

**ANALISIS SURVIVAL MENGGUNAKAN METODE KAPLAN-
MEIER DAN COX PROPORTIONAL HAZARD PADA WAKTU
KESEMBUHAN PENDERITA TUBERKULOSIS PARU
(Studi Kasus di Kota Probolinggo)**

SKRIPSI

**OLEH
AINI INTAN NUR LAILIYAH
NIM. 220601110106**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026**

**ANALISIS SURVIVAL MENGGUNAKAN METODE KAPLAN-
MEIER DAN COX PROPORTIONAL HAZARD PADA WAKTU
KESEMBUHAN PENDERITA TUBERKULOSIS PARU
(Studi Kasus di Kota Probolinggo)**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Aini Intan Nur Lailiyah
NIM. 220601110106**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2026**

**ANALISIS SURVIVAL MENGGUNAKAN METODE KAPLAN-
MEIER DAN COX PROPORTIONAL HAZARD PADA WAKTU
KESEMBUHAN PENDERITA TUBERKULOSIS PARU
(Studi Kasus di Kota Probolinggo)**

SKRIPSI

**Oleh
Aini Intan Nur Lailiyah
NIM. 220601110106**

Telah Disetujui untuk Diuji

Malang, 27 April 2026

Dosen Pembimbing I



Dr. Fachrur Rozi, M.Si
NIP. 19800527 200801 1 012

Dosen Pembimbing II



Ari Kusumastuti, M.Pd, M.Si
NIP. 19770521 200501 2 004

Mengetahui,
Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si
NIP. 19800527 200801 1 012

**ANALISIS SURVIVAL MENGGUNAKAN METODE KAPLAN-
MEIER DAN COX PROPORTIONAL HAZARD PADA WAKTU
KESEMBUHAN PENDERITA TUBERKULOSIS PARU
(Studi Kasus di Kota Probolinggo)**

SKRIPSI

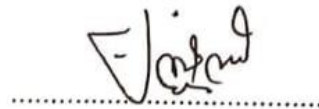
Oleh
Aini Intan Nur Lailiyah
NIM. 220601110106

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 2 Juni 2026

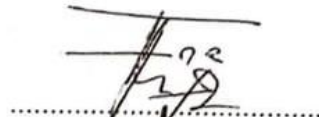
Ketua Penguji : Abdul Aziz, M.Si



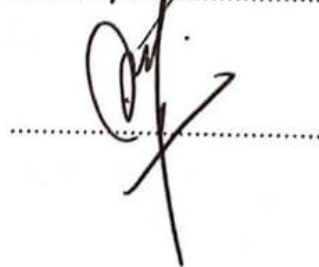
Anggota Penguji 1 : Ria Dhea Layla N.K., M.Si



Anggota Penguji 2 : Dr. Fachrur Rozi, M.Si



Anggota Penguji 3 : Ari Kusumastuti, M.Pd, M.Si



Mengetahui,
Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si

NIP. 19800527 200801 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Aini Intan Nur Lailiyah

NIM : 220601110106

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Analisis *Survival* Menggunakan Metode *Kaplan-Meier* dan *Cox Proportional Hazard* pada Waktu Kesembuhan Penderita Tuberkulosis Paru (Studi Kasus di Kota Probolinggo)

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar Pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 2 Juni 2026

Yang membuat pernyataan,

A handwritten signature in black ink is written over a blue rectangular official stamp. The stamp contains the text 'METERAI KEPER' and a serial number '16DANX425626568'.

Aini Intan Nur Lailiyah

NIM. 220601110106

MOTO

“ Maka sesungguhnya bersama kesulitan itu ada kemudahan.

Sesungguhnya bersama kesulitan itu ada kemudahan,”

(QS. Al-insyirah, 94:5-6)

“Allah tidak akan menaruh mimpi di hatimu tanpa memberimu kemampuan untuk mewujudkannya, selagi bisa diusahakan bertempurlah semaksimal mungkin”

“Jika kita gagal 1000x, maka pastikan kita bangkit 1001x”

PERSEMBAHAN

Skripsi ini penulis persembahkan kepada:

Kedua orang tua tercinta yang selalu memberikan doa, kasih sayang, dukungan, dan pengorbanan tanpa henti. Kepada keluarga dan teman-teman dekat yang selalu memberikan semangat serta menemani setiap proses perjuangan penulis.

Kepada kampus kebanggaan yang telah menjadi tempat penulis belajar dan berkembang. Dan yang paling istimewa, untuk diri sendiri yang telah mampu bertahan dan berjuang menyelesaikan perkuliahan hingga akhirnya dapat sampai pada titik ini. Terima kasih karena tidak menyerah dalam keadaan apa pun.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah Subhānahu wa Ta‘ālā atas limpahan rahmat, taufik, serta hidayah-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan judul “**Analisis *Survival* Menggunakan Metode *Kaplan-Meier* dan *Cox Proportional hazard* pada Waktu Kesembuhan Penderita Tuberkulosis Paru (Studi Kasus di Kota Probolinggo)**”. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad ﷺ ‘Alaihi Wasallam, keluarga, sahabat, dan seluruh pengikutnya hingga akhir zaman. Penyusunan naskah skripsi ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan arahan, motivasi, serta bantuan dalam bentuk apa pun. Oleh karena itu, dengan hormat penulis menyampaikan ucapan terima kasih kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah memberikan kesempatan kepada penulis untuk melaksanakan studi di UIN Malang.
2. Dr. H. Agus Mulyono, M.Si, selaku Dekan Fakultas Sains dan Teknologi, atas kebijakan serta dukungan akademik yang telah memberikan pelayanan yang baik kepada semua mahasiswa untuk kelancaran studi.
3. Dr. Fachrur Rozi, M.Si, selaku Ketua Program Studi Matematika, Dosen Pembimbing I, sekaligus Dosen Wali, yang telah memberikan bimbingan, arahan, dan motivasi yang sangat berarti bagi penulis sejak awal perkuliahan hingga proses penyelesaian skripsi ini.
4. Ari Kusumastuti, M.Pd, M.Si, selaku Dosen Pembimbing II, atas waktu, perhatian, serta saran-saran konstruktif yang membantu penulis memperbaiki dan menyempurnakan skripsi ini.
5. Abdul Aziz, M.Si dan Ria Dhea Layla N.K., M.Si selaku dosen penguji yang telah memberikan kritikan dan saran yang konstruktif demi kesempurnaan penulisan dan penyusunan proposal hingga skripsi.
6. Seluruh dosen Program Studi Matematika, yang telah memberikan ilmu, wawasan, dan inspirasi selama masa perkuliahan.

