

APLIKASI MODEL COX PROPORTIONAL HAZARD PADA KASUS DEMAM TIFOID

(Application of *Cox Proportional Hazard Model on The Case of Typhoid Fever*)

Nurul Azizah Baisaku, Jajang^{*}), Nunung Nurhayati

Jurusan Matematika, Fakultas MIPA, Universitas Jenderal Soedirman

Jl. DR. Soeparno No.61, Kabupaten Banyumas

e-mail: nurulabaisaku@gmail.com, *)jajang@unsoed.ac.id,

nunung.nurhayati@unsoed.ac.id

^{*}penulis korespondensi

Abstract. A common problem found in *survival* data is the presence of censored data. The length of hospitalization of Typhoid fever patients until declared cured is one of example of this data. Here, we use Cox regression model to analysis this data. *Partial likelihood* is one of the methods of estimating parameters for Cox regression model. In many cases of censored data, two objects (patients) have the same length of hospitalization (ties). Therefore, to estimate the parameters of the model must use the right method. Here we used *partial likelihood* Breslow, Efron, and Exact methods. The study was motivated by how the three methods performed for Cox regression model. The data used for the implementation of these methods is length of hospitalization of Typhoid fever patients at Mekar Sari Hospital-Bekasi in 2020. Based on AIC criteria, we found that exact method is the best model (minimum AIC) for parameter estimation of Cox regression model. Referring to the Cox regression model by using a significance level of 10%, there are five predictor variables that affects the length of patient hospitalization. The five variables are age, vomiting, dirty tongue, hemoglobin, and leukocyte.

Keywords: Typhoid fever, Cox regression, Breslow method, Efron method, exact method.

MSC2020: 62N02, 62N03

Received: 18-01-2022, accepted: 26-02-2022

1. Pendahuluan

Analisis *survival* merupakan metode statistik untuk menganalisis data *survival*. Data *survival* adalah data yang menyatakan lama waktu pengamatan hingga peristiwa kejadian. Kejadian dapat berupa kematian, kekambuhan, kesembuhan atau peristiwa lain yang dianggap menarik [1]. Masalah yang umum ditemukan pada data *survival* adalah adanya data tersensor, yaitu data *survival* yang memuat waktu *survival* dari individu yang belum mengalami kejadian. Data tersensor terdiri dari data tersensor kiri, data tersensor kanan dan data tersensor interval [2]. Salah satu metode analisis *survival* yang umum digunakan adalah regresi Cox [3]. Kajian model regresi Cox dalam analisis data telah banyak dilakukan [4–7].

Regresi Cox dapat digunakan untuk menyelesaikan data tersensor. Salah satu contoh data tersensor yang dalam hal ini tersensor kanan adalah data rawat inap pasien demam tifoid sampai dinyatakan sembuh. Demam Tifoid (*Typhoid fever*) merupakan penyakit yang menyerang saluran pencernaan yang disebabkan oleh bakteri *Salmonella typhi* yang mengkontaminasi makanan [8]. Inkubasi bakteri berlangsung pada masa 10-14 hari yang menyebabkan terjadinya demam yang meningkat secara bertahap disertai sakit kepala, *malaise* dan sesekali batuk. Gejala pada perut seperti nyeri, diare dan sembelit terjadi pada minggu pertama, sedangkan *hepatosplenomegali* ringan dan *rose spot* terjadi pada minggu kedua dan pada kasus berat dapat menyebabkan gangguan ginjal, *shock*, perubahan keadaan mental dan koma [9]. Demam Tifoid juga dapat menyebabkan pendarahan pada usus yang berakibat pada kematian [10]. Kajian terkait dengan Tifoid telah banyak dibahas dalam penelitian-penelitian sebelumnya [11–12]. Dari sisi metode yang digunakan, regresi *Cox proportional hazard* merupakan model yang telah banyak digunakan dalam penelitian medis [13].

Kasus demam tifoid dan kematian karena demam tifoid cukup tinggi. Peneliti menggunakan studi kasus demam Tifoid di Rumah Sakit Mekar Sari Bekasi menunjukkan ada tren naik. Tujuan dari penelitian ini adalah mengkaji faktor apa saja yang berpengaruh terhadap lama rawat inap pasien demam Tifoid dan bagaimana laju kesembuhannya. Dalam kasus ini dimungkinkan adanya kasus kejadian secara bersamaan (kasus kembar). Oleh karena itu, perlu dilakukan metode estimasi yang sesuai. Kejadian kembar dapat diselesaikan dengan beberapa metode, yaitu: metode Breslow, metode Efron dan metode *exact*. Oleh karena itu, dalam penelitian ini digunakan *partial likelihood* Breslow, *partial likelihood* Efron, dan *partial likelihood exact*. Perbandingan metode estimasi berdasarkan *partial likelihood* Breslow, Efron, dan *Exact* telah dikaji [14–15]. Hasil perbandingan menunjukkan bahwa metode *partial likelihood exact* dapat memebrikan hasil yang baik.

