
Aplikasi Model *Cox Proportional Hazard* pada Pasien Stroke RSD Balung Kabupaten Jember

Tutik Qomaria¹, Mohamad Fatekurohman², dan Dian Anggraeni³

^{1,2,3}Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Jember

tutikqomariyah12@gmail.com, mfatekurohman.fmipa@unej.ac.id,
dian_a.fmipa@unej.ac.id

Abstract. According to the World Health Organization (WHO) cardiovascular disease is a disease caused by impaired heart and blood vessel function. There are many types of cardiovascular disease, but the most common and most well-known are coronary heart disease and stroke. Stroke is a syndrome characterized by symptoms and / or rapidly developing clinical signs in the form of focal and global brain functional disorders lasting more than 24 hours (unless there are surgical interventions or bringing death), which are not caused by other causes besides vascular causes. The number of stroke patients in Indonesia in 2013 based on the diagnosis of health personnel (Nakes) was 1.236.825 (7,0%), while based on the diagnosis of symptoms was 2.137.941 (12,1%). In this study the factors that can affect the survival of stroke sufferers were analyzed using the Cox proportional hazard regression model, the dependent variable was the length of time the patient was treated and the independent variables were gender, age, hypertension status, cholesterol status, Diabetes Militus (DM) status, stroke type, and Body Mass Index (BMI). The result showed that age, DM status, and type of stroke were the most influential factors on the survival of stroke patients at Balung Regional Hospital.

Keywords : stroke disease, survival analysis, Cox proportional hazard model

1. Pendahuluan

Stroke merupakan suatu sindrom yang ditandai dengan gejala dan atau tanda klinis yang berkembang dengan cepat yang berupa gangguan fungsional otak fokal maupun global yang berlangsung lebih dari 24 jam (kecuali ada intervensi bedah atau membawa kematian), yang tidak disebabkan oleh sebab lain selain penyebab vaskuler. Ketahanan hidup pasien dipengaruhi oleh banyak faktor. Adapun dalam ilmu statistika ketahanan hidup pasien stroke dapat diprediksi dengan menggunakan analisis survival. Analisis survival atau analisis data ketahanan hidup adalah suatu metode statistik untuk menganalisis data dengan variabel terikat yang diperhatikan berupa waktu sampai terjadinya suatu kejadian [1]. Salah satu tujuan analisis survival adalah mengetahui hubungan antara waktu survival dengan variabel-variabel yang diduga mempengaruhi waktu survival. Hubungan tersebut dapat dimodelkan dengan model regresi *Cox proportional hazard*, yang mempunyai variabel terikat berupa waktu survival dan variabel bebas berupa variabel yang diduga berpengaruh terhadap waktu survival.

Penelitian kali ini menggunakan model *Cox proportional hazard* sebagai aplikasi untuk diterapkan pada kasus ketahanan hidup pasien stroke di Rumah Sakit Balung. Model terbaik akan dipilih berdasarkan nilai AIC terkecil menggunakan eliminasi *backward*. Penelitian ini dilakukan dengan tujuan untuk mengetahui bagaimana jenis kelamin, usia, status hipertensi, status kolesterol, status diabetes militus, jenis stroke, dan indeks masa tubuh dapat mempengaruhi ketahanan hidup pasien stoke.

2. Landasan Teori

2.1 Penyakit Stroke. Stroke adalah sindrom yang terdiri dari tanda atau gejala hilangnya fungsi sistem saraf pusat fokal atau global yang terjadi secara cepat dan mendadak (dalam menit atau pun detik) yang berlangsung lebih dari 24 jam atau menyebabkan kematian [2]. Faktor risiko penyakit stroke terdiri dari faktor risiko yang tidak dapat dimodifikasi dan faktor risiko yang dapat dimodifikasi. Faktor risiko yang tidak dapat dimodifikasi yaitu umur dan jenis kelamin. Sedangkan faktor risiko yang dapat dimodifikasi yaitu hipertensi, diabetes militus, kolesterol tinggi, dan obesitas [3].

2.2 Analisis Survival. Analisis survival adalah teknik statistik yang digunakan untuk menganalisis data yang bertujuan untuk mengetahui hasil dari variabel yang mempengaruhi suatu awal kejadian sampai akhir kejadian, contohnya waktu yang dicatat dalam hari, minggu, bulan, atau tahun. Untuk kejadian awal contohnya awal pasien terjangkit penyakit dan untuk kejadian akhir contohnya kematian pasien dan kesembuhan pasien [1]. Untuk T suatu variabel acak positif dan menunjukkan waktu survival setiap subjek, maka nilai-nilai yang mungkin untuk T yaitu $T \geq 0$. Distribusi dari T dapat dinyatakan dalam tiga cara yaitu [4]:

a. Fungsi Kepadatan Peluang

Fungsi kepadatan peluang atau *Probability Density Function* (PDF) adalah peluang suatu individu mati atau mengalami kejadian sesaat dalam interval waktu t sampai $t + \Delta t$. Fungsi kepadatan peluang $f(t)$ dirumuskan dengan [4],

