



PENGANTAR ANALISIS SURVIVAL (B) - SS234418

**ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI
SURVIVAL TIME PASIEN PENDERITA PRIMARY BILIARY
CIRRHOSIS**

MUHAMMAD RAFI INSANI

NRP 5003221155

RYANALDI ROBBY KUSUMA

NRP 5003221182

MUHAMMAD ILHAM RAMADHAN

NRP 5003221185

Dosen Pengampu

Jerry Dwi Trijoyo Purnomo, S.Si. M.Si, Ph.D

NIP 19810223 200812 1 003

Shofi Andari, S.Stat., M.Si., Ph.D.

NIP 198712072014042001

Program Studi Sarjana Statistika

Departemen Statistika

Fakultas Sains dan Analitika Data

Institut Teknologi Sepuluh Nopember

Surabaya

2024

DAFTAR ISI

DAFTAR ISI.....	2
DAFTAR TABEL.....	4
BAB I PENDAHULUAN	6
1.1 Latar Belakang.....	6
1.2 Rumusan Masalah.....	7
1.3 Tujuan Penelitian.....	7
1.4 Manfaat Penelitian.....	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Analisis Survival.....	9
2.1.1 Fungsi Survival dan Fungsi Hazard.....	9
2.1.2 Kurva Survival Kaplan-Meier.....	11
2.1.3 Uji Log Rank.....	11
2.2 Asumsi Proportional Hazard	12
2.3 Model Cox Proportional Hazard.....	15
2.4 Pengujian Residual	16
2.4.1 Schoenfeld Residual.....	16
2.4.2 Martingale Residual	16
2.5 Model Regresi Cox Time-Dependent Covariate	17
2.6 Estimasi Parameter Model Regresi Cox-Time Dependent Covariate	17
2.7 Primary Biliary Cholangitis.....	18
BAB III METODOLOGI PENELITIAN	19
3.1 Sumber Data	19
3.2 Variabel Penelitian	19
3.3 Langkah Analisis	23
BAB IV ANALISIS DAN PEMBAHASAN	24
4.1 Pre-processing Data.....	24
4.2 Uji Distribusi	24
4.3 Statistika Deskriptif.....	25
4.4 Kurva Kaplan-Meier dan Uji Log-Rank.....	29
4.4.1 Faktor Treatment Bagi Pasien PBC.....	29
4.4.2 Faktor Sex Bagi Pasien PBC.....	31

4.4.3 Faktor Ascites Bagi Pasien PBC.....	32
4.4.4 Faktor Hepatomegaly Bagi Pasien PBC.	33
4.4.5 Faktor Spiders Bagi Pasien PBC.....	35
4.4.6 Faktor Edema Bagi Pasien PBC.	37
4.4.7 Faktor Stage Bagi Pasien PBC.....	39
4.5 Pemodelan Cox Proportional Hazard	40
4.6 Evaluasi Model Cox Proportional Hazard.....	42
4.6.1 Backward Elimination	42
4.6.2 Log-Log Plot.....	43
4.6.3 Grambsch-Therneau Test.....	46
4.6.4 Schoenfeld Residual.....	47
4.6.5 Martingale Residual	48
4.7 Extended Cox Model dengan Time-Dependent Covariates	52
4.8 Evaluasi Model	53
BAB V KESIMPULAN.....	54
DAFTAR PUSTAKA	55
LAMPIRAN	56

DAFTAR TABEL

Tabel 3. 1 Struktur Data	22
Tabel 4. 1 Pengujian Distribusi Data	24
Tabel 4. 2 Statistika Deskriptif.....	25
Tabel 4. 3 Hasil Uji Log-Rank Berdasarkan Treatment.....	30
Tabel 4. 4 Hasil Uji Log-Rank Berdasarkan Sex	31
Tabel 4. 5 Hasil Uji Log-Rank Berdasarkan Ascites	32
Tabel 4. 6 Hasil Uji Log-Rank Berdasarkan Hepatomegaly.....	34
Tabel 4. 7 Hasil Uji Log-Rank Berdasarkan Spiders	36
Tabel 4. 8 Hasil Uji Log-Rank Berdasarkan Edema	38
Tabel 4. 9 Hasil Uji Log-Rank Berdasarkan Stage	40
Tabel 4. 10 Pengujian Cox Model.....	40
Tabel 4. 11 Pengujian Cox Model setelah Backward elimination	42
Tabel 4. 12 Pengujian Grambsch-Therneau untuk Asumsi PH.....	46
Tabel 4. 13 Pengujian Cox Model Time-Dependent.....	52
Tabel 4. 14 Pengujian AIC pada Cox Model	53

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Kurva Kaplan-Meier.....	11
Gambar 2. 2 Ilustrasi Kurva Kaplan-Meier.....	13
Gambar 4. 1 Diagram Pie Treatment.....	26
Gambar 4. 2 Diagram Pie Sex	26
Gambar 4. 3 Diagram Pie Ascites	26
Gambar 4. 4 Diagram Pie Hepatomegaly.....	27
Gambar 4. 5 Diagram Pie Spiders	27
Gambar 4. 6 Diagram Pie Edema.....	27
Gambar 4. 7 Diagram Pie Stage	28
Gambar 4. 8 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Treatment.....	30
Gambar 4. 9 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Sex	31
Gambar 4. 10 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Ascites	32
Gambar 4. 11 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Hepatomegaly	33
Gambar 4. 12 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Spiders	35
Gambar 4. 13 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Edema	37
Gambar 4. 14 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Stage	39
Gambar 4. 15 Plot Log-Log Survival Probability berdasarkan Status Edema.....	44
Gambar 4. 16 Plot Log-Log Survival Probability berdasarkan Tahap Histologis Penyakit.....	45
Gambar 4. 17 Kurva Schoenfeld Test	47
Gambar 4. 18 Kurva Schoenfeld Test	48
Gambar 4. 19 Martingale Residual pada Variable Serum Bilirubin	49
Gambar 4. 20 Martingale Residual pada Variable Serum Albumin	49
Gambar 4. 21 Martingale Residual pada Variable Urine Copper	50
Gambar 4. 22 Martingale Residual pada Variable Aspartate Aminotransferase	50
Gambar 4. 23 Martingale Residual pada Variable <i>Standardised Blood Clotting Time</i>	51

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam dunia kesehatan, memahami faktor-faktor yang memengaruhi kelangsungan hidup pasien merupakan aspek penting untuk meningkatkan kualitas perawatan dan hasil pengobatan. Seiring dengan kemajuan teknologi dan metode analisis data, pendekatan statistik semakin banyak digunakan untuk menggali lebih dalam tentang dinamika penyakit dan mendukung pengambilan keputusan klinis. Salah satu metode statistik yang efektif untuk menganalisis data kelangsungan hidup adalah *Cox Proportional Hazards* (Cox PH). Metode ini memungkinkan peneliti untuk mengevaluasi hubungan antara berbagai faktor klinis dengan risiko terjadinya suatu peristiwa, seperti kematian atau memburuknya kondisi pasien.

Salah satu contoh penyakit yang memerlukan analisis semacam ini adalah *Primary Biliary Cirrhosis* (PBC), sebuah penyakit hati autoimun kronis yang dapat menyebabkan kerusakan hati progresif dan meningkatkan risiko kematian jika tidak diatasi dengan baik. Penyakit ini terutama menyerang perempuan paruh baya dan memiliki perkembangan yang berbeda pada setiap individu. Oleh karena itu, diperlukan metode analitis yang tepat untuk memahami faktor-faktor yang memengaruhi kelangsungan hidup pasien dengan PBC, dan model Cox PH merupakan salah satu metode yang efektif untuk tujuan ini.

Model Cox PH digunakan untuk mempelajari hubungan antara berbagai variabel klinis dengan waktu hingga terjadinya suatu kejadian, misalnya kematian atau perburukan kondisi. Dengan model ini, kita dapat mengevaluasi pengaruh variabel seperti kadar bilirubin, albumin, dan usia terhadap risiko yang dihadapi pasien. Dengan memanfaatkan dataset PBC yang tersedia dalam *package survival*, analisis ini mampu memberikan informasi tentang faktor-faktor yang paling berpengaruh terhadap prognosis pasien.

Dataset PBC mencakup berbagai data klinis dan demografis dari pasien yang didiagnosis dengan PBC. Melalui analisis ini, tenaga medis dan peneliti dapat mengidentifikasi variabel-variabel penting yang berperan dalam perkembangan penyakit, sehingga dapat merancang strategi pengobatan yang lebih efektif. Selain itu, hasil dari model Cox PH dapat membantu memprediksi waktu kelangsungan hidup pasien berdasarkan karakteristik individu mereka, yang pada akhirnya dapat meningkatkan kualitas perawatan yang diberikan.

Dengan menggunakan pendekatan analitis seperti model Cox PH, penelitian ini diharapkan dapat memberikan wawasan mendalam tentang dinamika PBC. Hasil dari analisis ini dapat digunakan sebagai landasan untuk pengambilan keputusan klinis yang lebih baik, membantu mengidentifikasi pasien dengan risiko tinggi, dan merancang intervensi yang lebih tepat untuk meningkatkan peluang kelangsungan hidup pasien.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang diangkat, maka rumusan masalah yang terbentuk sebagai berikut.

1. Apa saja variabel-variabel klinis dan demografis yang signifikan memengaruhi risiko kematian atau perburukan kondisi pada pasien *Primary Biliary Cirrhosis* (PBC)?
2. Bagaimana hubungan antara kadar bilirubin, albumin, usia, dan faktor klinis lainnya dengan waktu kelangsungan hidup pasien berdasarkan hasil analisis regresi Cox?
3. Seberapa akurat model regresi Cox dalam memprediksi risiko dan waktu kejadian seperti kematian atau perburukan kondisi pada pasien PBC?
4. Apakah ada perbedaan tingkat risiko antar pasien yang dapat diidentifikasi berdasarkan karakteristik tertentu menggunakan model regresi Cox?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang diberikan, didapatkan tujuan dari penelitian sebagai berikut.

1. Mengidentifikasi faktor-faktor klinis dan demografis yang signifikan memengaruhi risiko kematian atau perburukan kondisi pasien dengan *Primary Biliary Cirrhosis* (PBC) melalui analisis model *Cox Proportional Hazards* (Cox PH).
2. Menganalisis pengaruh variabel-variabel klinis seperti kadar bilirubin, albumin, usia, dan faktor-faktor lainnya terhadap waktu kelangsungan hidup pasien dengan PBC, serta menentukan sejauh mana faktor-faktor tersebut memengaruhi risiko kejadian, seperti kematian atau perburukan kondisi.
3. Mengevaluasi efektivitas model *Cox PH* dalam memprediksi risiko kematian atau perburukan kondisi berdasarkan karakteristik individu pasien, serta menilai akurasi model dalam memprediksi waktu kelangsungan hidup pasien PBC.
4. Mengidentifikasi perbedaan risiko antara pasien dengan karakteristik tertentu berdasarkan hasil model regresi *Cox*, yang dapat digunakan untuk merancang strategi perawatan yang lebih spesifik dan efektif.

1.4 Manfaat Penelitian

Penelitian ini dapat memberikan manfaat bagi beberapa pihak. Berikut ini adalah pihak-pihak yang akan memperoleh manfaat dari penelitian ini.

1. Bagi Peneliti

Penelitian ini memberikan kontribusi penting dalam bidang analisis kelangsungan hidup, khususnya dalam penerapan model Cox PH untuk penyakit hati autoimun. Peneliti dapat menggunakan hasil penelitian ini sebagai referensi untuk penelitian lanjutan

mengenai analisis risiko pada penyakit kronis lainnya. Penelitian ini juga membuka peluang untuk mengembangkan pendekatan-pendekatan baru dalam pemodelan risiko, dengan menggunakan teknik statistik canggih untuk memperdalam pemahaman tentang faktor-faktor yang memengaruhi prognosis penyakit.

2. Bagi Pembaca

Penelitian ini memberikan manfaat dalam memperluas pemahaman mengenai penyakit Primary Biliary Cirrhosis (PBC) serta pendekatan analitis yang digunakan untuk menganalisis kelangsungan hidup pasien dengan penyakit kronis. Bagi masyarakat umum, penelitian ini dapat memberikan pemahaman dasar tentang pentingnya pemantauan dan pengelolaan penyakit hati autoimun serta bagaimana data klinis dapat digunakan untuk meramalkan perjalanan penyakit dan merancang intervensi yang lebih efektif.

3. Bagi Tenaga Medis

Penelitian ini memberikan manfaat besar bagi tenaga medis, terutama dalam meningkatkan pemahaman mereka mengenai faktor-faktor klinis yang memengaruhi kelangsungan hidup pasien dengan Primary Biliary Cirrhosis (PBC). Hal ini memungkinkan dokter untuk merancang perawatan yang lebih personal dan efektif, serta memfokuskan perhatian pada pasien dengan risiko tinggi, sehingga kualitas perawatan dapat meningkat.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Survival

Analisis kelangsungan hidup atau analisis survival merupakan sekumpulan prosedur statistik untuk analisis data di mana variabel hasil yang diinginkan adalah waktu hingga suatu peristiwa terjadi. Fokus utama analisis ini adalah waktu survival, dan data sering mengandung censoring. Censoring adalah kondisi di mana waktu pasti terjadinya peristiwa tidak diketahui, baik karena peristiwa belum terjadi pada akhir pengamatan (kiri censoring) maupun karena peristiwa terjadi sebelum pengamatan (kanan censoring). Kemudian fungsi utama seperti fungsi survival, yang menentukan kemungkinan bahwa seseorang akan bertahan lebih lama dari waktu tertentu, dan fungsi hazard yang menunjukkan tingkat risiko peristiwa tertentu terjadi pada waktu tertentu, digunakan untuk memahami pola waktu survival. Dimungkinkan untuk menggunakan berbagai teknik analisis, seperti kurva survival yang diestimasi oleh Kaplan Meier, Cox Proportional Hazards untuk menilai pengaruh variabel, atau model parametrik seperti distribusi Weibull untuk memprediksi waktu survival. Variabel waktu menunjukkan waktu yang telah "bertahan hidup" bagi seorang individu selama periode tindak lanjut tertentu, biasanya menyebut peristiwa tersebut sebagai *failure* dalam analisis survival, karena peristiwa yang dimaksud biasanya adalah kematian, kejadian penyakit, atau beberapa pengalaman negatif individu lainnya (Kleinbaum, 2012).

2.1.1 Fungsi Survival dan Fungsi Hazard

Nilai T pada fungsi survival merupakan notasi waktu survival dan merupakan variabel random yang memiliki fungsi distribusi peluang $f(t)$, maka fungsi kepadatan peluang dapat dinyatakan sebagai berikut.

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (2.1)$$

Lalu, untuk fungsi distribusi kumulatif dapat dinyatakan sebagai berikut.

$$F(t) = P(T < t) = \int_0^t f(u) du \quad (2.2)$$

Probabilitas bahwa suatu objek akan bertahan setelah waktu ke- t dikenal sebagai fungsi survival ($S(t)$), dinyatakan dengan persamaan pada berikut.

$$S(t) = P(T \geq t) = 1 - P(T < t) \quad (2.3)$$

Fungsi hazard $h(t)$ juga dapat disebut *conditional failure rate* merupakan laju kegagalan sesaat setelah individu bertahan sampai waktu ke- t . Dengan demikian, fungsi hazard adalah kebalikan dari fungsi survival. Fungsi hazard dapat dinyatakan sebagai berikut.

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \quad (2.4)$$

Dengan menggunakan teori probabilitas bersyarat $P(A | B) = \frac{P(A \cap B)}{P(B)}$ sebagai berikut.

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

Dengan asumsi bahwa $F(t) = 1 - S(t)$, maka $f(t) = \frac{d(F(t))}{dt} = \frac{d(1-S(t))}{dt}$, maka nilai $h(t)$ akan diperoleh sebagai berikut.

$$\begin{aligned}
 h(t) &= \frac{\left(\frac{d(1-S(t))}{dt} \right)}{S(t)} \\
 &= \frac{\left(\frac{-d(S(t))}{dt} \right)}{S(t)} \\
 h(t) dt &= \frac{-d(S(t))}{S(t)}
 \end{aligned}$$

Selanjutnya, dengan mengintegrasikan kedua bagian fungsi tersebut, diperoleh sebagai berikut.

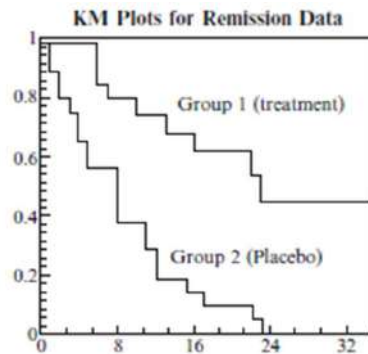
$$\begin{aligned}
 - \int_0^t h(u) du &= \int_0^t \frac{1}{S(u)} d(S(u)) \\
 &= \ln S(u) \Big|_0^t \\
 &= \ln S(u) - \ln S(0) \\
 &= \ln S(u)
 \end{aligned}$$

Dengan demikian, hubungan antara fungsi hazard dan survival dapat dinyatakan sebagai berikut.

$$H(t) = -\ln S(t) \quad (2.5)$$

2.1.2 Kurva Survival Kaplan-Meier

Dalam analisis survival, kurva survival Kaplan-Meier merupakan salah satu metode nonparametrik yang umum digunakan untuk mengestimasi fungsi survival ($S(t)$) berdasarkan data survival yang mungkin mengandung censoring. Metode ini membagi waktu pengamatan menjadi interval yang ditentukan oleh waktu peristiwa dan kemudian menghitung probabilitas survival kumulatif untuk masing-masing interval. Kurva Kaplan Meier memberikan visualisasi perubahan peluang survival dari waktu ke waktu menggunakan persamaan umum dari fungsi survival dengan contoh kurva sebagai berikut.