7. Bapak penulis, Indung Setya Pramana. Terima kasih atas segala perjuangan, pengorbanan, dan kasih sayang yang telah Bapak berikan sejak penulis kecil. Tanpa Bapak, penulis tidak akan menjadi pribadi yang seperti ini.
8. Pintu surgaku, Dewi Cameli Irani, S.Pd.,Gr yang selalu menjadi penyemangat penulis dan menjadi sandaran terkuat dari kerasnya dunia. Yang tidak henti-hentinya memberikan kasih sayang dan selalu memberikan motivasi yang luar biasa. Terima kasih untuk doa-doa yang selalu diberikan untuk penulis, terima kasih selalu berjuang untuk penulis, berkat doa serta dukungannya sehingga penulis bisa berada di titik ini.
9. Kepada sosok yang telah penulis anggap sebagai orang tua, Dr. Aliyah Hidayati, SpTHT-KL. Penulis mengucapkan terima kasih yang sebesar-besarnya tanpa adanya hubungan darah, namun dengan ketulusan yang begitu besar, telah memberikan segala perhatian, dukungan, dan bantuan materi yang telah diberikan sejak penulis menempuh pendidikan di bangku SMA hingga perguruan tinggi.
10. Kepada satu-satunya adik penulis, terima kasih telah menjadi salah satu alasan terbesar penulis untuk terus berjuang.
11. Kepada keluarga besar penulis, khususnya kakek dan nenek, om dan tante, kakak, keponakan, serta sepupu. Penulis mengucapkan terima kasih atas segala doa, dukungan, dan perhatian yang telah diberikan selama ini.
12. Kepada sahabat-sahabat penulis, terkhusus Sinjo yang selalu hadir dan siap menemani penulis di berbagai kondisi. Penulis juga mengucapkan terima kasih kepada seluruh teman-teman yang tidak dapat disebutkan satu per satu, baik teman ma'had, teman KKM, teman perkuliahan, teman PKL, maupun teman-teman sejak SMA. Terima kasih atas kebersamaan, dukungan, dan kenangan yang telah diberikan selama ini.
13. Rekan-rekan mahasiswa Matematika angkatan 2022, atas kebersamaan, serta dukungan moral yang turut memperkaya pengalaman belajar penulis.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi mendapatkan hasil yang terbaik. Semoga segala bentuk bantuan dan dukungan yang diberikan mendapatkan balasan terbaik dari Allah Subhānahu wa Ta‘ālā.

Malang, 2 Juni 2026

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTO	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR SIMBOL	xv
DAFTAR LAMPIRAN	xvii
ABSTRAK	xviii
ABSTRACT	xix
مستخلص البحث.....	xx
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah.....	8
BAB II KAJIAN TEORI	9
2.1 Teori Pendukung	9
2.1.1 Konsep Analisis <i>Survival</i>	9
2.1.2 Konsep Pensensoran (<i>Censoring</i>)	10
2.1.3 Analisis dan Penanganan <i>Missing Value</i>	12
2.1.4 Fungsi <i>Survival</i> dan Fungsi <i>Hazard</i>	17
2.1.5 Metode <i>Kaplan-Meier</i>	21
2.1.6 Uji <i>Log-rank</i>	24
2.1.7 Model <i>Cox Proportional Hazard</i>	26
2.1.8 Uji Asumsi <i>Cox Proportional Hazard</i>	30
2.1.9 Pengujian Parameter.....	32
2.1.10 Rasio <i>Hazard</i>	34
2.1.11 Tuberkulosis	35
2.2 Kajian Integrasi Topik dengan Al-Quran/Hadits	41
2.3 Kajian Topik dengan Teori Pendukung	45
BAB III METODE PENELITIAN	50
3.1 Jenis Penelitian.....	50
3.2 Data dan Sumber Data.....	50
3.3 Variabel Penelitian dan Struktur Data.....	51
3.4 Lokasi Penelitian	53
3.5 Teknik Analisis Data	53
BAB IV HASIL DAN PEMBAHASAN	57
4.1 <i>Preprocessing</i> Data	57
4.1.1 Identifikasi <i>Missing Value</i>	57
4.1.2 Identifikasi Mekanisme <i>Missing Data</i>	58

4.1.3	Penanganan <i>Missing Value</i> dengan Metode <i>Multiple Imputation</i>	59
4.1.4	Evaluasi Hasil Imputasi.....	62
4.2	Analisis Deskriptif.....	65
4.3	Analisis <i>Kaplan-Meier</i>	69
4.3.1	Kurva <i>Kaplan-Meier</i>	70
4.3.2	Uji <i>Log-rank</i>	77
4.4	Analisis Regresi <i>Cox Proportional Hazard</i>	82
4.4.1	Pembentukan Model <i>Cox Proportional Hazard</i>	82
4.4.2	Estimasi Parameter Model.....	83
4.4.3	Uji Asumsi <i>Cox Proportional Hazard</i>	85
4.4.4	<i>Pooling Rubin's Rule</i>	90
4.4.5	Uji Signifikansi Parameter Model	91
4.4.6	<i>Hazard Ratio</i>	93
BAB V	PENUTUP	97
5.1	Kesimpulan.....	97
5.2	Saran.....	98
DAFTAR PUSTAKA		99
LAMPIRAN		102
RIWAYAT HIDUP		123

DAFTAR TABEL

Tabel 3.1	Variabel Penelitian	51
Tabel 3.2	Pembentukan Variabel Dummy untuk X_5	52
Tabel 3.3	Struktur Data Penelitian	53
Tabel 4.1	Jumlah <i>Missing Value</i> pada Setiap Variabel	57
Tabel 4.2	Hasil Uji Mekanisme <i>Missing Data</i>	59
Tabel 4.3	Pola <i>Missing Data</i>	60
Tabel 4.4	Perbandingan Proporsi Kategori Sebelum & Sesudah Imputasi	64
Tabel 4.5	Hasil Uji <i>Log-rank</i>	78
Tabel 4.6	Jumlah Iterasi <i>Cox Proportional Hazard</i>	83
Tabel 4.7	Hasil Uji Asumsi <i>Cox Proportional Hazard</i> Model Pertama	86
Tabel 4.8	Hasil Uji Asumsi <i>Cox Proportional Hazard</i> Model Kedua	87
Tabel 4.9	Estimasi Parameter <i>Cox Proportional Hazard</i> Model Final	88
Tabel 4.10	Hasil Uji Asumsi <i>Cox Proportional Hazard</i> Model Final	89
Tabel 4.11	Hasil <i>Pooling Rubin's Rule</i> Model Final	90
Tabel 4.12	Hasil Uji Serentak Model Final pada Tiap Dataset Imputasi	92
Tabel 4.13	Hasil Uji Wald Model Final	92
Tabel 4.14	Nilai <i>Hazard Ratio</i> Tiap Variabel pada Model Final	93

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Kurva Fungsi <i>Survival</i> saat $t = 0$ dan $t = \infty$	18
Gambar 2.2 Ilustrasi Kurva <i>Kaplan-Meier</i> terhadap Waktu Kesembuhan	22
Gambar 3.1 <i>Flowchart</i> Teknik Analisis Data.....	56
Gambar 4.1 <i>Stripplot</i> Sebaran Nilai pada Lima Dataset Hasil Imputasi	62
Gambar 4.2 <i>Trace Plot</i> Konvergensi <i>Mean</i> dan Standar Deviasi Hasil Imputasi MICE pada Variabel X_4 , X_7 , dan X_8	63
Gambar 4.3 <i>Pie Chart</i> dari Masing-Masing Variabel Independen.....	66
Gambar 4.4 <i>Pie Chart</i> dari Variabel Status Kejadian	69
Gambar 4.5 Kurva <i>Kaplan-Meier</i> pada Dataset Imputasi 1	71

