fit* menggunakan *schoenfeld residual*.

Tabel 1. Pengujian Asumsi *Proportional Hazard* Metode *Breslow*

Variabel	r_{hitung}	*P-value*	Keputusan
X_1	0,1141	0,387	H_0 diterima
X_2	-0,0371	0,816	H_0 diterima
X_3	0,0267	0,837	H_0 diterima
X_4	0,0115	0,935	H_0 diterima
X_5	0,0053	0,968	H_0 diterima
X_6	0,0542	0,683	H_0 diterima
X_7	-0,0896	0,656	H_0 diterima
X_8	-0,0570	0,720	H_0 diterima

Pada taraf signifikansi α=0,05 dapat disimpulkan bahwa antara nilai *schoenfeld residual* dengan waktu *survival* tidak ada korelasi yang artinya semua variabel yang diduga mempengaruhi waktu *survival* pasien stroke memenuhi asumsi *proportional hazard*.

4.2.3. Pengujian Parameter

Setelah model awal Cox *proportional hazard* dengan metode *Breslow* diperoleh dan asumsi *proportional hazard* terpenuhi, maka dilakukan pengujian parameter dari model tersebut secara serentak maupun parsial menggunakan uji Rasio Likelihood dan uji *Wald*.

a. Pengujian Rasio Likelihood Model Awal

Dari model awal diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 49,6) > (χ^2(5%:8) = 15,507) atau (p-value = 4,881e-08) < (α = 0,05).

b. Pengujian *Wald* Model Awal

Tabel 2. Uji *Wald* Model Awal Metode *Breslow*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_1	-0,08414	0,32791	0,79749	H_0 diterima
X_2	0,12040	0,15455	0,43595	H_0 diterima
X_3	1,96797	0,33792	5,75e-09	H_0 ditolak
X_4	-1,38160	0,50995	0,00674	H_0 ditolak
X_5	0,63952	0,26683	0,01654	H_0 ditolak
X_6	-0,44002	0,23512	0,06128	H_0 diterima
X_7	0,21005	0,11663	0,07170	H_0 diterima
X_8	0,10943	0,13832	0,42886	H_0 diterima

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain variabel jenis stroke (X_3), riwayat hipertensi (X_4), dan tekanan darah sistolik (X_5). Karena ada variabel yang tidak signifikan terhadap model maka variabel yang tidak signifikan dikeluarkan dari model menggunakan seleksi *backward*.

c. Pengujian Rasio Likelihood Model Kedua

Berdasarkan seleksi *backward* yang dilakukan diperoleh model kedua Cox *proportional hazard* dengan pendekatan *Breslow* sebagai berikut:

$$h(t,X) = h_0(t)\exp(0{,}1249X_2 + 1{,}9773X_3 - 1{,}3817X_4 + 0{,}6552X_5 - 0{,}4579X_6 + 0{,}2191X_7 + 0{,}1137X_8).$$

Dari model kedua diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 49,533) > ($\chi^2_{(5\%:4)}$ = 14,067) atau (*p-value* = 1,784e-08) < (α = 0,05).

d. Pengujian *Wald* Model Kedua

Tabel 3. Uji *Wald* Model Kedua Metode *Breslow*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_2	0,1249	1,1330	0.41693	H_0 diterima
X_3	1,9773	0,3364	4,17e-09	H_0 ditolak
X_4	-1,3817	0,2511	0.00682	H_0 ditolak
X_5	0,6552	0,2602	0,01182	H_0 ditolak
X_6	-0,4579	0,2256	0,04237	H_0 ditolak
X_7	0,2191	0,1112	0,04881	H_0 ditolak
X_8	0,1137	0,1380	0,40992	H_0 diterima

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain variabel jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), dan tekanan darah diastolik (X_6), dan kadar gula darah sewaktu (X_7).