Gambar 2. 1 Ilustrasi Kurva Kaplan-Meier

Sumber : Buku Survival Analysis (Kleinbaum, 2012)

2.1.3 Uji Log Rank

Uji log-rank merupakan metode statistik yang digunakan untuk membandingkan dua atau lebih kurva survival untuk mengetahui apakah terdapat perbedaan yang signifikan dalam distribusi waktu survival antara kelompok yang berbeda. Menurut Kleinbaum (2012), Uji log-rank adalah uji chi-square sampel besar dengan kriteria statistik. Ini memberikan perbandingan kurva KM yang dibandingkan secara keseluruhan. Berikut merupakan hipotesis yang digunakan dalam pengujian Log-Rank.

Uji Hipotesis:

H_0 : Tidak terdapat perbedaan pada kurva survival antara kelompok yang berbeda dalam satu variabel

H_1 : Terdapat perbedaan pada kurva survival antara kelompok yang berbeda dalam satu variabel

Statistik uji:

$$\chi^2 = \sum_{i=1}^G \frac{(O_i - E_i)^2}{E_i} \quad (2.6)$$

dimana,

$$O_i - E_i = \sum_{i=1}^G \sum_{j=1}^n (m_{ij} - e_{ij}) ; e_{ij} = \left(\frac{n_{ij}}{\sum_{i=1}^G \sum_{j=1}^n n_{ij}} \right) (\sum_{i=1}^G \sum_{j=1}^n m_{ij}) \quad (2.7)$$

Keterangan:

O_i : nilai observasi individu kelompok ke-i

E_i : nilai ekspektasi individu kelompok ke-i

m_{ij} : jumlah objek yang mengalami event pada waktu ke- $t_{(j)}$ dan kelompok ke-i

n_{ij} : jumlah objek yang masih bertahan pada waktu ke- $t_{(j)}$ dan kelompok ke-i

e_{ij} : nilai ekspektasi pada waktu ke- $t_{(j)}$ dan kelompok ke-i

G : banyaknya kelompok

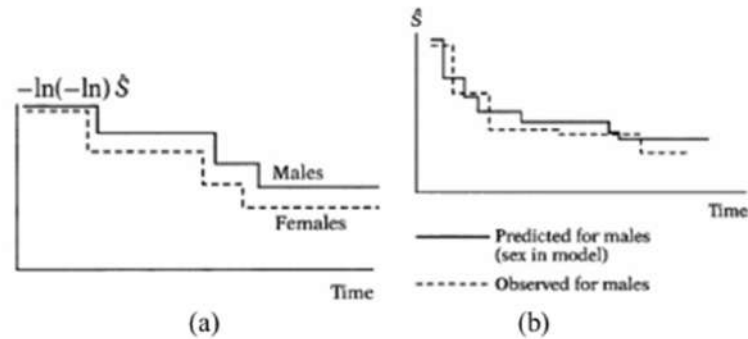
Kesimpulan: Tolak H_0 apabila $\chi^2_{hitung} > \chi^2_{(\alpha, G-1)}$

2.2 Asumsi Proportional Hazard

Suatu keadaan dikatakan dapat memenuhi asumsi proportional hazard apabila keadaan tersebut memiliki nilai hazard ratio yang konstan terhadap waktu (Kleinbaum & Klein, 2005). Untuk mengetahui apakah suatu keadaan memenuhi asumsi proportional hazard atau tidak dapat dilihat melalui tiga pendekatan sebagai berikut ini.

a. Pendekatan Grafik

Terdapat dua jenis grafik yang dapat digunakan untuk melakukan pengujian asumsi *proportional hazard*. Pendekatan grafik yang paling banyak digunakan dalam analisis survival adalah grafik $\ln [-\ln S(t)]$ dan grafik *observed versus expected* kurva survival yang dapat dijelaskan pada gambar berikut.



Gambar 2. 2 Ilustrasi Kurva Kaplan-Meier

Sumber : Buku Survival Analysis (Kleinbaum & Klein, 2005)

Diilustrasikan bahwa terdapat variabel jenis kelamin (*sex*) laki-laki (*males*) dan perempuan (*females*). Selanjutnya akan dilakukan analisis apakah variabel *sex* memenuhi asumsi *proportional hazard* atau tidak. Berdasarkan dengan gambar diatas (a), terlihat bahwa asumsi *proportional hazard* terpenuhi dikarenakan adanya garis yang mewakili data *males* sejajar dengan garis yang mewakili data *females*. Begitu juga apabila dilihat dari grafik *observed* versus *expected* pada gambar (b), variabel *sex* memenuhi asumsi *proportional hazard* karena kurva survival pengamatan (*observed*) dan prediksi (*expected*) saling berdekatan atau berhimpit (Kleinbaum & Klein, 2005).

b. Pendekatan Grafik

Pendekatan kedua yang digunakan dalam analisis asumsi *proportional hazard* adalah pendekatan dengan pengujian *goodness of fit (GOF)*. Pengujian ini dilakukan karena GOF menggunakan pendekatan statistik dan menghasilkan *p-value*. Terdapat beberapa langkah yang dalam pengujian *goodness of fit* ini.

1. Meregresikan waktu survival dengan prediktornya untuk mendapatkan nilai residual *Schoenfeld*.
2. Mencari korelasi antara variabel residual *Schoenfeld* dan waktu survival (diurutkan dari kecil ke besar).
3. Melakukan pengujian korelasi antara residual *Schoenfeld* dan waktu survival yang telah diurutkan dari besar ke kecil.

Hipotesis yang digunakan dalam pengujian korelasi ini adalah sebagai berikut.

Hipotesis :

$$H_0 : \rho = 0$$

$$H_0 : \rho \neq 0$$

Keputusan tolak H_0 ketika nilai *p-value* lebih dari taraf signifikan (α) yang artinya terdapat korelasi antara residual *Schoenfeld* dengan waktu survival yang telah diurutkan dari besar ke kecil. Oleh karena itu, asumsi *proportional* dapat terpenuhi ketika uji korelasi tidak signifikan (Kleinbaum & Klein, 2005).

c. Pendekatan Variabel *Time-Dependent*

Pendekatan untuk menguji asumsi proportional hazard selanjutnya adalah pendekatan dengan menggunakan variabel time-dependent. Menurut Kleinbaum & Klein (2005) terdapat tiga langkah yang dilakukan dalam pendekatan variabel *time-dependent*, yaitu :

1. Secara satu persatu.
2. Secara Simultan.
3. Secara penentuan variabel prediktor yang diduga tidak memenuhi asumsi proportional hazard.

2.3 Model Cox Proportional Hazard

Model Cox *proportional hazard* merupakan model matematis yang digunakan untuk menganalisis suatu data survival (Kleinbaum & Klein, 2015). Melalui model Cox *proportional hazard* dapat dilihat hubungan antara variabel independen terhadap variabel dependen yaitu waktu survival melalui fungsi hazardnya. Model Cox *proportional hazard* biasanya ditulis dalam formula model *hazard* seperti berikut.

$$h(t, X) = h_0(t) \exp \left(\sum_{j=1}^p \beta_j x_j \right) \quad (2.8)$$

Dimana $X = (X_1, X_2, \dots, X_p)$ adalah vektor dari suatu variabel prediktor, dan $h_0(t)$ adalah fungsi dari *baseline hazard*.

Model Cox *proportional hazard* menggambarkan *hazard* atau risiko suatu individu pada waktu t berdasarkan sekumpulan variabel independen X yang akan dianalisis. Model ini disebut sebagai model semiparametrik karena tidak memerlukan asumsi mengenai distribusi waktu kelangsungan hidup yang mendasari, dan fungsi *baseline hazard* tidak ditentukan secara eksplisit.

Estimasi parameter pada model Cox *proportional hazard* dilakukan dengan metode *Maximum Partial Likelihood Estimation* (MPLE), yaitu dengan memaksimalkan fungsi *partial likelihood*. Jika terdapat n sampel dan D_i adalah himpunan waktu yang mencakup semua waktu survival, maka fungsi *partial likelihood* dari model Cox *proportional hazard* dapat dinyatakan sebagai berikut.

$$L(\beta) = \prod_{i=1}^n \frac{\exp(\beta^T x_{(k)})}{\sum_{k \in D_i} \exp(\beta^T x_{(k)})} \quad (2.9)$$

Didapatkan fungsi *partial likelihood*, untuk langkah selanjutnya adalah memaksimalkan turunan pertama yang diperoleh implisit, maka digunakan oleh metode iterasi numerik, yaitu metode Newton-Raphson (Collet, 1994). Jika $g(\beta)$ adalah vektor berukuran $p \times 1$ yang merupakan turunan pertama fungsi $\ln L(\beta)$ terhadap parameter β . $H(\beta)$ sendiri merupakan matrik hessian berukuran $p \times p$ yang terisi turunan kedua dari fungsi $L(\beta)$, maka estimasi parameter pada iterasi tertentu sebagai berikut.

$$\beta_{hat}^{(m+1)} = \beta_{hat}^{(m)} - H^{-1}(\beta_{hat}^{(m)})g(\beta_{hat}^{(m)}) \quad (2.10)$$

Sebagai awal dari $\beta_{hat}^{(0)} = 0$, 0 merupakan iterasi yang akan berhenti jika $||\beta_{hat}^{m+1} - \beta_{hat}^m|| \leq \varepsilon$, dimana ε merupakan suatu bilangan yang sangat kecil.

2.4 Pengujian Residual

Dalam analisis data survival menggunakan model regresi Cox, istilah Cox Residual sering digunakan untuk merujuk pada berbagai jenis residual yang dihasilkan dari model tersebut. Residual-residual ini membantu mengevaluasi asumsi model, mendeteksi outlier, dan memeriksa kecocokan model secara keseluruhan. Berikut ini adalah jenis-jenis residual yang termasuk dalam kategori Cox Residual.

2.4.1 Schoenfeld Residual

Menurut Kleinbaum dan Klein (2005), uji dari Harrel dan Lee (1986) merupakan suatu variasi dari suatu uji asli yang diperkenalkan oleh Schoenfeld (1982). Uji Schoenfeld residual merupakan uji formal yang digunakan untuk menguji asumsi fungsi *proportional hazard* pada variabel independen yang diduga mempengaruhi waktu kegagalan. Pengujian ini dilakukan dengan mencari taksiran residual Schoenfeld untuk masing – masing individu pada tiap – tiap variabel dengan rumus sebagai berikut.

$$r_{ji} \text{ hat} = \delta_i (x_{ji} - a_{ji} \text{ hat}) \quad (2.11)$$

$$a_{ji} \text{ hat} = \frac{\sum_{l \in R_{(ti)}} x_{jl} e^{\beta x_l}}{\sum_{l \in R_{(ti)}} e^{\beta x_l}} \quad (2.12)$$

dimana

$r_{ji} \text{ hat}$ = Taksiran Schoenfeld residual dari variabel ke – j untuk individu ke – i

x_{ji} = Nilai variabel ke – j untuk individu ke – i

δ_i = Indikator *censoring*

$a_{ji} \text{ hat}$ = Rata – rata tertimbang dari nilai kovariat

2.4.2 Martingale Residual

Martingale residual merupakan metode yang digunakan untuk memeriksa bentuk fungsional dari kovariat yang diberikan membuat penggunaan terbaik untuk menjelaskan pengaruhnya pada Cox *proportional hazard*. Persamaan *martingale residual* dapat di jelaskan pada persamaan berikut ini.

$$M_i \text{ hat} (t) = N_i(t) - \int_0^t Y_i(s) \exp (\beta_{hat} Z_i(s)) d\Lambda_0 \text{ hat} (s) \quad (2.13)$$

dimana

$M_i\hat{h}(t)$ = Martingale Residual ke – i pada waktu ke – t

$N_i(t)$ = Proses perhitungan event (Data tidak tersensor di beri nilai 1 dan data tersensor diberi nilai 0) untuk data ke – i pada waktu ke – t

Y_i = Indikator, jika subjek ke – i adalah *under risk immediately* sebelum t

$\beta\hat{h}' Z_i$ = Vektor dimensi dari covariat

$\Lambda_0\hat{h}(s)$ = Estimasi Breslow dari fungsi *baseline hazard* kumulati

$$\Lambda_0\hat{h}(t) = \int_0^t \frac{\sum dN_i(s)}{\sum Y_j(s) \exp(\beta\hat{h}' Z_i(s))} \quad (2.14)$$

2.5 Model Regresi Cox Time-Dependent Covariate

Memodifikasi model regresi Cox *proportional hazard* merupakan model yang melibatkan variabel prediktor bebas oleh waktu dan variabel prediktor terikat oleh waktu (Kleinbaum & Klein, 2005). Seperti pada model regresi Cox *proportional hazard*, modifikasi model ini juga memuat fungsi *baseline hazard* $h_0(t)$ dikalikan dengan fungsi eksponensial. *Time-dependent* dilambangkan dengan $x(t)$ dan fungsi *hazard* tergantung pada waktu ke – t, sehingga hubungan dengan $x(t)$ dapat dinyatakan sebagai berikut.

$$h\{t | x(t)\} = h_0(t) \exp \sum_{j=1}^n \beta_j x_j \quad (2.15)$$

Probabilitas survival dinyatakan dengan $S(t | x(t)) = \exp[-H\{t|x(t)\}]$ sehingga,

$$H\{t | x(t)\} = \int_0^t h_0(u) \exp\{\beta^T x(u)\} du = H_0(t) \exp(\beta^T x(t))$$

Nilai hazard kumulatif $x(t)$ yaitu sebagai berikut :

$$H\hat{h}\{t | x(t)\} = \sum_{i=1}^n \int_0^t \frac{\exp\{\beta\hat{h}^T x(u)\} d\delta_i(u)}{\sum_{k=1}^n Y_k(u) \exp\{\beta\hat{h}^T x_k(u)\}} \quad (2.16)$$

Dimana $\delta_i(t)$ merupakan proses poisson sehingga $\delta_i(t) = I(T_i \geq t)$ dan I merupakan fungsi indikator yang bernilai 1 untuk status event meninggal dan 0 untuk status survive sedangkan $Y_k(t)$ menyatakan penderita yang survive pada waktu ke – t. Fungsi survival untuk variabel yang tergantung pada waktu ke – t adalah sebagai berikut (Thomas & Reyes, 2014).

$$S\hat{h}\{t | x(t)\} = \exp[-H\hat{h}\{t | x(t)\}] \quad (2.17)$$

2.6 Estimasi Parameter Model Regresi Cox-Time Dependent Covariate

Jika jumlah penderita dalam penelitian adalah n , dan data terdiri dari $\{t_i, \delta_i, x_i(t)\}$, dimana $i = 1, \dots, n$, dan dimana t_i adalah waktu pengamatan (data pencatatan terakhir) untuk penderita ke- i , δ_i menunjukkan status penderita yang bernilai 1 untuk status peristiwa meninggal dan 0 untuk status hidup pada waktu ke- t , dan x_i adalah kovariat dari penderita ke- i pada waktu ke- t . Sebagai hasil dari nilai fungsi partial likelihood, estimasi diberikan sebagai berikut.

$$L(\beta) = \sum_{i=1}^n \delta_i \left(x_i(t) \frac{\sum_{k \in D_i} \exp(\beta^T x_k(t)) (x_k(t))}{\sum_{k \in D_i} \exp(\beta^T x_k(t))} \right) \quad (2.18)$$

Dimana himpunan penderita D_i memiliki waktu survival yang lebih lama dari waktu dari pengamatan ke- i , nilai $\hat{\beta}$ dihitung dengan menggunakan estimasi maximum partial likelihood. Dengan matriks kovarians $I^{-1}(\hat{\beta})$, diketahui bahwa $\hat{\beta}$ konsisten dan asimtotik normal (bersifat kontinu), maka sebagai berikut.

$$I(\beta) = \sum_{i=1}^n \delta_i \left(\frac{\frac{\sum_{k \in D_i} \exp(\beta^T x_k(t)) (x_k(t)) (x_k(t))^T}{\sum_{k \in D_i} \exp(\beta^T x_k(t))}}{\frac{\{\sum_{k \in D_i} \exp(\beta^T x_k(t)) (x_k(t))\} \{\sum_{k \in D_i} \exp(\beta^T x_k(t)) (x_k(t))\}^T}{\sum_{k \in D_i} \exp(\beta^T x_k(t))}} \right) \quad (2.19)$$

(Fisher & Lin, 1999)

2.7 Primary Biliary Cholangitis

Kondisi peradangan yang dikenal sebagai *Primary Biliary Cholangitis* (PBC) atau diterjemahkan dalam Bahasa Indonesia yaitu Kongalitis Bilier Primer, menyebabkan saluran empedu kecil di hati rusak. Perkembangan yang lambat namun tak terhentikan akhirnya mengakibatkan dekompensasi hati dan sirosis. Penyakit ini dikenal setidaknya sejak tahun 1851 dan dijuluki "sirosis bilier primer" pada tahun 1949. Pada tahun 2014, kelompok advokasi pasien mengusulkan untuk mengubah istilah "sirosis" menjadi "kolangitis bilier primer" karena merupakan karakteristik penyakit lanjut. Menurut Tanaka (2024), Kondisi kolestatik autoimun kronis yang dikenal sebagai kolangitis bilier primer terutama menyerang wanita berusia antara 40 dan 70 tahun. Meskipun terdapat variabilitas geografis dan penurunan rasio wanita terhadap pria, studi epidemiologi terkini telah mengungkapkan peningkatan insiden secara global. Kolangitis bilier primer, seperti penyakit autoimun lainnya, berkembang pada orang yang rentan secara genetik saat mereka terpapar pemicu lingkungan, khususnya xenobiotik, merokok, dan mikrobiota usus.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini berasal dari dataset Mayo Clinic Primary Biliary Cholangitis (PBC), yang berasal dari uji klinis Mayo Clinic pada penyakit Primary Biliary Cholangitis (PBC) yang dilakukan antara tahun 1974 dan 1984. Uji klinis ini bertujuan untuk mengevaluasi efektivitas pengobatan D-penicillamine pada pasien dengan PBC, yang sebelumnya dikenal sebagai Primary Biliary Cirrhosis. Dataset mencakup pengamatan pasien selama lebih dari 10 tahun, baik pasien yang mendapatkan pengobatan maupun yang berada dalam kelompok kontrol.

