



TUGAS AKHIR - KS184822

***DIAGNOSTIC CHECKING RESIDUAL MODEL REGRESI  
COX PROPORTIONAL HAZARD PADA DATA PASIEN  
KANKER TIROID DI RUMAH SAKIT XYZ  
MENGUNAKAN COX-SNELL RESIDUAL DAN  
MARTINGALE RESIDUAL***

**RISKA DEVY APRILLYASARI  
NRP 062115 4000 0043**

**Dosen Pembimbing  
Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**TUGAS AKHIR - KS184822**

***DIAGNOSTIC CHECKING RESIDUAL MODEL REGRESI  
COX PROPORTIONAL HAZARD PADA DATA PASIEN  
KANKER TIROID DI RUMAH SAKIT XYZ  
MENGUNAKAN COX-SNELL RESIDUAL DAN  
MARTINGALE RESIDUAL***

**RISKA DEVY APRILLYASARI  
NRP 062115 4000 0043**

**Dosen Pembimbing  
Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**

*(Halaman ini sengaja dikosongkan)*



**FINAL PROJECT - KS184822**

**DIAGNOSTIC CHECKING RESIDUAL COX  
PROPORTIONAL HAZARD MODEL OF THYROID  
CANCER PATIENTS IN XYZ HOSPITAL USING COX-  
SNELL RESIDUAL AND MARTINGALE RESIDUAL**

**RISKA DEVY APRILLYASARI  
NRP 062115 4000 0043**

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

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**

*(Halaman ini sengaja dikosongkan)*

**LEMBAR PENGESAHAN**

**DIAGNOSTIC CHECKING RESIDUAL MODEL  
REGRESI COX PROPORTIONAL HAZARD PADA  
DATA PASIEN KANKER TIROID DI RUMAH SAKIT  
XYZ MENGGUNAKAN COX-SNELL RESIDUAL DAN  
MARTINGALE RESIDUAL**

**TUGAS AKHIR**

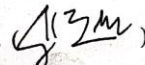
Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Statistika  
pada

Program Studi Sarjana Departemen Statistika  
Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember

Oleh :

**Riska Devy Aprillyasari**  
NRP. 062115 4000 0043

Disetujui oleh Pembimbing :

**Jerry Dwi Trijoyo Purnomo, S.Si, M.Si., Ph.D.**   
NIP. 19810223 200812 1 003



**Dr. Suhartono**  
NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

*(Halaman ini sengaja dikosongkan)*



**DIAGNOSTIC CHECKING RESIDUAL MODEL REGRESI  
COX PROPORTIONAL HAZARD PADA DATA PASIEN  
KANKER TIROID DI RUMAH SAKIT XYZ  
MENGUNAKAN COX-SNELL RESIDUAL DAN  
MARTINGALE RESIDUAL**

**Nama Mahasiswa** : Riska Devy Aprillyasari  
**NRP** : 062115 4000 0043  
**Departemen** : Statistika  
**Dosen Pembimbing** : Jerry Dwi Trijoyo Purnomo, S.Si.,  
M.Si., Ph.D.

**Abstrak**

*Kanker tiroid merupakan penyakit keganasan yang sering ditemukan pada kelenjar tiroid. Jumlah penderita kanker tiroid termasuk dalam 10 besar penyakit kanker dengan jumlah penderita terbanyak di Indonesia tahun 2018 berdasarkan data Globocan dengan kematian sebesar 19%, persentase yang lebih tinggi dibandingkan dunia yaitu sebesar 7,5%. Salah satu penyebab utama terjadinya kematian pada penderita kanker tiroid ialah metastasis. Metastasis adalah penyebaran sel kanker dari suatu organ ke organ yang lainnya. Regresi Cox proportional hazard merupakan metode survival yang paling umum digunakan. Cox-Snell residual dan Martingale residual merupakan alat yang digunakan dalam penelitian ini sebagai diagnostic checking untuk mengetahui linieritas variabel-variabel yang diduga berpengaruh terhadap hazard ratio pada model regresi Cox proportional hazard data pasien kanker tiroid di Rumah Sakit XYZ dari Januari 2008 hingga Februari 2019. Analisis regresi Cox Proportional Hazard menghasilkan model terbaik, yaitu survival time pasien kanker tiroid hingga mengalami metastasi di Rumah Sakit XYZ dipengaruhi secara signifikan oleh variabel jenis kelamin.*

**Kata Kunci** : *Cox-Snell residual, kanker tiroid, Martingale residual, metastasis, regresi Cox proportional hazard.*

*(Halaman ini sengaja dikosongkan)*

**DIAGNOSTIC CHECKING RESIDUAL COX  
PROPORTIONAL HAZARD MODEL OF THYROID  
CANCER PATIENTS IN XYZ HOSPITAL USING COX-  
SNELL RESIDUAL AND MARTINGALE RESIDUAL**

**Name** : Riska Devy Aprillyasari  
**Student Number** : 062115 4000 0043  
**Departement** : Statistics  
**Supervisor** : Jerry Dwi Trijoyo Purnomo, S.Si., M.Si.,  
Ph.D

**Abstract**

*Thyroid cancer is a malignant disease that is often found in the thyroid gland. The number of patients with thyroid cancer is included in the top 10 cancers with the highest number of sufferers in Indonesia in 2018 based on Globocan data with deaths of 19%, that is higher percentage than the world that is 7.5%. One of the main causes of death in patients with thyroid cancer is metastasis. Metastasis is the spread of cancer cells from one organ to another. Cox proportional hazard regression is the most commonly used survival method. Cox-Snell residual and Martingale residual is a tool used in this study as a diagnostic checking to determine the linearity of variables thought to affect hazard ratios in the Cox proportional hazard regression model in thyroid cancer patient datas at XYZ Hospital from January 2008 to February 2019. Cox proportional hazard regression analysis produces the best model, namely the survival time of thyroid cancer patients to experience metastasis in XYZ Hospital is significantly affected by sex.*

**Keywords** : *Cox proportional hazard regression, Cox-Snell residual, Martingale residual, metastasis, thyroid cancer.*

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan oleh Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “***Diagnostic Checking Residual Model Cox Proportional Hazard pada Data Pasien Kanker Tiroid di Rumah Sakit XYZ Menggunakan Cox-Snell Residual dan Martingale Residual***” dengan baik dan lancar, dan tepat waktu.

Penulis menyadari bahwa dalam menyelesaikan Tugas Akhir ini penulis telah banyak menerima bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada :

1. Kedua orang tua yang sangat berharga dan penulis sayangi, atas setiap do'a, nasihat, dan dukungan yang senantiasa diberikan, serta menjadi alasan utama penulis untuk menyelesaikan Tugas Akhir ini.
2. Bapak Jerry Dwi Trijoyo Purnomo, S.Si., M.Si., Ph.D. selaku pembimbing, yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, saran, serta dukungan kepada penulis selama proses penulisan Tugas Akhir ini.
3. Bapak Dr. Purnadi, M.Sc. dan Bapak Dr. Bambang Widjanarko O., S.Si., M.Si. selaku dosen penguji, yang telah memberikan saran dan motivasi demi kebaikan Tugas Akhir ini.
4. Bapak Dr. Suhartono selaku Kepala Departemen dan Ibu Dr. Santi Wulan Purnami, S.Si., M.Si. selaku Ketua Program Studi Sarjana yang telah memberikan saran dan arahan dalam proses belajar di Departemen Statistika ITS.
5. Bapak Prof. Drs. Nur Iriawan, MIKomp, Ph.D. selaku dosen wali penulis yang selalu memberikan pengetahuan baru, saran serta pengajaran selama berada di Institut Teknologi Sepuluh Nopember.
6. Seluruh teman-teman  $\Sigma 26$ , Statistika ITS angkatan 2015. Serta para sahabat yang berjuang bersama-sama di lab

Lingkes dan se-dosen pembimbing Sekar, Yani, Novita, dan Anindya. Dan juga teman-teman karib yang baik hatinya dinda, ani, nabila yang selama ini telah membantu, mendukung, dan mendengarkan keluh kesah penulis selama masa perkuliahan berlangsung.

7. Semua teman, relasi dan berbagai pihak yang tidak bisa penulis sebutkan namanya satu persatu yang telah membantu dalam penulisan laporan ini.

Penulis berharap Laporan Tugas Akhir ini dapat bermanfaat bagi masyarakat dan bagi ilmu pengetahuan. Kritik dan saran yang bersifat membangun akan sangat berguna bagi penelitian penulis yang lebih baik kedepannya, karena kesempurnaan hanya milik Allah SWT.

Surabaya, Mei 2019

Penulis

# DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>COVER PAGE</b> .....	iii
<b>LEMBAR PENGESAHAN</b> ..Error! Bookmark not defined.	
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR LAMPIRAN</b> .....	xix
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan .....	5
1.4 Manfaat .....	5
1.5 Batasan Masalah .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	7
2.1 Analisis <i>Survival</i> .....	7
2.2 Dasar Teori Analisis <i>Survival</i> .....	9
2.3 Kurva <i>Kaplan-Meier</i> .....	11
2.4 Uji <i>Log Rank</i> .....	13
2.5 Regresi Cox <i>Proportional Hazard</i> .....	14
2.5.1 Estimasi Parameter .....	14
2.5.2 Pengujian Signifikansi Parameter.....	15
2.6 Pemilihan Model Terbaik .....	16
2.7 <i>Hazard Ratio</i> .....	17
2.8 Asumsi <i>Proportional Hazard</i> .....	18
2.9 <i>Cox-Snell Residual</i> .....	19
2.10 <i>Martingale Residual</i> .....	19
2.11 Kanker Tiroid.....	21
2.12 Faktor Resiko Metastasis .....	21

<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>23</b>
3.1 Sumber Data.....	23
3.2 Kerangka Konsep.....	23
3.3 Variabel Penelitian .....	24
3.4 Struktur Data .....	25
3.5 Langkah Analisis.....	25
3.6 Diagram Alir .....	26
<b>BAB IV ANALISIS DAN PEMBAHASAN.....</b>	<b>29</b>
4.1 Karakteristik Data .....	29
4.2 Kurva <i>Kaplan Meier</i> dan Uji <i>Log-Rank</i> .....	33
4.2.1 Jenis Kelamin.....	33
4.2.2 Stadium.....	34
4.2.3 Operasi.....	35
4.2.4 Radioterapi.....	36
4.2.5 Terapi Hormonal.....	37
4.3 Regresi <i>Cox Proportional Hazard</i> .....	39
4.3.1 Model Terbaik.....	40
4.3.2 Uji Asumsi <i>Proportional Hazard</i> .....	41
4.4 <i>Cox-Snell Residual</i> .....	42
4.5 <i>Martingale Residual</i> .....	43
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>45</b>
5.1 Kesimpulan .....	45
5.2 Saran .....	46
<b>DAFTAR PUSTAKA .....</b>	<b>47</b>
<b>LAMPIRAN .....</b>	<b>49</b>
<b>BIODATA PENULIS .....</b>	<b>68</b>



## DAFTAR TABEL

	Halaman
<b>Tabel 3. 1</b> Variabel Penelitian .....	24
<b>Tabel 3. 2</b> Struktur Data Penelitian.....	25
<b>Tabel 4. 1</b> Karakteristik Data Penelitian.....	29
<b>Tabel 4. 2</b> Statistika Deskriptif Variabel Usia .....	30
<b>Tabel 4. 3</b> Hasil Uji <i>Log-Rank</i> Faktor Jenis Kelamin.....	34
<b>Tabel 4. 4</b> Hasil Uji <i>Log-Rank</i> Faktor Stadium .....	35
<b>Tabel 4. 5</b> Hasil Uji <i>Log-Rank</i> Faktor Operasi .....	36
<b>Tabel 4. 6</b> Hasil Uji <i>Log-Rank</i> Faktor Radioterapi.....	37
<b>Tabel 4. 7</b> Hasil Uji <i>Log-Rank</i> Faktor Terapi Hormonal ...	38
<b>Tabel 4. 8</b> Hasil Uji Serentak Regresi <i>Cox Proportional Hazard</i> .....	39
<b>Tabel 4. 9</b> Hasil Uji Parsial Regresi <i>Cox Proportional Hazard</i> .....	39
<b>Tabel 4. 10</b> Model Terbaik Berdasarkan Backward .....	40
<b>Tabel 4. 11</b> Hasil Uji Serentak Model Terbaik.....	40
<b>Tabel 4. 12</b> Hasil Uji Parsial Model Terbaik.....	41
<b>Tabel 4. 13</b> Uji Asumsi <i>Proportional Hazard</i> pada Model Terbaik.....	41

*(Halaman ini sengaja dikosongkan)*

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Ilustrasi Penyensoran Tipe Data .....	8
<b>Gambar 2.2</b> Ilustrasi Kurva <i>Kaplan Meier</i> .....	12
<b>Gambar 3.1</b> Kerangka Konsep Penelitian.....	23
<b>Gambar 3.2</b> Diagram Alir Penelitian .....	26
<b>Gambar 4.1</b> <i>Pie-Chart</i> Faktor Jenis Kelamin .....	30
<b>Gambar 4.2</b> <i>Pie-Chart</i> Faktor Stadium.....	31
<b>Gambar 4.3</b> <i>Pie-Chart</i> Faktor Operasi.....	31
<b>Gambar 4.4</b> <i>Pie-Chart</i> Faktor Radioterapi.....	32
<b>Gambar 4.5</b> <i>Pie-Chart</i> Faktor Terapi Hormonal.....	32
<b>Gambar 4.6</b> Kurva <i>Kaplan Meier</i> Faktor Jenis Kelamin .	33
<b>Gambar 4.7</b> Kurva <i>Kaplan Meier</i> Faktor Stadium.....	34
<b>Gambar 4.8</b> Kurva <i>Kaplan Meier</i> Faktor Operasi .....	35
<b>Gambar 4.9</b> Kurva <i>Kaplan Meier</i> Faktor Radioterapi .....	36
<b>Gambar 4.10</b> Kurva <i>Kaplan Meier</i> Faktor Terapi Hormonal .....	38
<b>Gambar 4.11</b> Grafik <i>Cox-Snell Residual</i> Variabel Jenis Kelamin .....	42
<b>Gambar 4.12</b> Grafik <i>Martingale Residual</i> Variabel Jenis Kelamin .....	43

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1</b> Data Penelitian .....	49
<b>Lampiran 2</b> Syntax R Kurva <i>Kaplan Meier</i> dan Uji <i>Log-Rank</i> .....	50
<b>Lampiran 3</b> Hasil Kurva <i>Kaplan Meier</i> dan Uji <i>Log-Rank</i> .....	52
<b>Lampiran 4</b> Syntax Regresi <i>Cox Proportional</i> <i>Hazard</i> .....	55
<b>Lampiran 5</b> Hasil Syntax R Regresi <i>Cox Proportional</i> <i>Hazard</i> .....	56
<b>Lampiran 6</b> Syntax R Pengujian Asumsi <i>Proportional</i> <i>Hazard</i> .....	59
<b>Lampiran 7</b> Hasil Uji Asumsi <i>Proportional Hazard</i> .....	62

*(Halaman ini sengaja dikosongkan)*

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Kanker merupakan pertumbuhan sel yang tidak terkontrol, mempunyai kemampuan untuk menginvasi dan bermetastasi (Murray, Granner, Mayes, dan Rodwell, 1997). Menurut American Cancer Society (2010) kanker termasuk penyakit ganas di dunia yang memiliki angka kematian tinggi dengan 13% kematian atau sekitar 7,4 juta penderita meninggal pada tahun 2004 di seluruh dunia. Lebih dari 70% kematian akibat kanker terjadi pada negara miskin dan berkembang. Prosentase ini diperkirakan akan terus meningkat hingga mencapai 12 juta kematian pada tahun 2030.

Kanker tiroid merupakan gangguan kelenjar tiroid yang berada pada bagian depan leher sedikit di bawah laring berbentuk kupu-kupu. Kelenjar tiroid merupakan bagian tubuh yang jarang mengalami keganasan, terjadi 0,85% dan 2,5% dari seluruh keganasan yang ada pada kelenjar tiroid. Kelenjar tiroid merupakan salah satu organ yang penting bagi tubuh manusia karena memiliki fungsi untuk menghasilkan hormon tiroksin yang mengatur tekanan darah, laju metabolisme dalam tubuh, mengatur denyut jantung, suhu tubuh, berat badan dan mengontrol tingkat kalsium dalam tubuh (Laseduw, 2012).