Proses dilakukan dengan analogi yang sama hingga didapatkan model akhir yaitu sebagai berikut:

$$h(t,X) = h_0(t)\exp(1{,}8186X_3 - 1{,}3931X_4 + 0{,}6525X_5 - 0{,}4569X_6).$$

e. Pengujian Rasio Likelihood Model Akhir

Dari model akhir diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 45,391) > ($\chi^2_{(5\%:4)}$ = 9,488) atau (*p-value* = 2,85e-09) < (α = 0,05).

f. Pengujian *Wald* Model Akhir

Tabel 4. Uji *Wald* Model Akhir Metode *Breslow*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_3	1,8186	0,3185	1,13e-08	H_0 ditolak
X_4	-1,3931	0,5017	0,00549	H_0 ditolak
X_5	0,6525	0,2536	0,01009	H_0 ditolak
X_6	-0,4569	0,2137	0,03252	H_0 ditolak

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain variabel jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), dan tekanan darah diastolik (X_6).

Karena seluruh variabel pada model telah signifikan maka diperoleh model akhir regresi Cox *proportional hazard* dengan pendekatan menggunakan metode *Breslow* yaitu sebagai berikut:

$$h(t,X) = h_0(t)\exp(1{,}8186X_3 - 1{,}3931X_4 + 0{,}6525X_5 - 0{,}4569X_6).$$

4.3. Pemodelan Regresi Cox *Proportional Hazard* dengan Metode *Efron*

4.3.1. Pemodelan Awal Regresi Cox *Proportional Hazard*

Model awal regresi Cox *proportional hazard* dengan metode *Efron* sebagai berikut:

$$h(t,X) = h_0(t)\exp(-0{,}0941X_1 + 0{,}1274X_2 + 2{,}0264X_3 - 1{,}4175X_4 + 0{,}6590X_5 - 0{,}4523X_6 + 0{,}2145X_7 + 0{,}1123X_8).$$

4.3.2. Pengujian Asumsi *Proportional Hazard*

Setelah diperoleh model awal selanjutnya dilakukan pengujian asumsi *proportional hazard* dengan uji *Goodness of Fit* menggunakan *schoenfeld residual*.

Tabel 5. Pengujian Asumsi *Proportional Hazard* Metode *Efron*

Variabel	r_{hitung}	*P-value*	Keputusan
X_1	0,11106	0,396	H_0 diterima
X_2	-0,03805	0,808	H_0 diterima
X_3	0,03119	0,807	H_0 diterima
X_4	0,01215	0,931	H_0 diterima
X_5	0,00599	0,963	H_0 diterima
X_6	0,05759	0,660	H_0 diterima
X_7	-0,09111	0,645	H_0 diterima
X_8	-0,05972	0,704	H_0 diterima

Pada taraf signifikansi α=0,05 dapat disimpulkan bahwa antara nilai *schoenfeld residual* dengan waktu *survival* tidak ada korelasi yang artinya semua variabel yang diduga mempengaruhi waktu *survival* pasien stroke memenuhi asumsi *proportional hazard*.

4.3.3. Pengujian Parameter

Setelah model awal Cox *proportional hazard* dengan metode *Efron* diperoleh dan asumsi *proportional hazard* terpenuhi, maka dilakukan pengujian parameter dari model tersebut secara serentak maupun parsial menggunakan uji Rasio Likelihood dan uji *wald*.

a. Pengujian Rasio Likelihood Model Awal

Dari model awal diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 51,78) > ($\chi^2_{(5\%:8)}$ = 15,507) atau (*p-value* = 1,857e-08) < (α = 0,05).

b. Pengujian *Wald* Model Awal

Tabel 6. Uji *Wald* Model Awal Metode *Efron*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_1	-0,0941	0,3280	0,77423	H_0 diterima
X_2	0,1274	0,1551	0,41146	H_0 diterima
X_3	2,0264	0,3410	2,8e-09	H_0 ditolak
X_4	-1,4175	0,5102	0,00547	H_0 ditolak
X_5	0,6590	0,2679	0,01388	H_0 ditolak
X_6	-0,4523	0,2361	0,05533	H_0 diterima
X_7	0,2145	0,1153	0,06288	H_0 diterima
X_8	0,1123	0,1381	0,41580	H_0 diterima