2. Metodologi

Misalkan T variabel acak yang menyatakan waktu sampai terjadinya kejadian. Fungsi *survival* pada saat t , dinotasikan $S(t)$, didefinisikan sebagai peluang terjadinya $\{T > t\}$ atau $S(t) = P(T > t)$ [2]. Diketahui fungsi distribusi kumulatif (*cumulative distribution function*, CDF) dari T , $F(t) = P(T \leq t)$. Dengan demikian $S(t)$ dapat dinyatakan dengan $S(t) = 1 - F(t)$. Lebih lanjut $S(t)$ dapat dinyatakan sebagai

$$S(t) = 1 - \int_0^t f(u)du, \quad (1)$$

dengan $f(\cdot)$ adalah fungsi densitas probabilitas (*probability density function*, pdf) dari variabel acak T .

Fungsi *hazard* menyatakan peluang bahwa variabel acak yang terkait dengan waktu *survival*, T , terletak diantara t dan $t + \Delta t$ dengan syarat T lebih besar dari t yang dapat dituliskan sebagai $P(t < T \leq t + \Delta t | T > t)$ [2]. Sehingga fungsi *hazard* dapat

dinyatakan dengan $h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \right\}$. Lebih lanjut fungsi ini dapat diuraikan sebagai

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t)}{P(T > t) \Delta t} \right\}. \quad (2)$$

Perhatikan fungsi *survival* pada persamaan (1) dan fungsi *hazard* pada persamaan (2). Berdasarkan persamaan (2),

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \right\} \\ &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(T \leq t + \Delta t) - P(T \leq t)}{S(t) \Delta t} \right\} \\ &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\} \frac{1}{S(t)} \\ &= \frac{f(t)}{S(t)}. \end{aligned}$$

Diketahui bahwa terdapat hubungan antara fungsi *survival* ($S(t)$) dan CDF ($F(t)$), dengan $S(t) = 1 - F(t)$. Dengan demikian jika kedua ruas diturunkan terhadap t , maka $F'(t) = -S'(t)$ atau $f(t) = S'(t)$. Lebih lanjut hubungan antara fungsi *hazard* dengan fungsi *survival* adalah [2]

$$h(t) = \frac{-S'(t)}{S(t)}. \quad (3)$$

Model Cox *proportional hazard* adalah model *hazard* pada persamaan berikut [16]

$$h(t, \mathbf{X}) = h_0(t) \exp \left(\sum_{j=1}^p \beta_j x_j \right), \quad (4)$$

dengan $h(t, \mathbf{X})$ adalah fungsi *hazard* pada waktu t untuk individu dengan variabel prediktor \mathbf{X} ; x_j adalah nilai dari variabel prediktor X_j ; $h_0(t)$ adalah fungsi *baseline hazard*; dan β_j adalah parameter ke j , $j = 1, 2, \dots, p$.

Metode maksimum likelihood merupakan salah satu metode estimasi parameter yang cukup sering digunakan dalam pemodelan statistika. Fungsi *likelihood* pada model Cox tidak mempertimbangkan probabilitas semua objek, oleh karena itu fungsi *likelihood* pada model Cox disebut *partial likelihood* [17].

Jika terdapat k waktu *survival* berbeda yang tidak tersensor, $t_{(1)} < t_{(2)} < \dots < t_{(k)}$, $k \leq n$ dengan $\mathbf{X}_{(1)}, \dots, \mathbf{X}_{(k)}$ vektor variabel yang bersesuaian, maka estimasi parameter model survival didasarkan pada metode maksimum *partial likelihood* [16]. Dalam masalah model Cox, juga perlu dipertimbangkan masalah waktu kejadian, khususnya ketika adanya kasus kejadian kembar. Kejadian kembar dapat diselesaikan dengan beberapa metode, yaitu: metode Breslow, metode Efron dan metode *exact*.

Misalkan $\mathbf{z}'_{\mathbf{u}^*_{(i)}} = (z_{1\mathbf{u}^*_{(i)}}, z_{2\mathbf{u}^*_{(i)}}, \dots, z_{p\mathbf{u}^*_{(i)}})'$ adalah vektor berukuran $1 \times p$ yang elemennya adalah penjumlahan nilai variabel prediktor ke p dari individu yang merupakan elemen dari $\mathbf{u}^*_{(i)}$, dengan $\mathbf{u}^*_{(i)}$ adalah himpunan individu yang mengalami kejadian pada saat $t_{(i)}$. $\boldsymbol{\beta}$ adalah parameter, \mathbf{x}'_i adalah vektor variabel prediktor, $R(t_{(i)})$ adalah himpunan risiko pada saat $t_{(i)}$, $m_{(i)}$ adalah banyaknya individu yang mengalami kejadian kembar pada waktu $t_{(i)}$.