Kanker tiroid termasuk dalam 10 peringkat dengan jumlah insiden terbanyak di dunia pada tahun 2018 yaitu menempati urutan ke-9 dengan jumlah insiden kanker terbanyak di seluruh dunia mencapai 567.233 insiden dengan tingkat kematian sebesar 7,5% atau sebanyak 41.071 jiwa yang meninggal. (International Agency for Research on Cancer , 2018) Berdasarkan data Globocan pada tahun 2018, jumlah insiden penderita kanker tiroid di Indonesia masuk kedalam 10 besar penyakit kanker dengan pengidap paling banyak dengan jumlah insiden yang mencapai 11.470 jiwa dengan tingkat mortalitasnya mencapai 2119 jiwa atau sebesar 19%. Ditinjau dari tingkat mortalitas pengidap kanker tiroid dunia jika dibandingkan dengan Indonesia, dapat diketahui

bahwa pengidap kanker tiroid di Indonesia memiliki tingkat kematian yang lebih tinggi yaitu sebesar 19% dan dunia hanya sebesar 7,5%. Oleh karena itu perlu dilakukan kajian tentang penyebab tingginya laju kematian penderita kanker tiroid di Indonesia untuk menekan angka morbiditas dan mortalitas penderita kanker tiroid di Indonesia. Salah satu penyebab utama kematian pada kanker tiroid ialah metastasis jauh. Arti dari metastasis itu sendiri merupakan penyebaran sel kanker dari suatu organ ke organ lainnya (organ yang dekat atau jauh). Sel kanker dapat menyebar melalui beberapa *rute* seperti aliran darah, sistem limfatik, atau menembus lapisan penutup rongga organ tubuh. Metastasis pada kanker tiroid biasanya menyebar pada paru-paru, limfa, liver dan tulang. Beberapa penyebab terjadinya metastasis pada penderita kanker tiroid menurut (Sung dkk., 2016) yang menjadi acuan faktor resiko penyebab metastasis pada laporan ini ialah usia dan *treatment* yang diberikan. Sedangkan menurut (Li dkk, 2016) faktor yang mendorong terjadinya metastasis ialah faktor jenis kelamin dan stadium. Pada variabel stadium digunakan pengecualian terhadap tingkat stadium 4 karena pada stadium 4 pasien tiroid pasti sudah mengalami metastasis sehingga dalam penelitian ini digunakan stadium 1,2, dan 3 guna mengetahui ketahanan pasien kanker tiroid pada stadium 1,2, dan 3 untuk mengalami metastasis.

Beberapa *treatment* yang umum diberikan kepada pasien kanker tiroid ialah operasi, radioterapi, dan terapi hormonal. Dalam penelitian kali ini diharapkan diperoleh jenis pengobatan terbaik yang secara signifikan dapat memperpanjang waktu *survival* seorang pasien kanker tiroid untuk mengalami metastasis.

Rumah Sakit XYZ merupakan salah satu rumah sakit yang memiliki spesialisasi dalam penanganan kanker di Indonesia. Rumah sakit XYZ sudah menjadi kepercayaan masyarakat dalam menangani kasus kanker dan menjadi pusat rujukan pasien kanker di seluruh Indonesia sampai saat ini. Salah satu kasus kanker yang ditangani di rumah sakit ini ialah kasus kanker tiroid. Dalam



penelitian kali ini data yang digunakan ialah data pasien kanker tiroid yang ada di Rumah Sakit XYZ.

Analisis survival atau analisis ketahanan hidup adalah analisis data yang berhubungan dengan waktu, mulai dari awal sampai terjadinya suatu peristiwa khusus (Collet, 2003). Jangka waktu dari awal dilakukan pengamatan pada suatu individu (*time origin*) sampai terjadinya suatu peristiwa khusus (*end point* atau *failure event*) disebut dengan waktu survival. Peristiwa khusus (*failure event*) tersebut dapat berupa kegagalan, kematian, kambuh atau sembuhnya dari suatu penyakit, respon dari suatu percobaan, atau peristiwa lain yang dipilih sesuai dengan kepentingan peneliti. Dalam penelitian ini, telah ditentukan *event* yang dianggap sukses ialah jika terjadi metastasis pada pasien yang terjangkit kanker tiroid.

Salah satu metode yang ada di analisis survival ialah regresi Cox *proportional hazard*. Salah satu tujuan model Cox *proportional hazard* adalah untuk memodelkan hubungan antara waktu survival dengan variabel-variabel yang diduga mempengaruhi waktu survival. Model Cox *proportional hazard* memiliki asumsi bahwa fungsi *hazard* dari individu yang berbeda adalah proporsional, atau rasio fungsi *hazard* dari dua individu yang berlainan adalah konstan (Lee dan Wang, 2003). Masalah yang dapat dipecahkan dengan metode ini adalah mengetahui faktor-faktor yang berpengaruh signifikan pada waktu ketahanan penderita kanker tiroid untuk mengalami metastasis di Rumah Sakit XYZ.

Pada penelitian ini, setelah melakukan analisis regresi Cox *proportional hazard* selanjutnya perlu dilakukan *diagnostic checking* terhadap residual model regresi Cox *proportional hazard* yang telah didapatkan karena model regresi Cox memiliki asumsi yang kuat bahwa *hazard ratio* linier terhadap prediktor. Dengan melakukan diagnostik ini, akan diketahui apakah model regresi Cox *proportional hazard* sudah sesuai, atau sebaliknya telah terjadi pelanggaran asumsi dari *hazard ratio*.

Berbagai metode *diagnostic checking* residual untuk model regresi *Cox proportional hazard*. Ada beberapa metode *diagnostic checking* yang dapat digunakan, antara lain *Cox-Snell residual* dan *Martingale residual*. *Cox-Snell residual* didasarkan pada penyesuaian kurva *Kaplan Meier* pada model regresi *Cox proportional hazard* lalu membandingkannya dengan distribusi eksponensial. *Martingale residual* menunjukkan ketidaksesuaian antara nilai yang diobservasi dengan nilai yang diduga dari indikator kegagalan pasien atau terjadinya *event*. Residual bernilai positif berarti pasien mengalami *event*, sedangkan bernilai negatif berarti pasien tidak mengalami *event* (Breheny, 2013). Metode *Cox-snell residual* dan *Martingale residual* akan dikaji secara mendalam pada penelitian ini.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, diharapkan adanya informasi mengenai faktor-faktor yang dapat mempengaruhi waktu *survival* bagi pasien kanker tiroid untuk mengalami metastasis, mengingat metastasis merupakan salah satu penyebab utama kematian pada kanker tiroid selanjutnya agar dapat dilakukan langkah yang tepat untuk memperpanjang waktu *survival* bagi penderita kanker tiroid khususnya di Rumah Sakit XYZ sampai terjadinya metastasis. Oleh karena itu perlu dilakukan analisis untuk mengetahui faktor-faktor risiko yang secara signifikan dapat mempengaruhi waktu *survival* pasien kanker tiroid untuk mengalami metastasis. Selain itu juga perlu dilakukan analisis terkait dengan *treatment* terbaik yang diberikan oleh pihak Rumah Sakit khususnya Rumah Sakit XYZ dalam upaya penyembuhan pasien kanker tiroid. Salah satu metode yang digunakan untuk menganalisis waktu *survival* ialah menggunakan regresi *Cox proportional hazard*. Hasil dari metode tersebut ialah suatu model yang didalamnya terdapat faktor-faktor atau hal-hal yang diduga dapat mempengaruhi waktu *survival* pasien kanker tiroid sampai terjadinya metastasis. Kemudian untuk memeriksa bentuk fungsional terbaik pada prediktor sesuai dengan asumsi *Cox proportional hazard*, bagaimana *diagnostic checking* model

regresi *Cox proportional hazard* data pasien kanker tiroid di Rumah Sakit XYZ dengan *Cox-Snell residual* dan *Martingale residual*.

### **1.3 Tujuan**

Tujuan yang ingin dicapai dalam penelitian ini ialah sebagai berikut:

1. Memodelkan data waktu *survival* hingga terjadi metastasis pasien kanker tiroid di Rumah Sakit XYZ dengan regresi *Cox proportional hazard*.
2. Mengetahui *treatment* terbaik dalam upaya meningkatkan waktu *survival* pasien kanker tiroid sampai terjadi metastasis di Rumah Sakit XYZ.
3. Mengaplikasikan metode *diagnostic checking* pada residual model regresi *Cox proportional hazard* data pasien tiroid di Rumah Sakit XYZ dengan *Cox-Snell Residual* dan *Martingale Residual*.

### **1.4 Manfaat**

Penelitian ini diharapkan dapat bermanfaat dalam kemajuan ilmu pengetahuan dan mampu dijadikan bahan pertimbangan bagi pihak Rumah Sakit XYZ dalam pengambilan keputusan mengenai cara penanganan pasien kanker tiroid agar dapat memperpanjang waktu *survival* sampai terjadinya metastasis. Bagi pembaca, penelitian ini diharapkan dapat digunakan sebagai bahan belajar dari aplikasi analisis *survival*.

### **1.5 Batasan Masalah**

Batasan masalah dalam penelitian ini yaitu data yang digunakan ialah data pasien kanker tiroid yang ada di Rumah Sakit XYZ pada jangka tahun tertentu dan juga digunakan metode *right censoring* dalam penentuan data yang akan dianalisis. Pada penelitian ini juga dibatasi bahwa metode analisis *survival* yang digunakan adalah regresi *Cox proportional hazard*.

*(Halaman ini sengaja dikosongkan)*

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Analisis Survival**

Analisis *Survival* atau analisis ketahanan adalah suatu metode statistik yang berkaitan dengan waktu dari suatu objek yang dimulai dari time origin atau start point sampai terjadinya kejadian tertentu yang telah ditetapkan (*failure event* atau *end point*). Analisis survival harus mempertimbangkan masalah utama dalam analisis yang disebut penyensoran. Penyensoran terjadi ketika kita mempunyai sebagian informasi dari ketahanan hidup seseorang, tetapi kita tidak mengetahui waktu ketahanan hidupnya secara pasti. Menurut (Kleinbaum dan Klein, 2012) secara umum ada tiga penyebab terjadinya penyensoran yaitu:

1. *The study ends* yaitu jika penelitian telah berakhir akan tetapi pasien belum mengalami *failure event*.
2. *Loss to follow up* yaitu jika seorang pasien tidak melanjutkan pengobatan atau karena pindah rumah sakit ketika penelitian berlangsung.
3. *Withdraws from the study* yaitu jika seorang pasien meninggal karena penyebab lain.

Menurut (Collet, 1994) ada tiga jenis sensor dalam analisis survival, yaitu:

##### 1. Sensor Kanan (*Right Censored*)

Jika objek yang diamati tidak mengalami *event* sampai penelitian selesai. Selain itu sensor kanan juga dapat terjadi karena objek hilang dari pengamatan (*lost to follow up*) atau objek mengalami *event* tetapi tidak disebabkan oleh faktor utama yang diteliti (*withdraws*) saat penelitian masih berlangsung. Dengan kata lain, data tersensor kanan terjadi jika waktu *survival* suatu objek tidak lengkap sampai dengan penelitian selesai.

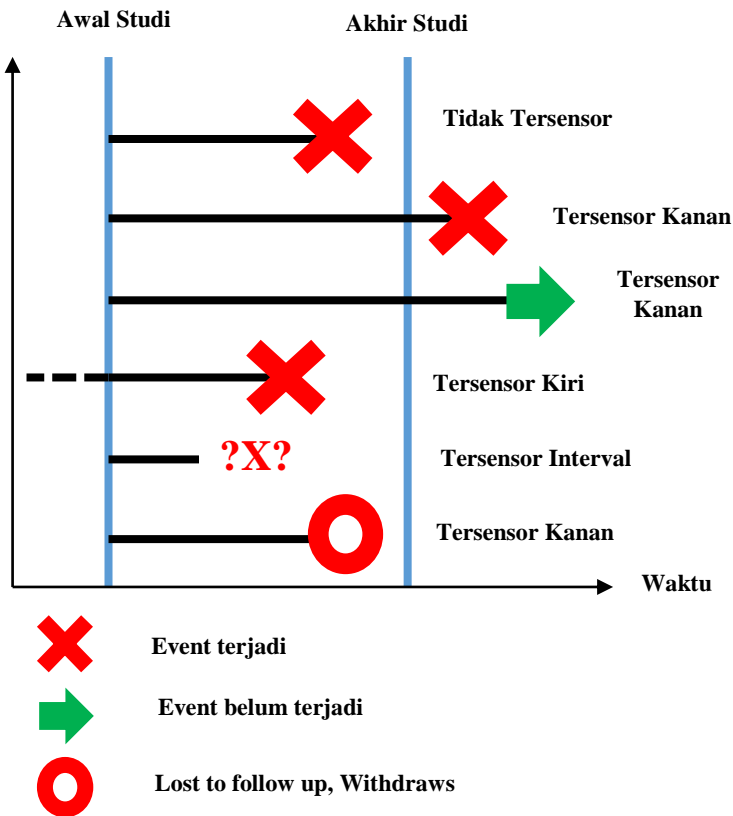
##### 2. Sensor Kiri (*Left Censored*)

Jika objek mengalami *event* sebelum data mulai dikumpulkan. Dengan kata lain, data tersensor kiri terjadi jika

waktu *survival* seorang individu yang sebenarnya kurang dari atau sama dengan waktu *survival* individu saat dilakukan penelitian. Atau dapat dikatakan bahwa saat penelitian dilakukan, individu sudah mengalami *event* tetapi tidak diketahui secara pasti kapan *event* itu terjadi.

### 3. Sensor Interval (*Interval Censored*)

Jika objek mengalami *event* diantara waktu  $t_1$  dan  $t_2$  tetapi tidak diketahui dengan pasti kapan *event* terjadi.



Gambar 2. 1 Ilustrasi Penyensoran Tipe Data

## 2.2 Dasar Teori Analisis Survival

Untuk  $T$  suatu variabel acak positif dan menunjukkan waktu survival setiap subjek, maka nilai-nilai yang mungkin untuk  $T$  yaitu  $\geq 0$ . Menurut (Lee dan Wang, 2003) distribusi dari  $T$  dapat dinyatakan dalam tiga cara yaitu sebagai berikut:

### 1. Fungsi Kepadatan Peluang

Fungsi kepadatan peluang atau PDF (*Probability Density Function*) adalah peluang suatu individu mati atau mengalami kejadian sesaat dalam interval waktu  $t$  sampai  $t + \Delta t$ . Fungsi kepadatan peluang  $f(t)$  dirumuskan sebagai berikut menurut (Lee dan Wang, 2003),

$$\begin{aligned} f(t) &= \lim_{\Delta t \rightarrow 0} \left[ \frac{P(t < T < (t + \Delta t))}{\Delta t} \right] \\ &= \lim_{\Delta t \rightarrow 0} \left[ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right] \end{aligned} \quad (2.1)$$

jika  $T$  merupakan variabel acak positif pada interval  $[0, \infty)$ , maka  $F(t)$  merupakan fungsi distribusi kumulatif kontinu dari  $T$ . Didefinisikan sebagai peluang suatu individu mengalami kejadian kurang dari sama dengan waktu  $t$  yaitu,

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad (2.2)$$

berdasarkan persamaan (2.2) diperoleh :

$$f(t) = \frac{d(F(t))}{dt} = F'(t) \quad (2.3)$$

### 2. Fungsi Survival

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

$$S(t) = P(T > t) = 1 - P(T \leq t) \quad (2.4)$$

dengan  $T$  adalah notasi waktu *survival* berupa variable random yang memiliki fungsi distribusi peluang  $f(t)$  sebagai berikut.

$$\begin{aligned}
 f(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < (t + \Delta t))}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}
 \end{aligned} \tag{2.5}$$

dengan

$$F(t) = P(T \leq t) = \int_0^t f(t) dt \tag{2.6}$$

### 3. Fungsi Hazard

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

$$h(t) = \lim_{\Delta t \rightarrow 0} \left[ \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \right] \tag{2.9}$$

Hubungan antara fungsi *hazard* dengan fungsi *survival* dapat diperoleh berdasarkan persamaan (2.4) dan (2.5) dengan menggunakan probabilitas bersyarat  $P(A \mid B) = P(A \cap B) / P(B)$  sebagai berikut

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

dengan  $F(t) = 1 - S(t)$  maka,



$$\begin{aligned} f(t) &= d[F(t)]/dt \\ &= d[1 - S(t)]/dt \end{aligned}$$

diperoleh nilai  $h(t)$  sebagai berikut.

$$h(t) = \frac{\left( \frac{d[1 - S(t)]}{dt} \right)}{S(t)}$$

Selanjutnya, mengintegrasikan kedua ruas fungsi tersebut sehingga didapatkan hasil sebagai berikut.

$$\begin{aligned} \int_0^t h(t)dt &= \int_0^t \frac{-1}{S(t)} d[S(t)] \\ &= -\ln S(t) \Big|_0^t \\ &= -\ln S(t) + \ln S(0) \\ &= -\ln S(t). \end{aligned}$$

Jadi hubungan antara fungsi *hazard* dengan fungsi *survival* dapat dinyatakan sebagai berikut.