3.2 Variabel Penelitian

Variabel yang digunakan pada penelitian ini terdiri atas variabel dependen dan variabel independen. Variabel dependen dalam penelitian ini adalah waktu survival (T) pasien penderita Primary Biliary Cholangitis (PBC). Waktu survival diukur sebagai waktu selama pasien menjalani perawatan hingga terjadi kegagalan (*failure event*) atau akhir masa pengamatan, yang dinyatakan dalam satuan hari.

Adapun ketentuan dari waktu survival dalam penelitian ini adalah waktu awal (*time origin*) ditetapkan sejak pasien pertama kali masuk dalam pengamatan penelitian klinis Mayo Clinic terkait Primary Biliary Cholangitis (PBC). Untuk skala pengukuran waktu *survival* dihitung dalam satuan hari, dimulai dari waktu awal hingga terjadinya kejadian kegagalan (*failure event*). Sedangkan kegagalan (*failure event*) didefinisikan sebagai kondisi di mana pasien meninggal akibat PBC atau penyebab lainnya selama periode penelitian.

Variabel dependen dalam penelitian ini adalah data survival, yaitu waktu yang diperlukan pasien penderita PBC (Primary Biliary Cholangitis) untuk bertahan hidup. Waktu ini diukur sejak pasien pertama kali masuk dalam pengamatan penelitian Mayo Clinic terkait PBC (start point) hingga waktu akhir, yaitu ketika pasien keluar dari rumah sakit atau meninggal akibat PBC atau penyebab lainnya (*failure event*), yang dilambangkan dengan T, dengan satuan waktu dalam hari. Perbedaan utama antara analisis survival dan analisis statistik lainnya adalah adanya data tersensor, di mana status pasien dikategorikan 0 sebagai tersensor dan 1 sebagai tidak tersensor.

Berikut adalah penjelasan tentang variabel independen dalam penelitian PBC (Primary Biliary Cholangitis):

1. Usia (*Age*)
Variabel usia menggambarkan usia pasien dalam tahun saat pengamatan dimulai. Usia pasien dalam penelitian ini diukur dalam satuan tahun, dan data ini merupakan data numerik yang menggunakan skala rasio.
2. Jenis Kelamin (*Sex*)

Jenis kelamin pasien dalam penelitian ini dikategorikan dalam dua kelompok, yaitu laki-laki dan perempuan. Variabel ini merupakan data kategori dengan skala nominal, karena kedua kategori ini tidak memiliki urutan atau perbedaan skala.

3. Edema

Edema merujuk pada adanya penumpukan cairan dalam tubuh pasien, yang dapat menyebabkan pembengkakan pada beberapa bagian tubuh. Variabel ini dikategorikan dalam dua kelompok: "ada edema" dan "tidak ada edema," yang merupakan data kategori dengan skala nominal. Penilaian ini memberikan gambaran tentang seberapa berat kondisi pasien terkait gangguan fungsi hati.

4. Bilirubin

Kadar bilirubin dalam darah adalah indikator penting dari fungsi hati. Bilirubin yang tinggi dapat menandakan kerusakan hati atau obstruksi saluran empedu. Variabel ini diukur dalam satuan miligram per desiliter (mg/dl), sehingga merupakan data numerik dengan skala rasio.

5. Albumin

Albumin adalah protein utama yang diproduksi oleh hati dan berfungsi untuk menjaga tekanan osmotik darah serta transportasi zat-zat penting. Kadar albumin dalam darah dapat mencerminkan status gizi pasien serta fungsi hati. Variabel ini diukur dalam satuan gram per desiliter (g/dL), dan menggunakan data numerik dengan skala rasio, di mana nol absolut berarti tidak ada albumin dalam darah.

6. Urine Copper

Kadar tembaga dalam darah juga berperan sebagai indikator kesehatan hati. Tembaga yang tinggi dalam tubuh dapat mengarah pada gangguan hati yang lebih serius, termasuk pada PBC. Variabel ini diukur dalam satuan mikrogram per hari (umol/day) dan termasuk dalam data numerik dengan skala rasio. Kadar tembaga yang lebih tinggi atau lebih rendah dapat memberikan wawasan terkait kondisi penyakit.

7. Aspartate Aminotransferase (AST)

Aspartate Aminotransferase adalah enzim yang ditemukan dalam hati dan otot, yang kadarnya meningkat ketika terjadi kerusakan pada hati atau otot. Variabel ini diukur dalam satuan unit per mililiter (U/ml) dan merupakan data numerik dengan skala rasio. Peningkatan kadar AST dapat menunjukkan adanya kerusakan pada jaringan hati pasien.

8. Prothrombin Time

Prothrombin Time mengukur waktu yang dibutuhkan untuk pembekuan darah dan dapat menunjukkan gangguan fungsi hati. Prothrombin Time yang lebih lama menunjukkan penurunan kemampuan darah untuk membeku, yang sering kali terkait dengan penyakit hati. Variabel ini diukur dalam detik dan menggunakan skala rasio, di mana waktu normal menggambarkan fungsi hati yang baik.

9. Stadium Penyakit (Stage)

Stadium penyakit menggambarkan tingkat keparahan penyakit berdasarkan hasil biopsi hati. Variabel ini diukur dalam kategori ordinal yang menunjukkan urutan stadium dari yang paling ringan hingga yang paling berat. Stadium ini memberi gambaran seberapa jauh penyakit berkembang pada pasien dan dapat mempengaruhi prognosis mereka.

10. Ascites

Ascites adalah kondisi penumpukan cairan dalam rongga perut yang sering terjadi pada pasien dengan penyakit hati lanjut. Variabel ini dikategorikan dalam dua kelompok: "ada ascites" dan "tidak ada ascites," yang merupakan data kategori dengan skala nominal. Kehadiran ascites mengindikasikan kondisi kesehatan yang lebih serius pada pasien.

11. Spider Veins

Spider veins adalah pembuluh darah kecil yang terlihat pada kulit pasien, sering kali muncul pada pasien dengan gangguan hati. Variabel ini dikategorikan dalam dua kelompok: "ada" atau "tidak ada," dengan skala nominal, yang menggambarkan apakah pasien menunjukkan gejala ini atau tidak. Keberadaan spider veins dapat menunjukkan perubahan terkait penyakit hati.

12. Alkali fosfatase (ALP)

Alkali fosfatase (ALP) adalah enzim yang ditemukan dalam hati, tulang, dan saluran empedu, diukur dalam satuan unit per liter (U/L). Peningkatan kadar ALP bisa menunjukkan adanya masalah pada hati atau saluran empedu, yang sering kali ditemukan pada pasien dengan penyakit hati kronis.

13. Serum kolesterol

Serum kolesterol mengukur kadar kolesterol dalam darah, diukur dalam miligram per desiliter (mg/dL). Kolesterol tinggi bisa meningkatkan risiko penyakit kardiovaskular. Pada pasien dengan PBC, kadar kolesterol sering kali lebih tinggi, yang dapat menunjukkan gangguan metabolisme lemak akibat kerusakan hati.

14. Hepatomegali

Hepatomegali, yaitu pembesaran hati, dinilai dengan nilai 1 (ada pembesaran hati) atau 0 (tidak ada pembesaran hati). Hepatomegali merupakan gejala yang sering terlihat pada pasien PBC, karena penyakit ini menyebabkan peradangan dan kerusakan hati.

15. Platelet

Jumlah trombosit dalam darah yang diukur dalam satuan per mikroliter. Trombosit berfungsi dalam pembekuan darah, dan rendahnya jumlah trombosit bisa menandakan masalah pada fungsi hati atau gangguan sumsum tulang.

16. Status

Status pasien pada akhir penelitian, yang dibagi menjadi tiga kategori: 0 untuk data tersensor, 1 untuk pasien yang menjalani transplantasi hati, dan 2 untuk pasien yang meninggal. Variabel ini digunakan untuk menganalisis waktu kelangsungan hidup atau kejadian yang terjadi selama pengamatan.

17. Time

Variabel ini mengukur waktu dalam satuan hari sejak pendaftaran pasien hingga kejadian yang relevan (kematian, transplantasi, atau pengamatan akhir pada 1986). Ini adalah variabel utama yang digunakan dalam analisis survival.

18. Treatment

Variabel ini menunjukkan jenis perawatan yang diterima oleh pasien, dengan nilai 1 untuk D-penicillamine, 2 untuk plasebo, dan NA untuk pasien yang tidak di-randomisasi. Variabel ini digunakan untuk menganalisis efek terapi yang berbeda terhadap pasien.

19. Trigliserida

Trigliserida mengukur kadar trigliserida dalam darah, diukur dalam miligram per desiliter (mg/dL). Trigliserida yang tinggi sering berhubungan dengan peningkatan risiko penyakit jantung dan kondisi metabolik lainnya.

Struktur data yang digunakan untuk penelitian ini adalah sebagai berikut.

Tabel 3. 1 Struktur Data

ID	Time	Status	X_1	X_2	...	X_m
1	T_1	0	X_{11}	X_{21}	...	X_{m1}
2	T_2	1	X_{12}	X_{22}	...	X_{m2}
3	T_3	1	X_{13}	X_{23}	...	X_{m3}
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
n	T_n	0	$X_{1,n}$	$X_{2,n}$...	$X_{m,n}$

Status pasien dikategorikan untuk menggambarkan keadaan pasien selama periode pengamatan. Di dalam penelitian ini, status dikodekan sebagai 0 atau 1, dengan 0 menunjukkan bahwa pasien dianggap tersensor, sedangkan 1 menunjukkan bahwa pasien tidak tersensor. Pasien dikatakan tersensor apabila keluar dari penelitian dengan sebab tertentu seperti meninggal atau mengundurkan diri ketika penelitian masih berlangsung dan pasien belum mengalami event padahal penelitian sudah berakhir.

3.3 Langkah Analisis

Langkah-langkah analisis dalam penelitian ini dilakukan secara sistematis untuk menjawab tujuan penelitian. Tahapan analisis meliputi beberapa langkah berikut:

1. Eksplorasi Data
Data diperiksa untuk memastikan kelengkapan, validitas, dan menghapus nilai hilang (missing values). Karakteristik awal variabel dieksplorasi untuk memahami distribusinya.
2. Analisis Deskriptif
Statistik deskriptif seperti rata-rata, median, dan distribusi data dihitung untuk menggambarkan karakteristik populasi penelitian.
3. Estimasi Kaplan-Meier
Kurva Kaplan-Meier digunakan untuk memperkirakan fungsi survival pasien secara keseluruhan dan dalam subkelompok tertentu.
4. Uji Log-rank
Uji log-rank dilakukan untuk membandingkan fungsi survival antar kelompok berdasarkan variabel independen.
5. Uji Asumsi Proportional Hazards (PH)
Asumsi PH diuji secara global dan untuk masing-masing variabel untuk memastikan validitas model Cox.
6. Permodelan Cox Proportional Hazards
Model Cox digunakan untuk mengidentifikasi variabel yang memengaruhi waktu survival dan menghitung risiko relatif (hazard ratio).
7. Pemodelan Extended Cox dengan *time-dependent covariates*
Untuk variabel yang melanggar asumsi PH, model extended Cox diterapkan dengan memasukkan variabel time-dependent.
8. Interpretasi hasil
Hasil dianalisis untuk memahami pengaruh variabel terhadap waktu survival, serta relevansinya dengan implikasi klinis.

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Pre-processing Data

Proses pre-processing data bertujuan untuk mempersiapkan data agar sesuai dengan kebutuhan analisis. Data yang digunakan dalam analisis ini adalah data Primary Biliary Cirrhosis (PBC), yang berisi informasi mengenai penyakit hati kronis pada pasien. Pada tahap awal, ditemukan adanya nilai missing (NA) pada beberapa variabel, yang kemudian dihilangkan karena metode analisis Cox Proportional Hazard tidak dapat menangani data dengan nilai missing. Setelah penghapusan entri dengan nilai missing, jumlah data yang semula sebanyak 312 menjadi 276 data.

Selain itu, pada data awal terdapat variabel status yang mencatat kategori berbeda, yaitu pasien yang tersensor, pasien yang meninggal akibat penyakit hati, dan pasien yang meninggal karena penyebab lain. Untuk tujuan analisis ini, dilakukan transformasi pada variabel status menjadi dua kategori: pasien yang tersensor dan pasien yang mengalami kejadian, yaitu kematian akibat penyakit hati. Transformasi ini bertujuan untuk memfokuskan analisis pada kejadian yang relevan, sehingga data menjadi lebih sederhana dan sesuai dengan model yang akan digunakan.

4.2 Uji Distribusi

Statistika Pengujian untuk mendapatkan distribusi yang tepat pada data menggunakan bantuan *software* Minitab yang dimuat pada tabel berikut.

Tabel 4. 1 Pengujian Distribusi Data

Distribution	AD	P	LRT P
Normal	1.990	<0.005	
Box-Cox Transformation	0.733	0.056	
Lognormal	8.259	<0.005	
3-Parameter Lognormal	0.645	*	0.000
Exponential	20.048	<0.003	
2-Parameter Exponential	18.816	<0.010	0.002
Weibull	0.890	0.022	
3-Parameter Weibull	0.514	0.204	0.052
Smallest Extreme Value	6.842	<0.010	
Largest Extreme Value	0.764	0.047	
Gamma	2.509	<0.005	
3-Parameter Gamma	0.582	*	0.000
Logistic	1.958	<0.005	
Loglogistic	3.771	<0.005	
3-Parameter Loglogistic	0.948	*	0.000
Johnson Transformation	0.409	0.343	

Pada Tabel 4.1, dapat dilihat bahwa nilai p-value dapat menyatakan bahwa data mengikuti beberapa distribusi tetapi dapat digunakan pada pembentukan model seperti 3-parameter Weibull dan Johnson. Hipotesis pengujian yang digunakan sebagai berikut.

H_0 = Waktu survival mengikuti distribusi tertentu.

H_1 = Waktu survival tidak mengikuti distribusi tertentu.

Pada taraf signifikansi sebesar 0.05 diperoleh nilai p-value sebesar <0.005 untuk distribusi yang dapat digunakan pada pembentukan model, sehingga didapatkan keputusan tolak H_0 , dimana memenuhi asumsi untuk melakukan regresi Cox.

4.3 Statistika Deskriptif

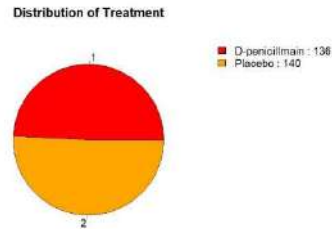
Statistika deskriptif merupakan metode yang digunakan untuk mengumpulkan dan menyajikan data agar dapat memberikan informasi yang bermanfaat (Walpole, 1995). Berikut ini adalah hasil analisis statistika deskriptif dari data yang digunakan.

Tabel 4. 2 Statistika Deskriptif

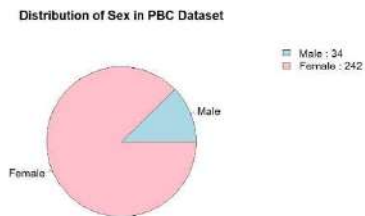
Variabel	N	Mean	StDev	Var	Min	Max
Time	276	1979.2	1112.4	12373389.9	41	4556
Age	276	49.283	0.4912	110.269	26	78
Bili	276	3.334	4.601	21.170	0.3	28
Chol	276	371.3	234.8	55125.6	120	1775
Albumin	276	3.5168	0.4048	0.1639	1.96	4.4
Copper	276	100.77	88.27	7791.37	4	588
Alk.phos	276	1997	2115	4475247	289	13862
Ast	276	124.12	56.72	3217.15	28.38	457.25
Trig	276	124.98	65.28	4261.58	33	598
Platelet	276	261.77	93.13	8672.98	62	563
Protime	276	10.736	1.008	1.017	9	17.1

Berdasarkan dari Tabel 4.2 terlihat bahwa rata-rata pasien berumur 49.283 dengan pasien termuda berumur 26 tahun dan yang paling tua berumur 78 tahun. ‘Bili’ yang menunjukkan kadar bilirubin serum yang dapat mencerminkan fungsi hati. Dapat dilihat pada tabel bahwa rata-rata dari kadar bilirubin adalah 3.334 dengan satuan mg/dL. Selanjutnya terdapat variabel ‘Chol’ yang menunjukkan tingkat kolesterol serum. Terlihat pada tabel bahwa rata-rata dari ‘Chol’ sebesar 371.3 dengan satuan mg/dL. Selanjutnya, untuk variabel ‘Albumin’ yang menunjukkan kadar albumin, suatu protein yang diproduksi oleh organ hati memiliki nilai rata-rata sebesar 3.5168 g/dL. Selanjutnya variabel ‘Copper’ merupakan kadar tembaga dalam urin dalam periode 24 jam, diperoleh rata-rata sebesar 100.77 $\mu\text{g}/24$ jam. Kemudian untuk variabel ‘Alk.phos’ menunjukkan kadar dari alkaline phosphatase serum yang berguna sebagai enzim yang meningkat pada gangguan hati, variabel tersebut memiliki rata-rata sebesar 1997 U/L. Selanjutnya untuk variabel ‘Ast’ menunjukkan kadar enzim aspartate aminotransferase (SGOT) yang berfungsi sebagai indikator kerusakan hati, variabel ‘Ast’ memiliki nilai rata-rata sebesar 124.12 U/mL.