DAFTAR SIMBOL

T	: Waktu ketahanan atau waktu hingga suatu <i>event</i> terjadi (variabel acak)
χ^2	: Nilai statistik uji <i>Chi-Square</i>
t	: Titik waktu tertentu dalam pengamatan
O	: Frekuensi observasi
E	: Frekuensi harapan
$\hat{\theta}_i$	: Estimasi parameter pada dataset imputasi ke- i
$\bar{\theta}$	: Rata-rata estimasi parameter dari seluruh dataset imputasi
U_i	: Variansi estimasi parameter pada dataset imputasi ke- i
$\bar{U}$	: Rata-rata variansi dalam imputasi
B	: Variansi antar imputasi
VT	: Variansi total (kombinasi variansi dalam dan antar imputasi)
$S(t)$	: Fungsi <i>survival</i>
$F(t)$	: Fungsi distribusi kumulatif/ CDF
$f(u)du$	: Peluang terjadinya <i>event</i> di sekitar waktu u di dalam integral
$P(T \leq t)$	: Peluang bahwa kejadian terjadi sebelum atau sama dengan waktu t
$h(t)$	: Fungsi <i>hazard</i> sesaat
Δ_t	: Panjang selang waktu yang sangat kecil setelah waktu t
$f(t)$	: Fungsi densitas probabilitas/ PDF
$S(u)$	: Nilai fungsi <i>survival</i> pada waktu sementara u di dalam integral
$H(t)$	: Fungsi <i>hazard</i> kumulatif
$\hat{S}(t_{(j)})$	: Estimasi fungsi <i>survival</i> pada waktu kejadian ke- j
$\hat{S}(t_{(j-1)})$	: Estimasi fungsi <i>survival</i> pada waktu kejadian sebelum $t_{(j)}$
d_j	: Banyaknya individu yang mengalami <i>event</i> pada waktu $t_{(j)}$
b_j	: Banyaknya individu yang belum mengalami <i>event</i> tepat sebelum $t_{(j)}$
C	: Banyaknya kategori
d_i	: Banyaknya individu yang mengalami kejadian pada kategori ke- i
E_i	: Nilai ekspektasi individu yang mengalami <i>event</i> pada kategori ke- i
J	: Banyaknya waktu kejadian
$h_0(t)$	: Fungsi dasar <i>hazard</i> (<i>baseline hazard function</i>), yaitu nilai fungsi <i>hazard</i> pada saat semua variabel prediktor bernilai nol
i	: Indeks kategori
j	: Indeks urutan kejadian
p	: Indeks kovariat
q	: Banyaknya kovariat
m	: Indeks individu dalam <i>risk set</i> / banyaknya imputasi (dalam konteks MICE)
$L(\beta)$	: Fungsi <i>partial likelihood</i> dari model <i>Cox Proportional Hazard</i>
β	: Vektor koefisien regresi
$X_{(j)}$	: Vektor kovariat dari individu yang mengalami <i>event</i> pada waktu $t_{(j)}$
X_m	: Vektor kovariat ke- m yang termasuk dalam <i>risk set</i>
$R(t_{(j)})$	: <i>Risk set</i>
δ	: Status kejadian
$u(\beta)$	: Turunan pertama dari fungsi <i>log partial likelihood</i>

$l(\beta)$	: Turunan kedua dari fungsi <i>log partial likelihood</i>
$\hat{\beta}_s$	: Nilai parameter iterasi ke-s
$\hat{\beta}_{s+1}$	: Nilai parameter setelah pembaruan pada iterasi ke-s+1
ε	: Ambang konvergensi
$\hat{\gamma}$	: Koefisien hubungan antara residual Schoenfeld dan waktu
LR	: Uji serentak (<i>likelihood ratio test</i>)
$L\hat{\omega}$	: Nilai likelihood dari model tanpa memasukkan variabel prediktor (model kosong),
$L\hat{\Omega}$	: Nilai likelihood dari model yang menyertakan seluruh variabel prediktor (model penuh).
W^2	: Uji Wald
var	: Varians
SE	: <i>Standard error</i>

DAFTAR LAMPIRAN

Lampiran 1	Data Penelitian.....	102
Lampiran 2	Model Logistik <i>X4</i> pada Setiap Dataset Imputasi	103
Lampiran 3	Model Logistik <i>X7</i> pada Setiap Dataset Imputasi	103
Lampiran 4	Model Logistik <i>X8</i> pada Setiap Dataset Imputasi	103
Lampiran 5	Hasil Imputasi Variabel <i>X4</i>	104
Lampiran 6	Hasil Imputasi Variabel <i>X7</i>	104
Lampiran 7	Hasil Imputasi Variabel <i>X8</i>	104
Lampiran 8	Estimasi <i>Kaplan-Meier</i> Berdasarkan Usia.....	105
Lampiran 9	Estimasi <i>Kaplan-Meier</i> Berdasarkan Jenis Kelamin	105
Lampiran 10	Estimasi <i>Kaplan-Meier</i> Berdasarkan Status Pengobatan	105
Lampiran 11	Estimasi <i>Kaplan-Meier</i> Berdasarkan Riwayat Pengobatan .	106
Lampiran 12	Estimasi <i>Kaplan-Meier</i> Berdasarkan Jenis Fasyankes	106
Lampiran 13	Estimasi <i>Kaplan-Meier</i> Berdasarkan Tipe Diagnosis.....	107
Lampiran 14	Estimasi <i>Kaplan-Meier</i> Berdasarkan Hasil Tes HIV.....	107
Lampiran 15	Estimasi <i>Kaplan-Meier</i> Berdasarkan Hasil Tes DM	107
Lampiran 16	Kurva <i>Kaplan Meier</i> Dataset Imputasi 2	108
Lampiran 17	Kurva <i>Kaplan-Meier</i> Dataset Imputasi 3.....	109
Lampiran 18	Kurva <i>Kaplan-Meier</i> Dataset Imputasi 4.....	110
Lampiran 19	Kurva <i>Kaplan-Meier</i> Dataset Imputasi 5.....	111
Lampiran 20	Hasil Uji <i>Log-rank</i> Dataset Imputasi 2	112
Lampiran 21	Hasil Uji <i>Log-rank</i> Dataset Imputasi 3	112
Lampiran 22	Hasil Uji <i>Log-rank</i> Dataset Imputasi 4	112
Lampiran 23	Hasil Uji <i>Log-rank</i> Dataset Imputasi 5	112
Lampiran 24	Hasil Estimasi Parameter Model Kedua	112
Lampiran 25	<i>Script R</i>	113

ABSTRAK

Lailiyah, Aini Intan Nur. 2026. **Analisis Survival Menggunakan Metode Kaplan-Meier dan Cox Proportional Hazard pada Waktu Kesembuhan Penderita Tuberkulosis Paru (Studi Kasus di Kota Probolinggo)**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
Pembimbing: (I) Dr. Fachrur Rozi, M.Si. (II) Ari Kusumastuti, M.Pd, M.Si.

Kata Kunci: Tuberkulosis, Analisis Survival, *Kaplan-Meier*, *Cox Proportional Hazard*, *Multiple Imputation by Chained Equations*

Analisis *survival* merupakan metode statistika yang digunakan untuk menganalisis waktu hingga terjadinya suatu peristiwa (*event*). Metode yang digunakan dalam penelitian ini adalah *Kaplan-Meier* dan regresi *Cox Proportional Hazard*. Metode *Kaplan-Meier* digunakan untuk mengestimasi fungsi *survival*, sedangkan regresi *Cox Proportional Hazard* digunakan untuk mengetahui pengaruh variabel bebas terhadap risiko terjadinya *event*. *Event* yang diamati adalah kesembuhan pasien tuberkulosis (TB). Tuberkulosis masih menjadi masalah kesehatan masyarakat di Indonesia dengan jumlah kasus yang terus meningkat, termasuk di Kota Probolinggo selama tahun 2021–2023. Perbedaan waktu kesembuhan antar pasien mengindikasikan adanya faktor-faktor yang memengaruhi waktu mencapai kesembuhan. Oleh karena itu, penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi waktu kesembuhan pasien TB di Kota Probolinggo tahun 2021–2023 menggunakan metode *Kaplan-Meier* dan regresi *Cox Proportional Hazard*. Tahapan penelitian dimulai dengan *preprocessing* data, meliputi pengkodean variabel, perhitungan waktu *survival*, serta identifikasi dan penanganan *missing value* menggunakan *Multiple Imputation by Chained Equations* (MICE). Selanjutnya dilakukan analisis deskriptif untuk menggambarkan karakteristik pasien. Analisis *Kaplan-Meier* dan uji *log-rank* untuk menggambarkan pola waktu menuju kesembuhan serta menguji perbedaan waktu kesembuhan antar kategori variabel. Analisis kemudian dilanjutkan dengan regresi *Cox Proportional Hazard* untuk memodelkan pengaruh faktor-faktor terhadap waktu kesembuhan pasien TB. Asumsi *proportional hazard* diuji menggunakan residual Schoenfeld. Hasil estimasi parameter dari seluruh data hasil imputasi kemudian digabungkan menggunakan *Rubin's Rule*. Signifikansi parameter diuji menggunakan *Likelihood Ratio Test* dan *Wald Test*, sedangkan *hazard ratio* digunakan untuk menginterpretasikan pengaruh masing-masing faktor terhadap laju kesembuhan. Hasil penelitian menunjukkan bahwa variabel usia, status pengobatan, jenis fasilitas layanan kesehatan, dan hasil tes HIV berpengaruh signifikan terhadap waktu kesembuhan. Pasien dengan status pengobatan sesuai standar memiliki laju kesembuhan 15.60 kali lebih tinggi dibandingkan pasien dengan pengobatan tidak sesuai standar. Pasien berusia produktif (15 - 64 tahun) memiliki laju kesembuhan 2.60 kali lebih tinggi dibandingkan pasien berusia tidak produktif, sedangkan pasien tanpa HIV memiliki laju kesembuhan 2.46 kali lebih tinggi dibandingkan pasien dengan HIV. Selain itu, pasien yang berobat di rumah sakit dan fasilitas kesehatan lainnya memiliki laju kesembuhan lebih rendah dibandingkan pasien yang berobat di puskesmas. Temuan ini dapat menjadi salah satu strategi penanganan pasien TB yang lebih terarah berdasarkan karakteristik pasien untuk meningkatkan keberhasilan pengobatan dan kecepatan kesembuhan.