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

### 2.3 Kurva Kaplan-Meier

Metode Kaplan Meier (1985) sangat populer untuk analisis survival yang paling cocok digunakan ketika ukuran sampel kecil. Analisis *Kaplan-Meier* menggunakan asumsi sebagai berikut :

1. Subyek yang menarik diri dari penelitian secara rata-rata memiliki “nasib” kesudahan variabel hasil (peristiwa) yang sama dengan subyek yang bertahan selama pengamatan.
2. Perbedaan waktu mulainya masuk dalam pengamatan antar subyek tidak mempengaruhi risiko (probabilitas) terjadinya variabel hasil (peristiwa). Probabilitas peristiwa untuk berbagai jangka waktu tersebut dapat digambarkan sebagai kurva analisis survival. (Murti, 1997).

*Kaplan-Meier* adalah komputasi untuk menghitung peluang survival. Metode *Kaplan-Meier* didasarkan pada waktu kelangsungan hidup individu dan mengasumsikan bahwa data

sensor adalah independen berdasarkan waktu kelangsungan hidup (yaitu, alasan observasi yang disensor tidak berhubungan dengan penyebab *failure time*) (Stevenson, 2009). Analisis Kaplan Meier digunakan untuk menaksir fungsi *survival* (Kleinbaum dan Klein, 2012). Metode *Kaplan Meier* dapat menunjukkan proporsi ketahanan hidup dari setiap objek. Persamaan umum dari fungsi *survival* yang digunakan untuk membentuk kurva *Kaplan-Meier* adalah sebagai berikut.

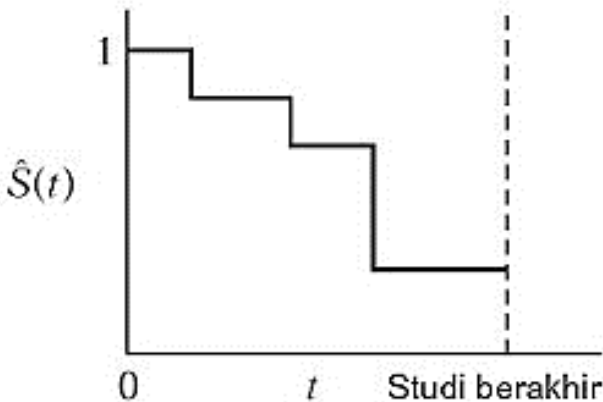
$$\hat{S}(t) = \prod_{t_i \leq t} \left( 1 - \frac{d_i}{n_i} \right) \quad (2.10)$$

keterangan :

$n_i$  : banyaknya individu dalam sampel yang hidup tidak mengalami metastasis dan terobservasi pada waktu ke-  $t_i$  ataupun yang mengalami metastasis pada waktu ke-  $t_i$ .

$d_i$  : banyaknya individu yang mengalami metastasis pada waktu ke-  $t_i$ .

Berikut merupakan ilustrasi visual dari kurva *Kaplan-Meier* yang terbentuk dari fungsi *survival*.



**Gambar 2. 2** Ilustrasi Kurva *Kaplan Meier*

## 2.4 Uji Log Rank

Pengujian yang digunakan untuk membandingkan kurva *survival Kaplan Meier* dalam kelompok yang berbeda disebut uji Log Rank. Hipotesis yang digunakan pada uji *log rank* untuk dua kelompok atau lebih adalah sebagai berikut (Kleinbaum dan Klein, 2012).

Hipotesis :

$H_0 : S_1(t) = S_2(t)$  (tidak ada perbedaan antar kategori pada kurva *survival*).

$H_1 : S_1(t) \neq S_2(t)$  (terdapat perbedaan antar kategori pada kurva *survival*).

Statistik uji :

$$\chi^2 = \sum_{i=1}^G \frac{(O_i - E_i)^2}{E_i} \quad (2.11)$$

dengan,

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

Keterangan :

$O_i$  : Nilai observasi individu kelompok ke- $i$

$E_i$  : Nilai ekspektasi individu kelompok ke- $i$

$m_{ij}$  : Jumlah objek pada kelompok ke- $i$  yang mengalami *event* saat waktu ke-  $t$

$n_{ij}$  : Jumlah objek pada kelompok ke- $i$  yang tidak mengalami *event* pada waktu ke-  $t$

$e_{ij}$  : Nilai ekspektasi kelompok ke- $i$  pada waktu ke-  $t$

$G$  : Banyaknya kelompok

$j$  : Failure event (1,2,...,  $h$  )

$i$  : 1,2,...,  $G$  .

Daerah penolakan :

Tolak  $H_0$  jika nilai  $\chi^2_{hitung} > \chi^2_{(\alpha, G-1)}$  atau nilai  $p\text{-value} < \alpha$ .

## 2.5 Regresi Cox *Proportional Hazard*

Analisis regresi *Cox proportional hazard* adalah metode yang dapat dipakai untuk memodelkan antara variabel respon yang berupa waktu *survival* dengan satu atau lebih variabel prediktor (Fernandes dan Solimun, 2016). Pada analisis regresi *cox*, data tidak perlu memenuhi distribusi apapun (Hosmer, Lameshow, dan May, 2008). Asumsi *proportional hazard* harus terpenuhi pada analisis regresi *Cox proportional hazard*. Pemodelan regresi *Cox proportional hazard* bertujuan untuk mengetahui hubungan variabel-variabel independen terhadap data *survival* secara bersamaan (Kleinbaum dan Klein, 2012) Adapun persamaan dari Regresi *Cox proportional hazard* adalah sebagai berikut:

$$h(t) = h_0(t) \exp(\beta_j' \mathbf{x}_j) \quad (2.12)$$

$$h(t) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad (2.13)$$

dengan,

$h_0(t)$  : *baseline hazard* pada waktu ke- $t$

$\beta_j$  : parameter regresi ke- $j$

$x_j$  : nilai dari variabel bebas ke- $j$

$j$  : 1,2,...,  $p$

### 2.5.1 Estimasi Parameter

Parameter dari variabel independen pada model regresi *Cox proportional hazard* dapat diestimasi dengan metode *Partial Likelihood Estimation*. Jika ada  $n$  pengamatan dengan urutan waktu *event* terjadi adalah  $t_{(1)} < t_{(2)} < \dots < t_{(n)}$  dan  $R(t_{(i)})$  adalah himpunan waktu dari semua pengamatan yang berisiko mengalami

*event* pada waktu  $t_{(l)}$ , maka memiliki fungsi *likelihood* sebagai berikut.

$$L(\boldsymbol{\beta}) = \prod_{t=1}^n \frac{\exp(\boldsymbol{\beta}^T \mathbf{x}_t)}{\sum_{I \in R_{(t)}} \exp(\boldsymbol{\beta}^T \mathbf{x}_I)} \quad (2.14)$$

dengan,

$x_{(t)}$  : vektor variabel dari pengamatan yang mengalami *event* pada saat waktu ke-  $t$

$x_{(l)}$  : vektor variabel dari pengamatan yang masih belum mengalami metastasis dan merupakan elemen dari  $R(t_{(l)})$ .

Iterasi *Newton-Raphson* digunakan pada metode regresi ini karena estimasi parameter yang didapatkan pada model ini implisit (Collet, 1994) . Jika  $g(\boldsymbol{\beta})$  adalah vektor berukuran  $p \cdot 1$  yang merupakan turunan pertama fungsi  $\ln L(\boldsymbol{\beta})$  terhadap parameter  $\boldsymbol{\beta}$  dan  $\mathbf{H}(\boldsymbol{\beta})$  adalah matrik hessian berukuran  $p \cdot p$  yang berisi turunan kedua dari fungsi  $\ln L(\boldsymbol{\beta})$  , maka estimasi parameter pada iterasi ke-  $(l + 1)$  adalah sebagai berikut

$$\boldsymbol{\beta}^{l+1} = \boldsymbol{\beta}^l - \mathbf{H}^{-1}(\boldsymbol{\beta}^l)g(\boldsymbol{\beta}^l) \quad (2.15)$$

Iterasi *Newton-Raphson* berhenti saat  $\|\boldsymbol{\beta}^{(l+1)} - \boldsymbol{\beta}^l\| \leq \varepsilon$  ,  $\varepsilon$  adalah bilangan yang sangat kecil.

### 2.5.2 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter terdiri dari dua tahap, yaitu uji serentak sebagai tahap pertama dan uji parsial sebagai tahap kedua.

a) Pengujian Parameter Secara Serentak

hipotesis :

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

$H_1$  : minimal terdapat satu dari  $\beta_j \neq 0$ , dengan  $j = 1, 2, \dots, p$

taraf signifikansi :  $\alpha = 0,05$

$$\text{statistik uji} \quad : G^2 = -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

Keterangan :

$L(\hat{\omega})$  : Nilai *likelihood* untuk model awal yaitu fungsi *likelihood* model regresi sebelum variabel prediktor dimasukkan.

$L(\hat{\Omega})$  : Nilai *likelihood* untuk model awal yaitu fungsi *likelihood* model regresi setelah variabel prediktor dimasukkan.

Daerah penolakan :

Tolak  $H_0$  jika  $G^2 > \chi_{p,\alpha}^2$  atau  $p\text{-value} < \alpha$  (0,05).

b) Pengujian Parameter Secara Parsial

hipotesis :

$H_0 : \beta_j = 0$ , (variabel prediktor tidak berpengaruh terhadap variabel respon).

$H_1 : \beta_j \neq 0$ , (variabel prediktor berpengaruh terhadap variabel respon) dengan  $j = 1, 2, \dots, p$

taraf signifikansi :  $\alpha = 0,05$

$$\text{statistik uji} \quad : W_k = \left[ \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \right]^2$$

keterangan :  $\hat{\beta}_k$  adalah penduga  $\beta_k$   
 $SE(\hat{\beta}_k)$  adalah *standar error* dari  $\beta_k$ .

Daerah penolakan :

Tolak  $H_0$  jika  $W_k > \chi_{\alpha;l}^2$  atau  $p\text{-value} < \alpha$

## 2.6 Pemilihan Model Terbaik

Model hazard dengan menggunakan banyak variabel independen akan menimbulkan permasalahan yaitu terjadinya kasus multikolinearitas. Oleh karena itu dibutuhkan suatu metode untuk memilih variabel independen yang dapat menghasilkan

model terbaik dan tidak terjadi kasus multikolinearitas. Salah satu metode untuk memilih variabel tersebut ialah metode *backward*. metode *backward* yaitu metode seleksi variabel yang bekerja dengan cara memasukkan semua variabel independen terlebih dahulu, kemudian variabel independen yang tidak berpengaruh signifikan dalam model akan dikeluarkan dalam model. Seleksi model terbaik yang digunakan dalam analisis survival dilihat berdasarkan kriteria *Akaike Information Criterion* (AIC) pada metode eliminasi *backward*. Nilai AIC terkecil adalah model terbaik. Berikut merupakan rumus untuk mendapatkan nilai AIC.

$$AIC = -2 \ln L(\hat{\beta}) + 2k \quad (2.16)$$

## 2.7 Hazard Ratio

Interpretasi terhadap perbandingan resiko dari suatu individu terhadap individu lainnya adalah dengan menggunakan *hazard ratio*. *Hazard ratio* juga digunakan untuk mengetahui adanya peningkatan atau penurunan resiko yang dialami oleh suatu individu yang dikenai perlakuan atau kondisi tertentu. Misal, terdapat 2 individu dengan karakteristik yang berbeda yaitu individu dengan karakteristik  $Z_i$  dan individu dengan karakteristik  $Z_j$ . Perbandingan terhadap tingkat resiko kegagalan yang dialami kedua individu tersebut dapat diformulasikan sebagai berikut :

$$HR = \frac{h(t | Z_i)}{h(t | Z_j)} = \frac{h_0(t) \exp(\beta^T z_i)}{h_0(t) \exp(\beta^T z_j)} = \exp(\beta^T (z_i - z_j)) \quad (2.17)$$

Persamaan (2.17) merupakan *hazard ratio* untuk variabel independen kontinu, sedangkan *hazard ratio* untuk variabel independen bertipe kategorik misal X adalah sebuah variabel prediktor dengan dua kategori yaitu 0 dan 1. Maka nilai HR dapat dihitung dengan rumus,

$$HR = \frac{h_0(t | x = 1)}{h_0(t | x = 0)} = \frac{h_0(t) \exp(\hat{\beta})}{h_0(t)} = \exp(\hat{\beta}) \quad (2.18)$$

## 2.8 Asumsi *Proportional Hazard*

Asumsi *proportional hazard* (PH) terpenuhi apabila keadaan tersebut memiliki nilai *hazard ratio* yang konstan terhadap waktu. Terdapat beberapa pendekatan yang dapat digunakan untuk mengetahui apakah keadaan memenuhi asumsi *proportional hazard* atau tidak (Kleinbaum & Klein, 2012), antara lain:

### 1. Pendekatan Grafis

Grafik yang dapat digunakan untuk mengetahui asumsi *proportional hazard* terpenuhi atau tidak adalah grafik  $\ln[-\ln S(t)]$

. Pada grafik  $\ln[-\ln S(t)]$ , asumsi terpenuhi jika pola setiap kategori dalam variabel sejajar atau tidak bersilangan satu sama lain. Pada grafik *observed versus expected*, asumsi terpenuhi jika pola setiap kategori dalam variabel saling berhimpitan.

### 2. Pengujian *Goodness of Fit*

Pendekatan *goodness of fit* dapat dilakukan menggunakan *Cox-Snell Residual*. Selain itu juga dapat dilakukan dengan korelasi pearson untuk menguji korelasi antara residual *Martingale* dengan *rank survival time* untuk masing-masing variabel prediktor. Pengujian asumsi *proportional hazard* terpenuhi untuk suatu variabel jika residual *Martingale* variabel tersebut tidak berkorelasi dengan waktu (Therneau dan Grambsch, 2000).

Hipotesis untuk pendekatan ini adalah sebagai berikut (Kleinbaum dan Klein, 2012).

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

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

Statistik uji :

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \quad (2.19)$$

$r$  adalah nilai korelasi *pearson* dengan rumus sebagai berikut :



$$r_{xy} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (2.20)$$

dengan,

$X$  : residual *martingale*.

$Y$  : *rank survival time*.

Daerah penolakan :

Tolak  $H_0$  jika  $|t| > t_{(\alpha/2, n-1)}$  atau nilai *p-value*  $< \alpha$ .