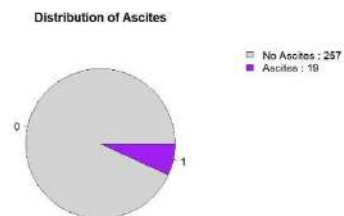
Selanjutnya variabel ‘trig’ menunjukkan kadar trigliserida serum yang merupakan jenis lemak dalam darah, variabel ‘trig’ sendiri memiliki nilai rata-rata sebesar 124.98 mg/dL. Selanjutnya variabel ‘platelet’ menunjukkan jumlah trombosit dalam darah yang memiliki rata-rata 261.77 dengan satuan $10^3/\mu L$. Terakhir terdapat variabel ‘Prottime’ yang menunjukkan waktu protrombin sebagai pengukur kemampuan pembekuan darah dalam tubuh yang memiliki rata-rata sebesar 10.736 detik. Untuk variabel kategorikal dapat dijelaskan dengan menggunakan diagram *pie* yang disajikan dalam gambar sebagai berikut.



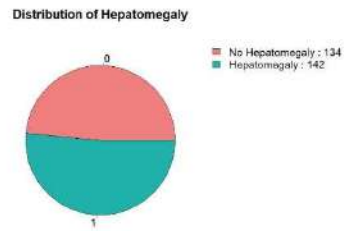
Gambar 4. 1 Diagram Pie Treatment



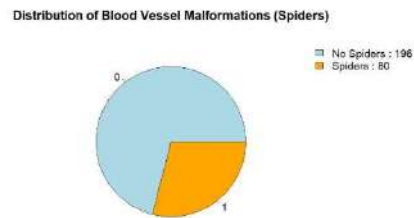
Gambar 4. 2 Diagram Pie Sex



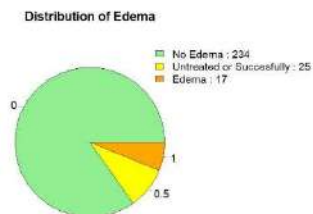
Gambar 4. 3 Diagram Pie Ascites



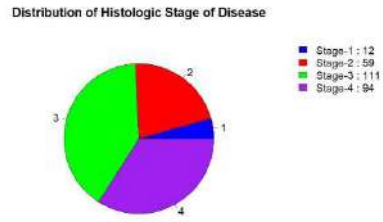
Gambar 4. 4 Diagram Pie Hepatomegaly



Gambar 4. 5 Diagram Pie Spiders



Gambar 4. 6 Diagram Pie Edema



Gambar 4. 7 Diagram Pie Stage

Pada tujuh diagram pie yang telah disajikan sebelumnya dapat diketahui bahwa autoimun kronis yang menyerang organ hati (PBC) paling banyak menggunakan treatment Placebo sebanyak 140 pasien sedangkan untuk treatment D – penicillamine sebanyak 136 pasien. Selanjutnya, PBC paling banyak diderita oleh pasien yang berjenis kelamin perempuan sebanyak 242 pasien, sedangkan untuk pasien laki-laki hanya sebanyak 34 pasien yang terpaut sangat jauh. Selanjutnya, pasien yang mempunyai Ascites lebih sedikit yakni sebanyak 19 pasien, jika dibandingkan dengan pasien yang tidak mempunyai Ascites sangatlah sangatlah karena jumlah dari pasien yang tidak memiliki Ascites adalah 257 pasien. Selanjutnya, untuk pasien PBC yang memiliki Hepatomegaly adalah sebanyak 142 pasien, hampir sama dengan pasien yang tidak memiliki Hepatomegaly yang berada di angka 134 pasien. Selanjutnya, pasien PBC yang mempunyai Spiders atau terjadinya kelainan pembuluh darah di kulit pasien sebanyak 80, sedangkan untuk pasien yang tidak mengalami Spiders berada di angka 196 pasien. Selanjutnya, pasien yang menderita PBC tetapi tidak mengalami Edema adalah sebanyak 234 pasien, untuk pasien yang mengalami Edema di tahap ringan sebanyak 25 pasien, dan untuk pasien PBC yang mengalami Edema di tahap parah adalah sebanyak 17 pasien. Selanjutnya, pasien penderita PBC yang berada di tahap histologis penyakit hati di stage 1 sebanyak 12 pasien, stage 2 sebanyak 59 pasien, untuk stage 3 sebanyak 111 pasien, dan untuk stage 4 berada di angka 94 pasien.

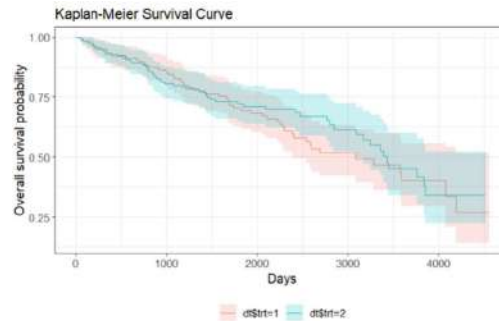
4.4 Kurva Kaplan-Meier dan Uji Log-Rank

Kurva Kaplan-Meier merupakan metode yang digunakan untuk memperkirakan fungsi kelangsungan hidup (Survival Function) pasien autoimun kronis yang menyerang organ hati berdasarkan data waktu hingga terjadi suatu kejadian (event) berdasarkan dengan faktor-faktor.

Uji Log-Rank merupakan sebuah metode yang digunakan untuk membandingkan suatu kurva survival dalam grup yang berbeda. Sebelum mengetahui karakteristik dan perbedaan kurva survival autoimun kronis yang menyerang hati berdasarkan faktor yang mempengaruhinya, dilakukan analisis deskriptif terlebih dahulu dengan menggunakan kurva Kaplan-Meier yang berfungsi untuk mengetahui gambaran karakteristik secara umum kurva survival. Berikut ini disajikan kurva Kaplan-Meier serta uji Log-Rank berdasarkan faktor-faktor yang diduga mempengaruhi ketahanan hidup pasien PBC.

4.4.1 Faktor Treatment Bagi Pasien PBC.

Terdapat dua kategori *treatment* yang dijalani pasien autoimun yang menyerang hati (PBC), diantaranya adalah D-penicillamine dengan pengkodean '1' dan Placebo dengan pengkodean '2'. Berikut ini merupakan kurva survival Kaplan-Meier berdasarkan faktor *treatment*.



Gambar 4. 8 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Treatment

Pada Gambar 4.8, dapat dilihat bahwa garis merah merupakan kategori dengan pengobatan ‘D-penicillamine dan garis biru merupakan kategori ‘Placebo’, dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cenderung serupa sepanjang waktu observasi. Hal ini, mengindikasikan bahwa tidak terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien yang mendapatkan D-penicillamine dan pasien yang menerima pengobatan Placebo. Dengan kata lain, pengobatan D-penicillamine kemungkinan tidak memberikan manfaat tambahan yang lebih baik dibandingkan dengan placebo dalam meningkatkan kelangsungan hidup pasien yang menderita PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan *treatment*.

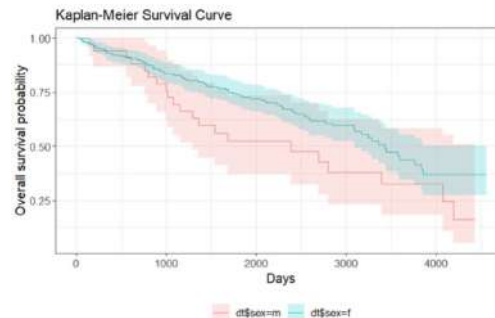
Tabel 4. 3 Hasil Uji Log-Rank Berdasarkan Treatment

Chi-Square	df	P-value
0.4	1	0.5

Pada pengujian Log-Rank faktor *treatment*, H_0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori D - penicillamine dan Placebo. H_1 memiliki arti bahwa terdapat perbedaan pada kurva survival antara kategori D - penicillamine dan Placebo. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.3, diperoleh nilai statistik uji sebesar 0.4 dengan derajat bebas 1 didapatkan *p-value* sebesar 0.5. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka *p-value* lebih dari α . Sehingga uji ini menghasilkan keputusan gagal tolak H_0 . Kesimpulan dari perhitungan *p-value* ini adalah tidak ada perbedaan kurva survival diantara pasien yang menderita PBC dengan *treatment* menggunakan D - penicillamine dan Placebo. Yang berarti waktu kelangsungan hidup antara pasien yang menderita PBC menggunakan treatment D – penicillamine sama dengan pasien yang menggunakan Placebo.

4.4.2 Faktor Sex Bagi Pasien PBC.

Terdapat dua kategori *Sex* yang dimiliki pasien autoimun yang menyerang hati (PBC), diantaranya adalah laki-laki (*male*) dengan pengkodean ‘m’ dan perempuan (*female*) dengan pengkodean ‘f’. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor *Sex*.



Gambar 4. 9 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Sex

Pada Gambar 4.9, dapat dilihat bahwa garis merah merupakan kategori dengan pasien berjenis kelamin ‘Laki-laki’ dan garis biru merupakan kategori pasien yang berjenis kelamin ‘Perempuan’. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cenderung serupa sepanjang waktu observasi. Hal ini, mengindikasikan bahwa tidak terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien yang berjenis kelamin Laki-laki dan pasien yang berjenis kelamin Perempuan. Dengan kata lain, jenis kelamin Laki-laki kemungkinan tidak memberikan manfaat tambahan yang lebih baik dibandingkan dengan Perempuan dalam meningkatkan kelangsungan hidup pasien yang menderita PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan jenis kelamin (*sex*).

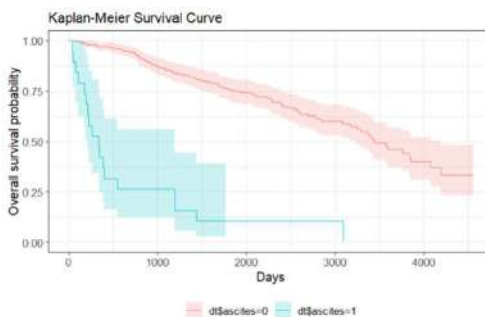
Tabel 4. 4 Hasil Uji Log-Rank Berdasarkan Sex

Chi-Square	df	P-value
4.5	1	0.03

Pada pengujian Log-Rank faktor *Sex*, H_0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori Laki-laki dan Perempuan. H_1 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori Laki-laki dan Perempuan. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.4, diperoleh nilai statistik uji sebesar 4.5 dengan derajat bebas 1 didapatkan *p-value* sebesar 0.03. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka *p-value* kurang dari α . Sehingga uji ini menghasilkan keputusan tolak H_0 . Kesimpulan dari perhitungan *p-value* ini adalah terdapat perbedaan kurva survival diantara pasien yang menderita PBC dengan faktor *Sex* yakni Laki-laki dan Perempuan. Yang berarti waktu kelangsungan hidup antara pasien berjenis kelamin Laki-laki berbeda dengan pasien yang berjenis kelamin Perempuan, dimana pasien yang berjenis kelamin perempuan mempunyai peluang bertahan hidup lebih lama daripada pasien yang berjenis kelamin Laki-laki.

4.4.3 Faktor Ascites Bagi Pasien PBC.

Terdapat dua kategori ‘Ascites’ yang dimiliki pasien yang menderita PBC, Ascites sendiri merupakan terjadinya penumpukan cairan di rongga perut pasien. Pasien dikelompokkan berdasarkan dua kategori yakni, tidak terdapat ‘Ascites’ dengan pengkodean ‘0’, sedangkan untuk pasien yang terdapat ‘Ascites’ diberi pengkodean ‘1’. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor Ascites.



Gambar 4. 10 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Ascites

Pada Gambar 4.10, dapat dilihat bahwa garis merah merupakan kategori dengan pasien yang tidak terdapat Ascites dan garis biru merupakan kategori pasien yang terdapat Ascites. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cukup signifikan sepanjang waktu observasi dan dapat dilihat pula bahwa kedua kurva tidak berhimpit satu sama lain. Hal ini, mengindikasikan bahwa terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien PBC yang tidak memiliki Ascites dan pasien yang memiliki Ascites. Dengan kata lain, faktor Ascites kemungkinan memberikan pengaruh yang signifikan terhadap penyakit PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan Ascites.

Tabel 4. 5 Hasil Uji Log-Rank Berdasarkan Ascites

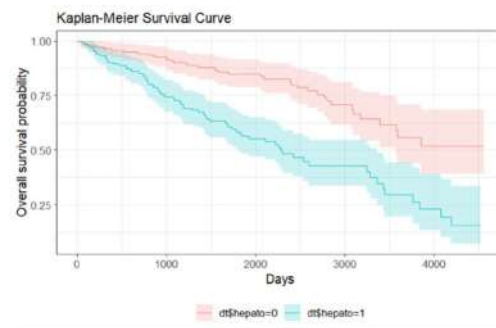
Chi-Square	df	P-value
------------	----	---------

110	1	< 2e-16
-----	---	---------

Pada pengujian Log-Rank faktor Ascites, H_0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori tidak terdapat Ascites di perut pasien dan terdapat Ascites pada pasien. H_1 memiliki arti bahwa terdapat perbedaan pada kurva survival antara kategori yang tidak terdapat Ascites di perut pasien dan terdapat Ascites pada pasien. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.5, diperoleh nilai statistik uji sebesar 110 dengan derajat bebas 1 didapatkan p -value sebesar $< 2e-16$. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka p -value kurang dari α . Sehingga uji ini menghasilkan keputusan tolak H_0 . Kesimpulan dari perhitungan p -value ini adalah terdapat perbedaan kurva survival diantara pasien yang menderita PBC dengan faktor Ascites yakni pasien yang tidak mempunyai Ascites dan pasien yang mempunyai Ascites. Yang berarti waktu kelangsungan hidup antara pasien yang mempunyai Ascites berbeda dengan pasien yang tidak mempunyai Ascites di perut mereka, dimana pasien yang tidak mempunyai Ascites memiliki peluang kelangsungan hidup yang lebih besar dibanding dengan pasien yang memiliki Ascites di tubuh mereka.

4.4.4 Faktor Hepatomegaly Bagi Pasien PBC.

Terdapat dua kategori ‘Hepatomegaly’ yang dimiliki pasien yang menderita PBC, Hepatomegaly sendiri merupakan terjadinya pembesaran hati pasien. Pasien dikelompokkan berdasarkan dua kategori yakni, tidak terdapat ‘Hepatomegaly’ dengan pengkodean ‘0’, sedangkan untuk pasien yang terdapat ‘Hepatomegaly’ diberi pengkodean ‘1’. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor Hepatomegaly.



Gambar 4. 11 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Hepatomegaly

Pada Gambar 4.11, dapat dilihat bahwa garis merah merupakan kategori dengan pasien yang tidak terdapat Hepatomegaly dan garis biru merupakan kategori pasien yang terdapat Hepatomegaly. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cukup signifikan sepanjang waktu observasi dan dapat dilihat pula bahwa kedua kurva tidak berhimpit satu sama lain. Hal ini, mengindikasikan bahwa terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien PBC yang tidak memiliki Hepatomegaly dan pasien yang memiliki Hepatomegaly. Dengan kata lain, faktor Hepatomegaly kemungkinan memberikan pengaruh yang signifikan terhadap penyakit PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan Hepatomegaly.

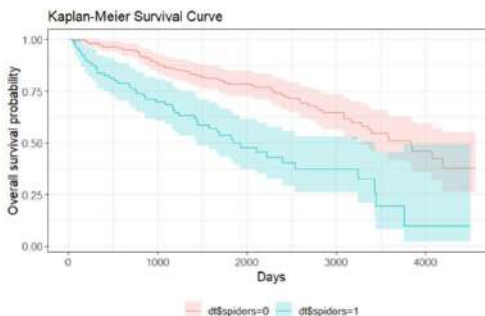
Tabel 4. 6 Hasil Uji Log-Rank Berdasarkan Hepatomegaly

Chi-Square	df	P-value
27.1	1	2e-07

Pada pengujian Log-Rank faktor Hepatomegaly, H0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori tidak terdapat Hepatomegaly di hati pasien dan terdapat Hepatomegaly pada pasien. H1 memiliki arti bahwa terdapat perbedaan pada kurva survival antara kategori yang tidak terdapat Hepatomegaly di hati pasien dan terdapat Hepatomegaly pada pasien. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.6, diperoleh nilai statistik uji sebesar 27.1 dengan derajat bebas 1 didapatkan p -value sebesar $2e-07$. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka p -value kurang dari α . Sehingga uji ini menghasilkan keputusan tolak H0. Kesimpulan dari perhitungan p -value ini adalah terdapat perbedaan kurva survival diantara pasien yang menderita PBC dengan faktor Hepatomegaly yakni pasien yang tidak terjadi Hepatomegaly dan pasien yang terjadi Hepatomegaly. Yang berarti waktu kelangsungan hidup antara pasien yang mempunyai Hepatomegaly berbeda dengan pasien yang tidak mempunyai Hepatomegaly di hati mereka, dimana pasien yang tidak memiliki Hepatomegaly memiliki peluang kelangsungan hidup yang lebih besar dibandingkan dengan pasien PBC yang mempunyai Hepatomegaly.