ABSTRACT

Lailiyah, Aini Intan Nur. 2026. **Survival Analysis Using the Kaplan–Meier Method and the Cox Proportional Hazards Model on the Recovery Time of Pulmonary Tuberculosis Patients (A Case Study in Probolinggo City)**. Thesis. Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
Advisors: (I) Dr. Fachrur Rozi, M.Si. (II) Ari Kusumastuti, M.Pd, M.Si.

Keywords: Tuberculosis, Survival Analysis, Kaplan–Meier, Cox Proportional Hazard, Multiple Imputation by Chained Equations

Survival analysis is a statistical method used to analyze the time until the occurrence of an event. The methods employed in this study were the Kaplan–Meier method and Cox Proportional Hazard regression. The Kaplan–Meier method was used to estimate the survival function, while Cox Proportional Hazard regression was used to determine the effect of independent variables on the risk of event occurrence. The event observed in this study was the recovery of tuberculosis (TB) patients. Tuberculosis remains a public health problem in Indonesia, with the number of cases continuing to increase, including in Probolinggo City during the period 2021–2023. Differences in recovery time among patients indicate the presence of factors that influence the time required to achieve recovery. Therefore, this study aimed to analyze the factors affecting the recovery time of TB patients in Probolinggo City during 2021–2023 using the Kaplan–Meier method and Cox Proportional Hazard regression. The analysis began with data preprocessing, including variable coding, calculation of survival time, and identification and treatment of missing values using Multiple Imputation by Chained Equations (MICE). Descriptive analysis was then conducted to describe patient characteristics. Kaplan–Meier analysis and the log-rank test were applied to describe recovery-time patterns and examine differences in recovery time across variable categories. The analysis was subsequently continued using Cox Proportional Hazard regression to model the effects of various factors on recovery time. The proportional hazard assumption was assessed using Schoenfeld residuals. Parameter estimates obtained from all imputed datasets were then combined using Rubin’s Rule. Parameter significance was evaluated using the Likelihood Ratio Test and Wald Test, while hazard ratios were used to interpret the effect of each factor on the recovery rate. The results showed that age, treatment status, type of healthcare facility, and HIV test results had a significant effect on recovery time. Patients receiving standard treatment had a recovery rate 15.60 times higher than those receiving non-standard treatment. Patients of productive age (15 – 64 years) had a recovery rate 2.60 times higher than those of non-productive age, while patients without HIV had a recovery rate 2.46 times higher than patients with HIV. In addition, patients receiving treatment at hospitals and other healthcare facilities had a lower recovery rate than those receiving treatment at community health centers. These findings may serve as a basis for developing more targeted TB management strategies based on patient characteristics in order to improve treatment success and accelerate recovery.

مستخلص البحث

ليلية، عيني إينتان نور. ٢٠٢٦. تحليل البقاء باستخدام أسلوب كابلان-ماير ونموذج كوكس للمخاطر النسبية لزمّن شفاء مرضى السل الرئوي (دراسة حالة في مدينة بروبولينغو). البحث العلمي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية في مالانج.
المشرفان: (١) الدكتور فخر الرازي، الماجستير في العلوم. (٢) آري كوسوماستوتي، الماجستير في التربية الماجستير في العلوم

الكلمات : السل الرئوي، تحليل البقاء، كابلان-ماير، نموذج كوكس للمخاطر النسبية، الإكمال المتعدد بالمعادلات المتسلسلة

يُعد تحليل البقاء أحد الأساليب الإحصائية المستخدمة لتحليل الزمن اللازم حتى وقوع حدث معين. وقد استُخدم في هذه الدراسة أسلوب كابلان-ماير ونموذج كوكس للمخاطر النسبية. ويُستخدم أسلوب كابلان-ماير لتقدير دالة البقاء، في حين يُستخدم نموذج كوكس للمخاطر النسبية لتحديد تأثير المتغيرات المستقلة على خطر وقوع الحدث. وكان الحدث المدروس في هذه الدراسة هو شفاء مرضى السل الرئوي. ولا يزال مرض السل يمثل مشكلة صحية عامة في إندونيسيا، حيث يستمر عدد الحالات في الارتفاع، بما في ذلك مدينة بروبولينغو خلال الفترة ٢٠٢١ - ٢٠٢٣. وتشير الاختلافات في زمن الشفاء بين المرضى إلى وجود عوامل تؤثر في الوقت اللازم للوصول إلى الشفاء. لذلك هدفت هذه الدراسة إلى تحليل العوامل المؤثرة في زمن شفاء مرضى السل الرئوي في مدينة بروبولينغو خلال الفترة ٢٠٢١-٢٠٢٣ باستخدام أسلوب كابلان-ماير ونموذج كوكس للمخاطر النسبية. بدأت الدراسة بمرحلة المعالجة الأولية للبيانات، والتي شملت ترميز المتغيرات، وحساب زمن البقاء، وتحديد القيم المفقودة ومعالجتها باستخدام أسلوب الإكمال المتعدد بالمعادلات المتسلسلة (*MICE*). بعد ذلك أُجري التحليل الوصفي لوصف خصائص المرضى. كما استُخدم تحليل كابلان-ماير واختبار لوغ-رانك لوصف أنماط زمن الشفاء واختبار الفروق بين فئات المتغيرات المختلفة. ثم استُخدم نموذج كوكس للمخاطر النسبية لدراسة تأثير العوامل المختلفة في زمن الشفاء. وتم اختبار افتراض المخاطر النسبية باستخدام بواقي شونيفيلد. وبعد ذلك جُمعت تقديرات المعلمات الناتجة عن جميع مجموعات البيانات المستكملة باستخدام قاعدة روبن. واختُبرت معنوية المعلمات باستخدام اختبار نسبة الإمكان واختبار والد، في حين استُخدمت نسبة المخاطر لتفسير تأثير كل عامل في معدل الشفاء. أظهرت نتائج الدراسة أن متغيرات العمر، وحالة العلاج، ونوع مرفق الرعاية الصحية، ونتيجة فحص فيروس نقص المناعة البشرية (*HIV*) لها تأثير معنوي في زمن الشفاء. وكان معدل الشفاء لدى المرضى الذين تلقوا علاجًا مطابقًا للمعايير أعلى بمقدار ١٥.٦٠ مرة مقارنة بالمرضى الذين تلقوا علاجًا غير مطابق للمعايير. كما كان معدل الشفاء لدى المرضى في سن العمل (١٥ - ٦٤ سنة) أعلى بمقدار ٢.٦٠ مرة مقارنة بالمرضى خارج سن العمل، في حين كان معدل الشفاء لدى المرضى غير المصابين بفيروس نقص المناعة البشرية أعلى بمقدار ٢.٤٦ مرة مقارنة بالمصابين به. بالإضافة إلى ذلك، كان معدل الشفاء لدى المرضى الذين تلقوا العلاج في المستشفيات ومرافق الرعاية الصحية الأخرى أقل مقارنة بالمرضى الذين تلقوا العلاج في المراكز الصحية الحكومية. ويمكن أن تسهم هذه النتائج في وضع استراتيجيات أكثر استهدافًا للتعامل مع مرضى السل الرئوي استنادًا إلى خصائص المرضى، بما يعزز نجاح العلاج ويسرّع عملية الشفاء.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Matematika memiliki peran penting dalam berbagai bidang ilmu, termasuk dalam analisis data kesehatan. Salah satu cabang matematika terapan yang berkembang pesat adalah statistika, yang menyediakan berbagai alat analisis untuk memahami dan menarik kesimpulan dari fenomena atau peristiwa yang diamati. Salah satu metode statistika yang digunakan untuk menganalisis kejadian yang berkaitan dengan waktu atau durasi hingga terjadinya peristiwa tertentu adalah analisis *survival* (Collett, 2015).