## 2.9 Cox-Snell Residual

Residual yang paling banyak diaplikasikan secara luas dalam data analisis survival adalah residual *Cox-snell*. Residual *Cox-snell* dapat digunakan sebagai alat *diagnostic checking* untuk mengetahui asumsi linieritas terpenuhi atau tidak. Residual *Cox-snell* untuk individu *ke-i* dengan  $i = 1, 2, 3, \dots, n$  diberikan sebagai berikut :

$$rc_i = \hat{H}_0(t_i) \exp(\hat{\boldsymbol{\beta}}' \mathbf{x}_i) \quad (2.21)$$

Dengan  $\hat{H}_0(t_i)$  merupakan estimasi fungsi *hazard* dasar kumulatif pada waktu  $t_i$ , residual *Cox-snell* merupakan nilai dari  $\hat{H}_i(t_i) = -\log \hat{S}_i(t_i)$  dengan  $i = 1, 2, 3, \dots, n$  dimana  $\hat{H}_i(t_i)$  dan  $\hat{S}_i(t_i)$  merupakan nilai estimasi dari fungsi *hazard* kumulatif dan fungsi *survival* untuk individu *ke-i* pada waktu  $t_i$ . Model dikatakan layak jika nilai residual *Cox-snell* berdistribusi eksponensial.

## 2.10 Martingale Residual

Residual merupakan suatu alat diagnostik yang digunakan untuk menilai suatu ketepatan model dan berfokus pada masing masing variabel secara grafikal. Penggunaan residual dalam analisis data survival digunakan untuk menilai hal-hal sebagai berikut :

1. Bentuk fungsional untuk mengetahui pengaruh dari variabel prediktor pada suatu model.
2. Ketepatan suatu model dengan memperhitungkan asumsi-asumsi pada pemodelan proportional hazard.
3. Ketepatan suatu model dengan memperhitungkan signifikansi dari masing-masing variabel.
4. Pengaruh yang diberikan untuk suatu variabel dalam estimasi parameter.

Model Cox PH pada persamaan dapat menghasilkan berbagai macam jenis residual, salah satunya adalah residual martingale. *Martingale residual* menyediakan ukuran perbedaan jumlah observasi yang mati dalam interval  $(0,t)$  dan jumlah prediksi dalam model. *Martingale residual* berguna untuk menentukan bentuk fungsional dari sebuah kovariat yang dimasukkan dalam sebuah model regresi *Cox proportional hazard*, dapat dikatakan memenuhi asumsi jika garis yang terbentuk pada *martingale residual* berada disekitar 0 dengan lebar toleransi  $\pm 0,5$  (Klein dan Moescheberger, 2003). Persamaan *martingale residual* adalah sebagai berikut (Tableman dan Kim, 2004).

$$\hat{M}_i = d_i - \left[ \hat{H}_0(t_i) \exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \right] \quad (2.22)$$

Dengan

$d_i = 1$ , jika pasien ke- $i$  mengalami *event*

$= 0$ , jika pasien ke- $i$  tersensor

$\hat{H}_0(t_i)$  = fungsi *baseline cumulative hazard*

Nilai residual *martingale* adalah antara  $-\infty$  hingga 1. Nilai tersebut negatif pada data tersensor. Residual *martingale* dapat menjadi gambaran mengenai perbedaan hasil pengamatan  $N_i(t)$  dengan angka prediksi pada kejadian  $(t)$ . Ketika perbedaan antara hasil pengamatan dengan angka prediksi untuk subjek ke- $i$  cukup besar, itu menunjukkan bahwa subjek ke- $i$  tidak sesuai dengan model dan mengakibatkan suatu nilai besar pada  $M_i(t)$ . Karena *range* dari  $(t)$  adalah  $(0, \infty)$  dan  $N_i(t)$  hanya bernilai 0 atau 1, maka dapat diambil kesimpulan residual *martingale*

bernilai  $(-\infty, 1)$  dan kesimetrisan dari distribusi residual *martingale* mendekati 0.

### **2.11 Kanker Tiroid**

Kanker tiroid adalah keganasan yang terjadi pada kelenjar tiroid dan termasuk dalam 10 besar angka kejadian kanker terbanyak di Indonesia pada tahun 2018, tetapi diantara kelenjar endokrin, kanker tiroid termasuk jenis keganasan paling sering ditemukan. Kelenjar tiroid merupakan organ tubuh yang relatif jarang mengalami keganasan, dengan angka kejadiannya meliputi 95% dari keseluruhan kanker endokrin. Kanker tiroid secara umum termasuk kelompok keganasan dengan prognosis relatif baik (Pasaribu, 2006)

Berikut merupakan beberapa tanda dan gejala yang mengindikasikan adanya kanker tiroid :

1. Benjolan pada leher, yang terkadang bertumbuh dengan cepat.
2. Benjolan pada leher, yang terkadang bertumbuh dengan cepat.
3. Nyeri di bagian depan leher, kadang-kadang menjalar ke telinga.
4. Suara serak atau perubahan suara lainnya yang menetap.
5. Sulit menelan.
6. Sulit bernapas.

### **2.12 Faktor Resiko Metastasis**

Metastasis adalah penyebaran sel kanker ke organ tubuh yang lainnya. Menurut (Sung dkk, 2016) dalam jurnalnya, terdapat beberapa faktor yang mempengaruhi terjadinya metastasis jauh pada pasien kanker tiroid adalah sebagai berikut

1. Jenis Kelamin
2. Usia
3. Surgery (*treatment*)
4. Ukuran tumor primer
5. *Lymph node dissection*
6. Invasi vaskular

7. Tiroiditis
8. Penyebaran Ekstratiroidal
9. Ablasi *Radioactive Iodine*

Berdasarkan faktor resiko metastasis oleh (Sung dkk, 2016), pada penelitian ini mengacu 2 variabel setelah disesuaikan dengan variabel yang ada yaitu variabel usia dan *treatment*.

Sedangkan menurut (Li dkk, 2016) terdapat beberapa faktor resiko yang dapat mendorong terjadinya metastasis pada penderita kanker tiroid yaitu antara lain :

1. Usia
2. Jenis kelamin
3. Stadium
4. Ukuran Tumor
5. *Stromal invasion*
6. *Angiolympathic invasion*
7. *Encapsulation*
8. *Capsule invasion*
9. *Extrathyroidal extension*

Berdasarkan faktor resiko metastasis oleh (Li dkk, 2016), dalam penelitian ini digunakan 2 variabel sesuai dengan data yang disediakan oleh Rumah Sakit XYZ yaitu variabel jenis kelamin dan stadium. Pada variabel stadium digunakan pengecualian terhadap tingkat stadium 4 karena pada stadium 4 pasien tiroid pasti sudah mengalami metastasis sehingga dalam penelitian ini digunakan stadium 1,2, dan 3 guna mengetahui ketahanan pasien kanker tiroid pada stadium 1,2, dan 3 untuk mengalami metastasis.

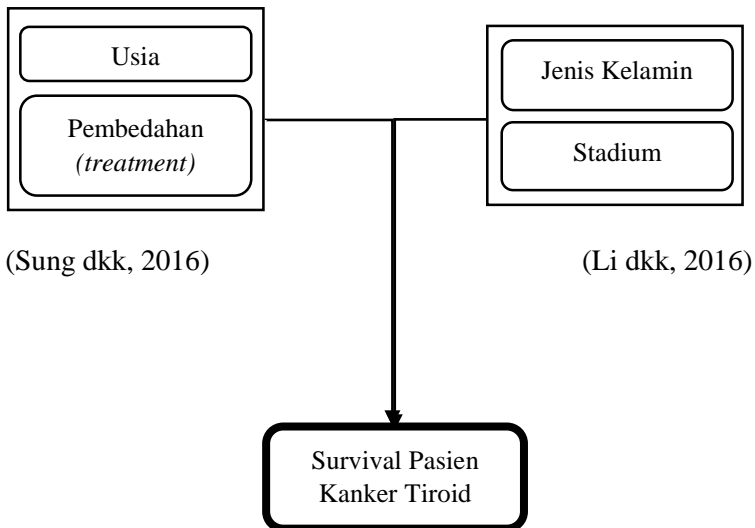
## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder pasien kanker tiroid yang diperoleh dari Rumah Sakit XYZ. Data yang digunakan ialah data rekam medis pasien kanker tiroid pada bulan Januari 2008 hingga Februari 2019 yang terdiri dari 167 pasien. Namun, setelah dilakukan *preprocessing* (menghapus data pasien yang informasinya tidak lengkap dan data eksklusi) data yang digunakan dalam analisis ini ialah sebanyak 103 pasien.

### 3.2 Kerangka Konsep

Kerangka konsep dari variabel yang diduga mempengaruhi waktu *survival* pasien kanker tiroid di Rumah Sakit XYZ ialah sebagai berikut.



**Gambar 3. 1** Kerangka Konsep Penelitian

### 3.3 Variabel Penelitian

Variabel dari data kanker tiroid yang digunakan pada penelitian ini adalah sebagai berikut.

**Tabel 3.1** Variabel Penelitian

Variabel	Keterangan	Definisi Operasional	Skala
T	Waktu <i>survival</i>	Waktu pertama kali pasien melakukan pemeriksaan pengobatan di RS XYZ hingga terjadinya <i>event</i> (metastasis) atau tersensor.	Rasio
D	Status <i>event</i>	Pasien mengalami <i>event</i> atau tidak.	Nominal
X <sub>1</sub>	Jenis kelamin	Jenis kelamin pasien	Nominal
X <sub>2</sub>	Usia Pasien	Usia pasien saat pertama kali melakukan pengobatan di Rumah Sakit XYZ.	Rasio
X <sub>3</sub>	Stadium 0 = Stadium 1 1 = Stadium 2 2 = Stadium 3	Stadium klinis pasien kanker tiroid saat pertama kali pasien diperiksa atau berobat di Rumah Sakit XYZ.	Ordinal
X <sub>4</sub>	Operasi 0 = Tidak 1 = Operasi	Pasien diberikan operasi atau tidak selama perawatan kanker tiroid di Rumah Sakit XYZ.	Nominal
X <sub>5</sub>	Radioterapi 0 = Tidak 1 = Radioterapi	Pasien diobati dengan radioterapi atau tidak selama perawatan kanker tiroid di Rumah Sakit XYZ.	Nominal

**Tabel 3.1** Variabel Penelitian (Lanjutan)

$X_6$	Terapi hormonal 0 = Tidak 1 = Terapi hormonal	Pasien diobati dengan terapi hormonal atau tidak selama perawatan kanker tiroid di Rumah Sakit XYZ.	Nominal
-------	---	---	---------

### 3.4 Struktur Data

Struktur data dari variabel-variabel yang digunakan dalam penelitian ini dijelaskan dalam tabel berikut.

**Tabel 3. 2** Struktur Data Penelitian

Pasien ke-	T	d	$X_1$	$X_2$	...	$X_j$	...	$X_p$
1	$T_1$	$d_1$	$x_{11}$	$x_{12}$	...	$x_{1j}$	...	$x_{1p}$
2	$T_2$	$d_2$	$x_{21}$	$x_{22}$	...	$x_{2j}$	...	$x_{2p}$
3	$T_3$	$d_3$	$x_{31}$	$x_{32}$	...	$x_{3j}$	...	$x_{3p}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$i$	$T_i$	$d_i$	$x_{i1}$	$x_{i2}$	...	$x_{ij}$	...	$x_{ip}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$n$	$T_n$	$d_n$	$x_{n1}$	$x_{n2}$	...	$x_{nj}$	...	$x_{np}$

### 3.5 Langkah Analisis

Langkah analisis dalam penelitian ini adalah sebagai berikut.

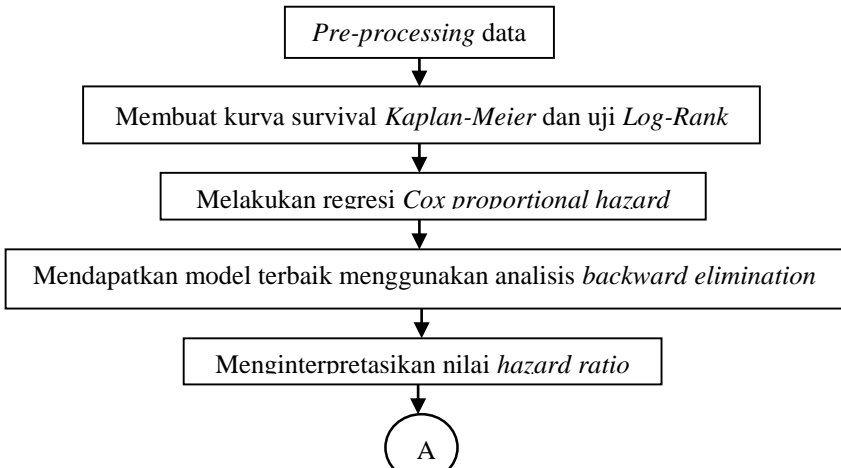
1. *Preprocessing* data (menghapus data yang hilang, menghitung waktu *survival* (T), menentukan status (d), menyaring data eksklusi).
2. Membuat kurva *survival* dan melakukan uji *Log-Rank* pada masing-masing variabel prediktor menggunakan analisis *Kaplan-Meier* untuk mendeskripsikan kategori pada setiap variabel tersebut.
3. Membuat model regresi *Cox proportional hazard* dari data pasien kanker tiroid untuk mengetahui hubungan antara variabel-variabel prediktor yang diduga mempengaruhi waktu *survival*.
4. Mendapatkan model *Cox proportional hazard* terbaik menggunakan *backward elimination* dengan kriteria nilai

AIC terkecil kemudian melakukan uji signifikansi untuk mengetahui variabel-variabel yang mempengaruhi waktu *survival* secara signifikan.

5. Menginterpretasikan nilai *hazard ratio* dari variabel-variabel yang signifikan terhadap waktu *survival*.
6. Melakukan uji asumsi *proportional hazard* dengan *goodness of fit*. Jika asumsi tidak terpenuhi maka membuat model regresi lain. Jika asumsi terpenuhi, menggunakan regresi *Cox proportional hazard*.
7. Melakukan *diagnostic checking* residual dari model regresi *Cox proportional hazard* dengan metode *Cox-Snell* dan *Martingale residual*.
  - Jika residual tidak memenuhi asumsi, membuat model regresi lain.
  - Jika residual memenuhi asumsi, model regresi *Cox proportional hazard* adalah model regresi terbaik.
8. Menarik kesimpulan dan saran.

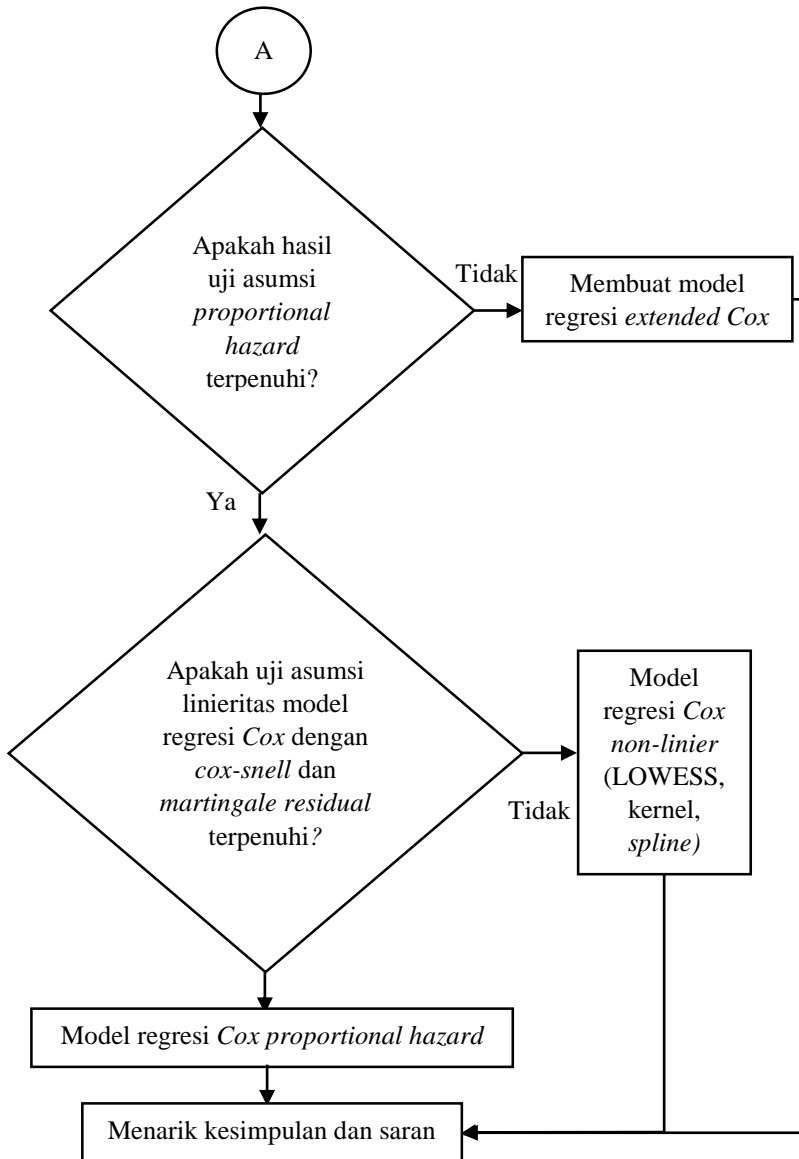
### 3.6 Diagram Alir

Berdasarkan langkah analisis sebelumnya, berikut merupakan diagram alir dari penelitian ini.