4.4.5 Faktor Spiders Bagi Pasien PBC.

Terdapat dua kategori 'Spiders' yang dimiliki pasien yang menderita PBC, Spiders sendiri merupakan terjadinya kelainan pembuluh darah di kulit. Pasien dikelompokkan berdasarkan dua kategori yakni, tidak terdapat 'Spiders' dengan pengkodean '0', sedangkan untuk pasien yang terdapat 'Spiders' diberi pengkodean '1'. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor Spiders.



Gambar 4. 12 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Spiders

Pada Gambar 4.12, dapat dilihat bahwa garis merah merupakan kategori dengan pasien yang tidak terdapat Spiders dan garis biru merupakan kategori pasien yang terdapat Spiders. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cukup signifikan sepanjang waktu observasi dan dapat dilihat pula bahwa kedua kurva tidak berhimpit satu sama lain. Hal ini, mengindikasikan bahwa terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien PBC yang tidak memiliki Spiders dan pasien yang memiliki Spiders. Dengan kata lain, faktor Spiders kemungkinan memberikan pengaruh yang signifikan terhadap penyakit PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan Spiders.

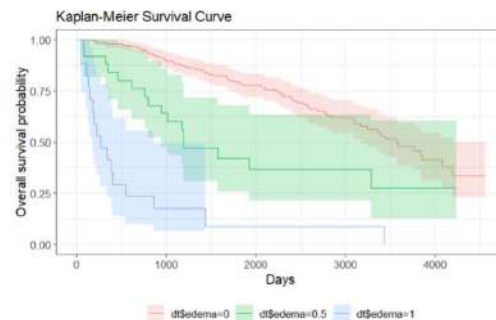
Tabel 4. 7 Hasil Uji Log-Rank Berdasarkan Spiders

Chi-Square	df	P-value
27.2	1	2e-07

Pada pengujian Log-Rank faktor Spiders, H_0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara kategori tidak terdapat Spiders di kulit pasien dan terdapat Spiders pada pasien. H_1 memiliki arti bahwa terdapat perbedaan pada kurva survival antara kategori yang tidak terdapat Spiders di kulit pasien dan terdapat Spiders pada pasien. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.7, diperoleh nilai statistik uji sebesar 27.2 dengan derajat bebas 1 didapatkan p -value sebesar $2e-07$. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka p -value kurang dari α . Sehingga uji ini menghasilkan keputusan tolak H_0 . Kesimpulan dari perhitungan p -value ini adalah terdapat perbedaan kurva survival diantara pasien yang menderita PBC dengan faktor Spiders yakni pasien yang tidak terjadi Spiders dan pasien yang terjadi Spiders. Yang berarti waktu kelangsungan hidup antara pasien yang mempunyai Spiders berbeda dengan pasien yang tidak mempunyai Spiders di kulit mereka, dimana pasien yang tidak memiliki Spiders mempunyai peluang kelangsungan hidup yang lebih besar dibanding dengan yang mempunyai Spiders.

4.4.6 Faktor Edema Bagi Pasien PBC.

Terdapat tiga kategori 'Edema' yang dimiliki pasien yang menderita PBC, Edema sendiri merupakan terjadinya pembengkakan akibat penumpukan darah. Pasien dikelompokkan berdasarkan tiga kategori yakni, tidak terdapat 'Edema' dengan pengkodean '0', untuk pasien yang memiliki 'Edema' ringan diberi pengkodean '0.5', sedangkan untuk pasien yang terdapat 'Edema' parah diberi pengkodean '1'. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor Edema.



Gambar 4. 13 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Edema

Pada Gambar 4.13, dapat dilihat bahwa garis merah merupakan kategori dengan pasien yang tidak memiliki Edema, garis hijau menunjukkan pasien dengan kategori yang mempunyai Edema ringan, dan garis biru merupakan kategori pasien yang terdapat Edema parah. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cukup signifikan sepanjang waktu observasi dan dapat dilihat pula bahwa ketiga kurva tidak berhimpit satu sama lain. Hal ini, mengindikasikan bahwa terdapat perbedaan yang signifikan dalam kelangsungan hidup antara pasien PBC yang tidak memiliki Edema dan pasien yang memiliki Edema secara ringan maupun parah. Dengan kata lain, faktor Edema kemungkinan memberikan pengaruh yang signifikan terhadap penyakit PBC. Namun, untuk memastikan apakah perbedaan ini signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan Edema.

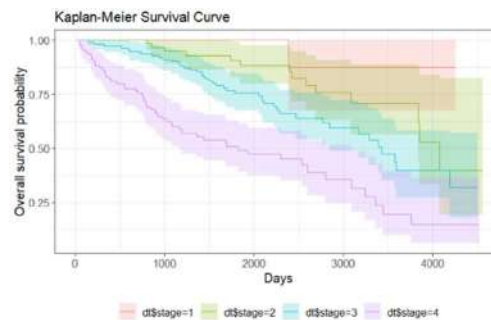
Tabel 4. 8 Hasil Uji Log-Rank Berdasarkan Edema

Chi-Square	df	P-value
113	2	< 2e-16

Pada pengujian Log-Rank faktor Edema, H0 memiliki arti bahwa tidak terdapat perbedaan pada kurva survival antara ketiga kategori. H1 memiliki arti bahwa terdapat perbedaan pada kurva survival ketiga kategori. Berdasarkan hasil uji Log-Rank yang disajikan pada Tabel 4.8, diperoleh nilai statistik uji sebesar 113 dengan derajat bebas 2 didapatkan *p-value* sebesar $< 2e-16$. Jika dibandingkan dengan nilai α yakni sebesar 0.05, maka *p-value* kurang dari α . Sehingga uji ini menghasilkan keputusan tolak H0. Kesimpulan dari perhitungan *p-value* ini adalah terdapat perbedaan kurva survival diantara pasien yang menderita PBC dengan faktor Edema yakni pasien yang tidak terjadi Edema, dan pasien yang terjadi Edema secara ringan maupun parah. Yang berarti waktu kelangsungan hidup antara pasien yang mempunyai Edema berbeda dengan pasien yang tidak mempunyai Edema, dimana pasien yang tidak memiliki Edema memiliki peluang kelangsungan hidup lebih besar daripada kategori yang lain.

4.4.7 Faktor Stage Bagi Pasien PBC.

Terdapat empat kategori ‘Stage’ yang dimiliki pasien yang menderita PBC, Stage sendiri merupakan tahapan histologis penyakit hati. Pasien dikelompokkan berdasarkan empat kategori yakni tahap 1,2,3, dan 4. Stage 1 diberi pengkodean ‘1’, Stage 2 diberi pengkodean ‘2’, Stage 3 diberi pengkodean ‘3’, sedangkan untuk Stage 4 diberi pengkodean ‘4’. Berikut ini disajikan kurva survival Kaplan-Meier berdasarkan faktor Stage.



Gambar 4. 14 Kurva Kaplan-Meier Pasien PBC Berdasarkan Faktor Stage

Pada Gambar 4.14, dapat dilihat bahwa garis merah merupakan kategori dengan pasien yang berada di stage 1, garis hijau menunjukkan pasien dengan kategori yang berada di stage 2, garis biru menunjukkan pasien yang berada di stage 3, dan garis ungu merupakan kategori pasien yang berada di stage 4. dari kurva tersebut didapat juga bahwa terjadi penurunan probabilitas survival pada kedua kategori yang cukup signifikan sepanjang waktu observasi dan dapat dilihat pula bahwa kurva stage 2 dan 3 berhimpit serta stage 1 dan 2 berhimpit pula. Hal ini, mengindikasikan semakin tinggi tahapan penyakit (Stage 3 dan 4), semakin rendah peluang kelangsungan hidup pasien. Sebaliknya, pasien pada Stage 1 dan 2 memiliki prognosis yang lebih baik. Grafik ini menegaskan bahwa tahapan penyakit hati berhubungan kuat dengan kelangsungan hidup, dan deteksi dini pada tahapan awal dapat memberikan hasil yang lebih baik bagi pasien. Namun, untuk memastikan apakah terdapat perbedaan yang signifikan secara statistik, diperlukan analisis lebih lanjut, maka akan dilakukan pengujian Log-Rank untuk mengetahui perbedaan yang berarti antara kurva survival berdasarkan Stage.

Tabel 4. 9 Hasil Uji Log-Rank Berdasarkan Stage

Chi-Square	df	P-value
44.6	3	1e-09

Berdasarkan hasil uji Log-Rank yang ditampilkan pada Tabel 4.9, perbedaan kelangsungan hidup antar kelompok yang dikategorikan berdasarkan stagemenunjukkan hasil yang signifikan secara statistik. Nilai Chi-Square sebesar 44,6 dengan derajat kebebasan (df) sebesar 3, dan *p-value* sebesar 1e-09 menunjukkan bahwa terdapat perbedaan yang sangat signifikan antara kurva survival untuk kelompok stage yang berbeda. Dengan kata lain, tingkat keparahan stage mempengaruhi probabilitas dari kelangsungan hidup pasien secara signifikan. Pasien dengan stage yang lebih parah cenderung memiliki prognosis yang lebih buruk dibandingkan dengan pasien tanpa stage atau dengan stage yang lebih ringan. Hasil ini juga menekankan pentingnya untuk memantau dan menangani tahapan histologis penyakit hati sebagai faktor klinis yang memengaruhi hasil kelangsungan hidup pasien.

4.5 Pemodelan Cox Proportional Hazard

Pada bagian ini, dilakukan pemodelan menggunakan metode Cox Proportional Hazard (Cox PH) untuk menganalisis hubungan antara variabel-variabel independen dengan waktu survival pasien. Model ini dipilih karena mampu menangani data survival dengan waktu pengamatan yang bervariasi serta mempertimbangkan efek tersensor. Pemodelan ini bertujuan untuk mengidentifikasi faktor-faktor yang berpengaruh signifikan terhadap waktu survival pasien penderita PBC. Berikut adalah hasil pemodelan Cox Proportional Hazard berdasarkan data penelitian.

Tabel 4. 10 Pengujian Cox Model

Variabel	coef	exp(coef)	se(coef)	z	Pr(> z)
trt	-1.242e-01	8.832e-01	2.147e-01	-0.579	0.56290

age	2.890e-02	1.029e+00	1.164e-02	2.482	0.01305
sex	-3.656e-01	6.938e-01	3.113e-01	-1.174	0.24022
ascites	8.833e-02	1.092e+00	3.872e-01	0.228	0.81955
hepato	2.552e-02	1.026e+00	2.510e-01	0.102	0.91900
spiders	1.012e-01	1.107e+00	2.435e-01	0.416	0.67760
edema	1.011e+00	2.749e+00	3.941e-01	2.566	0.01029
bili	8.001e-02	1.083e+00	2.550e-02	3.138	0.00170
chol	4.918e-04	1.000e+00	4.442e-04	1.107	0.26829
albumin	-7.408e-01	4.767e-01	3.078e-01	-2.407	0.01608
copper	2.490e-03	1.002e+00	1.170e-03	2.128	0.03337
alk.phos	1.048e-06	1.000e+00	3.969e-05	0.026	0.97893
ast	4.070e-03	1.004e+00	1.958e-03	2.078	0.03767
trig	-9.758e-04	9.990e-01	1.333e-03	-0.732	0.46414
platelet	9.019e-04	1.001e+00	1.184e-03	0.762	0.44629
protime	2.324e-01	1.262e+00	1.061e-01	2.190	0.02850
stage	4.545e-01	1.575e+00	1.754e-01	2.591	0.00958
Concordance			0.849 (se = 0.018)		
Likelihood Ratio Test			167.7		
Wald Test			174.1		
Score (Logrank) Test			283		

Berdasarkan hasil pemodelan Cox Proportional Hazard, beberapa variabel independen menunjukkan pengaruh signifikan terhadap risiko kejadian (*failure event*) pada pasien PBC. Signifikansi variabel ditentukan berdasarkan nilai p-value dari uji Wald, di mana nilai p-value < 0.05 mengindikasikan bahwa variabel tersebut berpengaruh secara signifikan. Variabel pada model ini yang berpengaruh signifikan terhadap risiko kejadian antara lain adalah age, edema, bili, albumin, copper, ast, protime, dan stage. Variabel-variabel ini memiliki kontribusi penting dalam memprediksi risiko kejadian (*failure event*) pada pasien PBC berdasarkan model Cox Proportional Hazard yang telah dibangun.

4.6 Evaluasi Model Cox Proportional Hazard

Setelah membangun model Cox Proportional Hazard, langkah selanjutnya adalah melakukan evaluasi untuk memastikan bahwa model yang dihasilkan sesuai dengan data dan memenuhi asumsi yang diperlukan. Evaluasi ini bertujuan untuk memastikan model dapat menggambarkan hubungan antara variabel kovariat dengan risiko kejadian secara akurat dan dapat diinterpretasikan dengan baik.

4.6.1 Backward Elimination

Backward elimination adalah metode seleksi variabel yang bertujuan untuk menyederhanakan model Cox Proportional Hazard dengan menghapus variabel independen yang tidak signifikan secara bertahap. Hasil akhir dari backward elimination adalah model yang lebih sederhana namun tetap mampu menjelaskan hubungan antara variabel independen dan risiko kejadian secara signifikan. Dengan bantuan software Rstudio didapatkan model Cox Proportional Hazard setelah backward elimination sebagai berikut.

Tabel 4. 11 Pengujian Cox Model setelah Backward elimination

Variabel	coef	exp(coef)	se(coef)	z	Pr(> z)
age	0.0313836	1.0318812	0.0102036	3.076	0.00210
edema	0.8217952	2.2745795	0.3471465	2.367	0.01792
bili	0.0851214	1.0888492	0.0193352	4.402	1.07e-05
albumin	-0.7185954	0.4874364	0.2724486	-2.638	0.00835
copper	0.0028535	1.0028576	0.0009832	2.902	0.00370
ast	0.0043769	1.0043865	0.0018067	2.423	0.01541
protime	0.2275175	1.2554794	0.1013729	2.244	0.02481
stage	0.4327939	1.5415584	0.1456307	2.972	0.00296
Concordance				0.845 (se = 0.019)	
Likelihood Ratio Test				163.8	
Wald Test				176.1	
Score (logrank) Test				257.5	

Berdasarkan hasil pemodelan ulang menggunakan backward elimination, delapan variabel signifikan dipertahankan dalam model Cox Proportional Hazard, yaitu age, edema, bili, albumin, copper, ast, protime, dan stage. Model yang terbentuk adalah sebagai berikut.

$$h(t|x) = h_0(t) \exp(\beta_1(age) + \beta_2(edema) + \beta_3(bili) + \beta_4(albumi) + \beta_5(copper) + \beta_6(ast) + \beta_7(protime) + \beta_8(stage))$$

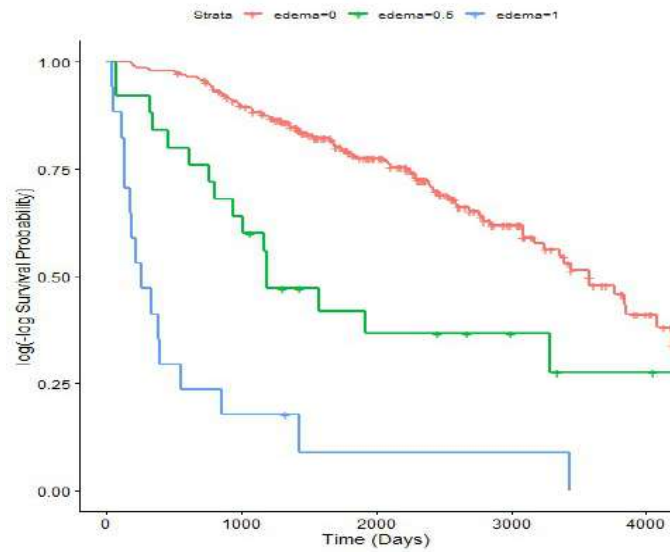
Setiap variabel ini menunjukkan hubungan signifikan dengan risiko kejadian pada tingkat kepercayaan 95%. Dari hasil estimasi model, dapat diperoleh nilai Hazard Ratio masing-masing variabel yang dapat menggambarkan hubungan antara variabel independen dan waktu kejadian (survival time).

Berdasarkan hasil estimasi model, variabel age memiliki hazard ratio (HR) sebesar 1.031, yang berarti setiap peningkatan satu tahun usia meningkatkan risiko kejadian sebesar 3.1%. Pasien dengan edema memiliki HR sebesar 2.275, menunjukkan bahwa pasien dengan edema memiliki risiko kejadian 2.28 kali lebih tinggi dibandingkan pasien tanpa edema. Variabel bili memiliki HR sebesar 1.089, yang berarti setiap peningkatan satu unit pada serum bilirubin meningkatkan risiko kejadian sebesar 8.9%. protime memiliki HR sebesar 1.255, menunjukkan bahwa setiap peningkatan satu unit pada waktu pembekuan darah meningkatkan risiko kejadian sebesar 25.5%. Terakhir, stage memiliki HR sebesar 1.542, yang berarti setiap peningkatan satu tingkat histologic stage meningkatkan risiko kejadian sebesar 54.2%.