Analisis *survival* merupakan metode khusus yang digunakan untuk menganalisis data waktu hingga terjadinya peristiwa, seperti kematian, kekambuhan, atau pemulihan (Kleinbaum & Klein, 2012). Waktu dalam analisis ini dihitung sejak awal pengamatan hingga peristiwa terjadi. Selain untuk mengetahui distribusi waktu *survival*, analisis ini juga digunakan untuk melihat pengaruh faktor-faktor tertentu terhadap waktu tersebut (Faisal dkk., 2020). Salah satu karakteristik unik dari analisis *survival* adalah adanya data tersensor, yaitu ketika informasi tentang waktu kejadian tidak lengkap. Hal ini terjadi jika subjek belum mengalami peristiwa hingga akhir masa observasi atau hilang dalam proses pengamatan (Klein & Moeschberger, 2003).

Dua metode utama yang digunakan dalam analisis *survival* adalah metode non-parametrik *Kaplan-Meier* dan model semiparametrik *Cox Proportional Hazard*. Model Cox digunakan untuk mengetahui pengaruh variabel bebas terhadap risiko terjadinya peristiwa dalam asumsi bahwa *hazard ratio* bersifat konstan sepanjang

waktu (Collett, 2015). Sementara itu, metode *Kaplan-Meier* digunakan untuk mengestimasi fungsi *survival* dan membandingkan antar kategori (Hari dkk., 2018).

Sebelum menetapkan metode analisis yang digunakan, penelitian ini diawali dengan *pra-research* berupa eksplorasi awal terhadap data pasien tuberkulosis yang diperoleh dari Dinas Kesehatan Kota Probolinggo. Tahap awal ini dilakukan untuk mengidentifikasi karakteristik data, khususnya keberadaan variabel waktu hingga terjadinya peristiwa kesembuhan serta adanya data tersensor, yaitu pasien yang belum sembuh hingga akhir masa pengamatan atau mengalami putus pengobatan. Hasil eksplorasi awal menunjukkan bahwa data memiliki struktur waktu *survival* yang jelas dan mengandung sensor, sehingga pendekatan analisis *survival* menjadi relevan. Selanjutnya, dilakukan pemodelan awal menggunakan metode *Kaplan-Meier* untuk menggambarkan distribusi waktu menuju kesembuhan, serta penerapan model *Cox Proportional Hazard* untuk mengkaji pengaruh kovariat terhadap laju kesembuhan. Berdasarkan hasil *pra-research* tersebut, model *Cox Proportional Hazard* dinilai sesuai untuk digunakan karena mampu mengakomodasi data tersensor dan memberikan interpretasi pengaruh faktor-faktor penentu kesembuhan melalui rasio *hazard*. Oleh karena itu, pemilihan model Cox dalam penelitian ini didasarkan pada kesesuaian karakteristik data dengan asumsi dan tujuan model, bukan semata-mata pertimbangan teoritis.

Variabel yang digunakan dalam penelitian ini merupakan variabel standar yang dicatat oleh Dinas Kesehatan sesuai dengan panduan Kementerian Kesehatan Republik Indonesia (Kemenkes RI). Variabel usia, Diabetes Melitus (DM), dan *Human Immunodeficiency Virus* (HIV) digunakan karena berkaitan dengan daya tahan tubuh yang memengaruhi respons pengobatan dan kecepatan kesembuhan.

Variabel jenis kelamin, jenis fasilitas layanan kesehatan, dan status pengobatan dicatat dalam formulir penemuan terduga TB sebagai bagian dari pemantauan pasien, di mana ketidakteraturan pengobatan berisiko menyebabkan kegagalan terapi dan TB resisten obat. Selain itu, tipe diagnosis dan riwayat pengobatan digunakan untuk menentukan penatalaksanaan dan regimen terapi yang sesuai (Kementerian Kesehatan, 2025). Oleh karena itu, pemilihan variabel dalam penelitian ini memiliki dasar klinis dan epidemiologis yang jelas serta tidak bersifat acak.

Tuberkulosis (TB) adalah penyakit infeksi menular yang disebabkan oleh *Mycobacterium Tuberculosis*, suatu basil tahan asam (BTA) yang umumnya menyerang paru-paru, namun dapat juga mengenai organ lain di luar paru. Penularan terjadi terutama melalui udara, ketika penderita TB paru mengeluarkan droplet berisi bakteri saat batuk, bersin, atau berbicara. TB dapat menimbulkan komplikasi berbahaya dan kematian jika tidak diobati secara tuntas, sehingga menjadi masalah kesehatan global yang terus mendapat perhatian. Secara global, TB masih merupakan salah satu dari 10 penyakit menular penyebab kematian terbesar di dunia dan berada pada urutan ketiga penyebab kematian setelah penyakit kardiovaskuler dan infeksi saluran pernapasan akut (ISPA). Menurut *Global Tuberculosis Report 2022* dalam penelitian Avy dkk. (2024), hampir 10 juta orang jatuh sakit akibat TB setiap tahunnya, dengan angka kematian mencapai 1,4 juta kasus per tahun. WHO juga melaporkan bahwa Indonesia menempati peringkat kedua jumlah penderita TB paru terbanyak di dunia setelah India. Di Indonesia, beban TB masih menjadi tantangan besar meskipun berbagai upaya pengendalian telah dilakukan. Pada tahun 2020 ditemukan sebanyak 351.936 kasus TB, dengan

angka kematian mencapai 150.000 jiwa. Kasus terbanyak berasal dari provinsi berpenduduk padat seperti Jawa Barat, Jawa Timur, dan Jawa Tengah.