Partial likelihood pada metode Breslow dan metode Efrom dapat dilihat pada persamaan (5) dan persamaan (6) berikut

$$L(\boldsymbol{\beta}_{Breslow}) = \prod_{i=1}^k \frac{\exp(\mathbf{z}'_{\mathbf{u}^*_{(i)}} \boldsymbol{\beta})}{[\sum_{l \in R(t_{(i)})} \exp(\mathbf{x}'_l \boldsymbol{\beta})]^{m_{(i)}}}, \quad (5)$$

$$L(\boldsymbol{\beta}_{Efron}) = \prod_{i=1}^k \frac{\exp(\mathbf{z}'_{\mathbf{u}^*_{(i)}} \boldsymbol{\beta})}{\prod_{j=1}^{m_{(i)}} [\sum_{l \in R(t_{(i)})} \exp(\mathbf{x}'_l \boldsymbol{\beta}) - \frac{j-1}{m_{(i)}} \sum_{l \in \mathbf{u}^*_{(i)}} \exp(\mathbf{x}'_l \boldsymbol{\beta})]}. \quad (6)$$

Partial likelihood metode *exact* mempunyai pola yang sedikit berbeda dengan metode Breslow maupun Efron. *Partial likelihood* pada metode *exact* adalah

$$L(\boldsymbol{\beta}_{exact}) = \prod_{i=1}^k \frac{\exp(\mathbf{z}'_{\mathbf{u}^*_{(i)}} \boldsymbol{\beta})}{[\sum_{\mathbf{u}_{(j)} \in \mathbf{u}_i} \exp(\mathbf{z}'_{\mathbf{u}_{(j)}} \boldsymbol{\beta})]}, \quad (7)$$

dengan $\mathbf{u}_{(j)}$ adalah himpunan individu yang dipilih secara acak sebanyak $m_{(i)}$ dari r_i individu di $R(t_{(i)})$.

Akaike's Information Criterion (AIC) merupakan salah satu kriteria yang sering digunakan untuk memilih model terbaik dari dua atau lebih model yang dibandingkan. $AIC = 2k - 2 \ln(L)$, dengan k banyaknya parameter model yang diestimasi dan L adalah nilai maksimum likelihood model. Model terbaik adalah model dengan nilai *AIC* terkecil.

Model *proportional hazard* mengasumsikan bahwa *hazard ratio* dari dua individu adalah independen terhadap waktu dan mengharuskan variabel prediktor tidak bergantung pada waktu. Jika salah satu variabel prediktor bergantung pada waktu, maka asumsi *proportional hazard* dilanggar [16]. Uji asumsi *proportional hazard* dapat dilakukan dengan melihat plot *log minus log survival* dan uji *goodness of fit (GOF)* [18]. Pengujian asumsi *proportional hazard* dengan GOF lebih sering digunakan sebab tidak subjektif seperti pendekatan plot. Uji GOF lebih global karena tidak dapat mendeteksi penyimpangan spesifik dari asumsi PH dibandingkan metode lain [17].

Hazard ratio dibagi menjadi dua, yaitu *hazard ratio* pada data kategorik dan *hazard ratio* pada data kontinu [17]. Untuk *hazard ratio* pada data kategorik, misalnya bernilai 0 dan 1, maka interpretasi dengan $x = 1$ dibandingkan $x = 0$. *Hazard ratio* pada data kontinu diinterpretasikan untuk setiap penambahan variabel x sebesar 1 unit, individu dengan nilai x yang bertambah 1 unit.

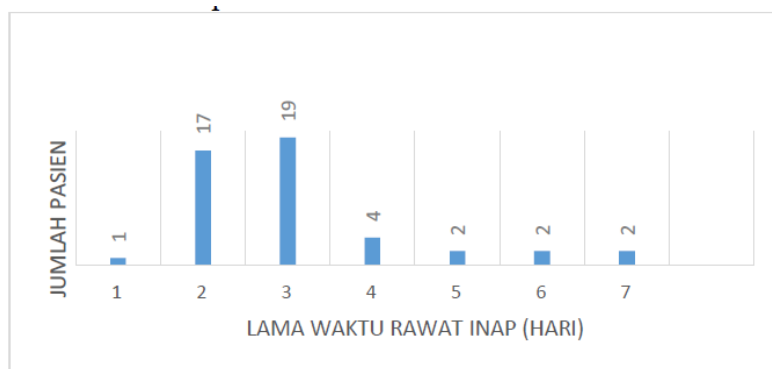
Penelitian dilakukan dengan metode *Cox proportional hazard* untuk menganalisis data durasi rawat inap pasien demam Tifoid di RS Mekar Sari Bekasi tahun 2020 yang berjumlah 47 pasien. Variabel penelitian yang digunakan dalam penelitian ini adalah durasi rawat inap pasien demam Tifoid sampai dinyatakan sembuh (Y), umur (X_1), jenis Kelamin (X_2), suhu (X_3), mual (X_4), muntah (X_5), pusing (X_6), lemas (X_7), demam (X_8), nyeri ulu hati (X_9), diare (X_{10}), lidah kotor (X_{11}), batuk (X_{12}), hemoglobin (X_{13}), leukosit (X_{14}), hematokrit (X_{15}), trombosit (X_{16}). Pengolahan data dan perhitungan pada penelitian ini dilakukan dengan bantuan *software R*.