**Gambar 3.2** Diagram Alir Penelitian





**Gambar 3.2** Diagram Alir Penelitian (lanjutan)

*(Halaman ini sengaja dikosongkan)*

## BAB IV ANALISIS DAN PEMBAHASAN

### 4.1 Karakteristik Data

Data pasien kanker tiroid di Rumah Sakit XYZ yang terlampir pada lampiran 1 terdiri dari 103 pasien dengan jumlah pasien yang mengalami *event* (metastasis) yaitu sebanyak 6 pasien dan 97 pasien lainnya tersensor. Karakteristik dari data pasien kanker tiroid yang dirawat di Rumah Sakit XYZ adalah sebagai berikut.

**Tabel 4. 1** Karakteristik Data Penelitian

<b>Variabel</b>	<b><i>Event</i> (n=6)</b>	<b><i>Censored</i> (n=97)</b>
Jenis Kelamin ( $x_1$ )		
Laki-laki	5	24
Perempuan	1	73
Stadium ( $x_3$ )		
Stadium 1	3	52
Stadium 2	2	19
Stadium 3	1	26
Operasi ( $x_4$ )		
Ya	4	68
Tidak	2	29
Radioterapi ( $x_5$ )		
Ya	2	24
Tidak	4	73
Terapi Hormonal ( $x_6$ )		
Ya	2	29
Tidak	4	68

Berdasarkan tabel 4.1 dapat diketahui bahwa jumlah pasien laki-laki yang mengalami metastasis sebanyak 5 pasien dan untuk pasien perempuan ialah sebanyak 1 orang pasien. Pasien kanker tiroid pada stadium 1, 2, dan 3 yang mengalami metastasis masing-

masing ialah sebanyak 3, 2, dan 1 pasien. Pada pasien yang melakukan operasi terdapat 4 pasien yang mengalami metastasis dan pasien yang tidak melakukan operasi terdapat 2 pasien yang mengalami metastasis. Untuk pasien yang melakukan radioterapi terdapat 2 pasien yang mengalami metastasis dan 4 pasien yang tidak melakukan radioterapi mengalami metastasis. Sama seperti pasien yang melakukan radioterapi, jumlah pasien yang melakukan terapi hormonal terdapat 2 orang yang mengalami metastasis dan 4 orang yang tidak melakukan terapi hormonal dan mengalami metastasis.

Selain itu terdapat variabel usia dan berikut merupakan analisis deskriptif untuk variabel usia.

**Tabel 4. 2** Statistika Deskriptif Variabel Usia

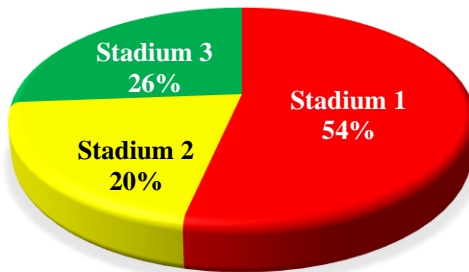
<i>Mean</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Varians</i>	<i>Min</i>	<i>Max</i>	<i>Range</i>
42,680	31,000	41,000	53,000	229,32	16,000	86,000	70,000

Tabel 4.2 menunjukkan bahwa rata-rata usia pasien kanker tiroid di Rumah Sakit XYZ sekitar 42,680 tahun, 50% dari pasien kanker tiroid berusia di bawah 41 tahun sedangkan 50% sisanya berusia diatas 41 tahun. 25% pasien kanker tiroid berusia diatas dibawah 31 tahun sedangkan 25% lainnya berusia diatas 53 tahun. Usia pasien kanker tiroid termuda ialah berusia 16 tahun sedangkan usia pasien kanker tiroid yang paling tua ialah 86 tahun. Selisih antara pasien yang paling muda dengan pasien yang paling tua ialah 70 tahun.



**Gambar 4. 1** Pie-Chart Faktor Jenis Kelamin

Berdasarkan Gambar 4.1 dapat diketahui bahwa sebesar 72% dari penderita kanker tiroid di Rumah Sakit XYZ ialah lebih banyak yang berjenis kelamin perempuan yaitu sebanyak 74 pasien sedangkan 28% sisanya berjenis kelamin laki-laki yaitu sebanyak 29 pasien.



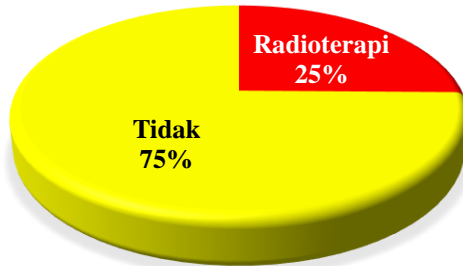
**Gambar 4. 2** *Pie-Chart* Faktor Stadium

Pada variabel stadium pasien, persentase jumlah pasien dengan masing-masing stadium kanker tiroid dijelaskan dalam Gambar 4.2. Gambar 4.2 menunjukkan bahwa penderita kanker tiroid di Rumah Sakit XYZ paling banyak ialah pada stadium 1 yaitu sebesar 54% lebih tepatnya sebanyak 55 pasien. Pasien dengan stadium 3 yaitu sebesar 26% atau sebanyak 27 pasien dan yang terendah ialah pasien stadium 2 yaitu sebesar 20% dengan jumlah 21 pasien.



**Gambar 4. 3** *Pie-Chart* Faktor Operasi

Gambar 4.3 menunjukkan bahwa jumlah pasien penderita kanker tiroid di Rumah Sakit XYZ yang melakukan operasi ialah sebesar 70% atau sebanyak 72 pasien, sedangkan 30% sisanya tidak melakukan operasi yaitu sebanyak 31 pasien.



**Gambar 4. 4** Pie-Chart Faktor Radioterapi

Sedangkan untuk *treatment* radioterapi, pada Gambar 4.4 dapat dilihat bahwa hanya sekitar 25% atau 26 pasien yang melakukan radioterapi dan 75% sisanya tidak melakukan *treatment* radioterapi yaitu sebanyak 77 pasien.

Salah satu *treatment* untuk pasien kanker tiroid selain operasi dan radioterapi terdapat juga terapi hormonal. Gambar 4.5 menunjukkan persentase pasien kanker tiroid di Rumah Skait XYZ yang melakukan terapi hormonal dan yang tidak melakukan terapi hormonal.



**Gambar 4. 5** Pie-Chart Faktor Terapi Hormonal

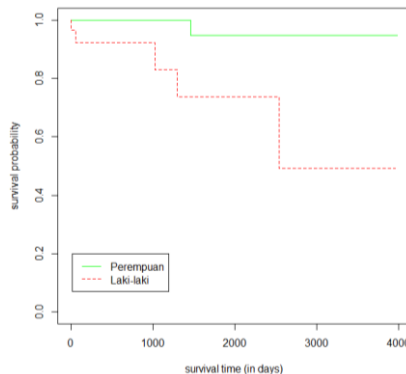
Pasien kanker tiroid di Rumah Sakit XYZ sebesar 30% pasien kanker tiroid yang melakukan terapi hormonal yaitu sebanyak 31 pasien. Sedangkan sebesar 70% pasien lainnya tidak melakukan terapi hormonal atau sebanyak 72 pasien tidak melakukan terapi hormonal.

## 4.2 Kurva Kaplan Meier dan Uji Log-Rank

Kurva *Kaplan-Meier* dan uji *log-rank* faktor-faktor yang diduga mempengaruhi waktu *survival* pasien kanker tiroid adalah sebagai berikut.

### 4.2.1 Jenis Kelamin

Kurva *Kaplan Meier* faktor jenis kelamin pasien kanker tiroid adalah sebagai berikut.



**Gambar 4. 6** Kurva *Kaplan Meier* Faktor Jenis Kelamin

Berdasarkan Gambar 4.6, kurva berwarna hijau menunjukkan kurva *survival* bagi pasien perempuan dan kurva berwarna merah menunjukkan kurva *survival* bagi pasien laki-laki. Dari kedua kurva tersebut dapat dilihat bahwa tidak terjadi perpotongan antara kurva satu dengan yang lainnya, sehingga dapat disimpulkan bahwa secara visual, waktu *survival* antara pasien perempuan dan laki-laki ialah berbeda. Selanjutnya, kedua

kurva *survival* tersebut akan diuji menggunakan uji *log-rank* untuk mengetahui apakah kurva *survival* antara pasien perempuan dan laki-laki berbeda secara signifikan. Berikut merupakan hasil pengujian *log-rank* faktor jenis kelamin.

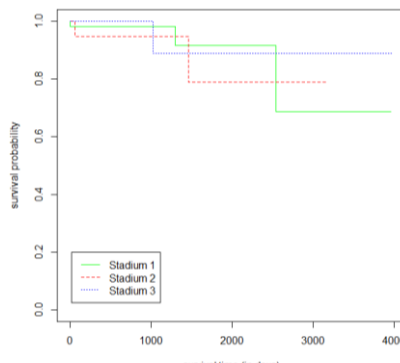
**Tabel 4. 3** Hasil Uji *Log-Rank* Faktor Jenis Kelamin

$\chi^2$	df	<i>p-value</i>
10,200	1	0,001

Berdasarkan hasil uji *Log-Rank* pada Tabel 4.3, diperoleh nilai  $\chi^2$  sebesar 10,200 dengan derajat bebas 1 dimana nilai tersebut lebih besar dari 3,841 ( $\chi^2_{(0,05, 1)}$ ) selain itu juga diketahui nilai *p-value* sebesar 0,001 lebih kecil dari nilai  $\alpha = 0,05$  sehingga dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara kurva *survival* pasien kanker tiroid laki-laki dan perempuan dengan kesimpulan bahwa peluang *survival* pasien perempuan lebih besar dibandingkan pasien laki-laki untuk mengalami metastasis.

#### 4.2.2 Stadium

Kurva *Kaplan Meier* untuk stadium pada pasien kanker tiroid adalah sebagai berikut.



**Gambar 4. 7** Kurva *Kaplan Meier* Faktor Stadium



Berdasarkan Gambar 4.7, kurva *survival* stadium pasien kanker tiroid secara visual terlihat saling berpotongan antara stadium 1,2, dan 3. Sehingga secara visual dapat disimpulkan bahwa kurva *survival* antar stadium tidak berbeda signifikan. Namun, untuk memperkuat dugaan tersebut selanjutnya dilakukan uji *log-rank* untuk mengetahui apakah kurva *survival* antar stadium tersebut berbeda secara signifikan atau tidak.

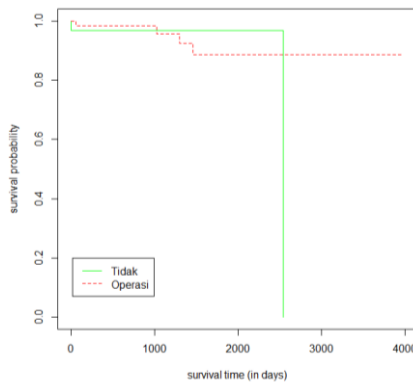
**Tabel 4. 4** Hasil Uji *Log-Rank* Faktor Stadium

$\chi^2$	df	<i>p-value</i>
0,500	2	0,800

Pada Tabel 4.4 setelah dilakukan pengujian *log-rank* diperoleh nilai  $\chi^2$  sebesar 35,600 dengan derajat bebas 3 1 dimana nilai tersebut lebih kecil dari 5,991 ( $\chi^2_{(0,05, 2)}$ ) dan nilai *p-value* sebesar 0,8 dimana nilai tersebut lebih besar dibandingkan nilai  $\alpha = 0,05$  yang artinya kurva *survival* antara stadium 1, 2 maupun 3 ialah tidak berbeda signifikan antar satu dengan yang lain.

### 4.2.3 Operasi

Kurva *Kaplan Meier* pada faktor operasi disajikan pada Gambar 4.8.



**Gambar 4. 8** Kurva *Kaplan Meier* Faktor Operasi

Gambar 4.8 menunjukkan bahwa kurva *survival* antara kelompok pasien tidak operasi dan operasi terlihat berpotongan, sehingga secara visual dapat dikatakan bahwa kurva *survival* antara kelompok pasien yang tidak dioperasi dengan pasien yang dioperasi tidak berbeda secara signifikan. Namun, untuk mendukung pernyataan tersebut diperlukan adanya suatu pengujian statistik yaitu dengan menggunakan uji *log-rank* dan berikut merupakan hasil dari uji *log-rank* faktor operasi.

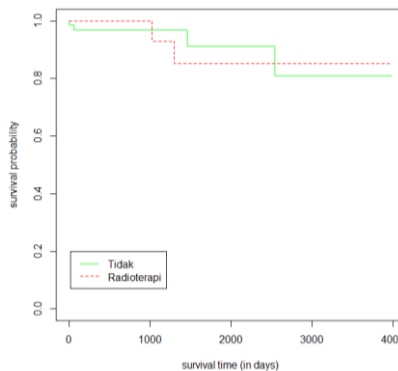
**Tabel 4. 5** Hasil Uji *Log-Rank* Faktor Operasi

$\chi^2$	Df	<i>p-value</i>
2,300	1	0,100

Berdasarkan Tabel 4.5, hasil pengujian *log-rank* diperoleh nilai  $\chi^2$  sebesar 2,300 dengan derajat bebas 1 dan nilai *p-value* sebesar 0,100 dan nilai tersebut lebih besar dibandingkan dengan nilai  $\alpha$  (0,05) sehingga dapat dikatakan bahwa kurva *survival* antara kelompok pasien yang dioperasi dan yang tidak dioperasi berbeda secara signifikan.

#### 4.2.4 Radioterapi

Kurva *survival* antara pasien yang melakukan radioterapi dan yang tidak ditampilkan pada Gambar 4.9.



**Gambar 4. 9** Kurva *Kaplan Meier* Faktor Radioterapi

Kurva *Survival* antara kelompok pasien yang melakukan radioterapi dengan pasien yang tidak melakukan radioterapi terlihat saling berpotongan dan diduga tidak terdapat perbedaan yang signifikan antara kedua kelompok tersebut. Selanjutnya akan dilakukan pengujian *log-rank* untuk mengetahui secara pasti apakah kedua kelompok pasien tersebut (yang melakukan radioterapi dan yang tidak melakukan radioterapi) berbeda secara signifikan atau tidak. Berikut merupakan hasil pengujian *log-rank* terhadap variabel radioterapi.