$$f(t) = \frac{d(F(t))}{dt} = F'(t)$$

b. Fungsi Survival

Fungsi survival $S(t)$ didefinisikan sebagai peluang suatu individu dapat bertahan hidup dengan waktu survival sampai dengan waktu $t(t > 0)$, yaitu sebagai berikut:

$$f(t) = F'(t) = -S'(t)$$

c. Fungsi Hazard

Fungsi hazard $h(t)$ didefinisikan sebagai kelajuan suatu individu mengalami kejadian dalam interval waktu dari t sampai $t + \Delta t$ dengan syarat individu tersebut masih bertahan hidup sampai dengan waktu t , fungsi kumulatif hazard $H(t)$ dapat dinyatakan dengan persamaan berikut:

$$H(t) = -\ln S(t)$$

2.3 Tipe-tipe Penyensoran

a. Tipe I

Tersensor tipe I yaitu dilakukan pengamatan terhadap objek-objek selama waktu tertentu yang telah ditentukan untuk mengakhiri semua n individu yang masuk pada waktu yang sama. Sering terdapat objek yang mengalami kejadian setelah masa pengamatan selesai dan sebagian lagi mengalami kejadian di luar waktu yang ditetapkan dalam penelitian. Sehingga individu tersebut belum mengalami kejadian hingga akhir periode penelitian, sedangkan waktu awal dari objek penelitian dapat diamati secara penuh.

b. Tipe II

Tersensor tipe II adalah tipe penyensoran dimana sampel ke- r merupakan penelitian terkecil dalam sampel acak berukuran $n(1 \leq r \leq n)$. Total sampel berukuran n dengan n berlanjut sampai mengalami kejadian meninggal atau gagal maka percobaan

akan dihentikan sampai r dari n unit penelitian yang mengalami kejadian. Semua objek yang masih hidup atau belum mengalami kejadian setelah diperoleh r kegagalan maka dikatakan tersensor. Semua unit penelitian masuk pada waktu yang sama.

c. Tipe III

Tersensor tipe III yaitu individu atau objek masuk ke dalam penelitian pada waktu yang berlainan selama periode waktu tertentu. Beberapa objek memiliki kemungkinan gagal atau meninggal sebelum pengamatan berakhir sehingga waktu tahan hidupnya dapat diketahui secara pasti. Kemungkinan kedua adalah objek keluar sebelum pengamatan berakhir dan kemungkinan ketiga adalah objek tetap hidup hingga batas berakhirnya penelitian. [4]

2.4 Maksimum Partial Likelihood Estimation (MPLE). Estimasi parameter dalam pemodelan *Cox proportional hazard* adalah dengan cara memaksimalkan fungsi *partial likelihood* atau biasa disebut *maksimum partial likelihood estimation* (MPLE). Misalkan terdapat n individu dengan r individu mengalami event, sehingga $n - r$ merupakan jumlah individu yang tersensor dan diasumsikan hanya terdapat satu individu yang mengalami event pada suatu waktu tertentu. Waktu survival terurut dari r individu yang mengalami event dinotasikan $t_{(1)} < t_{(2)} < \dots < t_{(r)}$. Himpunan dari individu yang mengalami event sebelum waktu $t_{(l)}$ dinotasikan sebagai $R(t_{(l)})$ sehingga fungsi *partial likelihood* dari model *Cox proportional hazard* dapat dirumuskan dengan,

$$L(\beta) = \prod_{l=1}^r \frac{\exp(\beta^T x_{(l)})}{\sum_{f \in R(t_{(l)})} \exp(\beta^T x_f)}$$

$x_{(l)}$ merupakan vektor variabel dari individu yang gagal pada waktu ke- l dengan waktu $t_{(l)}$. Notasi $R(t_{(l)})$ adalah seluruh individu yang memiliki risiko gagal pada waktu ke- l . Setelah didapatkan fungsi *partial likelihood*, langkah berikutnya yaitu memaksimalkan turunan pertama fungsi $\ln L(\beta)$. Karena estimasi parameter yang diperoleh implisit, maka digunakan metode iterasi numerik, yaitu metode Newton-Rhapson [5].

Jika $g(\beta)$ merupakan vektor yang berukuran $p \times l$ yang merupakan turunan pertama fungsi $\ln L(\beta)$ terhadap parameter β . $H(\beta)$ adalah matrik hessian berukuran $p \times p$ yang berisi turunan kedua dari fungsi $\ln L(\beta)$, estimasi parameter pada iterasi ke $(l + 1)$ adalah sebagai berikut :

$$\beta^{(l+1)} = \beta^l - H^{-1}(\beta^{(l)})g(\beta^l)$$

Sebagai awalan $\beta^{(0)}$ iterasi akan berhenti jika, $\|\beta^{(l+1)} - \beta^l\| \leq \varepsilon$ dimana ε merupakan suatu bilangan yang sangat kecil.