Tahapan penelitiannya adalah sebagai berikut:

1. Analisis deskriptif untuk mengidentifikasi apakah terdapat kejadian kembar.
2. Estimasi parameter dengan *partial likelihood* Breslow, *partial likelihood* Efron dan *partial likelihood exact*.
3. Identifikasi variabel-variabel yang signifikan dari setiap model yang dihasilkan pada tahap (2).
4. Memilih model terbaik dari model-model yang dihasilkan dari tahap (3) berdasarkan nilai *AIC* terkecil.
5. Melakukan pengujian asumsi *proportional hazard*.
6. Melakukan interpretasi model regresi *Cox proportional hazard*.

3. Hasil dan Pembahasan

Tahapan awal dalam model Cox adalah identifikasi waktu kejadian apakah terdapat waktu kejadian kembar atau tidak. Gambar 1 menyajikan diagram batang dari variabel waktu lama rawat inap pasien sampai sembuh dan keluar dari rumah sakit.



Gambar 1. Lama waktu rawat inap pasien

Berdasarkan Gambar 1 terlihat bahwa lama rawat inap beberapa pasien dapat mempunyai waktu yang sama. Oleh karena itu, analisis berikutnya adalah estimasi parameter model Cox untuk mendapatkan nilai *hazard* persamaan (4). Metode yang digunakan adalah metode Breslow, Efron dan *exact*.

3.1 Estimasi parameter dengan metode Breslow, Efron dan Exact

Tahapan dalam estimasi model adalah dengan menspesifikasi model awal (melibatkan semua variabel prediktor) dari hasil estimasi dengan **Breslow, Efron dan Exact**. Berikutnya adalah diidentifikasi signifikansi setiap variabel prediktor dan juga diidentifikasi model terbaik dari masing-masing metode (**Breslow, Efron dan Exact**). Terakhir mengambil model terbaik dari model yang diperoleh dari metode masing-masing tersebut.

Tabel 1. Nilai statistik G Model Cox Awal dari Metode Breslow, Efron, dan Exact

Metode	Statistik G	$\chi^2_{0,05;16}$	Keputusan
Breslow	11,3774	26,2962	Tidak signifikan
Efron	19,6656	26,2962	Tidak signifikan
Exact	22,2560	26,2962	Tidak signifikan

Dapat dilihat bahwa model dari masing-masing metode dengan melibatkan semua variabel prediktor tidak menunjukkan signifikan. Oleh karena itu model tersebut perlu diperbaiki. Untuk berikutnya dalam setiap model akan dilakukan seleksi variabel. Seleksi variabel yang digunakan adalah metode *backward*.

3.2 Pemilihan model dan seleksi variabel dengan kriteria AIC

Pada subbab sebelumnya, terlihat bahwa model pada persamaan (5), (6) dan (7) yang hasilnya disajikan pada Tabel 1 merupakan model awal dari metode Breslow, Efron, dan *exact* tidak layak digunakan. Untuk itu, dilakukan pencarian calon model terbaik lain dengan menggunakan eliminasi *backward* pada tiap metode serta nilai *AIC*-nya.

Tabel 2. Calon model terbaik Breslow dan nilai AIC

No	Model	AIC
1	Semua variabel prediktor	313,6247
2	Semua variabel prediktor kecuali trombosit	311,6438
3	Semua variabel prediktor kecuali trombosit dan suhu	309,6816
4	Semua variabel prediktor kecuali trombosit, suhu, dan pusing	307,7507
5	Semua variabel prediktor kecuali trombosit, suhu, pusing, dan hematokrit	305,7974
6	Semua variabel prediktor kecuali trombosit, suhu, pusing, hematokrit dan batuk	303,9298
7	Semua variabel prediktor kecuali trombosit, suhu, pusing, hematokrit, batuk dan diare	302,1630
8	Semua variabel prediktor kecuali trombosit, suhu, pusing, hematokrit, batuk, diare dan leukosit	300,4676
9	Semua variabel prediktor kecuali trombosit, suhu, pusing, hematokrit, batuk, diare, leukosit dan jenis kelamin	298,9554
10	Variabel umur, mual, muntah, demam, nyeri ulu hati, lidah kotor dan hemoglobin	297,4218
11	Variabel umur, muntah, demam, nyeri ulu hati, lidah kotor dan hemoglobin	296,0255
12	Variabel umur, muntah, nyeri ulu hati, lidah kotor dan hemoglobin	294,5105
13	Variabel umur, muntah, nyeri ulu hati dan hemoglobin	293,5995
14	Variabel umur, muntah dan hemoglobin	293,1015
15	Variabel umur dan muntah	292,1732
16	Variabel muntah	293,8987