**Tabel 4. 6** Hasil Uji *Log-Rank* Faktor Radioterapi

$\chi^2$	df	<i>p-value</i>
0,000	1	1,000

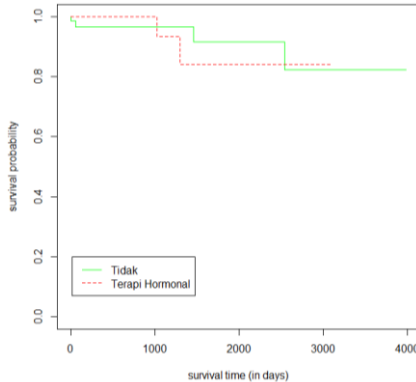
Hasil pengujian *log-rank* yang disajikan dalam Tabel 4.6 menunjukkan nilai  $\chi^2$  sebesar 0,000 dengan derajat bebas 1 dimana nilai tersebut lebih kecil dari 3,841 ( $\chi^2_{(0,05, 1)}$ ) selain itu nilai *p-value* yang diperoleh ialah 1,000 dimana nilai tersebut lebih besar daripada  $\alpha$  (0,05) sehingga dapat disimpulkan bahwa tidak terdapat perbedaan yang signifikan antara kurva *survival* pasien yang melakukan radioterapi dengan pasien yang tidak melakukan radioterapi.

#### 4.2.5 Terapi Hormonal

Variabel terapi hormonal terdiri dari dua kategori yaitu pasien yang melakukan terapi hormonal dan tidak melakukan terapi hormonal. Untuk mengetahui apakah kedua kategori tersebut berbeda secara signifikan, digunakan kurva *Kaplan-Meier* dan selanjutnya diuji *log-rank*. Berikut merupakan kurva *Kaplan-Meier* variabel terapi hormonal.

Gambar 4.10 menunjukkan bahwa kurva *survival* antara kelompok pasien yang melakukan terapi hormonal dengan pasien yang melakukan terapi hormonal saling berpotongan yang mengidentifikasi tidak terdapat perbedaan yang signifikan

antara kurva *survival* pasien yang melakukan terapi hormonal dan yang tidak melakukan terapi hormonal.



**Gambar 4. 10** Kurva *Kaplan Meier* Faktor Terapi Hormonal

Untuk memperkuat dugaan bahwa tidak terdapat perbedaan antara pasien yang melakukan terapi hormonal dengan pasien yang melakukan terapi hormonal selanjutnya akan diuji menggunakan *log-rank* dan didapatkan hasil sebagai berikut.

**Tabel 4. 7** Hasil Uji *Log-Rank* Faktor Terapi Hormonal

$\chi^2$	df	<i>p-value</i>
0,000	1	0,900

Berdasarkan Tabel 4.7 dapat dilihat bahwa nilai  $\chi^2$  yang diperoleh ialah sebesar 0,000 dengan derajat bebas sebesar 1 dimana nilai tersebut lebih kecil dari 3,841 ( $\chi^2_{(0,05, 1)}$ ) dan nilai *p-value* sebesar 0,900 yang merupakan lebih besar dibandingkan 0,05. Dari hasil tersebut dapat disimpulkan bahwa kurva *survival* antara pasien yang melakukan terapi hormonal dengan pasien yang tidak melakukan terapi hormonal tidak memiliki perbedaan yang signifikan.

### 4.3 Regresi Cox Proportional Hazard

Regresi *Cox proportional hazard* digunakan untuk mengetahui faktor-faktor apa saja yang diduga mempengaruhi waktu *survival* pasien kanker tiroid. Berikut merupakan hasil pengujian serentak regresi *Cox proportional hazard*.

**Tabel 4. 8** Hasil Uji Serentak Regresi *Cox Proportional Hazard*

<i>Likelihood Ratio Test</i>	<i>df</i>	<i>p-value</i>
13,010	7	0,070

Pada uji serentak didapatkan nilai *likelihood ratio test* sebesar 41,860 dimana nilai tersebut lebih kecil dari  $14,067 = \chi^2_{(0,05, 7)}$  dengan nilai *p-value* 0,000 dimana nilai tersebut lebih besar dari  $\alpha = 0,05$ . Nilai *p-value* lebih besar dari  $\alpha = 0,05$  berarti pada model penuh tidak terdapat minimal satu variabel yang berpengaruh signifikan terhadap waktu *survival* pasien kanker tiroid. Untuk lebih pastinya, dilakukan pula uji parsial untuk mengetahui variabel mana saja yang signifikan berpengaruh terhadap waktu *survival* pasien kanker tiroid untuk mengalami metastasis. Berikut merupakan hasil pengujian secara parsial.

**Tabel 4. 9** Hasil Uji Parsial Regresi *Cox Proportional Hazard*

<b>Variabel</b>	<b>Koefisien</b>	<b>Standard Error</b>	<b>P-Value</b>
Laki-laki	3,100	1,329	0,020*
Usia	-0,032	0,039	0,409
Stadium 2	1,183	1,285	0,357
Stadium 3	0,397	1,619	0,806
Operasi	-0,423	1,416	0,765
Radioterapi	-1,670	1,478	0,259
Terapi Hormonal	1,120	1,224	0,360

\* signifikan pada  $\alpha = 0,05$

Berdasarkan Tabel 4.9 dapat dilihat bahwa pada model ini terdapat 1 variabel dari 6 variabel yang memiliki nilai *p-value* dibawah  $\alpha (0,05)$  yaitu variabel jenis kelamin. Selanjutnya akan dilakukan analisis untuk mendapatkan model terbaik.

### 4.3.1 Model Terbaik

Pemilihan model regresi *Cox* terbaik dilakukan dengan menggunakan analisis *backward elimination*. Cara kerja *backward elimination* ini pertama dengan memasukkan semua variabel kedalam model dan dihitung nilai AIC-nya. Kemudian satu per satu variabel dikeluarkan dari model dan dihitung AIC nya kembali sampai ditemukan model dengan nilai AIC terkecil. Hasil analisis menggunakan *backward elimination* didapatkan model terbaik sebagai berikut.

**Tabel 4. 10** Model Terbaik Berdasarkan Backward

Variabel	AIC
Jenis Kelamin, Usia, Stadium, Operasi, Radioterapi, Terapi Hormonal	43,41
Jenis Kelamin, Usia, Stadium, Radioterapi, Terapi Hormonal	41,42
Jenis Kelamin, Usia, Radioterapi, Terapi Hormonal	39,66
Jenis Kelamin, Usia, Radioterapi	38,18
Jenis Kelamin, Usia	36,96

Berdasarkan model terbaik yang diperoleh dari *backward elimination* dalam Tabel 4.12, didapatkan model yang terbaik dengan kriteria nilai AIC terkecil yaitu pada model jenis kelamin dan usia. Selanjutnya dilakukan uji signifikansi model terbaik secara serentak maupun parsial. Pengujian signifikansi secara serentak untuk model terbaik ialah sebagai berikut.

**Tabel 4. 11** Hasil Uji Serentak Model Terbaik

<i>Likelihood Ratio Test</i>	<i>df</i>	<i>p-value</i>
10,820	2	0,004

Hasil pengujian serentak untuk model terakhir didapatkan nilai *likelihood ratio test* sebesar 10,820 dengan derajat bebas 2 dimana nilai tersebut lebih besar dari 5,991 ( $\chi^2_{(0,05, 2)}$ ) dan nilai *p*-

*value* sebesar 0,004 dimana nilai tersebut lebih kecil dari  $\alpha$  (0,05) yang berarti terdapat minimal 1 variabel prediktor yang mempengaruhi waktu *survival* pasien kanker tiroid. Untuk mengetahui variabel mana saja yang secara signifikan mempengaruhi waktu *survival* pasien kanker tiroid, perlu dilakukan pengujian secara parsial. Berikut merupakan hasil pengujian model terbaik secara parsial.

**Tabel 4. 12** Hasil Uji Parsial Model Terbaik

Variabel	Koefisien	Hazard Ratio	<i>p-value</i>
Laki-laki	2,575	13,136	0,019*
Usia	-0,047	0,954	0,161

\* signifikan pada  $\alpha = 0,05$

Berdasarkan Tabel 4.12, dapat dilihat bahwa variabel operasi memiliki nilai *p-value* lebih kecil dibandingkan  $\alpha$  (0,05), sehingga diperoleh model regresi *Cox proportional hazard* sebagai berikut.

$$\hat{h}(t, \mathbf{X}) = \hat{h}_0(t) \exp(2,575x_{1(t)} - 0,047x_2)$$

Nilai *Hazard Ratio* dari variabel yang signifikan yaitu variabel jenis kelamin yaitu pada jenis kelamin laki-laki dimana pasien kanker tiroid berjenis kelamin laki-laki memiliki resiko untuk mengalami metastasis 13,136 kali dibandingkan pasien kanker tiroid berjenis kelamin perempuan.

#### 4.3.2 Uji Asumsi *Proportional Hazard*

Pengujian asumsi *proportional hazard* dari model regresi *Cox proportional hazard* terbaik yang telah diperoleh dari *backward elimination* adalah sebagai berikut.

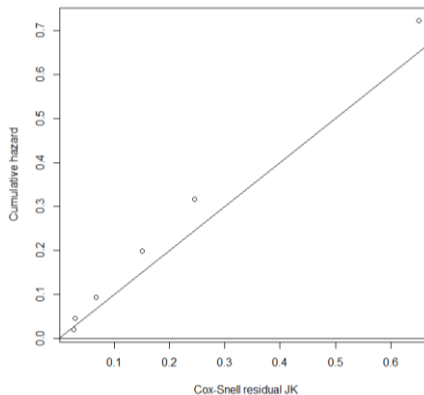
**Tabel 4. 13** Uji Asumsi *Proportional Hazard* pada Model Terbaik

Variabel	Korelasi	<i>p-value</i>
Laki-laki	-0,398	0,342
Usia	-0,128	0,808
Global		0,625

Tabel 4.13 menunjukkan hasil pengujian asumsi *proportional hazard*. Tabel 4.13 menunjukkan bahwa semua variabel memiliki nilai *p-value* yang lebih besar dibandingkan  $\alpha$  (0,05). Hal ini menandakan bahwa semua variabel telah memenuhi asumsi *proportional hazard*. Nilai *p-value* secara Global ialah sebesar 0,625 dimana nilai tersebut lebih besar dibandingkan  $\alpha$  (0,05) sehingga dapat dikatakan bahwa model regresi telah sesuai.

#### 4.4 Cox-Snell Residual

Grafik *cox-snell residual* dapat digunakan untuk mengetahui apakah suatu variabel telah memenuhi asumsi *proportional hazard* atau tidak. *cox-snell* residual dari variabel-variabel yang diduga berpengaruh pada model terbaik yang diduga mempengaruhi waktu *survival* pasien kanker tiroid ditunjukkan berdasarkan Gambar 4.11.



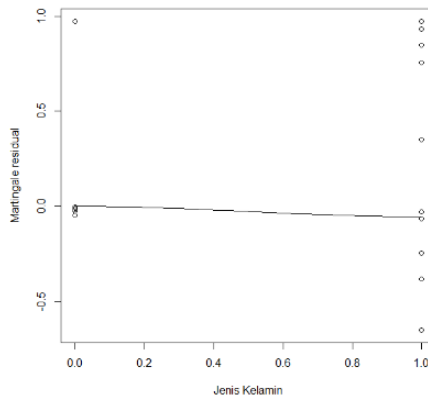
**Gambar 4. 11** Grafik *Cox-Snell Residual* Variabel Jenis Kelamin

Grafik *Cox-Snell* pada Gambar 4.11 menunjukkan bahwa plot residual masing-masing variabel berada disekitar garis diagonal yang menandakan variabel jenis kelamin dan usiatelah memenuhi uji asumsi *proportional hazard*.



#### 4.5 *Martingale Residual*

Selain menggunakan *cox-snell*, dapat pula menggunakan *martingale residual*. Pada *martingale residual*, suatu variabel dianggap telah memenuhi asumsi *proportional hazard* apabila membentuk garis lurus disekitar 0 dengan toleransi  $\pm 0.5$  (Klein dan Moeschberger, 2003). *Martingale residual* pada variabel-variabel yang terbentuk pada model terbaik adalah sebagai berikut.



**Gambar 4. 12** Grafik *Martingale Residual* Variabel Jenis Kelamin

Gambar 4.12 menunjukkan variabel jenis kelamin dan usia membentuk garis lurus pada angka 0 dan sekitarnya (batas toleransi  $\pm 0.5$ ) yang menandakan telah linier terhadap *hazard ratio* pada model regresi, sehingga asumsi linieritas pada regresi *Cox proportional hazard* terpenuhi.

*(Halaman ini sengaja dikosongkan)*

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan analisis dan pembahasan yang telah dilakukan, dapat diambil kesimpulan sebagai berikut.

1. Variabel yang memiliki perbedaan waktu survival sampai dengan terjadinya metastasis berdasarkan kurva *survival* secara signifikan ialah variabel jenis kelamin, yakni perbedaan waktu survival antara laki-laki dan perempuan ialah berbeda signifikan dengan perempuan memiliki peluang untuk bertahan dari terjadinya metastasis lebih tinggi dibandingkan dengan laki-laki.
2. Model terbaik yang terbentuk berdasarkan *backward elimination* ialah model variabel jenis kelamin dan usia dengan nilai AIC terkecil yaitu 36,96. Pada uji signifikansi secara serentak dan parsial diperoleh hasil variabel jenis kelamin berpengaruh signifikan terhadap model dan diperoleh nilai *hazard ratio* untuk pasien berjenis kelamin laki-laki memiliki laju resiko untuk mengalami metastasis 13,136 kali dibandingkan pasien kanker tiroid berjenis kelamin perempuan.
3. Pengujian asumsi *proportional hazard* menggunakan *cox-snell residual* menghasilkan kesimpulan bahwa variabel prediktor pada model terbaik telah memenuhi asumsi *proportional hazard* karena membentuk plot residual disekitar garis diagonal.
4. Pengujian asumsi *proportional hazard* dan asumsi *hazard ratio* linier terhadap variabel prediktor pada model regresi terbaik menggunakan *martingale residual* telah terpenuhi karena membentuk garis di sumbu 0 dan sekitarnya.

## 5.2 Saran

Berdasarkan analisis yang telah dilakukan, pembahasan dan kesimpulan yang telah dipaparkan, saran untuk penelitian selanjutnya adalah sebagai berikut.

1. Menambah jumlah data pasien lebih banyak lagi agar hasil yang diolah dapat lebih representatif dan lebih akurat serta memperinci penggolongan pasien sesuai dengan ilmu kesehatan agar data yang digunakan lebih homogen.
2. Apabila asumsi *Proportional Hazard* tidak terpenuhi maka lebih baik menggunakan model *Extended Cox*.
3. Apabila asumsi linieritas tidak terpenuhi pada *Cox-Snell* dan *Martingale Residual* maka lebih baik menggunakan model LOWESS, Kernel atau *Spline*.
4. Hasil pada penelitian ini hanya difokuskan untuk pembacaan data secara statistika saja dan tidak dapat digunakan sebagai acuan atau dasar pertimbangan untuk penanganan penyakit khususnya di bidang kanker.

## DAFTAR PUSTAKA

- Breheeny, P. (2013). *Residuals and Model Diagnostics*. myweb.uiowa.edu: <http://myweb.uiowa.edu/pbreheeny/721-0/f15/notes/11-10.pdf>. (diakses pada 01 Februari 2019).
- Collet, D. (1994). *Modelling Survival Data in Medical Research*. London: Chapman & Hall.
- Collet, D. (2003). *Modelling Survival Data in Medical Research*. New York: Chapman & Hall/Crc.
- Fernandes, A. R., & Solimun. (2016). *Pemodelan Statistika Pada Analisis Reliabilitas dan Survival*. Malang: UB Press.
- Hosmer, D. W., Lameshow, J. S., & May, S. (2008). *Applied Survival Analysis*. Canada: John Wiley & Sons, Inc.
- International Agency for Research on Cancer . (2018). *Estimated cancer incidence, mortality, and prevalence world wide in 2018*. GLOBOCAN.
- Klein, J. P., & Moeschberger, M. L. (2003). *Analysis Techniques for Censored and Truncated Data Second Edition*. New York: Springer.
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis*. New York: Springer.
- Laseduw, J. (2012). Kanker Tiroid. <http://kankertiroid.wordpress.com/2013/06/21/penyebab> (diunduh tanggal 01 Februari 2019).
- Lee, E. T., & Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis Third Edition*. New Jersey: John Wiley & Son, Inc.
- Li, Y.-R., Chen, S.-T., Hseuh, C., Chao, T.-C., Ho, T.-Y., & Lin, J.-D. (2016). Risk Factors of Distant Metastasis in The Follicular Variant of Papillary Thyroid Carcinoma. *Journal of the Formosan Medical Association*, 115, 665-671.
- Murray, R. K., Granner, D. K., Mayes, P. A., & Rodwell, V. W. (1997). *Biokimia Harper*. Jakarta: EGC Penerbit Buku Kedokteran.