Provinsi Jawa Timur merupakan wilayah dengan prevalensi TB terbanyak kedua setelah Jawa Barat (Rahem dkk., 2024). Kondisi di tingkat daerah juga menunjukkan situasi yang mengkhawatirkan. Berdasarkan data dari Dinas Kesehatan Kota Probolinggo, jumlah kasus tuberkulosis di wilayah ini terus mengalami peningkatan setiap tahunnya. Pada tahun 2021 tercatat 357 kasus penderita tuberkulosis yang menjalani pengobatan. Angka ini meningkat tajam pada tahun 2022 menjadi 559 kasus, dan kembali naik pada tahun 2023 hingga mencapai 610 kasus. Tren peningkatan tersebut menunjukkan bahwa tuberkulosis masih menjadi masalah kesehatan serius di Kota Probolinggo dan membutuhkan kajian lebih mendalam. Selain itu, penelitian berbasis data lokal mengenai analisis *survival* pasien tuberkulosis di kota ini masih sangat terbatas, sehingga menjadikan penelitian ini memiliki nilai kebaruan (*novelty*) yang kuat. Dukungan dari Dinas Kesehatan Kota Probolinggo, baik berupa kemudahan akses data maupun sambutan positif terhadap kegiatan penelitian, turut memperkuat alasan pemilihan wilayah ini sebagai lokasi penelitian. Dengan demikian, hasil penelitian diharapkan dapat memberikan masukan strategis bagi upaya pengendalian dan penanggulangan tuberkulosis di Kota Probolinggo.

Pendekatan analisis *survival* dalam penelitian kesehatan, khususnya pada kasus tuberkulosis juga memiliki makna spiritual yang mendalam dalam perspektif islam. Dalam Islam, setiap upaya manusia, termasuk penelitian ilmiah di bidang kesehatan, dapat bernilai ibadah apabila dilandasi dengan niat yang ikhlas dan bertujuan untuk kemaslahatan umat. Allah SWT menegaskan dalam firman-Nya:

وَمَا خَلَقْتُ الْجِنَّ وَالْإِنْسَ إِلَّا لِيَعْبُدُونِ

Artinya: “*Tidaklah Aku menciptakan jin dan manusia kecuali untuk beribadah kepada-Ku.*” (QS.Adz-dzariyat, 51:56).

Ayat ini mengandung pesan mendasar bahwa hakikat keberadaan manusia di dunia adalah untuk beribadah kepada Allah SWT. Menurut penafsiran Ibnu Katsir, Allah tidak membutuhkan ibadah manusia, namun manusialah yang membutuhkan Allah dalam setiap aspek kehidupannya. Ibadah merupakan bentuk pengakuan atas ketergantungan manusia kepada Sang Pencipta, serta jalan untuk memperoleh rahmat dan pertolongan-Nya (Abdullah, 2007). Dengan demikian, kegiatan penelitian dan pengembangan ilmu pengetahuan yang bertujuan menjaga kesehatan dan menyelamatkan kehidupan manusia juga termasuk bagian dari ibadah, karena sejalan dengan amanah Allah untuk menjaga kehidupan dan menegakkan kemaslahatan di bumi. Dalam konteks penelitian ini, analisis *survival* digunakan untuk memahami faktor-faktor yang memengaruhi laju kesembuhan penderita tuberkulosis. Upaya ini mencerminkan bentuk ikhtiar manusia untuk mengenali sebab-sebab kesembuhan yang Allah tetapkan. Sebagaimana dalam kehidupan beragama, di antara berbagai bentuk ibadah terdapat amalan yang paling utama dan paling berpengaruh dalam mendekatkan diri kepada Allah SWT, demikian pula dalam penelitian ini, di antara berbagai faktor penyebab kesembuhan terdapat faktor yang paling dominan. Keduanya memiliki makna yang paralel, sama-sama berusaha menemukan hal yang paling berpengaruh untuk mencapai hasil terbaik, yaitu kesembuhan bagi tubuh dan kedekatan bagi ruh.

Berdasarkan uraian tersebut, tuberkulosis masih menjadi masalah kesehatan yang serius baik secara global maupun di Indonesia. Tingginya prevalensi TB serta adanya data tersensor pada pasien yang belum sembuh atau putus obat menjadikan

analisis *survival*, khususnya metode *Kaplan-Meier* dan *Cox Proportional Hazard*, relevan untuk digunakan. Dari perspektif Islam, penelitian ini tidak hanya memberikan kontribusi ilmiah, tetapi juga menjadi bentuk pengabdian spiritual yang mencerminkan kesadaran bahwa menjaga kesehatan dan menolong sesama adalah bagian dari ibadah kepada Allah SWT.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, rumusan masalah pada penelitian ini sebagai berikut:

1. Bagaimana distribusi kesembuhan pasien tuberkulosis (TB) di Kota Probolinggo berdasarkan faktor-faktor yang berpengaruh menggunakan metode *Kaplan-Meier*?
2. Bagaimana model regresi *Cox Proportional Hazard* serta nilai *hazard ratio* dari faktor-faktor yang memengaruhi waktu kesembuhan pasien tuberkulosis (TB) di Kota Probolinggo?

1.3 Tujuan Penelitian

Tujuan penelitian yang dicapai berdasarkan rumusan masalah yang telah dibentuk sebagai berikut:

1. Menganalisis distribusi kesembuhan pasien tuberkulosis di Kota Probolinggo berdasarkan faktor-faktor yang berpengaruh dengan menggunakan metode *Kaplan-Meier*.

2. Mengestimasi model regresi *Cox Proportional Hazard* serta menghitung nilai *hazard ratio* dari faktor-faktor yang memengaruhi waktu kesembuhan pasien tuberkulosis (TB) di Kota Probolinggo

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberi manfaat sebagai berikut:

1. Bagi peneliti, penelitian ini menjadi sarana untuk mengasah kemampuan dalam menerapkan teori statistika, khususnya analisis *survival* dalam konteks nyata di bidang kesehatan, serta memperluas wawasan dan pemahaman tentang peran statistika dalam mendukung pengambilan keputusan berbasis data pada kasus tuberkulosis.
2. Bagi dinas kesehatan dan puskesmas, hasil analisis dalam penelitian ini dapat memberikan gambaran mengenai pola waktu kesembuhan pasien TB berdasarkan berbagai faktor yang diteliti. Temuan ini diharapkan dapat membantu dalam perencanaan intervensi yang lebih tepat sasaran, serta menjadi masukan dalam evaluasi efektivitas program pengendalian TB.
3. Bagi pasien dan masyarakat, penelitian ini dapat memberikan pemahaman mengenai faktor-faktor yang dapat memengaruhi keberhasilan pengobatan TB, sehingga meningkatkan kesadaran dan partisipasi dalam mendukung program pemberantasan TB.
4. Bagi kampus, penelitian ini diharapkan turut berkontribusi dalam memperkuat peran Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang dalam mengembangkan ilmu matematika di bidang kesehatan.

1.5 Batasan Masalah

Adapun batasan masalah yang ditetapkan oleh peneliti antara lain:

1. Data yang digunakan mencakup pasien tuberkulosis yang tercatat oleh Dinkes Kota probolinggo dan menjalani pengobatan pada tahun 2021–2023.
2. Waktu *survival* didefinisikan sebagai lama waktu (dalam hari) yang dibutuhkan pasien sejak memulai pengobatan hingga dinyatakan sembuh.
3. Jenis data tersensor yang digunakan dalam penelitian ini adalah data tersensor kanan (*right-censored*), di mana pasien yang belum sembuh hingga akhir periode pengamatan atau tidak menyelesaikan pengobatan dicatat sebagai data tersensor.
4. Faktor-faktor yang dianalisis terbatas pada variabel-variabel yang tersedia dalam data, yaitu usia, jenis kelamin, status pengobatan, riwayat pengobatan, jenis fasilitas layanan kesehatan, tipe diagnosis TB, hasil tes HIV, dan hasil tes DM, yang di mana masing-masing faktor tersebut akan dijelaskan dalam sub bab tersendiri.
5. Penelitian ini tidak mencakup pengembangan metode alternatif untuk menangani pelanggaran asumsi, karena fokus utama penelitian adalah memperoleh model *Cox Proportional Hazard* yang valid dan sesuai dengan asumsi yang mendasarinya.