Tabel 3. Calon model terbaik Efron dan nilai AIC

No	Model	AIC
1	Semua variabel prediktor	281,1440
2	Semua variabel prediktor kecuali trombosit	279,1568
3	Semua variabel prediktor kecuali trombosit dan diare	277,2928
4	Semua variabel prediktor kecuali trombosit, diare, dan leukosit	275,3918
5	Semua variabel prediktor kecuali trombosit, diare, leukosit, dan pusing	273,4691
6	Semua variabel prediktor kecuali trombosit, diare, leukosit, pusing dan hematokrit	271,5574
7	Semua variabel prediktor kecuali trombosit, diare, leukosit, pusing, hematokrit dan suhu	269,7870
8	Semua variabel prediktor kecuali trombosit, diare, leukosit, pusing, hematokrit, suhu dan batuk	268,1563
9	Semua variabel prediktor kecuali trombosit, diare, leukosit, pusing, hematokrit, suhu, batuk dan jenis kelamin	267,1034
10	Variabel umur, mual, muntah, demam, nyeri ulu hati, lidah kotor dan hemoglobin	265,7538
11	Variabel umur, muntah, demam, nyeri ulu hati, lidah kotor dan hemoglobin	264,5304
12	Variabel umur, muntah, nyeri ulu hati, lidah kotor dan hemoglobin	263,3986
13	Variabel umur, muntah, nyeri ulu hati dan hemoglobin	263,3002
14	Variabel umur, muntah dan hemoglobin	264,4462
15	Variabel umur dan muntah	264,4235
16	Variabel muntah	269,0928

Tabel 4. Calon model terbaik exact dan nilai AIC

No	Model	AIC
1	Semua variabel prediktor	128,2375
2	Semua variabel prediktor kecuali suhu	126,2900
3	Semua variabel prediktor kecuali suhu dan pusing	124,3356
4	Semua variabel prediktor kecuali suhu, pusing dan trombosit	122,5058
5	Semua variabel prediktor kecuali suhu, pusing, trombosit dan hematokrit	120,7155
6	Semua variabel prediktor kecuali suhu, pusing, trombosit, hematokrit dan batuk	119,1853
7	Semua variabel prediktor kecuali suhu, pusing, trombosit, hematokrit, batuk dan lemas	117,4009
8	Semua variabel prediktor kecuali suhu, pusing, trombosit, hematokrit, batuk, lemas dan diare	115,9661
9	Semua variabel prediktor kecuali suhu, pusing, trombosit, hematokrit, batuk, lemas, diare dan mual	114,8272
10	Variabel umur, jenis kelamin, muntah, nyeri ulu hati, lidah kotor, hemoglobin dan leukosit	113,6289
11	Variabel umur, jenis kelamin, muntah, lidah kotor, hemoglobin dan leukosit	112,2841
12	Variabel umur, muntah, lidah kotor, hemoglobin dan leukosit	111,0841
13	Variabel umur, muntah, hemoglobin dan leukosit	112,0078
14	Variabel umur, muntah dan leukosit	112,6821
15	Variabel umur dan muntah	113,6974
16	Variabel muntah	118,4705

3.3 Pemilihan Model Terbaik Secara Keseluruhan

Berdasarkan analisis pada subbab 3.2, telah diperoleh model yang memberikan nilai AIC terkecil di tiap metode. Pemilihan model terbaik secara keseluruhan dari ketiga metode estimasi parameter yaitu metode Breslow, Efron, dan exact dapat dilakukan dengan membandingkan nilai AIC dari tiap metode sebagai berikut:

Tabel 5. Model terbaik tiap metode dan nilai AIC

Metode	Variabel Prediktor	AIC
Breslow	umur, muntah	292,1732
Efron	umur, muntah, nyeri ulu hati, hemoglobin	263,3002
Exact	umur, muntah, lidah kotor, hemoglobin, leukosit	111,0841

Model yang memberikan nilai AIC terkecil merupakan model terbaik secara keseluruhan yang akan digunakan sebagai model akhir. Pada Tabel 5 terlihat bahwa model terbaik dengan metode *exact* dengan variabel prediktor umur, muntah, lidah kotor, hemoglobin dan leukosit memberikan nilai AIC terkecil dibandingkan dengan model terbaik pada metode Breslow dan metode Efron. Model regresi Cox dengan variabel prediktor umur, muntah, lidah kotor, hemoglobin dan leukosit yang merupakan model terbaik dari metode *exact* memberikan hasil sebagai berikut.