2.5 Uji Log-rank. Pengujian keberartian koefisien kovariat bertujuan untuk mengetahui ada atau tidaknya pengaruh dari suatu kovariat terhadap variabel respons. Pengujian keberartian koefisien data penelitian ini menggunakan uji Log-Rank. Ekspetasi dan varians banyaknya objek yang mengalami peristiwa dirumuskan sebagai berikut:

$$E(d_{jA}) = \frac{n_{jA}d_j}{n_j}$$

$$\chi^2 = \frac{(\sum d_{jA} - \sum E(d_{jA}))^2}{\sum E(d_{jA})} + \frac{(\sum d_{jB} - \sum E(d_{jB}))^2}{\sum E(d_{jB})}$$

dengan kriteria keputusan H_0 adalah ditolak jika $\chi_{hitung}^2 > \chi_{\alpha}^2(db)$.

2.6 Uji Asumsi *Proportional Hazard*

a. Pendekatan Grafik

Asumsi *proportional hazard* menyatakan bahwa fungsi hazard dari individu yang berlainan adalah proporsional atau rasio dari fungsi hazard dua individu yang berlainan adalah konstan [6]. Salah satu cara untuk menguji asumsi proportional hazard adalah dengan membandingkan kurva estimasi $\log(-\log(S(t)))$ In antara kategori dari variabel yang diteliti. Kurva yang sejajar antar kategori dan tidak berpotongan mengindikasikan terpenuhinya asumsi proportional hazard [1].

Fungsi kumulatif hazard dilogaritmakan sehingga dapat dituliskan

$$\log[-\log S_1(t)] - \log[-\log S_0(t)] = \beta$$

b. Uji Statistik Menggunakan *Goodness Of Fit* (GOF)

Pemeriksaan asumsi proportional hazard menggunakan pendekatan grafik akan subjektif. Kleinbaum dan Klein (2012) merekomendasikan untuk menggunakan metode grafik dan uji statistik dalam memeriksa asumsi tersebut. Untuk pemeriksaan asumsi menggunakan uji statistik, Grambsch dan Therneau (1994) memodifikasi metode residual yang diperkenalkan oleh Schoenfeld (1982). Langkah-langkah pengujian asumsi proportional hazard ini adalah:

1. Menggunakan model *Cox proportional hazard* untuk mendapatkan residual Schoenfeld untuk setiap variabel prediktor. Residual Schoenfeld ada pada setiap variabel prediktor pada model dan pada setiap objek yang mengalami event.
2. Membuat variabel rank waktu survival yang telah diurutkan berdasarkan waktu survival mulai dari individu yang mengalami event pertama kali.
3. Menguji korelasi antara variabel residual Schoenfeld dan rank waktu survival.

Tolak H_0 jika $t_{hit} > t_{\frac{\alpha}{2}, n-2}$ atau p -value kurang dari α . Yang berarti terdapat korelasi antara residual Schoenfeld dengan rank waktu survival.

c. Pendekatan Variabel *Time Dependent*

Uji asumsi *proportional hazard* dengan menggunakan variabel *time dependent* menggunakan model *Cox extended* yang melibatkan fungsi waktu. Fungsi waktu yang digunakan dapat berupa t , $\ln t$, dan fungsi waktu yang lain yang mengandung t . Hipotesis yang digunakan adalah :

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_q = 0$$

$$H_1 : \text{minimal terdapat satu } \delta_m \neq 0; m = 1, 2, 3, \dots, q$$

dimana δ_m merupakan koefisien dari q variabel bebas yang tidak memenuhi asumsi *proportional hazard*. Pemilihan fungsi waktu dapat berdasarkan $p - value$ yang dihasilkan dari variabel yang tidak memenuhi asumsi *proportional hazard*. Fungsi waktu yang digunakan adalah fungsi waktu yang menghasilkan $p - value$ terkecil.

2.7 Model Cox proportional hazard. Dalam analisis survival terdapat model parametrik dan model semiparametrik. Model *Cox proportional hazard* merupakan model berdistribusi semiparametrik karena dalam model *Cox proportional hazard* tidak memerlukan informasi tentang distribusi yang mendasari waktu survival dan untuk mengestimasi parameter regresi dari model *Cox proportional hazard* tanpa harus menentukan fungsi hazard dasar (*baseline hazard*) [4].