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain variabel variabel jenis stroke (X_3), riwayat hipertensi (X_4), dan tekanan darah sistolik (X_5). Karena ada variabel yang tidak signifikan terhadap model maka variabel yang tidak signifikan dikeluarkan dari model menggunakan seleksi *backward*.

c. Pengujian Rasio Likelihood Model Kedua

Berdasarkan seleksi *backward* yang dilakukan diperoleh model kedua Cox *proportional hazard* dengan pendekatan *Efron* sebagai berikut:

$$h(t,X) = h_0(t)\exp(0{,}1324X_2 + 2{,}0368X_3 - 1{,}4185X_4 + 0{,}6768X_5 - 0{,}4725X_6 + 0{,}2247X_7 + 0{,}1174X_8).$$

Dari model kedua diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 51,6974) > ($\chi^2_{(5\%:4)}$ = 14,067) atau (*p-value* = 6,698e-09) < (α = 0,05).

d. Pengujian *Wald* Model Kedua

Tabel 7. Uji *Wald* Model Kedua Metode *Efron*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_2	0,1324	0,1545	0,39136	H_0 diterima
X_3	2,0368	0,3396	2e-09	H_0 ditolak
X_4	-1,4185	0,5111	0,00551	H_0 ditolak
X_5	0,6768	0,2610	0,00951	H_0 ditolak
X_6	-0,4725	0,2265	0,03696	H_0 ditolak
X_7	0,2247	0,1097	0,04056	H_0 ditolak
X_8	0,1174	0,1377	0,39390	H_0 diterima

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), dan tekanan darah diastolik (X_6), dan kadar gula darah sewaktu (X_7).

Proses dilakukan dengan analogi yang sama hingga didapatkan model akhir yaitu sebagai berikut:

$$h(t,X) = h_0(t)\ \exp(1{,}9851X_3 - 1{,}3377X_4 + 0{,}6569X_5 - 0{,}4510X_6 + 0{,}2118X_7).$$

e. Pengujian Rasio Likelihood Model Akhir

Dari model akhir diperoleh hasil bahwa paling sedikit ada satu variabel bebas dari persamaan yang berpengaruh secara signifikan terhadap waktu *survival* pasien stroke karena (G = 50,57) > ($\chi^2_{(5\%:5)}$ = 11,070) atau (*p-value* = 1,058e-09) < (α = 0,05).

f. Pengujian *Wald* Model Akhir

Tabel 8. Uji *Wald* Model Akhir Metode *Efron*

Variabel	$\hat{\beta}_j$	$SE\hat{\beta}_j$	*P-value*	Keputusan
X_3	1,9851	0,3350	3,1e-09	H_0 ditolak
X_4	-1,3377	0,5097	0,00868	H_0 ditolak
X_5	0,6569	0,2560	0,01028	H_0 ditolak
X_6	-0,4510	0,2173	0,03793	H_0 ditolak
X_7	0,2118	0,1076	0,04893	H_0 ditolak

Pada uji *Wald* variabel bebas yang signifikan terhadap model antara lain variabel jenis stroke (X_3), riwayat hipertensi (X_4), tekanan darah sistolik (X_5), tekanan darah diastolik (X_6), dan kadar gula darah sewaktu (X_7).

Karena seluruh variabel pada model telah signifikan maka diperoleh model akhir regresi Cox *proportional hazard* dengan pendekatan menggunakan metode *Efron* yaitu sebagai berikut:

$$h(t,X) = h_0(t)\ \exp(1{,}9851X_3 - 1{,}3377X_4 + 0{,}6569X_5 - 0{,}4510X_6 + 0{,}2118X_7).$$

4.4. Perbandingan Model Regresi Cox *Proportional Hazard* dengan Metode *Breslow* dan *Efron*

Tabel 9. Perbandingan Nilai AIC Model *Breslow* dan *Efron*

Metode	AIC
Breslow	457,4977
Efron	453,1567

Berdasarkan tabel 9, dapat dilihat bahwa nilai AIC yang lebih kecil ada pada model regresi Cox *proportional hazard* menggunakan metode *Efron* yaitu 453,1567, sehingga dapat disimpulkan berdasarkan kriteria AIC terkecil maka model terbaik yang diperoleh yaitu model regresi Cox *proportional hazard* menggunakan metode *Efron* yang terdiri dari lima variabel bebas yaitu jenis stroke (X_3), riwayat stroke (X_4), tekanan darah sistolik (X_5), tekanan darah diastolik (X_6), dan kadar gula darah sewaktu (X_7).