- Murti, B. (1997). *Prinsip dan Metodologi Riset Epidemiologi*. Yogyakarta: UGM.
- Pasaribu, E. T. (2006). Epidemiologi dan Gambaran Klinis Kanker Tiroid. *Majalah Kedokteran Nusantara*, Vol 39. No 3, 270-273 .
- Stevenson, M. (2009). *An Introduction to Survival Analysis*. New Zealand: Massey University.
- Sung, M. K., Lee, Y.-M., Sung, T.-Y., Yoon, J. H., Chung, K.-W., & Hong, S. J. (2016). Risk Factors for Distant Metastasis in Patients with Follicular Thyroid Carcinoma. *The Korean Journal of Endocrine Surgery* , 16:25-30.
- Tableman, T. M., & Kim, J. S. (2004). *Survival Analysis Using S*: Boca Raton: Chapman &Hall/CRC.
- Therneau, T. M., & Grambsch, P. M. (2000). *Modelling Survival Data: Extending the Cox Model*. New York: Springer-Verlag.

## LAMPIRAN

### Lampiran 1 Data Penelitian

No	T	d	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	1072	0	0	16	0	1	0	1
2	56	1	1	17	1	1	0	0
3	3005	0	0	17	0	1	0	1
4	2541	1	1	18	0	0	0	0
5	395	0	0	18	2	1	1	0
6	214	0	0	18	0	1	0	0
7	485	0	0	20	0	0	0	0
8	126	0	0	21	0	0	0	0
9	3745	0	0	22	0	1	0	0
10	61	0	0	23	0	0	0	0
11	12	0	0	25	0	1	0	0
12	1755	0	0	25	0	1	0	0
13	834	0	1	27	1	1	0	0
14	972	0	0	27	0	1	0	0
15	368	0	0	28	0	1	0	1
16	1004	0	0	28	0	1	1	1
17	1096	0	0	28	0	0	0	0
18	1346	0	1	29	1	1	1	1
19	11	0	0	29	0	0	0	0
20	54	0	0	29	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
93	121	0	0	63	2	1	0	0
94	208	0	1	65	2	1	1	0
95	35	0	1	65	0	0	0	0
96	965	0	0	66	2	1	0	1
97	14	0	0	66	1	1	0	0
98	7	0	1	68	2	0	0	0
99	2578	0	0	68	2	1	0	0
100	920	0	0	73	2	1	0	1
101	314	0	1	76	1	1	0	0
102	357	0	1	77	2	0	0	0
103	7	0	1	79	2	1	0	1

**Lampiran 2** Syntax R Kurva *Kaplan Meier* dan Uji *Log-Rank*

```

library(survival)
library(KMsurv)
library(ggplot2)
library(MASS)

data=read.csv("D:/TA/TA Riska/Olah
              Data/Bismillah/data1.csv",sep=";",header=T)
data

#Kaplan Meier JK
win.graph()
KMJK=survfit(Surv(Survt,Status)~JK,data=data)
plot(KMJK, col=c("green","red"), lty=1:3, xlab="survival time (in days)",
      ylab="survival probability")
legend(17, 0.2, c("Perempuan", "Laki-laki"), col=c("green","red"), lty = 1:3)
#LogRank JK
LRJK=survdiff(Surv(Survt, Status)~JK, data=data, rho=1)
LRJK

#Kaplan Meier Stage
win.graph()
KMStage=survfit(Surv(Survt,Status)~Stage,data=data)
plot(KMStage, col=c("green","red","blue"), lty=1:3, xlab="survival time (in
days)", ylab="survival probability")
legend(17, 0.2, c("Stadium 1", "Stadium 2", "Stadium 3"),
      col=c("green","red","blue "), lty = 1:3)
#LogRank Stage
LRStage=survdiff(Surv(Survt, Status)~Stage, data=data, rho=1)
LRStage

#Kaplan Meier Hormon Therapy
win.graph()
KMHT=survfit(Surv(Survt,Status)~Hormonaltherapy,data=data)
plot(KMHT, col=c("green","red"), lty=1:3, xlab="survival time (in days)",
      ylab="survival probability")
legend(17, 0.2, c("Tidak", "Terapi Hormonal"), col=c("green","red"), lty =
1:3)

```

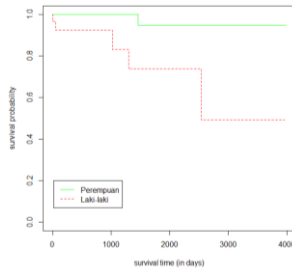


```
#LogRank HT
LRHT=survdiff(Surv(Survt, Status)~HormonalTherapy, data=data, rho=1)
LRHT

#Kaplan Meier Surgery
win.graph()
KMS=survfit(Surv(Survt,Status)~Surgery,data=data)
plot(KMS, col=c("green","red"), lty=1:3, xlab="survival time (in days)",
      ylab="survival probability")
legend(17, 0.2, c("Tidak", "Operasi"), col=c("green","red"), lty = 1:3)
#LogRank Surgery
LRSur=survdiff(Surv(Survt, Status)~Surgery, data=data, rho=1)
LRSur

#Kaplan Meier Radiotherapy
win.graph()
KMR=survfit(Surv(Survt,Status)~Radiotherapy,data=data)
plot(KMR, col=c("green","red"), lty=1:3, xlab="survival time (in days)",
      ylab="survival probability")
legend(17, 0.2, c("Tidak", "Radioterapi"), col=c("green","red"), lty = 1:3)
#LogRank Radiotherapy
LRRad=survdiff(Surv(Survt, Status)~Radiotherapy, data=data, rho=1)
LRRad
```

### Lampiran 3 Hasil Kurva *Kaplan Meier* dan Uji *Log-Rank*



Call:

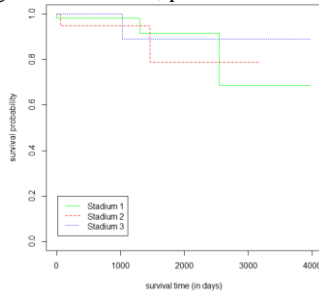
```
survdif(formula = Surv(Survt, Status) ~ JK, data = data, rho = 1)
```

N Observed Expected (O-E)<sup>2</sup>/E (O-E)<sup>2</sup>/V

JK=0 74 0.922 4.22 2.58 10.2

JK=1 29 4.805 1.51 7.21 10.2

Chisq= 10.2 on 1 degrees of freedom, p= 0.001



Call:

```
survdif(formula = Surv(Survt, Status) ~ Stage, data = data,  
rho = 1)
```

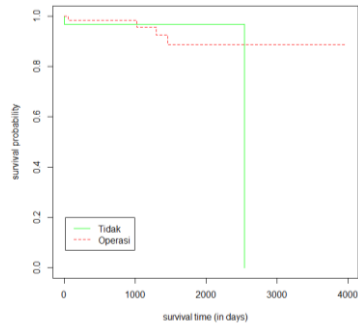
N Observed Expected (O-E)<sup>2</sup>/E (O-E)<sup>2</sup>/V

Stage=0 55 2.837 2.65 0.0135 0.0268

Stage=1 21 1.913 1.39 0.1940 0.2690

Stage=2 27 0.978 1.69 0.2981 0.4552

Chisq= 0.5 on 2 degrees of freedom, p= 0.8



Call:

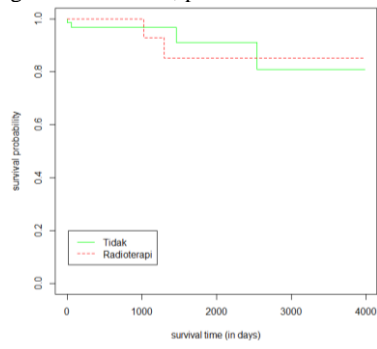
```
survdif(formula = Surv(Survt, Status) ~ Surgery, data = data,
rho = 1)
```

N Observed Expected (O-E)<sup>2</sup>/E (O-E)<sup>2</sup>/V

Surgery=0 31 1.89 0.731 1.825 2.33

Surgery=1 72 3.84 4.997 0.267 2.33

Chisq= 2.3 on 1 degrees of freedom, p= 0.1



Call:

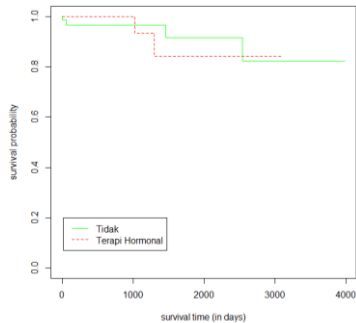
```
survdif(formula = Surv(Survt, Status) ~ Radiotherapy, data = data,
rho = 1)
```

N Observed Expected (O-E)<sup>2</sup>/E (O-E)<sup>2</sup>/V

Radiotherapy=0 77 3.80 3.86 0.00115 0.00373

Radiotherapy=1 26 1.93 1.86 0.00238 0.00373

Chisq= 0 on 1 degrees of freedom, p= 1



Call:

```
survdif(formula = Surv(Survt, Status) ~ Hormonaltherapy, data = data,
rho = 1)
```

	N	Observed	Expected	$(O-E)^2/E$	$(O-E)^2/V$
Hormonaltherapy=0	72	3.80	3.95	0.00617	0.0211
Hormonaltherapy=1	31	1.93	1.77	0.01375	0.0211

Chisq= 0 on 1 degrees of freedom, p= 0.9

**Lampiran 4** Syntax Regresi Cox Proportional Hazard

```
library(survival)
library(KMsurv)
library(ggplot2)
library(MASS)

data=read.csv("D:/TA/TA Riska/Olah
              Data/Bismillah/data1.csv",sep=";",header=T)
data

#Regresi cox PH Model Penuh
coxph1 <- coxph(formula=Surv(Survt, Status) ~ as.factor(JK) + Usia +
                as.factor(Stage) + as.factor(Surgery) + as.factor(Radiotherapy) +
                as.factor(Hormonaltherapy),data=data)
summary(coxph1)

#Mendapatkan model terbaik
finalfixx = step(coxph(formula=Surv(Survt, Status) ~ JK + Usia + Stage +
                    Surgery + Radiotherapy + Hormonaltherapy, data=data), direction
                  = 'backward')

#Regresi cox PH Model Terbaik
coxph2 <- coxph(formula=Surv(Survt,
                             Status)~as.factor(JK)+as.factor(JK)+Usia,data=data)
summary(coxph2)
```

**Lampiran 5** Hasil Syntax R Regresi Cox Proportional Hazard

Call:

```
coxph(formula = Surv(Survt, Status) ~ as.factor(JK) + Usia1 +
  as.factor(Stage) + as.factor(Surgery) + as.factor(Radiotherapy) +
  as.factor(Hormonaltherapy), data = data)
```

n= 103, number of events= 6

	coef	exp(coef)	se(coef)	z	Pr(> z )
as.factor(JK)1	3.10009	22.19991	1.32887	2.333	0.0197 *
Usia1	-0.03194	0.96857	0.03865	-0.826	0.4086
as.factor(Stage)1	1.18323	3.26491	1.28449	0.921	0.3570
as.factor(Stage)2	0.39717	1.48761	1.61881	0.245	0.8062
as.factor(Surgery)1	-0.42316	0.65497	1.41554	-0.299	0.7650
as.factor(Radiotherapy)1	-1.66945	0.18835	1.47826	-1.129	0.2588
as.factor(Hormonaltherapy)1	1.12081	3.06734	1.22433	0.915	0.3600

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
as.factor(JK)1	22.1999	0.04505	1.64141	300.252
Usia1	0.9686	1.03245	0.89790	1.045
as.factor(Stage)1	3.2649	0.30629	0.26334	40.479
as.factor(Stage)2	1.4876	0.67222	0.06231	35.516
as.factor(Surgery)1	0.6550	1.52678	0.04086	10.499
as.factor(Radiotherapy)1	0.1883	5.30927	0.01039	3.414
as.factor(Hormonaltherapy)1	3.0673	0.32602	0.27836	33.800

Concordance= 0.923 (se = 0.033 )

Likelihood ratio test= 13.01 on 7 df, p=0.07

Wald test = 10.12 on 7 df, p=0.2

Score (logrank) test = 16.64 on 7 df, p=0.02

```
> finalfixusia1 = step(coxph(formula=Surv(Survt, Status)~JK+Usia+Stage+Surgery+Radiotherapy+Hormonaltherapy, data=data), direction = 'backward')
Start: AIC=43.41
```

```
Surv(Survt, Status) ~ JK + Usia1 + Stage + Surgery + Radiotherapy +
  Hormonaltherapy
```

	Df	AIC
- Surgery	1	41.416
- Stage	1	41.649
- Hormonaltherapy	1	41.834
- Radiotherapy	1	42.613
- Usia	1	43.039
<none>		43.410
- JK	1	49.566

```
Step: AIC=41.42
```

```
Surv(Survt, Status) ~ JK + Usia + Stage + Radiotherapy + Hormonaltherapy
```

	Df	AIC
- Stage	1	39.655
- Hormonaltherapy	1	39.857
- Radiotherapy	1	40.852
- Usia	1	41.219
<none>		41.416
- JK	1	48.598

```
Step: AIC=39.66
```

```
Surv(Survt, Status) ~ JK + Usia + Radiotherapy + Hormonaltherapy
```

	Df	AIC
- Hormonaltherapy	1	38.184
- Radiotherapy	1	38.869
- Usia	1	39.250
<none>		39.655
- JK	1	46.898

```
Step: AIC=38.18
```

```
Surv(Survt, Status) ~ JK + Usia + Radiotherapy
```

	Df	AIC
- Radiotherapy	1	36.960
- Usia1	1	37.688
<none>		38.184
- JK	1	44.904

Step: AIC=36.96

Surv(Survt, Status) ~ JK + Usia1

	Df	AIC
<none>		36.960
- Usia1	1	37.184
- JK	1	43.022

Call:

```
coxph(formula = Surv(Survt, Status) ~ as.factor(JK) + Usia1,
      data = data)
```

n= 103, number of events= 6

	coef	exp(coef)	se(coef)	z	Pr(> z )
as.factor(JK)1	2.57532	13.13557	1.09773	2.346	0.019 *
Usia1	-0.04676	0.95432	0.03334	-1.402	0.161

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
as.factor(JK)1	13.1356	0.07613	1.528	112.937
Usia1	0.9543	1.04787	0.894	1.019

Concordance= 0.871 (se = 0.055 )

Likelihood ratio test= 10.82 on 2 df, p=0.004

Wald test = 7.82 on 2 df, p=0.02

Score (logrank) test = 12.5 on 2 df, p=0.002



## Lampiran 6 *Syntax R Pengujian Asumsi Proportional Hazard*

```

library(survival)
library(MASS)

data=read.csv("D:/TA/TA Riska/Olah
              Data/Bismillah/data1.csv",sep=";",header=T)
data

#Grambsch Therneau
ks.cph=coxph(formula=Surv(Survt,Status)~JK+Stage+Surgery, data=data)
gof=cox.zph(ks.cph)
gof

#Cox-Snell Residual JK
win.graph()
coxph3<-coxph(Surv(Survt,Status)~JK,data=data)
rc3<-abs(data$Status-residuals(coxph3))
km.rc3<-survfit(Surv(rc3,Status)~1,data=data)
summary.km.rc3<-summary(km.rc3)
rcu3<-summary.km.rc3$time
surv.rc3 <- summary.km.rc3$surv
plot(rcu3,-log(surv.rc3),type="p",xlab="Cox-Snell
      residual",ylab="Cumulative hazard", main="Jenis Kelamin")
abline(a=0,b=1); abline(v=0); abline(h=0)

#Cox-Snell Residual Stage
win.graph()
coxph4<-coxph(Surv(Survt,Status)~Stage,data=data)
rc4<-abs(data$Status-residuals(coxph4))
km.rc4<-survfit(Surv(rc4,Status)~1,data=data)
summary.km.rc4<-summary(km.rc4)
rcu4<-summary.km.rc4$time
surv.rc4 <- summary.km.rc4$surv
plot(rcu4,-log(surv.rc4),type="p",xlab="Cox-Snell residual
      Stadium",ylab="Cumulative hazard",main="Stage")
abline(a=0,b=1); abline(v=0); abline(h=0)