Model ini menunjukkan kinerja yang baik dengan nilai concordance sebesar 0.845, yang mengindikasikan kemampuan model untuk membedakan risiko antara individu cukup tinggi. Selain itu, hasil dari likelihood ratio test, Wald test, dan log-rank test menunjukkan bahwa model secara keseluruhan signifikan dalam menjelaskan hubungan antara variabel independen dan risiko kejadian.

4.6.2 Log-Log Plot

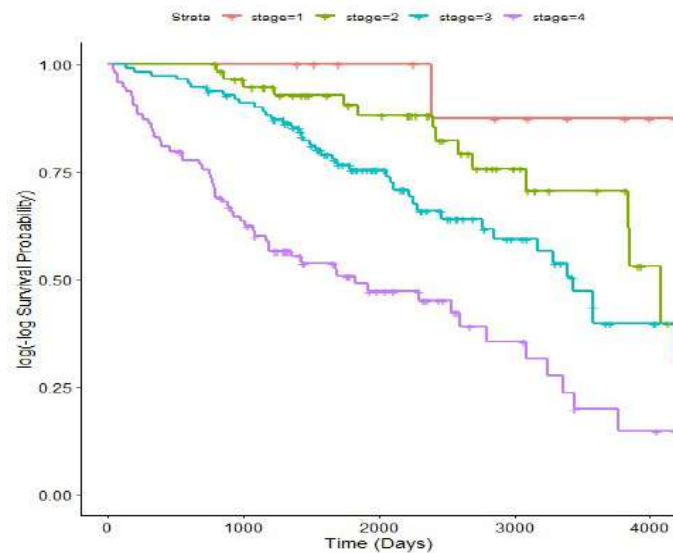
Dalam analisis ini, pendekatan Log-Log Survival Probability digunakan untuk memeriksa asumsi proportional hazard pada data survival berdasarkan status edema. Metode ini melibatkan transformasi fungsi survival menjadi Log-Log(Survival Probability), yang kemudian diplot terhadap waktu pengamatan untuk setiap kelompok. Jika asumsi proportional hazard terpenuhi, kurva Log-Log dari setiap kelompok akan menghasilkan garis yang sejajar atau mendekati paralel. Proses analisis dilakukan menggunakan perangkat lunak RStudio dengan data yang telah diolah sebelumnya. Grafik yang dihasilkan memberikan gambaran visual mengenai validitas asumsi proportional hazard untuk status edema. Berikut adalah hasil visualisasi dari analisis tersebut.



Gambar 4. 15 Plot Log-Log Survival Probability berdasarkan Status Edema

Gambar ini menunjukkan kurva Log-Log survival probability untuk tiga kelompok berdasarkan status edema, yaitu edema = 0 (merah), edema = 0.5 (hijau), dan edema = 1 (biru). Kurva untuk kelompok edema = 0 terlihat paling stabil dengan penurunan yang paling lambat, menunjukkan bahwa kelompok ini memiliki risiko kejadian paling rendah. Sementara itu, kurva untuk edema = 0.5 menunjukkan penurunan yang lebih cepat dibandingkan edema = 0, menandakan risiko yang lebih tinggi tetapi tidak secepat edema = 1. Kurva untuk kelompok edema = 1 memiliki penurunan paling tajam, mencerminkan risiko kejadian tertinggi dan tingkat survival terendah di antara ketiga kelompok. Ketidaksejajaran kurva antar kelompok, terutama kelompok edema = 1, mengindikasikan kemungkinan adanya pelanggaran asumsi proportional hazard.

Selanjutnya, dilakukan analisis Log-Log survival probability untuk memeriksa asumsi proportional hazard berdasarkan tahap histologis penyakit (Histologic Stage of Disease). Grafik Log-Log survival probability digunakan untuk memvisualisasikan pola probabilitas survival yang ditransformasikan terhadap waktu pengamatan, terpisah berdasarkan setiap tahap histologis. Kurva yang dihasilkan membantu menentukan apakah rasio hazard antar kelompok tetap konstan sepanjang waktu pengamatan. Gambar berikut menunjukkan hasil analisis Log-Log survival probability untuk tahap histologis penyakit.



Gambar 4. 16 Plot Log-Log Survival Probability berdasarkan Tahap Histologis Penyakit

Gambar ini menggambarkan kurva Log-Log survival probability untuk empat kelompok berdasarkan tahap histologis penyakit, yaitu stage 1 (merah), stage 2 (hijau), stage 3 (biru), dan stage 4 (ungu). Kurva untuk stage 1 menunjukkan pola yang stabil dengan tingkat penurunan paling lambat, menandakan risiko kejadian paling rendah dan tingkat survival tertinggi. Pada stage 2, kurva menurun lebih cepat dibandingkan stage 1, tetapi tetap lebih lambat dibandingkan stage 3 dan stage 4. Kurva untuk stage 3 menunjukkan penurunan yang signifikan, mencerminkan peningkatan risiko kejadian dibandingkan stage sebelumnya. Sementara itu, stage 4 memiliki penurunan kurva paling tajam, menandakan risiko kejadian tertinggi dan tingkat survival terendah di antara semua kelompok. Ketidaksejajaran kurva antar kelompok, terutama pada stage 4, dapat mengindikasikan adanya pelanggaran asumsi proportional hazard.

4.6.3 Grambsch-Therneau Test

Pada tahap ini, evaluasi asumsi proportional hazard dilakukan menggunakan Grambsch-Therneau Test. Uji ini bertujuan untuk mengidentifikasi apakah hubungan antara covariates dan hazard bersifat konstan seiring waktu. Dengan bantuan R, pengujian ini akan memberikan output berupa nilai p untuk setiap variabel dalam model, yang dapat digunakan untuk menentukan apakah asumsi proportional hazard terpenuhi atau tidak. Output berikut menyajikan hasil pengujian Grambsch-Therneau untuk model yang telah disusun.

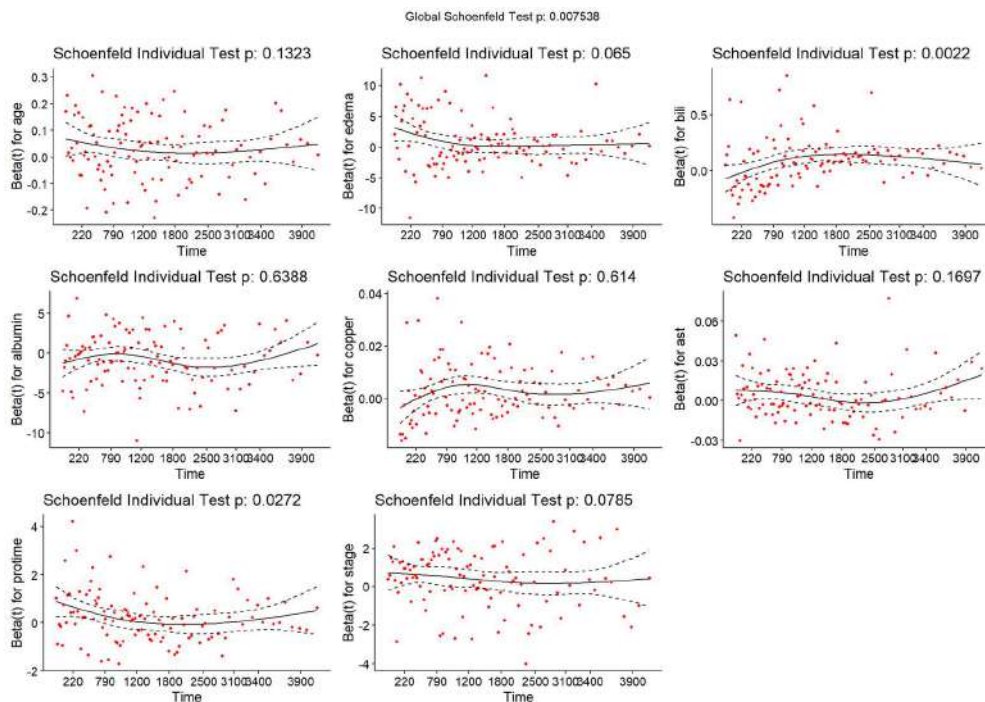
Tabel 4. 12 Pengujian Grambsch-Therneau untuk Asumsi PH

Variabel	chisq	df	p
age	2.265	1	0.1323
edema	3.404	1	0.0650
bili	9.391	1	0.0022
albumin	0.220	1	0.6388
copper	0.254	1	0.6140
ast	1.886	1	0.1697
protime	4.878	1	0.0272
stage	3.096	1	0.0785
GLOBAL	20.857	8	0.0075

Suatu variabel dikatakan memenuhi asumsi proportional hazard apabila nilai p-value lebih besar dari tingkat signifikansi tertentu. Berdasarkan hasil Grambsch-Therneau Test, dengan menggunakan taraf signifikansi 0.05 didapatkan variabel bili dan protime tidak memenuhi asumsi proportional hazard, yang berarti hubungan antara variabel-variabel ini dengan risiko kejadian berubah seiring waktu. Sebaliknya, variabel age, edema, albumin, copper, ast, dan stage memenuhi asumsi tersebut. Pengujian global menunjukkan bahwa model secara keseluruhan melanggar asumsi proportional hazard. Dengan demikian, pelanggaran pada variabel tertentu, terutama bili dan protime, mengindikasikan perlunya penyesuaian model, misalnya dengan menambahkan komponen waktu (time-dependent covariates) untuk meningkatkan keandalan model.

4.6.4 Schoenfeld Residual

Pada bagian ini, dilakukan pengujian asumsi Proportional Hazard (PH) menggunakan residual Schoenfeld. Grafik residual ini membantu memvisualisasikan stabilitas koefisien regresi Cox terhadap waktu. Jika asumsi PH terpenuhi, maka residual Schoenfeld akan tersebar secara acak di sekitar garis horizontal (beta tetap konstan) dan p-value dari uji individual maupun global akan lebih besar dari tingkat signifikansi tertentu (misalnya 0.05). Berikut ini adalah grafik Schoenfeld Residual untuk setiap variabel yang digunakan dalam model.

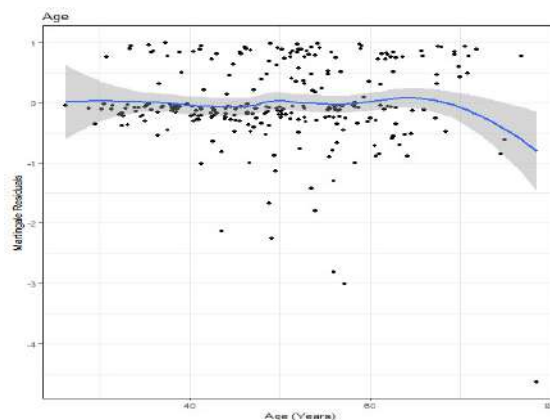


Gambar 4. 17 Kurva Schoenfeld Test

Pengujian asumsi Proportional Hazard (PH) menggunakan residual Schoenfeld menunjukkan bahwa secara global asumsi PH ditolak dengan p-value sebesar 0.0075, yang mengindikasikan adanya pelanggaran pada model. Secara individual, variabel bilirubin memiliki p-value sebesar 0.0022 dan protime sebesar 0.0272, keduanya signifikan di bawah taraf 0.05, yang berarti koefisien regresi untuk kedua variabel ini berubah seiring waktu dan tidak memenuhi asumsi PH. Hal ini terlihat pada grafik residual Schoenfeld, di mana garis untuk kedua variabel tersebut menunjukkan pola tren yang tidak horizontal. Sementara itu, variabel lainnya seperti age, edema, albumin, copper, ast, dan stage memiliki p-value lebih besar dari 0.05, sehingga koefisiennya dapat dianggap stabil terhadap waktu dan memenuhi asumsi PH. Dengan demikian, pelanggaran asumsi PH pada variabel bilirubin dan protime memerlukan penyesuaian model, seperti menambahkan komponen time-dependent covariates agar hasil analisis menjadi lebih akurat.

4.6.5 Martingale Residual

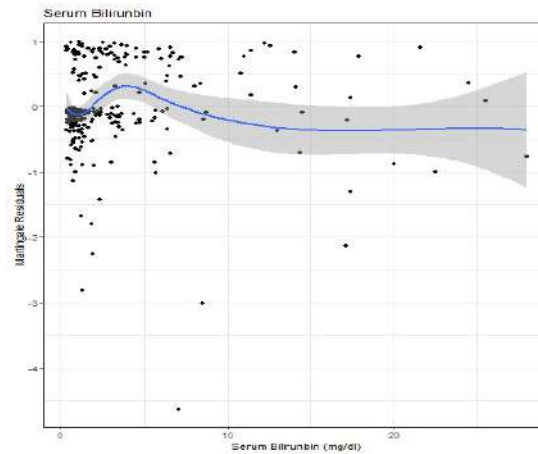
Martingale residual digunakan untuk mengevaluasi kesesuaian model regresi cox, terutama dalam memeriksa hubungan antara variabel kontinu dengan risiko kejadian. Residual ini membantu mengidentifikasi pola atau deviasi dari asumsi yang diharapkan dalam model. Plot residual Martingale terhadap variabel kontinu dapat menunjukkan apakah hubungan antara variabel tersebut dengan log hazard bersifat linear atau tidak. Jika pola residual menunjukkan tren non-linear, maka transformasi variabel atau pendekatan pemodelan lain mungkin diperlukan. Analisis ini dilakukan menggunakan perangkat lunak RStudio, dimana residual Martingale di-plot terhadap beberapa variabel, yakni Age Variable, Serum Bilirubin Variable, Serum Albumin Variable, Urine Copper Variable, Aspartate Aminotransferase Variable, dan Standardised Blood Clotting Time Variable. Berikut adalah hasil plot Martingale residual untuk masing-masing variabel.



Gambar 4. 18 Kurva Schoenfeld Test

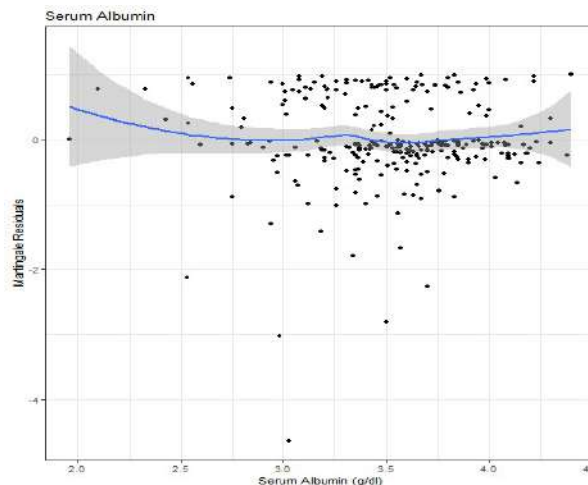
Plot Martingale residuals terhadap usia menunjukkan pola non-linear, terutama pada usia di atas 60 tahun di mana garis regresi non-parametrik menurun tajam, mengindikasikan residual cenderung lebih negatif dan asumsi hubungan linier antara usia dan outcome mungkin tidak terpenuhi. Sebagian besar residual berada di sekitar nol untuk usia antara 40 hingga 60 tahun, menunjukkan kecocokan model yang cukup baik di rentang ini, namun penyebaran residual yang lebih negatif pada usia tua menunjukkan perlunya transformasi variabel usia, seperti

polinomial atau splines, untuk menangkap pola ini dengan lebih baik. Beberapa titik dengan residual ekstrem (sekitar -3 hingga -4) berpotensi menjadi outlier dan memerlukan evaluasi lebih lanjut untuk memastikan keakuratan model.



Gambar 4. 19 Martingale Residual pada Variable Serum Bilirubin

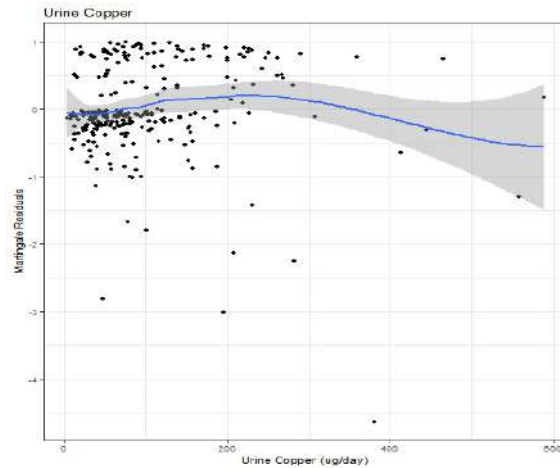
Plot Martingale residuals terhadap kadar serum bilirubin menunjukkan pola non-linear, di mana garis regresi non-parametrik awalnya naik tajam pada kadar bilirubin rendah lalu menurun dan stabil di kadar yang lebih tinggi, mengindikasikan kemungkinan hubungan non-linier antara serum bilirubin dan outcome. Penyebaran residu lebih terlihat pada kadar bilirubin rendah, dengan beberapa titik ekstrem pada nilai residual negatif di bawah -2, yang menunjukkan potensi outlier atau data yang kurang cocok dengan asumsi model. Tren ini menyarankan perlunya mempertimbangkan transformasi non-linear untuk serum bilirubin agar model lebih akurat dalam menangkap hubungan antara variabel ini dan outcome.