BAB II

KAJIAN TEORI

2.1 Teori Pendukung

2.1.1 Konsep Analisis *Survival*

Analisis *survival* merupakan salah satu metode statistika yang digunakan untuk mengevaluasi data dengan variabel respons yang berbentuk waktu hingga terjadinya suatu peristiwa tertentu (*time until an event occurs*) (Harlan, 2017). Tujuan utama dari metode ini adalah untuk mengetahui bagaimana pengaruh berbagai faktor terhadap lamanya waktu yang dibutuhkan sebelum suatu kejadian terjadi, dengan pengukuran waktu yang dapat dinyatakan dalam satuan hari, minggu, bulan, atau tahun. Penerapan analisis *survival* cukup luas dalam berbagai bidang. Misalnya, pada pasien leukemia, waktu yang dihitung adalah lamanya pasien berada dalam masa remisi sebelum terjadi kekambuhan, yang dihitung dalam minggu. Dalam penelitian mengenai penyakit jantung, yang diamati adalah berapa tahun hingga seseorang mengalami kondisi tersebut. Dalam konteks sosial, analisis ini dapat digunakan untuk mengetahui berapa minggu dibutuhkan sampai mantan narapidana kembali ditangkap (*residivisme*). Sementara itu, dalam studi pasca transplantasi, peristiwa yang diamati adalah kematian, dengan waktu yang dihitung sejak transplantasi dilakukan hingga kematian, biasanya dalam satuan bulan. Ragam penerapan ini menunjukkan bahwa analisis *survival* sangat relevan untuk mengukur waktu menuju peristiwa penting dalam berbagai konteks, baik di bidang medis maupun non-medis (Kleinbaum & Klein, 2012).

Dalam analisis *survival*, terdapat tiga konsep utama yang perlu dipahami (Zukhaila, 2016). Pertama, *survival time (T)* merujuk pada durasi waktu yang

mencerminkan berapa lama seorang individu bertahan sebelum mengalami peristiwa yang menjadi fokus studi dalam periode pengamatan tertentu. Kedua, *event* atau peristiwa, dalam analisis ini dan dilambangkan dengan δ , mencakup kejadian seperti kematian, timbulnya penyakit, atau bahkan kejadian positif seperti peningkatan status gizi atau kesembuhan. Jika $\delta = 1$, maka kejadian tersebut telah terjadi, sedangkan $\delta = 0$ menunjukkan data tersebut bersifat tersensor karena kejadian belum terjadi. Ketiga, *censoring* atau sensor adalah keadaan ketika waktu pasti terjadinya *event* tidak dapat diketahui, misalnya karena partisipan belum mengalami *event* hingga studi berakhir, tidak lagi terpantau, atau mengalami kejadian lain yang tidak sesuai dengan fokus penelitian (Kleinbaum & Klein, 2012). Ketiga unsur ini menjadi fondasi penting dalam pemodelan dan interpretasi hasil analisis *survival*.

2.1.2 Konsep Pensensoran (*Censoring*)

Pensensoran dalam analisis *survival* merupakan cara untuk menangani ketidaklengkapan data ketika tidak semua subjek mengalami kejadian yang diamati selama periode penelitian. Terdapat beberapa penyebab umum data menjadi tersensor (Kleinbaum & Klein, 2012):

1. Subjek hilang dari pengamatan (*lost to follow-up*)

Artinya, subjek tidak dapat lagi diikuti dalam pengamatan karena berpindah tempat tinggal, menolak, putus berobat atau alasan lain sehingga tidak tersedia informasi kejadian hingga akhir studi.

2. Subjek tidak mengalami *event* selama penelitian

Subjek tetap berada dalam studi hingga akhir, tetapi tidak mengalami *event* yang menjadi fokus penelitian selama masa pengamatan.

3. Subjek dihentikan karena sebab lain

Misalnya, subjek meninggal dunia akibat sebab yang tidak terkait dengan kejadian yang diamati, atau pengamatan dihentikan karena batas waktu penelitian telah tercapai.

Jenis sensor tidak ditentukan dari model *survival*, tetapi dari cara pendefinisian waktu kejadian dan kondisi subjek yang belum mengalami *event*. Terdapat tiga jenis utama pensensoran yang perlu dikenali dalam analisis *survival* (Collett, 2015):

1. Sensor kanan (*right censoring*)

Terjadi ketika subjek belum mengalami *event* hingga studi berakhir atau keluar dari studi sebelum *event* terjadi. Contohnya adalah partisipan yang pindah sebelum kejadian diamati.

2. Sensor kiri (*left censoring*)

Keadaan di mana *event* sudah terjadi sebelum subjek diamati dalam studi, namun waktu tepat kejadiannya tidak diketahui. Misalnya, pasien telah mengidap kanker saat pertama kali diperiksa, tetapi waktu pasti awal penyakit tidak diketahui.

3. Sensor interval (*interval censoring*)

Event diketahui terjadi dalam suatu rentang waktu, tetapi waktu pastinya tidak dapat ditentukan. Contoh kasusnya adalah seseorang yang diperiksa sehat di usia 45 tahun dan didiagnosis kanker pada usia 50 tahun, sehingga kejadian diperkirakan terjadi antara usia 45–50 tahun.

2.1.3 Analisis dan Penanganan *Missing Value*

Missing value merupakan kondisi ketika terdapat data yang tidak teramati atau tidak tersedia pada suatu variabel dalam dataset penelitian. Keberadaan *missing value* dapat memengaruhi kualitas hasil analisis karena berpotensi menimbulkan bias serta menurunkan tingkat akurasi dalam pengambilan kesimpulan. Kondisi ini dapat disebabkan oleh berbagai faktor, seperti kesalahan manusia dalam pengolahan data, kerusakan alat pengumpulan data, responden yang menolak menjawab pertanyaan tertentu, serta adanya *dropout* dalam studi longitudinal. Oleh karena itu, sebelum dilakukan analisis lebih lanjut, diperlukan identifikasi terhadap pola dan mekanisme terjadinya *missing data* agar dapat ditentukan metode penanganan yang tepat dan sesuai dengan karakteristik data (Rubin, 1987).

Salah satu pendekatan yang digunakan untuk menganalisis pola *missing data* adalah melalui uji statistik, yaitu *Chi-Square* dan *Fisher Exact*. Uji *Chi-Square* merupakan metode statistika inferensial yang digunakan untuk menguji hubungan antara dua variabel kategorik dengan membandingkan frekuensi observasi dan frekuensi harapan. Agar hasil pengujian valid, terdapat beberapa asumsi yang harus dipenuhi, yaitu data diperoleh secara acak, setiap pengamatan bersifat independen, serta frekuensi harapan tidak terlalu kecil (tidak ada sel yang kurang dari 1 dan maksimal 20% sel kurang dari 5). Selain itu, ukuran sampel yang cukup besar (umumnya lebih dari 40) diperlukan agar pendekatan distribusi *Chi-Square* menjadi lebih baik. Apabila asumsi tersebut tidak terpenuhi, terutama pada ukuran sampel kecil atau tabel kontingensi sederhana seperti 2×2 dengan frekuensi harapan rendah, maka digunakan uji *Fisher Exact* sebagai alternatif. Uji ini tidak bergantung pada

pendekatan distribusi sehingga lebih akurat dalam kondisi data terbatas atau tidak memenuhi asumsi *Chi-Square*. Dengan demikian, pemilihan metode uji bergantung pada karakteristik data yang dianalisis.