```

```

#Cox-Snell Residual Hormonal Therapy
win.graph()
coxph5<-coxph(Surv(Survt,Status)~Hormonaltherapy,data=data)
rc5<-abs(data$Status-residuals(coxph5))
km.rc5<-survfit(Surv(rc5,Status)~1,data=data)
summary.km.rc5<-summary(km.rc5)
rcu5<-summary.km.rc5$time
surv.rc5 <- summary.km.rc5$urv
plot(rcu5,-log(surv.rc5),type="p",xlab="Cox-Snell
      residual",ylab="Cumulative hazard",main="Hormonal Therapy")
abline(a=0,b=1); abline(v=0); abline(h=0)

#Cox-Snell Residual Surgery
win.graph()
coxph6<-coxph(Surv(Survt,Status)~Surgery,data=data)
rc6<-abs(data$Status-residuals(coxph6))
km.rc6<-survfit(Surv(rc6,Status)~1,data=data)
summary.km.rc6<-summary(km.rc6)
rcu6<-summary.km.rc6$time
surv.rc6 <- summary.km.rc6$urv
plot(rcu6,-log(surv.rc6),type="p",xlab="Cox-Snell residual
      Operasi",ylab="Cumulative hazard",main="Surgery")
abline(a=0,b=1); abline(v=0); abline(h=0)

#Cox-Snell Residual Radiotherapy
win.graph()
coxph7<-coxph(Surv(Survt,Status)~Radiotherapy,data=data)
rc7<-abs(data$Status-residuals(coxph7))
km.rc7<-survfit(Surv(rc7,Status)~1,data=data)
summary.km.rc7<-summary(km.rc7)
rcu7<-summary.km.rc7$time
surv.rc7 <- summary.km.rc7$urv
plot(rcu7,-log(surv.rc7),type="p",xlab="Cox-Snell
      residual",ylab="Cumulative hazard",main="Radiotherapy")
abline(a=0,b=1); abline(v=0); abline(h=0)

```

```
#Martingale Residual JK
win.graph()
fit3 <- coxph(Surv(Survt,Status)~JK,data=data)
scatter.smooth(data$JK,resid(fit3),type="p",xlab="Jenis
Kelamin",ylab="Martingale residual")

#Martingale Residual Stage
win.graph()
fit4 <- coxph(Surv(Survt,Status)~Stage,data=data)
scatter.smooth(data$Stage,resid(fit4),type="p",xlab="Stadium",ylab="Marting
ale residual")

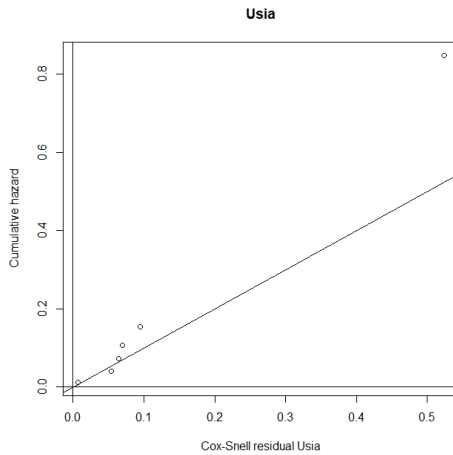
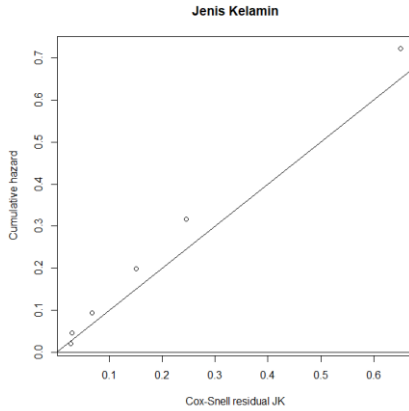
#Martingale Residual Hormonal Therapy
win.graph()
fit5 <- coxph(Surv(Survt,Status)~Hormonaltherapy,data=data)
scatter.smooth(data$Hormonaltherapy,resid(fit5),type="p",xlab="Hormonalth
erapy",ylab="Martingale residual")

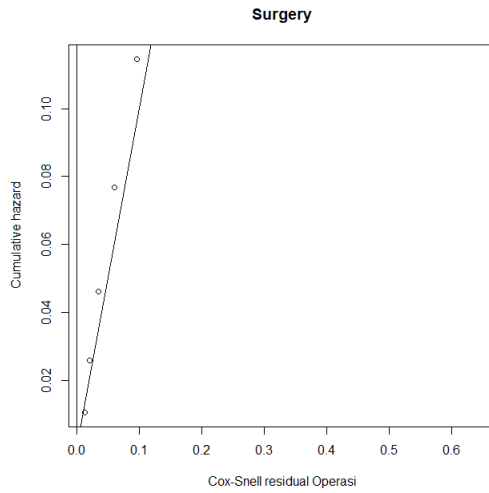
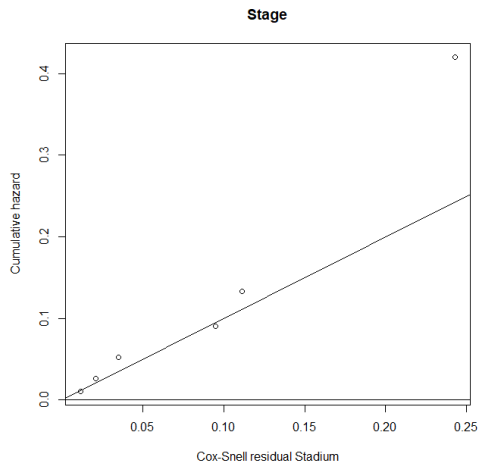
#Martingale Residual Surgery
win.graph()
fit6 <- coxph(Surv(Survt,Status)~Surgery,data=data)
scatter.smooth(data$Surgery,resid(fit6),type="p",xlab="Operasi",ylab="Marti
ngale residual")

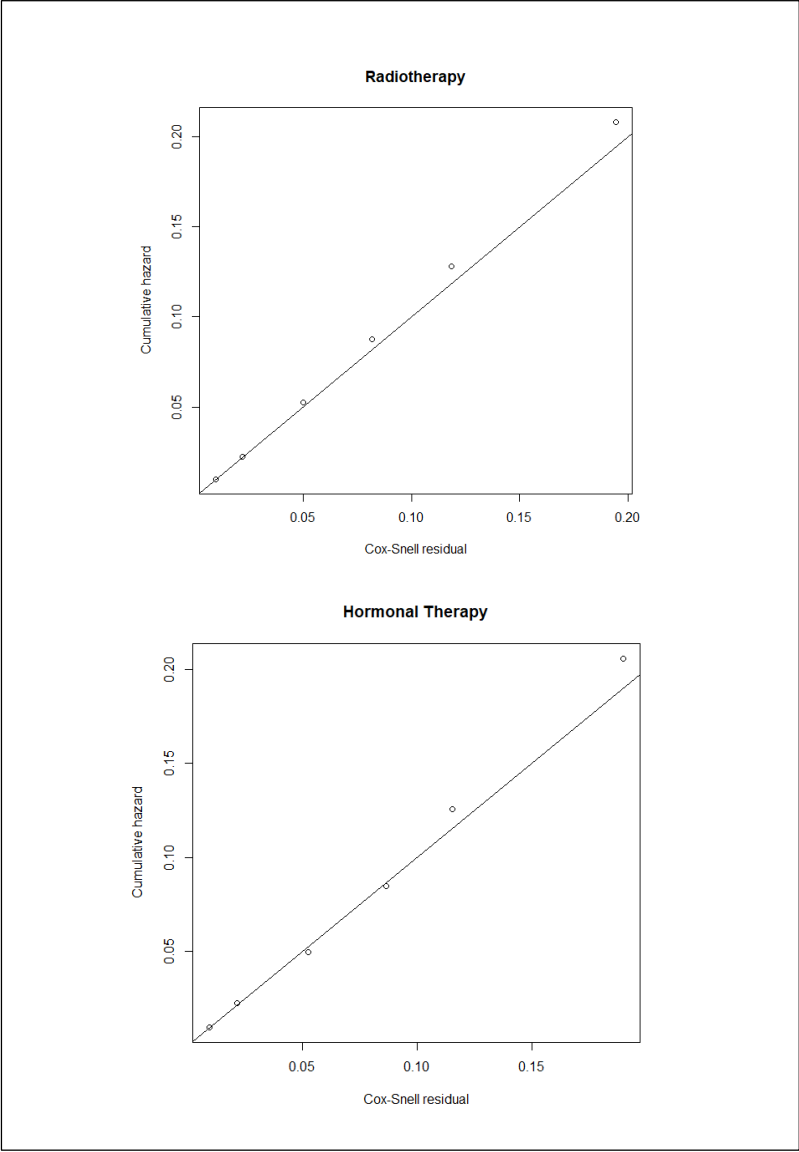
#Martingale Residual Radiotherapy
win.graph()
fit7 <- coxph(Surv(Survt,Status)~Radiotherapy,data=data)
scatter.smooth(data$Radiotherapy,resid(fit7),type="p",xlab="Radiotherapy",y
lab="Martingale residual")
```

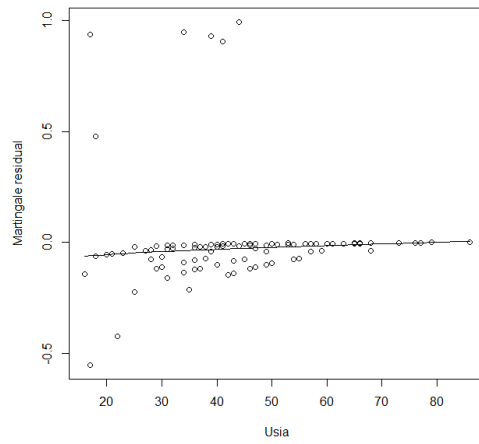
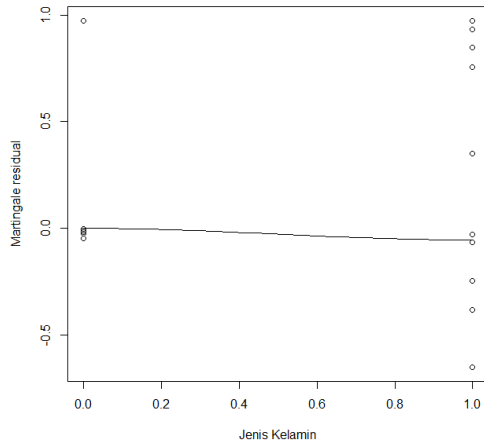
**Lampiran 7** Hasil Uji Asumsi *Proportional Hazard*

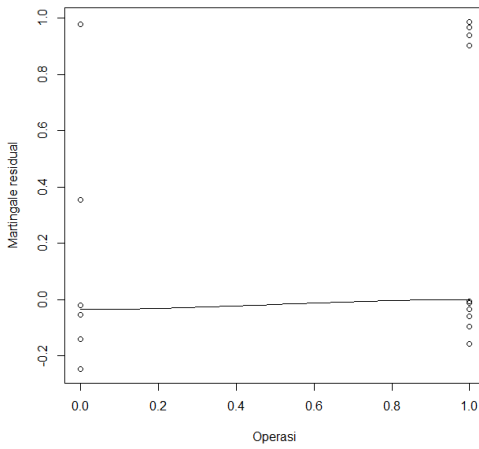
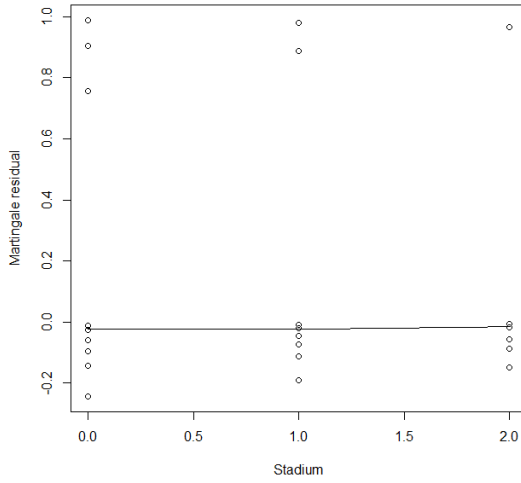
```
rho chisq p
Usia1 -0.128 0.0589 0.808
as.factor(JK)1 -0.398 0.9011 0.342
GLOBAL NA 0.9391 0.625
```



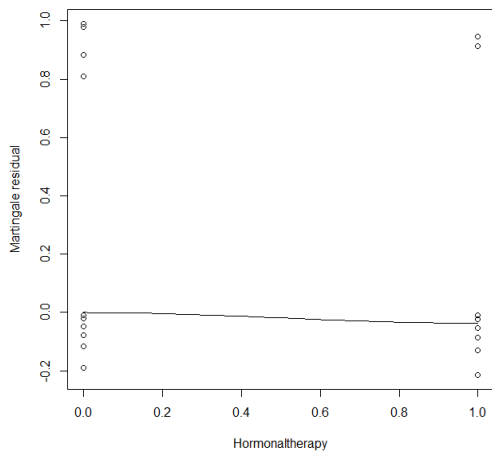
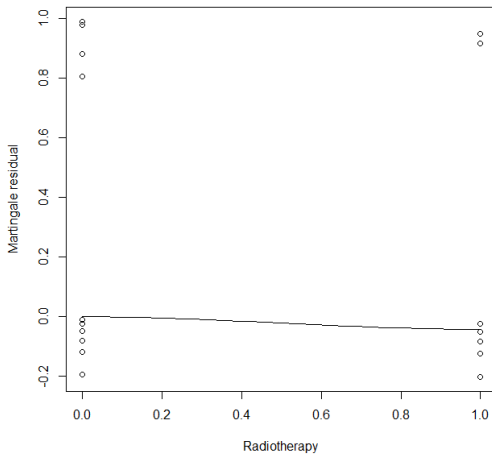












## BIODATA PENULIS



Penulis lahir di Surabaya pada tanggal 24 April 1997. Pendidikan formal yang telah ditempuh penulis ialah berawal di SD Muhammadiyah 21 tahun 2003 hingga 2009, Setelah menamatkan pendidikan sekolah dasar, penulis melanjutkan pendidikan di SMP Negeri 1 Surabaya dari tahun 2009 hingga 2012. Setelah itu penulis melanjutkan pendidikannya di SMA Negeri 2 Surabaya pada tahun 2012 hingga 2015 dan sampailah pada saat ini penulis menempuh pendidikan di Departemen Statistika Institut Teknologi Sepuluh Nopember sejak 2015 hingga sekarang. Pada masa perkuliahan di 3 semester awal, penulis aktif di organisasi luar tepatnya mengikuti bisnis secara *online* yaitu Oriflame dikarenakan penulis menyukai wirausaha dan ingin menjadi mahasiswa yang mandiri. Selain itu saat penulis menginjak semester 3, penulis juga bergabung dalam *volunteer* Humas Rektorat ITS sampai dengan pertengahan tahun 2018. Dengan bergabung di *volunteer* Humas Rektorat ITS, penulis mendapatkan banyak pengalaman dikarenakan sering mengikuti acara-acara yang diadakan oleh Humas dan Rektorat seperti menjadi panitia wisuda, kuliah tamu, dll. Pada tahun 2017/2018 penulis juga pernah menjadi anggota dari Dewan Perwakilan Mahasiswa ITS. Penulis pernah melakukan kerja praktik di Kantor Wilayah Direktorat Jenderal Pajak 1 Jawa Timur. Jika ada pembaca yang ingin berdiskusi lebih lanjut dengan penulis, dapat menghubungi alamat email [riskadevy24@gmail.com](mailto:riskadevy24@gmail.com).