Model *Cox proportional hazard* disebut dengan model *cox* karena asumsi *proportional hazardnya* yaitu fungsi *hazard* dari individu yang berbeda adalah *proportional* atau rasio dari fungsi *hazard* dua individu yang berbeda adalah konstan [7]. Melalui model *Cox* dapat dilihat hubungan antara variabel bebas (variabel independen) terhadap variabel terikat (variabel dependen) yaitu waktu *survival* melalui fungsi *hazardnya*. Untuk variabel X yang ber-Covariate, maka persamaan yang digunakan adalah

$$h(t, X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

dengan memisalkan,

t : Waktu Survival

$h_0(t)$: Fungsi dasar *hazard*,

$\beta_1, \beta_2, \dots, \beta_p$: Parameter regresi,

X_1, X_2, \dots, X_p : Variabel bebas

i : $1, 2, \dots, p$

2.8 Model Cox Non Proportional Hazard. Jika asumsi *proportional hazard* tidak terpenuhi berarti komponen linier dari model berubah-ubah tergantung waktu dan dikatakan *non proportional hazard* [3]. Adapun cara untuk mengatasi masalah *non*

proportional hazard yaitu dengan menggunakan model Regresi *Cox Stratified* dan Model Regresi *Cox extended*.

a. Model Regresi *Cox Stratified*

Model stratified Cox merupakan perluasan dari model *Cox proportional hazard* untuk mengatasi variabel bebas yang tidak memenuhi asumsi proporsional hazard. Asumsi proportional hazard menyatakan bahwa rasio fungsi hazard dari dua individu konstan dari waktu ke waktu atau ekuivalen dengan pernyataan bahwa fungsi hazard suatu individu terhadap fungsi hazard individu lain adalah proporsional [6]. Modifikasi dilakukan dengan menstratifikasi variabel bebas yang tidak memenuhi asumsi proportional hazard. Variabel bebas yang memenuhi asumsi proportional hazard masuk ke dalam model, sedangkan variabel bebas yang tidak memenuhi asumsi, yang sedang distratifikasi, tidak masuk dalam model [1].

Bentuk umum fungsi hazard dari model stratified Cox tanpa interaksi adalah [1]:

$$h_s(t, X) = h_{0s}(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$

dengan

s : strata yang didefinisikan dari Z^* ,

s : 1,2,3,..., m^*

$h_{0s}(t)$: fungsi dasar hazard untuk setiap strata

$\beta_1, \beta_2, \dots, \beta_k$: parameter regresi

Strata didefinisikan sebagai kategori yang berbeda dari variabel stratifikasi Z^* dan m^* merupakan banyaknya strata. Dalam model *Cox stratified*, fungsi dasar hazard $h_{0s}(t)$ berbeda untuk setiap strata. Parameter regresi $\beta_1, \beta_2, \dots, \beta_k$ untuk model ini sama untuk setiap strata sehingga perkiraan rasio hazard sama untuk masing-masing strata. Estimasi parameter pada model *Cox stratified* sama halnya dengan estimasi parameter

pada model *Cox proportional hazard*, yaitu menggunakan *Maximum Partial Likelihood Estimation (MPLE)*.

b. Model Regresi *Cox Extended*

Dalam model regresi Cox ada variabel yang melibatkan waktu t . Variabel ini disebut variabel *time-dependent* (bergantung waktu). Variabel *time-dependent* didefinisikan sebagai variabel yang mempunyai nilai berubah sepanjang waktu (t). Jika ada variabel *time-dependent* dalam model, model regresi Cox dapat digunakan tetapi tidak memenuhi asumsi *proportional hazard*. Sehingga perlu digunakan model regresi *Cox extended*. Dalam model ini, model regresi Cox diperluas dengan model yang mengandung kovariat *time-dependent*. Jika x_1, x_2, \dots, x_{p_1} adalah kovariat *time-independent* yang memenuhi asumsi *proportional hazard*, $x_{p_1+1}, x_{p_1+2}, \dots, x_{p_2}$ adalah kovariat *time-independent* yang tidak memenuhi asumsi *proportional hazard* dan $x_1(t_j), x_2(t_j), \dots, x_{p_2}(t_j)$ adalah kovariat *time-dependent* maka model regresi *Cox Extended* didefinisikan sebagai berikut:

$$h(t, x(t)) = h_0(t) \exp \left[\sum_{\alpha=1}^{p_1} \beta_{\alpha} X_{\alpha} + \sum_{b=p_1+1}^{p_2} \beta_b X_b + \sum_{b=p_1+1}^{p_2} \delta_b X_b(t_j) \right]$$

di mana β dan δ adalah vektor koefisien dari kovariat, p_1 adalah jumlah kovariat yang memenuhi asumsi *proportional hazard* dan p_2 adalah jumlah kovariat yang tidak memenuhi asumsi *proportional hazard*. Estimasi parameter pada model *Cox extended* sama halnya dengan estimasi parameter pada model *Cox proportional hazard*, yaitu menggunakan *Maximum Partial Likelihood Estimation (MPLE)*.

2.9 Pengujian Signifikansi Parameter. Setelah mendapatkan variabel predictor yang termasuk ke dalam model, maka langkah selanjutnya adalah uji signifikansi parameter model. Uji yang dilakukan ada dua, yaitu uji serentak dan uji parsial.