4.5. Interpretasi Model

Interpretasi *hazard ratio* dari model terbaik yaitu risiko pasien stroke gagal bertahan hidup dari masing-masing variabel bebas yang berpengaruh terhadap model. Penderita stroke jenis hemoragik memiliki risiko untuk gagal bertahan hidup lebih besar daripada penderita dengan jenis stroke iskemik yaitu sebesar 7,280 kali. Sedangkan untuk penderita stroke yang memiliki riwayat hipertensi memiliki risiko gagal bertahan hidup lebih kecil daripada penderita stroke yang tidak memiliki riwayat hipertensi sebesar 0,262 kali. Pada variabel tekanan darah sistolik, semakin besarnya tekanan darah sistolik maka semakin menambah besar risiko yang dimiliki oleh pasien stroke untuk gagal bertahan hidup sebesar 92,9%. Sedangkan pada variabel tekanan darah diastolik, menurunnya tekanan darah diastolik akan menambah risiko yang dimiliki pasien stroke untuk gagal bertahan hidup sebesar 36,3%. Dan pada variabel kadar gula darah sewaktu, semakin besar kadar gula darah sewaktu maka semakin menambah besar risiko pasien stroke untuk gagal bertahan hidup sebesar 23,6%.

5. KESIMPULAN

Berdasarkan hasil penelitian mengenai regresi Cox *proportional hazard* menggunakan metode *Breslow* dan *Efron* pada data *survival* pasien stroke di RSUD Tugurejo Kota Semarang maka dapat ditarik kesimpulan sebagai berikut:

1. Pada penerapan data dengan menggunakan metode pendekatan *Breslow*, diperoleh model akhir Cox *proportional hazard* sebagai berikut.
   $h(t,X) = h_0(t)\exp(1{,}8186X_3 - 1{,}3931X_4 + 0{,}6525X_5 - 0{,}4569X_6)$.
2. Pada penerapan data dengan menggunakan metode pendekatan *Efron*, diperoleh model akhir Cox *proportional hazard* sebagai berikut.
   $h(t,X) = h_0(t)\exp(1{,}9851X_3 - 1{,}3377X_4 + 0{,}6569X_5 - 0{,}4510X_6 + 0{,}2118X_7)$.
3. Dari kriteria AIC terkecil didapatkan model terbaik regresi Cox *proportional hazard* dengan pendekatan metode *Efron* yang terdiri dari lima variabel bebas yaitu jenis stroke (X_3), riwayat stroke (X_4), tekanan darah sistolik (X_5), tekanan darah diastolik (X_6), dan kadar gula darah sewaktu (X_7).

DAFTAR PUSTAKA

Bhandare, S.K. dan Jain Y.K. 2014. *Min Max Normalization Based Data Pertubation Method for Privacy Protection.* International Journal of Computer & Communication Technology Vol. 3, No. 4: Hal 45-50.

Collett, D. 2015. *Modelling Survival Data in Medical Research third edition*. CRC Press.

Hosmer, D.W. *et al.* 2008. *Applied Survival Analysis: Regression Modelling of Time to Event Data second edition*. New Jersey: Jhon Wiley.

Kleinbaum, D.G., dan Klein, M. 2005. *Survival Analysis A Self-Learning Text 2nd Edition*. Landon : Springer.

Lee, E. T. dan Wang, J. W. 2003. *Statsitical Methods for Survival Data Analysis*. Canada: JohnWiley & Sons, Inc.

Lingga, Lanny. 2013. *All About Stroke Hidup Sebelum dan Pasca Stroke*. Jakarta: PT Elex Media Komputindo.