Tabel 6. ANOVA model terbaik *exact*

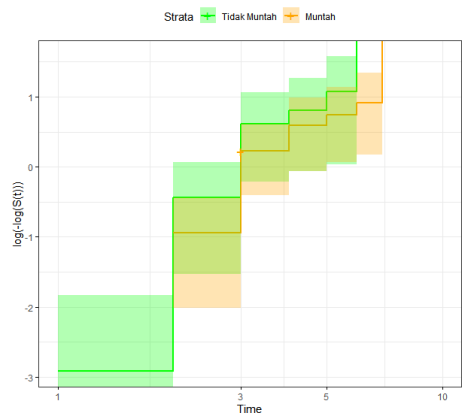
Variabel	Coef	exp(coef)	Z	P-value
umur	-0,0633	0,9386	-2,9190	0,0035
muntah	-1,8435	0,1583	-2,8420	0,0045
lidah kotor	1,8926	6,6368	1,6650	0,0959
hemoglobin	1,0531	2,8665	1,9000	0,0575
leukosit	0,9753	2,6521	1,8610	0,0628

Berdasarkan Tabel 6 dapat dilihat bahwa ada lima variabel prediktor dengan nilai AIC terbaik. Berikutnya akan dilakukan pengujian terhadap signifikansi model dengan statistik uji *G*. Statistik uji *G* bergantung pada nilai \ln *likelihood* model tereduksi dan \ln *likelihood* model *full*. Berdasarkan Tabel 6, dengan lima variabel prediktor diperoleh nilai $\ln L_R$ (*reduction*) dan $\ln L_f$ (*full*) secara berturut-turut sebesar -59,24681 dan -50,54205. Dengan demikian diperoleh nilai $G = -2(\ln L_R - \ln L_f) = 17,4095 > \chi_{0,05;5}^2 = 11,0705$ yang berarti model layak digunakan.

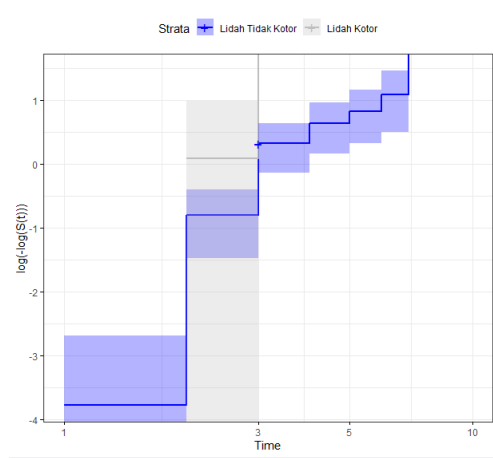
Pengujian selanjutnya adalah uji parsial untuk melihat variabel prediktor apa saja yang berpengaruh signifikan terhadap variabel respon. Berdasarkan Tabel 6 terlihat bahwa ada dua variabel prediktor yang signifikan pada taraf signifikansi 5%, yaitu umur dan muntah. Sedangkan ketika menggunakan taraf signifikansi 10%, maka semua variabel prediktor pada Tabel 6 adalah signifikan.

3.4 Pengujian Asumsi Proportional Hazard

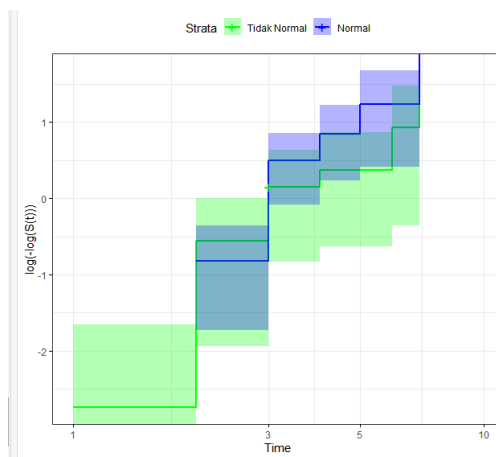
Pengujian asumsi *proportional hazard* dapat dilakukan dengan plot *log minus log* dan uji *Goodness Of Fit* (GOF). Pengujian dengan plot *log minus log* hanya dapat dilakukan pada variabel prediktor yang bersifat kategorik. Hasil pengujian dengan plot *log minus log* dapat digunakan sebagai dugaan sementara kemudian dilakukan uji GOF untuk memperoleh kesimpulan yang lebih pasti.