1. Uji Serentak

Hipotesis : $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$

H_1 : minimal terdapat satu $\beta_k \neq 0; k = 1, 2, 3, \dots, p$

Statistik ujinya adalah :

$$G^2 = -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

dengan :

$L(\hat{\omega})$: nilai *likelihood* untuk model tanpa menyertakan variabel prediktor.

$L(\hat{\Omega})$: nilai *likelihood* untuk model dengan menyertakan variabel prediktor.

k : banyak parameter dalam model

Tolak H_0 jika $G^2 \geq X^2_{(a,p)}$ atau *p - value* < α

2. Uji Parsial

Hipotesis pada uji parsial adalah sebagai berikut.

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0; k = 1, 2, 3, \dots, p$

Statistik ujinya adalah :

$$W^2 = \frac{\hat{\beta}_k^2}{(SE(\hat{\beta}_k))^2}$$

Tolak H_0 jika $W^2 > \chi^2_{\alpha,p}$ atau *p - value* < α

2.10 Seleksi Model Terbaik. Seleksi model terbaik digunakan untuk mendapatkan model terbaik yang dapat menggambarkan hubungan antara waktu survival dengan beberapa variabel independen secara tepat. Salah satu prosedur yang digunakan untuk menentukan model terbaik adalah menggunakan eliminasi backward. Langkah-langkah eliminasi backward sebagai berikut [8].

1. Membuat model regresi yang berisi semua variabel independen yang tersedia.
2. Memilih satu variabel independen yang merupakan variabel terakhir untuk dimasukkan dalam model.
3. Melakukan pengujian pada variabel independen yang terpilih pada langkah 2 dan memutuskan untuk menghilangkan variabel tersebut atau tidak.
4. Mengulangi langkah 2 dan 3 untuk setiap variabel yang terdapat pada model. Apabila tidak ada kriteria yang sesuai berdasarkan langkah 3 maka proses telah selesai karena tidak ada lagi variabel independen yang dihilangkan dari model.

2.11 Pemilihan Model Terbaik. *Akaike's Information Criterion* (AIC) adalah metode yang berguna untuk mendapatkan model terbaik yang ditemukan oleh Akaike. Menurut Stasinopolous, *et al.* besarnya AIC dapat dilihat pada persamaan :

$$AIC = -2l(\hat{\beta}) + 2df$$

dengan $l(\hat{\beta})$ adalah fungsi log (likelihood) dan df adalah total derajat bebas yang digunakan dalam model. Model regresi ataupun distribusi terbaik adalah model regresi yang memiliki nilai AIC terkecil. Kelebihan AIC terletak pada pemilihan model regresi terbaik untuk tujuan (*forecasting*) yaitu dapat menjelaskan kecocokan model dengan data yang ada [9].

2.12 Model Hazard Ratio. Hazard ratio merupakan hazard relatif dari individu ke-i dengan kovariat mengalami peristiwa dibandingkan individu ke-i dengan kovariat yang konstan atau bebas terhadap waktu [1]. *Hazard ratio* juga menunjukkan adanya peningkatan atau penurunan risiko individu yang dikenai perlakuan tertentu (Lee, 1992). Misalkan terdapat dua individu dengan karakteristik tersebut maka dari model umum *Cox proportional hazard* diperoleh formula untuk menduga *hazard ratio*, yaitu:

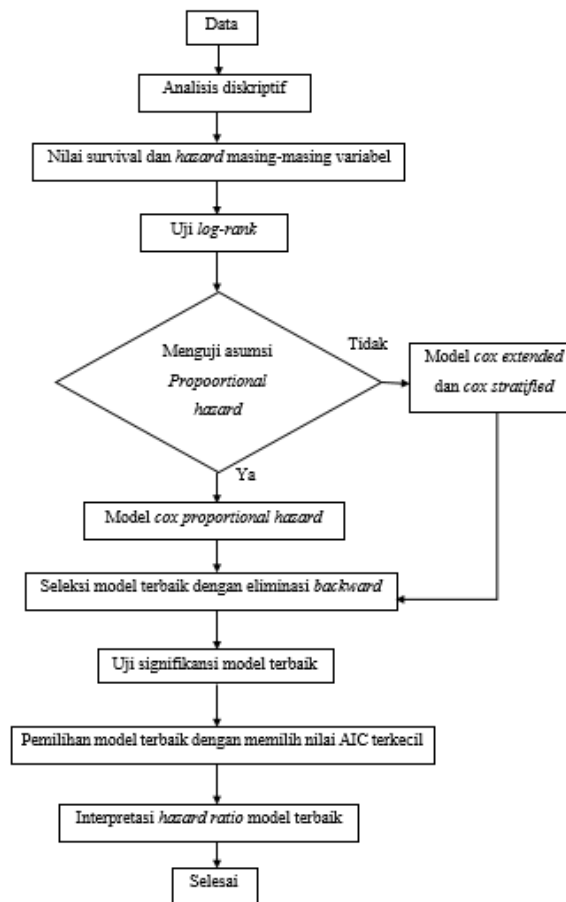
$$\begin{aligned}\widehat{HR} &= \frac{h_i(t)}{h_j(t)} \\ &= \frac{h_0(t)e^{(\beta_1x_{1i}+\beta_2x_{2i}+\dots+\beta_px_{pi})}}{h_0(t)e^{(\beta_1x_{1j}+\beta_2x_{2j}+\dots+\beta_px_{pj})}} \\ &= \exp\{\beta_1(x_{1i} - x_{1j}) + \dots + \beta_p(x_{pi} - x_{pj})\}\end{aligned}$$