Gambar 4. 20 Martingale Residual pada Variable Serum Albumin

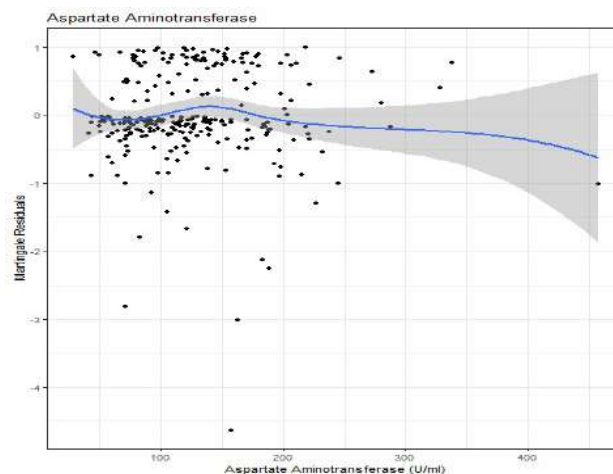
Plot Martingale residuals terhadap kadar serum albumin menunjukkan pola yang relatif mendekati linier dengan sedikit penyimpangan di sekitar nilai 3.0 g/dl, di mana garis regresi non-parametrik turun dan kemudian naik kembali. Sebagian besar residual tersebar secara

merata di sekitar nol, menunjukkan bahwa model cukup baik menangkap hubungan antara serum albumin dan outcome. Namun, terdapat beberapa titik dengan residual ekstrem di bawah -2, yang mengindikasikan potensi outlier atau data yang kurang sesuai dengan asumsi model. Meskipun pola ini tidak menunjukkan penyimpangan yang signifikan, evaluasi lebih lanjut terhadap data ekstrem dan asumsi linieritas mungkin diperlukan untuk memastikan kecocokan model yang lebih baik.



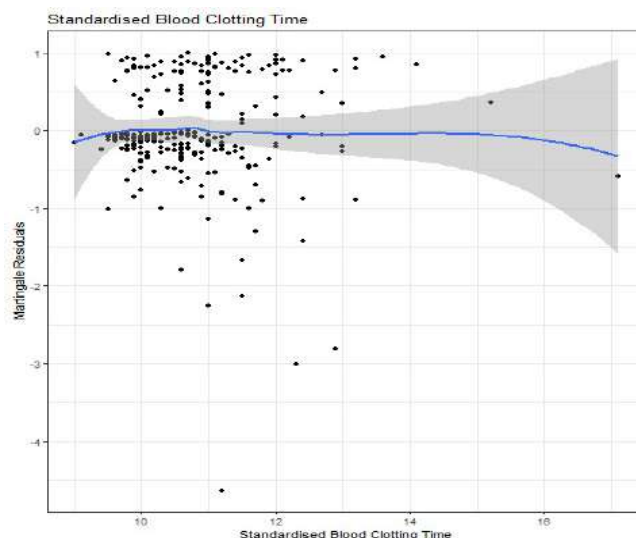
Gambar 4. 21 Martingale Residual pada Variable *Urine Copper*

Plot residual martingale terhadap variabel *Urine Copper* menunjukkan adanya indikasi hubungan non-linear antara *Urine Copper* dan model regresi Cox. Pada nilai *Urine Copper* yang rendah hingga menengah, residual cenderung stabil di sekitar nol, namun pada nilai yang lebih tinggi (di atas 300 $\mu\text{g/day}$), garis *smoothing* berwarna biru menunjukkan tren menurun. Hal ini mengisyaratkan bahwa peningkatan *Urine Copper* cenderung berhubungan dengan penyimpangan dari asumsi model yang linier, terutama pada rentang nilai yang ekstrem, sehingga perlu dipertimbangkan transformasi atau pendekatan non-linear untuk memperbaiki kesesuaian model.



Gambar 4. 22 Martingale Residual pada Variable Aspartate Aminotransferase

Plot residual martingale terhadap variabel *Aspartate Aminotransferase* (U/ml) menunjukkan adanya tren non-linear dalam hubungan antara variabel ini dan model regresi Cox. Pada nilai *Aspartate Aminotransferase* yang rendah hingga menengah, residual cenderung stabil di sekitar nol, namun garis *smoothing* berwarna biru menunjukkan penurunan yang lebih jelas seiring peningkatan nilai *Aspartate Aminotransferase* di atas 200 U/ml. Tren ini mengindikasikan potensi ketidaksesuaian asumsi linieritas dalam model, terutama pada nilai ekstrem, yang dapat menyebabkan penyimpangan. Area abu-abu yang melebar di bagian kanan plot juga menunjukkan tingkat ketidakpastian yang lebih tinggi pada nilai *Aspartate Aminotransferase* yang besar, sehingga perlu dievaluasi lebih lanjut, misalnya dengan transformasi variabel atau pendekatan non-linear, untuk meningkatkan kesesuaian model.



Gambar 4. 23 Martingale Residual pada Variable *Standardised Blood Clotting Time*

Plot residual martingale terhadap variabel *Standardised Blood Clotting Time* menunjukkan hubungan yang relatif stabil di sebagian besar rentang nilai variabel ini. Garis *smoothing* berwarna biru cenderung mendatar pada nilai *Blood Clotting Time* yang rendah hingga menengah, menunjukkan bahwa model cukup sesuai pada rentang ini. Namun, pada nilai yang lebih tinggi (di atas 15), terdapat tren penurunan residual yang mengindikasikan potensi hubungan non-linear atau efek variabel yang tidak sepenuhnya tertangkap oleh model linier. Area abu-abu yang melebar di ujung kanan menunjukkan tingkat ketidakpastian yang lebih tinggi akibat data yang lebih jarang pada nilai ekstrem. Hal ini mengisyaratkan perlunya evaluasi lebih lanjut untuk memastikan bahwa asumsi linieritas tetap terpenuhi atau mempertimbangkan pendekatan alternatif seperti transformasi variabel.

4.7 Extended Cox Model dengan Time-Dependent Covariates

Extended Cox Model dengan Time-Dependent Covariates digunakan untuk variabel yang tidak memenuhi asumsi proporsionalitas hazard (PH), yang mengindikasikan perubahan pengaruh variabel terhadap risiko kejadian seiring waktu. Untuk mengatasi hal ini, variabel yang tidak memenuhi asumsi PH dibagi menjadi dua komponen: sebelum dan setelah cut point. Dalam model ini, variabel bilirubin dan protime dibagi berdasarkan cut point, menghasilkan dua variabel baru yang kemudian dimasukkan dalam model Cox untuk menggambarkan pengaruh yang bervariasi sepanjang waktu.

Dengan menggunakan cut point 1825 hari, model Cox dengan variabel time-dependent yang terbentuk dalam bentuk hazard function adalah sebagai berikut.

$$h(t|x(t)) = h_0(t) \exp(\beta_1(age) + \beta_2(edema) + \beta_3(albumi) + \beta_4(copper) + \beta_5(ast) + \beta_6(stage) + \gamma_1(bili \times g_1(t)) + \gamma_2(bili \times g_2(t)) + \gamma_3(ptime \times g_1(t)) + \gamma_4(ptime \times g_2(t)))$$

Berdasarkan cut point 1825 hari, fungsi Heaviside didefinisikan menjadi dua komponen, yaitu $g_1(t)$ dan $g_2(t)$ yang digunakan untuk membagi waktu menjadi dua periode. Fungsi $g_1(t)$ bernilai 1 jika $t \leq 1825$ dan bernilai 0 jika $t > 1825$ yang menunjukkan periode sebelum atau pada titik waktu 1825 hari. Sebaliknya, $g_2(t)$ bernilai 1 jika $t > 1825$ dan bernilai 0 jika $t \leq 1825$ yang merepresentasikan periode setelah titik waktu 1825 hari. Dengan pemisahan ini, pengaruh variabel terhadap risiko kejadian dapat dianalisis secara terpisah untuk periode sebelum dan sesudah cut point, sehingga memungkinkan model untuk menangkap perubahan pengaruh variabel seiring berjalannya waktu.

Kemudian dilakukan permodelan Cox dengan variabel time-dependent yang digunakan untuk menangkap perubahan pengaruh kovariat terhadap risiko kejadian seiring waktu. Berikut adalah hasil pemodelan dengan cut point 1825 hari untuk variable bili dan protime yang tidak memenuhi asumsi PH.

Tabel 4. 13 Pengujian Cox Model Time-Dependent

Variabel	coef	exp(coef)	se(coef)	z	Pr(> z)
Age	0.009505	1.009551	0.009723	0.978	0.328259
Edema	0.699045	2.011830	0.326267	2.143	0.032149
Albumin	-0.463860	0.628852	0.310080	-1.496	0.134670
Copper	0.002323	1.002326	0.001089	2.134	0.032847
Ast	0.002234	1.002236	0.001936	1.154	0.248446
Stage	0.294085	1.341897	0.153294	1.918	0.055056
Billi: gt1	0.031393	1.031891	0.022165	1.416	0.156673
Billi: gt2	0.201569	1.223321	0.057399	3.512	0.000445
Protime: gt1	0.604112	1.829626	0.125442	4.816	1.47e-06
Protime: gt2	-0.061181	0.940653	0.164121	-0.373	0.709310
Concordance				0.895 (se = 0.1013)	
Likelihood Ratio Test				267.2	
Wald Test				126.9	
Score (logrank) Test				320.1	

Hasil output menunjukkan bahwa variabel protime sebelum 1825 hari (protime_bef) memiliki hazard ratio sebesar 1,83 dengan p-value signifikan yang mengindikasikan peningkatan risiko kejadian sebesar 83% sebelum titik waktu tersebut. Namun, setelah 1825 hari (protime_aft), pengaruhnya tidak signifikan. Untuk variabel bilirubin setelah 1825 hari (bili_aft), hazard ratio sebesar 1,223 dengan p-value signifikan yang menunjukkan peningkatan risiko kejadian sebesar 22,3% setelah titik waktu tersebut, sementara pengaruh bilirubin sebelum titik waktu (bili_bef) tidak signifikan. Hal ini menunjukkan bahwa bilirubin memiliki pengaruh yang lebih kuat terhadap risiko kejadian setelah 1825 hari, sementara pengaruh protime signifikan hanya pada periode sebelum 1825 hari.

Model Cox dengan time-dependent covariates menunjukkan kinerja yang sangat baik dengan nilai concordance sebesar 0.895, yang mengindikasikan kemampuan model dalam membedakan risiko individu mencapai 89,5%. Uji keseluruhan model, termasuk Likelihood ratio test, Wald test, dan Score test semuanya signifikan. Hasil ini mengonfirmasi bahwa model secara keseluruhan signifikan dan mampu menjelaskan hubungan antara variabel independen dengan waktu kejadian secara akurat.

4.8 Evaluasi Model

Pada tahap evaluasi model, kriteria informasi Akaike (AIC) digunakan untuk membandingkan performa model backward elimination dengan model Cox yang mencakup variabel time-dependent. AIC memberikan penilaian terhadap keseimbangan antara kompleksitas model dan kualitas fit terhadap data, di mana nilai AIC yang lebih rendah menunjukkan model yang lebih baik. Berikut adalah hasil perhitungan AIC untuk kedua model.

Tabel 4. 14 Pengujian AIC pada Cox Model

Model	AIC
Cox Model Backward Elimination	952.5814
Time-Dependent Cox Model	853.2173

Dari hasil perbandingan AIC, model Cox dengan variabel time-dependent memiliki nilai AIC yang lebih rendah (853.2173) dibandingkan dengan model backward elimination (952.5814). Hal ini menunjukkan bahwa model dengan variabel time-dependent memberikan fit yang lebih baik terhadap data tanpa mengorbankan efisiensi. Dengan demikian, model ini lebih optimal dalam menjelaskan hubungan antara variabel independen dan waktu kejadian.

BAB V

KESIMPULAN

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi waktu kelangsungan hidup pasien penderita *Primary Biliary Cirrhosis* (PBC) menggunakan model *Cox Proportional Hazards*. Hasil analisis menunjukkan bahwa beberapa faktor klinis, seperti ascites, hepatomegali, spiders, dan edema, memiliki pengaruh signifikan terhadap waktu kelangsungan hidup pasien, di mana pasien tanpa kondisi tersebut memiliki peluang kelangsungan hidup yang lebih tinggi dibandingkan pasien yang mengalaminya. Selain itu, faktor jenis kelamin juga memengaruhi kelangsungan hidup, dengan pasien perempuan memiliki peluang bertahan hidup lebih lama dibandingkan laki-laki. Namun, faktor perawatan, yaitu penggunaan D-penicillamine dibandingkan placebo, tidak menunjukkan perbedaan signifikan terhadap kelangsungan hidup pasien. Melalui analisis kurva Kaplan-Meier dan uji log-rank, ditemukan bahwa beberapa variabel memberikan perbedaan signifikan pada kurva survival antar kelompok pasien. Model *Cox Proportional Hazards* yang digunakan juga mampu mengidentifikasi variabel-variabel signifikan yang memengaruhi risiko kematian pasien PBC, sehingga memungkinkan prediksi waktu survival berdasarkan karakteristik individu. Penelitian ini menggunakan data klinis dari Mayo Clinic PBC dataset yang memberikan wawasan penting tentang dinamika penyakit PBC. Hasil penelitian ini dapat dimanfaatkan oleh tenaga medis untuk mengidentifikasi pasien dengan risiko tinggi dan merancang strategi perawatan yang lebih efektif, sehingga memberikan kontribusi dalam peningkatan kualitas perawatan pasien.

DAFTAR PUSTAKA

- Kleinbaum, D., & Klein, M. (2005). *Survival Analysis, A SelfLearning*. New York: Springer.
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text* (3rd ed.). New York: Springer.
- Kleinbaum, D. G., & Klein, M. (2015). *Logistic Regression: A Self-Learning Text* (3rd ed.). New York: Springer.
- Tanaka, A. (2024). *Primary Biliary Cholangitis: Pathogenesis, Diagnosis, and Management*. Tokyo: Springer.

LAMPIRAN

Lampiran. Link dataset PBC

<https://its.id/m/DatasetuncleanPBC>

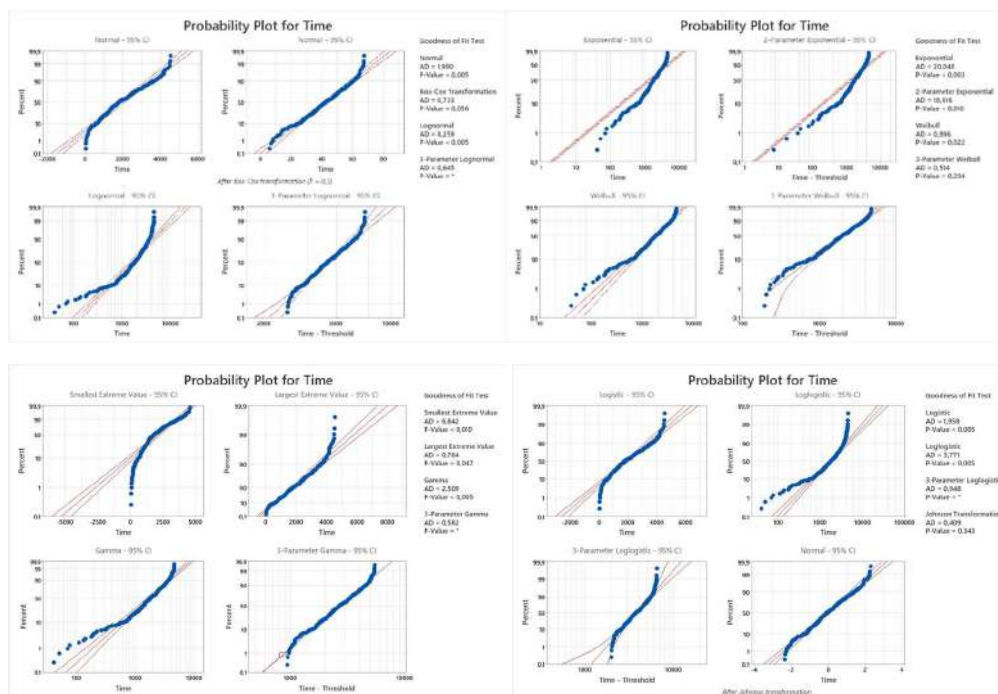
Lampiran. Link dataset cleaned

<https://its.id/m/DatasetcleannPBC>

Lampiran. Link Rstudio

https://its.id/m/syntaxR_FP

Lampiran . Output Pengujian Distribusi



Goodness of Fit Test

Distribution	AD	P	LRT	P
Normal	1,990	<0,005		
Box-Cox Transformation	0,733	0,056		
Lognormal	8,259	<0,005		
3-Parameter Lognormal	0,645	*	0,000	
Exponential	20,048	<0,003		
2-Parameter Exponential	18,816	<0,010	0,002	
Weibull	0,896	0,022		
3-Parameter Weibull	0,514	0,204	0,052	
Smallest Extreme Value	6,842	<0,010		
Largest Extreme Value	0,764	0,047		
Gamma	2,509	<0,005		
3-Parameter Gamma	0,582	*	0,000	
Logistic	1,958	<0,005		
Loglogistic	3,771	<0,005		
3-Parameter Loglogistic	0,948	*	0,000	
Johnson Transformation	0,409	0,343		