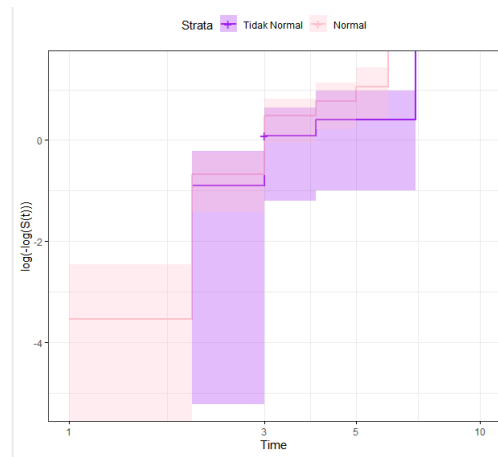
Gambar 2. Plot *log minus log* pada variabel muntah



Gambar 3. Plot *log minus log* pada variabel lidah kotor



Gambar 4. Plot *log minus log* pada variabel hemoglobin



Gambar 5. Plot *log minus log* pada variabel leukosit

Pada Gambar 2 – 5, terlihat bahwa kedua garis pada masing-masing variabel cenderung sejajar, yang berarti bahwa asumsi *proportional hazard* pada variabel muntah, lidah kotor, hemoglobin dan leukosit terpenuhi. Selanjutnya, untuk variabel prediktor kontinu seperti variabel umur dilakukan uji GOF sekaligus melihat secara objektif variabel prediktor yang memenuhi asumsi *Proportional Hazard* (PH). Hasil pengujian dengan uji GOF dapat dilihat pada Tabel 7. Dengan menggunakan taraf signifikansi $\alpha = 5\%$, maka semua variabel prediktor memenuhi asumsi PH.

Tabel 7. Hasil Uji GOF

Variabel	P-value	Keputusan	Kesimpulan
Umur	0,5000	Gagal menolak H_0	Memenuhi asumsi PH
Muntah	0,4800	Gagal menolak H_0	Memenuhi asumsi PH
Lidah Kotor	0,7300	Gagal menolak H_0	Memenuhi asumsi PH
Hemoglobin	0,3600	Gagal menolak H_0	Memenuhi asumsi PH
Leukosit	0,2900	Gagal menolak H_0	Memenuhi asumsi PH

Berdasarkan pengujian asumsi *proportional hazard* dengan melihat plot *log minus log* dan uji GOF, diperoleh variabel umur, muntah, lidah kotor, hemoglobin dan leukosit memenuhi asumsi *proportional hazard*.

Tabel 8. Hasil estimasi model terbaik secara keseluruhan

Variabel	Coef	Exp(coef)
Umur	-0,0633	0,9386
Muntah	-1,8435	0,1583
Lidah Kotor	1,8926	6,6368
Hemoglobin	1,0531	2,8665
Leukosit	0,9753	2,6521

Berdasarkan Tabel 8, koefisien variabel umur sebesar $-0,0633$, sehingga *hazard ratio* untuk variabel umur sama dengan $0,9386$. Hal ini berarti, jika ada dua pasien yang usianya berbeda setahun tetapi kondisinya sama (variabel selain umur dianggap bernilai sama) maka pasien yang lebih muda mempunyai laju kesembuhan $\frac{1}{e^{-0,0633}} = e^{0,0633} = 1,0653$ kali lebih besar dibandingkan pasien yang lebih tua.

Koefisien variabel muntah sebesar $-1,8435$, sehingga *hazard ratio* untuk variabel muntah sama dengan $0,1583$. Hal ini berarti, jika ada dua pasien yang satu mengalami muntah dan yang satunya lagi tidak (kondisi lainnya dianggap sama), maka pasien yang tidak mengalami muntah mempunyai laju kesembuhan (kesempatan sembuh) $\frac{1}{0,1583} = 6,3171$ kali lebih besar dibandingkan pasien yang mengalami muntah.

Koefisien variabel lidah kotor sebesar $1,8926$, sehingga *hazard ratio* variabel lidah kotor sama dengan $6,6368$. Hal ini berarti, jika ada dua pasien yang satu mengalami lidah kotor dan yang satunya lagi tidak (kondisi lainnya dianggap sama), maka pasien yang mengalami lidah kotor mempunyai laju kesembuhan (kesempatan sembuh) $6,6368$ kali lebih besar dibandingkan pasien yang tidak mengalami lidah kotor.

Koefisien variabel hemoglobin sebesar $1,0531$, sehingga *hazard ratio* sama dengan $2,8665$. Hal ini berarti, jika ada dua pasien yang satu memiliki kadar hemoglobin normal dan yang satunya lagi tidak (kondisi lainnya dianggap sama), maka pasien dengan kadar hemoglobin normal memiliki laju kesembuhan (kesempatan sembuh) lebih besar $2,8665$ kali dibandingkan pasien dengan kadar hemoglobin tidak normal.