Untuk kovariat yang bersifat kategorik dengan variabel dummy bernilai 1 dan 0 maka *hazard ratio* dapat diinterpretasikan sebagai *ratio* dari penduga hazard untuk individu yang bernilai 1 terhadap penduga hazard untuk individu yang bernilai 0. Sedangkan untuk kovariat yang bersifat kuantitatif, lebih bermakna jika hazard ratio dikurangi 1 lalu dikalikan dengan 100% yang menyatakan perubahan persentase hazard penduga untuk penambahan 1 unit variabel tersebut [10].

3. Metode Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder, yaitu data pasien stroke yang diambil dari RSD Balung. Data yang diambil merupakan data pasien yang menderita stroke pada Januari 2016 sampai Desember 2017 sebanyak 145 sampel pasien dengan 72 data tidak tersensor (1=meninggal) dan 73 data tersensor (0=hidup). Terdapat 7 variabel untuk membentuk model terbaik dari data stroke, yaitu jenis kelamin (X_1), usia (X_2), status hipertensi (X_3), status kolesterol (X_4), status dm (X_5), jenis stroke (X_6), dan imt (X_7).

Langkah-langkah yang perlu dilakukan dalam penelitian yaitu mengumpulkan data sekunder kemudian melakukan analisis deskriptif variabel-variabel penjelas serta nilai dan plotting fungsi survival dan fungsi hazard untuk setiap variabel. Kemudian dilakukan uji *log-rank*. Selanjutnya penyusunan model *Cox proportional hazard* yang didahului dengan uji asumsi *Proportional Hazard*. Dilanjutkan dengan pemilihan model terbaik dengan menggunakan eliminasi *backward* dan dipilih model yang memiliki nilai AIC terkecil. Terakhir, dilakukan interpretasi *hazard ratio* dari model terbaik.



Gambar 1. Metode Penelitian

4. Hasil Penelitian

Penelitian dilakukan dengan menggunakan program aplikasi olah data R 3.4.3 dengan jumlah 145 data pasien, 73 data tersensor dan 72 data tidak tersensor.

- a. Hasil analisis deskriptif dari semua variabel yaitu variabel jenis kelamin, usia, status hipertensi, status kolesterol dan status DM dengan jumlah pasien meninggal dunia dan persentase kematian pasien.

Tabel 1 Analisis Deskriptif semua variabel bebas

| | Jumlah Pasien | Jumlah | % |
|-------------------------------|---------------|--------|--------|
| Jenis Kelamin | | | |
| Laki-laki | 84 | 43 | 51,19% |
| Perempuan | 61 | 29 | 47,54% |
| Usia | | | |
| Usia < 45 th | 18 | 10 | 55,56% |
| $45 \leq \text{usia} < 55$ th | 41 | 24 | 58,54% |
| $55 \leq \text{usia} < 65$ th | 44 | 18 | 40,91% |
| $65 \leq \text{usia} < 75$ th | 30 | 10 | 33,33% |
| Usia ≥ 75 th | 12 | 10 | 83,33% |
| Stat. Hipertensi | | | |
| Tidak Hipertensi | 70 | 34 | 48,57% |
| Hipertensi | 75 | 38 | 50,67% |
| Status Kolesterol | | | |
| Tidak | 95 | 43 | 45,26% |
| Kolesterol | | | |
| Kolesterol | 50 | 29 | 58% |

Tabel 1 Analisis Deskriptif semua variabel bebas (lanjutan)

| Status DM | | | |
|-----------------------------|-----|----|--------|
| Tidak DM | 106 | 53 | 50% |
| DM | 39 | 19 | 48,72% |
| Jenis Stroke | | | |
| Stroke Hemoragik | 106 | 46 | 43,4% |
| Stroke Iskemik | 39 | 26 | 66,67% |
| IMT | | | |
| $18,5 \leq \text{IMT} < 25$ | 65 | 33 | 50,77% |
| $25 \leq \text{IMT} < 30$ | 23 | 10 | 43,48% |
| $30 \leq \text{IMT} < 35$ | 30 | 13 | 43,33% |
| $35 \leq \text{IMT} < 40$ | 25 | 15 | 60% |
| $\text{IMT} \geq 40$ | 2 | 1 | 50% |