Lampiran . Output Check Missing Value (NA)

```
## Check Missing Value (NA)
colSums(is.na(pbc))

##      id      time      status      trt      age      sex      ascites      hep
ato
##      0      0      0      106      0      0      106
106
## spiders      edema      bili      chol      albumin      copper      alk.phos
ast
##      106      0      0      134      0      108      106
106
##      trig platelet      protime      stage
##      136      11      2      6
```

Lampiran . Output Statistika Deskriptif Survival Event

```
## Call: survfit(formula = y ~ 1)
##
##      n events median 0.95LCL 0.95UCL
## [1,] 276      111 3358      2847      3839
```

Lampiran .Output Log-rank test Treatment

```
## Call:
## survdiff(formula = y ~ dt$trt)
##
##      N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$trt=1 136      57      53.7      0.209      0.405
## dt$trt=2 140      54      57.3      0.195      0.405
```

```
##
## Chisq= 0.4 on 1 degrees of freedom, p= 0.5
```

Lampiran .Output Log-rank test Gender (sex)

```
## Call:
## survdiff(formula = y ~ dt$sex)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$sex=m  34         21   13.7      3.878      4.47
## dt$sex=f 242         90   97.3      0.546      4.47
##
## Chisq= 4.5 on 1 degrees of freedom, p= 0.03
```

Lampiran .Output Log-rank test Presence of Ascites

```
## Call:
## survdiff(formula = y ~ dt$ascites)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$ascites=0 257         93  108.69      2.26     110
## dt$ascites=1  19         18   2.31    106.56     110
##
## Chisq= 110 on 1 degrees of freedom, p= <2e-16
```

Lampiran .Output Log-rank test Presence of Hepatomegaly

```
## Call:
## survdiff(formula = y ~ dt$hepato)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$hepato=0 134         36   62.9      11.5     27.1
## dt$hepato=1 142         75   48.1      15.0     27.1
##
## Chisq= 27.1 on 1 degrees of freedom, p= 2e-07
```

Lampiran .Output Log-rank test Blood Vessel Malformations in the Skin

```
## Call:
## survdiff(formula = y ~ dt$spiders)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$spiders=0 196         65   87.3      5.69     27.2
## dt$spiders=1  80         46   23.7     20.96     27.2
##
## Chisq= 27.2 on 1 degrees of freedom, p= 2e-07
```

Lampiran .Output Log-rank Edema Status

```
## Call:
## survdiff(formula = y ~ dt$edema)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
```

```
## dt$edema=0    234      79   100.95      4.77    53.09
## dt$edema=0.5   25      16     8.06      7.83     8.46
## dt$edema=1     17      16     1.99     98.74   101.59
##
##  Chisq= 113  on 2 degrees of freedom, p= <2e-16
```

Lampiran .Output Log-rank Histologic Stage of Disease

```
## Call:
## survdiff(formula = y ~ dt$stage)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## dt$stage=1  12         1    7.66    5.790    6.29
## dt$stage=2  59        14   28.88    7.668   10.45
## dt$stage=3 111        41   47.53    0.896    1.57
## dt$stage=4  94        55   26.93   29.250   39.28
##
##  Chisq= 44.6  on 3 degrees of freedom, p= 1e-09
```

Lampiran .Output Cox Model

```
## Call:
## coxph(formula = y ~ trt + age + sex + ascites + hepato + spiders +
##       edema + bili + chol + albumin + copper + alk.phos + ast +
##       trig + platelet + protime + stage, data = dt)
##
## n= 276, number of events= 111
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## trt      -1.242e-01  8.832e-01  2.147e-01 -0.579  0.56290
## age       2.890e-02  1.029e+00  1.164e-02  2.482  0.01305 *
## sexf     -3.656e-01  6.938e-01  3.113e-01 -1.174  0.24022
## ascites   8.833e-02  1.092e+00  3.872e-01  0.228  0.81955
## hepato    2.552e-02  1.026e+00  2.510e-01  0.102  0.91900
## spiders   1.012e-01  1.107e+00  2.435e-01  0.416  0.67760
## edema     1.011e+00  2.749e+00  3.941e-01  2.566  0.01029 *
## bili      8.001e-02  1.083e+00  2.550e-02  3.138  0.00170 **
## chol      4.918e-04  1.000e+00  4.442e-04  1.107  0.26829
## albumin  -7.408e-01  4.767e-01  3.078e-01 -2.407  0.01608 *
## copper    2.490e-03  1.002e+00  1.170e-03  2.128  0.03337 *
## alk.phos  1.048e-06  1.000e+00  3.969e-05  0.026  0.97893
## ast       4.070e-03  1.004e+00  1.958e-03  2.078  0.03767 *
## trig     -9.758e-04  9.990e-01  1.333e-03 -0.732  0.46414
## platelet  9.019e-04  1.001e+00  1.184e-03  0.762  0.44629
## protime   2.324e-01  1.262e+00  1.061e-01  2.190  0.02850 *
## stage     4.545e-01  1.575e+00  1.754e-01  2.591  0.00958 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## trt              0.8832    1.1323    0.5798    1.3453
## age              1.0293    0.9715    1.0061    1.0531
```

```

## sexf      0.6938      1.4414      0.3769      1.2771
## ascites   1.0924      0.9155      0.5114      2.3332
## hepato    1.0259      0.9748      0.6273      1.6777
## spiders   1.1066      0.9037      0.6865      1.7835
## edema     2.7487      0.3638      1.2697      5.9505
## bili      1.0833      0.9231      1.0305      1.1388
## chol      1.0005      0.9995      0.9996      1.0014
## albumin   0.4767      2.0977      0.2608      0.8714
## copper     1.0025      0.9975      1.0002      1.0048
## alk.phos  1.0000      1.0000      0.9999      1.0001
## ast       1.0041      0.9959      1.0002      1.0079
## trig      0.9990      1.0010      0.9964      1.0016
## platelet  1.0009      0.9991      0.9986      1.0032
## protime   1.2617      0.7926      1.0247      1.5534
## stage     1.5754      0.6348      1.1170      2.2219
##
## Concordance= 0.849 (se = 0.018 )
## Likelihood ratio test= 167.7 on 17 df, p=<2e-16
## Wald test              = 174.1 on 17 df, p=<2e-16
## Score (logrank) test = 283.7 on 17 df, p=<2e-16

```

Lampiran

```

##                      Start:                      AIC=966.66
## y ~ trt + age + sex + ascites + hepato + spiders + edema + bili
+
##      chol + albumin + copper + alk.phos + ast + trig + platelet
+
##                      protime      +      stage
##
##                      Df      AIC
##      -      alk.phos      1      964.66
##      -      hepato      1      964.67
##      -      ascites      1      964.72
##      -      spiders      1      964.84
##      -      trt      1      965.00
##      -      trig      1      965.22
##      -      platelet      1      965.24
##      -      chol      1      965.82
##      -      sex      1      965.99
##      <none>      966.66
##      -      ast      1      968.69
##      -      copper      1      968.85
##      -      protime      1      968.99
##      -      albumin      1      970.35
##      -      age      1      970.84
##      -      edema      1      971.00

```

```

##          -          stage                      1          971.83
##          -          bili                      1          973.34
##
##                      Step:                      AIC=964.66
## y ~ trt + age + sex + ascites + hepato + spiders + edema + bili
+
##      chol + albumin + copper + ast + trig + platelet + protime
+
##                      stage
##
##                      Df                      AIC
##          -          hepato                      1          962.68
##          -          ascites                      1          962.72
##          -          spiders                      1          962.84
##          -          trt                      1          963.00
##          -          trig                      1          963.22
##          -          platelet                      1          963.27
##          -          chol                      1          963.82
##          -          sex                      1          963.99
##          <none>                      964.66
##          -          ast                      1          966.70
##          -          protime                      1          967.04
##          -          copper                      1          967.36
##          -          albumin                      1          968.55
##          -          age                      1          968.86
##          -          edema                      1          969.06
##          -          stage                      1          969.87
##          -          bili                      1          971.35
##
##                      Step:                      AIC=962.68
## y ~ trt + age + sex + ascites + spiders + edema + bili + chol
+
##      albumin + copper + ast + trig + platelet + protime + stage
##
##                      Df                      AIC
##          -          ascites                      1          960.72
##          -          spiders                      1          960.86
##          -          trt                      1          961.04
##          -          trig                      1          961.24
##          -          platelet                      1          961.27
##          -          chol                      1          961.84
##          -          sex                      1          962.01
##          <none>                      962.68

```

```

##      -      ast                      1      964.72
##      -      protime                  1      965.08
##      -      copper                    1      965.40
##      -      albumin                  1      966.69
##      -      age                      1      966.88
##      -      edema                    1      967.07
##      -      bili                     1      969.47
##      -      stage                    1      970.09
##
##                               Step:                               AIC=960.72
## y ~ trt + age + sex + spiders + edema + bili + chol + albumin
+
##      copper + ast + trig + platelet + protime + stage
##
##                               Df                               AIC
##      -      spiders                      1      958.91
##      -      trt                          1      959.08
##      -      trig                         1      959.24
##      -      platelet                     1      959.28
##      -      chol                         1      959.86
##      -      sex                         1      960.10
##      <none>                             960.72
##      -      ast                         1      962.72
##      -      protime                     1      963.32
##      -      copper                      1      963.88
##      -      edema                       1      965.12
##      -      age                         1      965.14
##      -      albumin                     1      965.83
##      -      bili                        1      967.65
##      -      stage                       1      968.30
##
##                               Step:                               AIC=958.91
## y ~ trt + age + sex + edema + bili + chol + albumin + copper +
##      ast + trig + platelet + protime + stage
##
##                               Df                               AIC
##      -      trt                          1      957.27
##      -      trig                         1      957.38
##      -      platelet                     1      957.39
##      -      chol                         1      958.16
##      -      sex                         1      958.18
##      <none>                             958.91
##      -      ast                         1      960.84

```

```

##          -          protime          1          961.59
##          -          copper           1          962.56
##          -          age             1          963.33
##          -          edema           1          963.65
##          -          albumin          1          963.90
##          -          bili             1          966.22
##          -          stage            1          967.87
##
##                      Step:                      AIC=957.27
## y ~ age + sex + edema + bili + chol + albumin + copper + ast +
##          trig + platelet + protime + stage
##
##                      Df                      AIC
##          -          trig              1          955.55
##          -          platelet           1          955.66
##          -          sex                1          956.51
##          -          chol               1          956.66
##          <none>                      957.27
##          -          ast                1          958.96
##          -          protime            1          959.83
##          -          copper             1          961.03
##          -          albumin            1          962.44
##          -          edema              1          962.49
##          -          age                1          962.55
##          -          bili               1          964.23
##          -          stage              1          965.91
##
##                      Step:                      AIC=955.55
## y ~ age + sex + edema + bili + chol + albumin + copper + ast +
##          platelet + protime + stage
##
##                      Df                      AIC
##          -          platelet           1          953.85
##          -          sex                1          954.72
##          -          chol               1          954.97
##          <none>                      955.55
##          -          ast                1          957.42
##          -          protime            1          958.28
##          -          copper             1          959.12
##          -          albumin            1          960.46
##          -          edema              1          961.26
##          -          age                1          961.46
##          -          bili               1          962.82

```

```

##          -          stage                      1          963.95
##
##          Step:                                AIC=953.85
## y ~ age + sex + edema + bili + chol + albumin + copper + ast +
##                               protime          +          stage
##
##                               Df                      AIC
##          -          sex                      1          952.81
##          -          chol                      1          953.62
##          <none>                                953.85
##          -          ast                      1          955.43
##          -          protime                  1          956.44
##          -          copper                  1          957.68
##          -          albumin                 1          958.55
##          -          edema                   1          959.30
##          -          age                     1          959.56
##          -          bili                    1          961.60
##          -          stage                   1          961.99
##
##          Step:                                AIC=952.81
## y ~ age + edema + bili + chol + albumin + copper + ast + protime
+
##                               stage
##
##                               Df                      AIC
##          -          chol                      1          952.58
##          <none>                                952.81
##          -          ast                      1          954.90
##          -          protime                  1          955.42
##          -          albumin                 1          956.85
##          -          edema                   1          957.65
##          -          copper                  1          959.37
##          -          bili                    1          960.07
##          -          stage                   1          960.85
##          -          age                     1          961.50
##
##          Step:                                AIC=952.58
## y ~ age + edema + bili + albumin + copper + ast + protime +
stage
##
##                               Df                      AIC
##          <none>                                952.58
##          -          protime                  1          955.06

```



```
##      -      ast                      1      955.79
##      -      edema                    1      955.95
##      -      albumin                  1      957.27
##      -      copper                    1      958.18
##      -      age                      1      959.97
##      -      stage                    1      960.11
## - bili      1 966.57
```

Lampiran

```
## Call:
## coxph(formula = y ~ age + edema + bili + albumin + copper + ast +
##       protime + stage, data = dt)
##
##      n= 276, number of events= 111
##
##              coef exp(coef)    se(coef)      z Pr(>|z|)
## age      0.0313836  1.0318812  0.0102036  3.076  0.00210 **
## edema     0.8217952  2.2745795  0.3471465  2.367  0.01792 *
## bili     0.0851214  1.0888492  0.0193352  4.402 1.07e-05 ***
## albumin -0.7185954  0.4874364  0.2724486 -2.638  0.00835 **
## copper    0.0028535  1.0028576  0.0009832  2.902  0.00370 **
## ast       0.0043769  1.0043865  0.0018067  2.423  0.01541 *
## protime   0.2275175  1.2554794  0.1013729  2.244  0.02481 *
## stage     0.4327939  1.5415584  0.1456307  2.972  0.00296 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## age      1.0319      0.9691      1.0114      1.0527
## edema     2.2746      0.4396      1.1519      4.4915
## bili     1.0888      0.9184      1.0484      1.1309
## albumin   0.4874      2.0515      0.2858      0.8314
## copper    1.0029      0.9972      1.0009      1.0048
## ast       1.0044      0.9956      1.0008      1.0079
## protime   1.2555      0.7965      1.0292      1.5314
## stage     1.5416      0.6487      1.1588      2.0508
##
## Concordance= 0.845 (se = 0.019 )
## Likelihood ratio test= 163.8 on 8 df,  p=<2e-16
## Wald test              = 176.1 on 8 df,  p=<2e-16
## Score (logrank) test = 257.5 on 8 df,  p=<2e-16
```

Lampiran. Grambsch-Therneau Test or Goodness of Fit Test

```
gt_test
##      chisq df      p
## age      2.265  1 0.1323
## edema     3.404  1 0.0650
## bili      9.391  1 0.0022
```

```
## albumin 0.220 1 0.6388
## copper 0.254 1 0.6140
## ast 1.886 1 0.1697
## protime 4.878 1 0.0272
## stage 3.096 1 0.0785
## GLOBAL 20.857 8 0.0075
```

Lampiran. Cox Model Time-dependent

```
## Call:
## coxph(formula = y ~ age + edema + albumin + copper + ast + stage +
##       bili_bef + bili_aft + protime_bef + protime_aft, data = dt)
##
## n= 276, number of events= 111
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## age           0.009505  1.009551  0.009723  0.978 0.328259
## edema          0.699045  2.011830  0.326267  2.143 0.032149 *
## albumin       -0.463860  0.628852  0.310080 -1.496 0.134670
## copper          0.002323  1.002326  0.001089  2.134 0.032847 *
## ast            0.002234  1.002236  0.001936  1.154 0.248446
## stage          0.294085  1.341897  0.153294  1.918 0.055056 .
## bili_bef        0.031393  1.031891  0.022165  1.416 0.156673
## bili_aft        0.201569  1.223321  0.057399  3.512 0.000445 ***
## protime_bef     0.604112  1.829626  0.125442  4.816 1.47e-06 ***
## protime_aft    -0.061181  0.940653  0.164121 -0.373 0.709310
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## age           1.0096      0.9905      0.9905      1.029
## edema          2.0118      0.4971      1.0614      3.813
## albumin        0.6289      1.5902      0.3425      1.155
## copper          1.0023      0.9977      1.0002      1.004
## ast            1.0022      0.9978      0.9984      1.006
## stage          1.3419      0.7452      0.9937      1.812
## bili_bef        1.0319      0.9691      0.9880      1.078
## bili_aft        1.2233      0.8174      1.0932      1.369
## protime_bef     1.8296      0.5466      1.4308      2.340
## protime_aft     0.9407      1.0631      0.6819      1.298
##
## Concordance= 0.895 (se = 0.013 )
## Likelihood ratio test= 267.2 on 10 df,  p=<2e-16
## Wald test               = 126.9 on 10 df,  p=<2e-16
## Score (logrank) test = 320.1 on 10 df,  p=<2e-16
```

Lampiran. Grambsch-Therneau Test Cox time-dependent

```
ph_test
##               chisq df      p
## age           0.7041  1 0.401
```

```
## edema      2.4506  1 0.117
## albumin    0.2318  1 0.630
## copper      0.0288  1 0.865
## ast        0.3045  1 0.581
## stage      0.1798  1 0.672
## bili_bef   2.2579  1 0.133
## bili_aft   5.4691  1 0.019
## protime_bef 0.3112  1 0.577
## protime_aft 1.3575  1 0.244
## GLOBAL     15.9395 10 0.101
```

Lampiran. Model Comparison Cox Model Backward vs Cox Model Time-dependent

```
AIC(backward)
```

```
## [1] 952.5814
```

```
AIC(timedep_cox)
```

```
## [1] 853.2173
```