Koefisien variabel leukosit sebesar $0,9753$, sehingga *hazard ratio* variabel leukosit sama dengan $2,6521$. Hal ini berarti, jika ada dua pasien yang satu memiliki kadar leukosit normal dan yang satunya tidak (kondisi lainnya dianggap sama), maka pasien dengan kadar leukosit normal memiliki laju kesembuhan (kesempatan sembuh) lebih besar $2,6521$ kali dibandingkan pasien dengan kadar leukosit tidak normal.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, diperoleh kesimpulan bahwa umur, muntah, lidah kotor, hemoglobin dan leukosit berpengaruh terhadap laju kesembuhan pasien demam tifoid. Model Cox pada kasus demam Tifoid di Kabupaten Bekasi adalah

$$h(t, \mathbf{X}) = h_0(t) \exp(-0,0633 \text{ umur} - 1,8435 \text{ muntah} + 1,8926 \text{ lidah kotor} + 1,0531 \text{ hemoglobin} + 0,9753 \text{ leukosit})$$

Laju kesembuhan pasien dengan umur yang lebih muda $1,0653$ kali dibandingkan dengan satu tahun usia lebih tua. Laju kesembuhan pasien yang tidak mengalami muntah, $6,3171$ kali dibandingkan dengan pasien yang muntah. Laju kesembuhan pasien yang mengalami lidah kotor $6,6368$ kali dibandingkan dengan yang tidak kotor. Laju kesembuhan pasien dengan kadar hemoglobin normal $2,8665$ kali dibandingkan dengan pasien dengan kadar hemoglobin yang tidak normal. Laju kesembuhan pasien dengan leukosit normal memiliki laju kesembuhan $2,6521$ kali dibandingkan pasien dengan kadar leukosit tidak normal.

Daftar Pustaka

- [1] Kleinbaum, D.G. dan Klein, M., (2004), *Survival Analysis*, Springer.
- [2] Collett, D., (2015), *Modelling Survival Data in Medical Research*, CRC press.
- [3] Guo, S., (2010), *Survival Analysis*, Oxford University Press.
- [4] Brembilla, A., Olland, A., Puyraveau, M., Massard, G., Mauny, F., dan Falcoz, P.E., (2018), Use of the Cox regression analysis in thoracic surgical research, *J. Thorac. Dis.* **10**, 3891-3896. <https://doi.org/10.21037/jtd.2018.06.15>
- [5] Ozenne, B., Sorensen, A.L., Scheike, T., Torp-Pedersen, C., dan Gerds, T.A., (2017), riskRegression: Predicting the risk of an event using Cox regression models, *The R Journal* **9**, 440-460.
- [6] Atlam, M., Torkey, H., El-Fishawy, N., dan Salem, H., (2021), Coronavirus disease 2019 (COVID-19): survival analysis using deep learning and Cox regression model, *Pattern Analysis and Applications* **24**, 993-1005. <https://doi.org/10.1007/s10044-021-00958-0>
- [7] Hughey, J.J., Rhoades, S.D., Fu, D.Y., Bastarache, L., Denny, J.C., dan Chen, Q., (2019), Cox regression increases power to detect genotype-phenotype associations in genomic studies using the electronic health record, *BMC genomics* **20**, 1–7. <https://doi.org/10.1186/s12864-019-6192-1>
- [8] Wijayakusuma, H.H., (2006), *Tanaman Obat untuk Penyakit Anak*. Pustaka Populer Obor.
- [9] Davey, P., (2015), *Medicine at a Glance*, John Wiley & Sons.
- [10] Ulfa, F. dan Handayani, O.W.K., (2018), Kejadian demam Tifoid di wilayah kerja Puskesmas Pagiyanten. *HIGEIA (Journal of Public Health Research and Development)* **2**, 227–238.
- [11] Ramaningrum, G., Anggraheny, H.D., dan Putri, T.P., (2017), Faktor-faktor yang mempengaruhi kejadian demam Tifoid pada anak di RSUD Tugurejo Semarang, *Jurnal Kedokteran Muhammadiyah* **5**.
- [12] Priskila, E., Silangit, T., dan Rimbun, S., (2021), Hubungan antara total leukosit dan trombosit pada pasien demam Typhoid. *JKM* **14**, 41–48.
- [13] Chao, D.L., Park, J.K., Marks, F., Ochiai, R.L., Longini, I.M., dan Halloran, M.E., (2015), The contribution of neighbours to an individual's risk of Typhoid outcome, *Epidemiology & Infection* **143**, 3520–3527.
- [14] Fatekurohman, M., Nurmala, N., dan Anggraeni, D., (2018), Comparison of exact, Efron and Breslow parameter approach method on hazard ratio and stratified cox regression model. *In Journal of Physics: Conference Series* vol **1008** p 012007. IOP Publishing.

- [15] Borucka, J., (2014), Methods of handling tied events in the Cox proportional *hazard* model, *Studia Oeconomica Posnaniensia* **2**, 91–106.
- [16] Lee, E.T. and Wang, J., (2003), *Statistical Methods for Survival Data Analysis*, 3rd ed., John Wiley & Sons, Inc.
- [17] Kleinbaum, D.G. dan Klein, M., (2005), *Survival Analysis: A self-learning text*, 2nd ed., *New York: Springer Science and Business Media, Inc.*
- [18] Kuitunen, I., Ponkilainen, V.T., Uimonen, M.M., Eskelinen, A., dan Reito, A., (2021), Testing the proportional *hazards* assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: methodological perspectives and review, *BMC Musculoskeletal Disorders* **22**, 489.