Berdasarkan analisis deskriptif dari variabel-variabel tersebut menunjukkan bahwa pasien perempuan, usia ≥ 75 th, kategori hipertensi, kategori kolesterol, tidak diabetes militus, stroke iskemik, dan $35 \leq \text{IMT} < 40$ mempunyai proporsi meninggal yang paling tinggi daripada kategori lainnya.

b. Nilai dan plotting fungsi survival dan fungsi hazard yang memberikan gambaran katahanan hidup dan resiko kematian pasien.

c. Uji asumsi model *Cox proportional hazard* yang menyatakan bahwa semua variabel yang digunakan dalam penelitian memenuhi asumsi *Proportional Hazard*

dengan $p\text{-value} > 0.05$ dengan nilai *hazard ratio* dan koefisien dari model *Cox proportional hazard* seperti berikut.

Tabel 2 Estimasi parameter dan *hazard ratio* model *Cox proportional hazard*

| Variabel | <i>Hazard Ratio</i> | <i>Koef.</i> | <i>p-value</i> |
|-------------------------|---------------------|--------------|----------------|
| Jenis Kelamin | 0,884825 | 0,122365 | 0,623877 |
| Usia | 1,202999 | 0,184818 | 0,160876 |
| Status Hipertensi | 1,244192 | 0,218486 | 0,484655 |
| Status Kolesterol | 1,475181 | 0,388781 | 0,447851 |
| Status DM | 0,570416 | 0,561390 | 0,071121 |
| Jenis Stroke | 3,075299 | 1,123402 | 0,000277 |
| IMT | 0,997478 | 0,002525 | 0,990109 |
| <i>Likelihood ratio</i> | | | 0.01 |

Berdasarkan hasil estimasi parameter pada Tabel 2 diperoleh model *Cox proportional hazard* adalah :

$$h(t) = h_0(t) \exp(-0,122365X_1 + 0,184818X_2 + 0,218486X_3 + 0,388781X_4 - 0,561390X_5 + 1,123402X_6 - 0,002525X_7)$$

Setelah didapatkan model, maka dilakukan uji serentak untuk mengetahui apakah variabel independen yang digunakan pada model berpengaruh signifikan atau tidak secara bersama-sama. Pada Tabel 2 diperoleh $p\text{-value}$ *likelihood ratio* sebesar 0,01, nilai tersebut kurang dari $\alpha = 0,05$ sehingga menghasilkan keputusan tolak H_0 . Artinya, bahwa minimal ada satu variabel independen yang berpengaruh signifikan terhadap model. Dengan demikian, model *Cox proportional hazard* yang terbentuk secara bersama-sama telah berpengaruh signifikan. Selain melakukan uji serentak, juga dilakukan uji parsial. Selain melakukan uji serentak, juga dilakukan uji parsial. Berdasarkan tabel 4.31 terdapat satu

variabel independen yang signifikan yakni jenis stroke (X_6). Hal ini dibuktikan dengan p -value sebesar 0,000277 yang lebih kecil dari $\alpha = 0,05$. Karena masih banyak variabel yang tidak signifikan, maka perlu dilakukan eliminasi *backward* untuk menentukan model *Cox proportional hazard* terbaik. Eliminasi *backward* dilakukan dengan cara memilih *Akaike's Information Criterion* (AIC) terbaik atau dalam kasus ini adalah nilai AIC terkecil. Berikut ini merupakan hasil dari eliminasi *backward*.

Tabel 3 Model Terbaik Sesuai Kriteria AIC

| Tahap ke- | Variabel | AIC |
|-----------|-----------------------------------------------------------------------------------------|--------|
| 1 | IMT, jenis kelamin, status hipertensi, status kolesterol, usia, status Dm, jenis stroke | 591,04 |
| 2 | Jenis kelamin, status hipertensi, usia, status kolesterol, status DM, jenis stroke | 589,04 |
| 3 | Status hipertensi, usia, status kolesterol, status Dm, jenis stroke | 587,28 |
| 4 | Status kolesterol, usia, status DM, jenis stroke | 585,74 |
| 5 | Usia, status Dm, jenis stroke | 585,61 |

Tabel 3 menunjukkan bahwa nilai AIC terkecil adalah 585,61. Variabel yang terbentuk adalah Usia (X_2), Status Dm (X_5) dan jenis stroke (X_6). Variabel tersebut merupakan faktor yang berpengaruh terhadap waktu ketahanan hidup pasien stroke di RSD Balung.

d. Estimasi Parameter Model *Cox proportional hazard* Terbaik

Setelah diperoleh model terbaik, langkah selanjutnya adalah melakukan estimasi parameter untuk model terbaik. Berikut estimasi parameter model terbaik.

Tabel 4 Estimasi Parameter Model *Cox proportional hazard* Terbaik

| Variabel | Koef | Hazard Ratio | Standard Error | p-value |
|-------------------------|--------|--------------|----------------|-------------------------|
| Usia | 0,1901 | 1,2094 | 0,1213 | 0,1169 |
| Status DM | 0,4821 | 0,6175 | 0,2924 | 0,0992 |
| Jenis Stroke | 1.1045 | 3,0178 | 0,2776 | 6,93 x 10 ⁻⁵ |
| <i>Likelihood ratio</i> | | | | 0.001 |

Berdasarkan hasil estimasi parameter pada Tabel 4 diperoleh model :

$$h(t) = h_0(t)\exp(0,1901X_2 - 0,4821X_5 + 1,1045X_6)$$

Dari model *Cox proportional hazard* terbaik, dilakukan uji serentak untuk mengetahui apakah model telah signifikan atau belum. Uji serentak dilakukan dengan melihat nilai *p-value likelihood ratio*. Pada Tabel 4 diperoleh *p-value likelihood ratio* sebesar 0,01, nilai tersebut kurang dari $\alpha = 0,05$ sehingga menghasilkan keputusan tolak H_0 . Artinya, bahwa minimal ada satu variabel yang signifikan dalam model *Cox proportional hazard* terbaik, sehingga model *Cox proportional hazard* terbaik telah signifikan.

e. Interpretasi Model Terbaik

Interpretasi model *Cox proportional hazard* dilakukan dengan melihat nilai *hazard ratio* (HR) berdasarkan variabel independen yang disajikan pada Tabel 5.

Tabel 5 Hazard Ratio Model *Cox proportional hazard* Terbaik

| Variabel | Hazard Rasio |
|--------------|--------------|
| Usia | 1,2094 |
| Status DM | 0,6175 |
| Jenis Stroke | 3,0178 |

Berdasarkan Tabel 5 nilai *hazard ratio* (HR) untuk variabel usia adalah 1,2094. Nilai tersebut berarti bahwa semakin bertambahnya usia maka akan mengalami peningkatan risiko kematian sebesar 1,2094 kali. Nilai *hazard ratio* (HR) untuk variabel status DM adalah 0,6175. Nilai tersebut berarti bahwa pasien yang tidak memiliki penyakit penyerta DM memiliki risiko kematian 0,6175 kali lebih kecil daripada pasien yang memiliki penyakit penyerta DM. Nilai *hazard ratio* (HR) untuk variabel jenis stroke adalah 3,0178. Nilai tersebut berarti bahwa pasien dengan stroke hemoragik memiliki risiko kematian 3,0178 kali lebih besar dari pada pasien stroke iskemik.

5. Kesimpulan

Faktor-faktor yang mempengaruhi ketahanan hidup pasien stroke di RSD Balung yaitu usia, status diabetes militus, dan jenis stroke. Semakin bertambahnya usia maka akan mengalami peningkatan risiko kematian sebesar 1,2094 kali. Pasien yang tidak memiliki penyakit penyerta DM memiliki risiko kematian 0,6175 kali lebih kecil daripada pasien yang memiliki penyakit penyerta DM. Pasien dengan stroke hemoragik memiliki risiko kematian 3,0178 kali lebih besar dari pada pasien stroke iskemik.

Pada penelitian ini menggunakan data dengan sensor tipe III. Untuk penelitian selanjutnya bisa menggunakan jenis data sensor interval atau data terpotong. Selanjutnya dimodelkan dengan menggunakan model *Cox proportional hazard*. Untuk penelitian selanjutnya dapat menggunakan model regresi yang lain.

Daftar Pustaka

- [1] Kleinbaum, D.G. dan Klein.M. *Survival Analysis a Self-Learning Text*. Springer. New York. 2012.
- [2] Ginsberg, Lionel. *Lecture Notes: Neurology*. Erlangga. Jakarta. 2007.
- [3] Wahjoepramono,E.J. *Stroke Tata Laksana Fase Akut*. Universitas Pelita Harapan. Jakarta. 2005.
- [4] Lee, E. T. dan Wang, J. W. *Statistical Methods for Survival Data Analysis Third Edition*. John Wiley & Sons, Inc. New Jersey. 2003.
- [5] Collett, D. *Modeling Survival Data In Medical Research*. Chapman&Hall/Crc. London. 2003.
- [6] Guo, S. *Survival Analysis*. Oxford University Press, Inc. New York. 2010.
- [7] Iskandar, M. *Model Cox Proportional Hazard Pada Kejadian Bersama*. Skripsi. Universitas Negeri Yogyakarta. Yogyakarta. 2015.
- [8] Le, C. T. *Applied Survival Analysis*. John Willey And Sons, Inc. New York. 1997.
- [9] Fathutrahman, M. Pemilihan Model Regresi Terbaik Menggunakan Metode Akaike's Information Critettrion Dan Schwart Information Criterion. *Jurnal Informatika Mulawarman*. 4(3) :37-41. Universitas Mulawarman. 2009.
- [10] Allison, Pd. *Survival Analysis Using Sas: A Practical Guide Ed Ke-2*. Sas Institute Inc. North Carolina (Us). 2010.