Dalam konteks analisis *missing* data, kedua uji tersebut digunakan untuk mengetahui ada tidaknya hubungan antara kejadian *missing* pada suatu variabel dengan variabel lain. Hipotesis yang digunakan adalah (Fitri dkk., 2023):

1. H_0 : Tidak terdapat hubungan antara *missing* data suatu variabel dengan variabel lain yang diuji
2. H_1 : Terdapat hubungan antara *missing* data suatu variabel dengan variabel lain yang diuji

Secara matematis, pengujian *Chi-Square* dilakukan dengan membandingkan frekuensi observasi dan frekuensi harapan yang dihitung menggunakan rumus pada persamaan (2.1):

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (2.1)$$

Keterangan:

χ^2 : Nilai statistik *Chi-Square*

O : Frekuensi observasi

E : Frekuensi harapan

Sedangkan uji *Fisher Exact* secara matematis pada persamaan (2.2) digunakan untuk menghitung probabilitas kemunculan suatu tabel kontingensi 2×2 berdasarkan distribusi peluang dengan asumsi total baris dan kolom tetap. Probabilitas tersebut dinyatakan dalam bentuk fungsi faktorial, yaitu:

$$P = \frac{(a + b)! (c + d)! + (a + c)! (b + d)!}{a! b! c! d! n!} \quad (2.2)$$

dengan a, b, c , dan d menyatakan frekuensi pada masing-masing sel tabel, serta $n = a + b + c + d$ sebagai jumlah total observasi. Simbol “!” menyatakan operasi faktorial. Kriteria pengujian ditentukan berdasarkan nilai $p - value$ berikut:

1. Jika $p - value < \alpha = 0.05$, maka H_0 ditolak
2. Jika $p - value \geq \alpha = 0.05$, maka H_0 gagal ditolak

Berdasarkan hasil pengujian, mekanisme *missing data* dapat diklasifikasikan menjadi tiga jenis, yaitu *Missing Completely At Random* (MCAR), *Missing At Random* (MAR), dan *Missing Not At Random* (MNAR). Data dikatakan MCAR apabila tidak terdapat hubungan yang signifikan antara kejadian *missing* dengan variabel lain, sehingga kehilangan data terjadi secara acak. Jika terdapat hubungan signifikan dengan minimal satu variabel, maka data tidak memenuhi asumsi MCAR dan cenderung mengikuti mekanisme MAR, yaitu *missing data* bergantung pada variabel lain yang dapat diamati. Adapun MNAR merupakan kondisi di mana *missing data* bergantung pada nilai variabel itu sendiri yang tidak teramati, sehingga tidak dapat diidentifikasi secara langsung melalui uji statistik (Little & Rubin, 2020).

Dalam penanganan *missing value*, terdapat dua pendekatan utama, yaitu penghapusan data (*deletion*) dan imputasi (*imputation*). Metode penghapusan, seperti *listwise deletion*, dilakukan dengan menghapus seluruh observasi yang memiliki nilai hilang, namun dapat mengurangi ukuran sampel dan berpotensi menimbulkan bias, terutama jika data tidak bersifat MCAR. Oleh karena itu, metode imputasi lebih disarankan karena dapat mempertahankan jumlah data dengan menggantikan nilai yang hilang menggunakan estimasi tertentu. Metode

imputasi tunggal (*single imputation*) seperti *mean* atau regresi umumnya tidak direkomendasikan karena dapat menurunkan variansi data dan menghasilkan *standard error* yang terlalu kecil. Sebagai alternatif, digunakan metode *Multiple Imputation* yang mampu menghasilkan estimasi lebih valid dengan mempertimbangkan ketidakpastian imputasi. Metode ini terdiri dari tiga tahap, yaitu imputasi (membentuk beberapa dataset lengkap), analisis pada masing-masing dataset, dan *pooling* menggunakan aturan Rubin untuk memperoleh estimasi akhir. *Multiple Imputation* direkomendasikan terutama pada data dengan mekanisme MAR karena memanfaatkan informasi dari variabel yang teramati dalam memperkirakan nilai yang hilang (Heymans & Twisk, 2022).

Salah satu implementasi yang paling umum digunakan dalam *Multiple Imputation* adalah metode *Multiple Imputation by Chained Equations* (MICE). Metode ini bekerja secara iteratif dengan memanfaatkan hubungan antar variabel dalam dataset untuk mengestimasi nilai yang hilang. Dalam implementasinya, setiap variabel yang memiliki nilai hilang dimodelkan sebagai fungsi dari variabel lain menggunakan model regresi yang sesuai dengan tipe datanya. Untuk variabel kategorik biner digunakan model regresi logistik:

$$\text{logit}(P(Y = 1|\mathbf{X})) = \ln\left(\frac{P(Y = 1|\mathbf{X})}{1 - P(Y = 1|\mathbf{X})}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (2.3)$$

dengan $\mathbf{X} = (X_1, X_2, \dots, X_p)'$.

Parameter model diestimasi berdasarkan data yang tidak mengandung *missing*, kemudian digunakan untuk memperoleh prediksi. Nilai imputasi dibangkitkan secara acak dari distribusi probabilitas yang sesuai dan dilakukan secara iteratif hingga mencapai kondisi konvergen. Dalam penerapannya, terdapat dua parameter penting, yaitu banyak dataset yang dibentuk (m) dan jumlah iterasi ($maxit$).

Parameter *maxit* menunjukkan banyaknya iterasi yang dilakukan hingga nilai imputasi mencapai kondisi stabil (konvergen). Konvergensi ini penting untuk memastikan hasil imputasi akurat, dan dapat dievaluasi melalui *trace plot*, di mana nilai imputasi akan berfluktuasi kecil di sekitar suatu titik. Secara umum, jumlah iterasi yang diperlukan relatif sedikit, yaitu sekitar 20–30 atau kurang sebagai aturan praktis. Sementara itu, parameter *m* menyatakan jumlah dataset hasil imputasi yang dibentuk. Semakin besar nilai *m*, maka semakin baik dalam merepresentasikan ketidakpastian imputasi, meskipun secara praktis nilai $m = 5$ telah dianggap cukup (Rubin, 1987).

Selanjutnya, hasil analisis dari masing-masing dataset imputasi digabungkan menggunakan aturan Rubin (*Rubin's Rules*) melalui fungsi *pool()* pada *package mice*. Estimasi parameter diperoleh dari rata-rata hasil analisis (Utami & Danardono, 2019):

$$\bar{\theta} = \frac{1}{m} \sum_{i=1}^m \hat{\theta}_i \quad (2.4)$$

Dengan variansi dalam imputasi (*within imputation variance*) dinyatakan sebagai:

$$\bar{U} = \frac{1}{m} \sum_{i=1}^m U_i; U_i = (SE_i)^2 \quad (2.5)$$

Dan variansi antar imputasi (*between imputation variance*) sebagai:

$$B = \frac{1}{m-1} \sum_{i=1}^m (\hat{\theta}_i - \bar{\theta})^2 \quad (2.6)$$

Kombinasi persamaan (2.5) dan (2.6) menghasilkan variansi total:

$$VT = \bar{U} + \left(1 + \frac{1}{m}\right) B \quad (2.7)$$

Keterangan:

m : Banyaknya imputasi

$\hat{\theta}_i$: Estimasi parameter pada dataset imputasi ke- i

$\bar{\theta}$: Rata-rata estimasi parameter dari seluruh dataset imputasi

U_i : Variansi estimasi parameter pada dataset imputasi ke- i , dimana

$\bar{U}$: Rata-rata variansi dalam imputasi, yang merepresentasikan ketidakpastian dari model dalam masing-masing imputasi

B : Variansi antar imputasi, yang merepresentasikan ketidakpastian antar dataset imputasi

VT : Variansi total, yang merupakan kombinasi variansi dalam dan antar imputasi

Dengan demikian, metode MICE tidak hanya menghasilkan estimasi parameter, tetapi juga mempertimbangkan ketidakpastian akibat data yang hilang, sehingga sangat sesuai digunakan pada data dengan mekanisme *Missing At Random* (MAR). Implementasi metode ini dalam penelitian dilakukan menggunakan *package mice* pada perangkat lunak R, dengan proses penggabungan hasil melalui fungsi *pooling()* (Pramudita dkk., 2025).

2.1.4 Fungsi *Survival* dan Fungsi *Hazard*

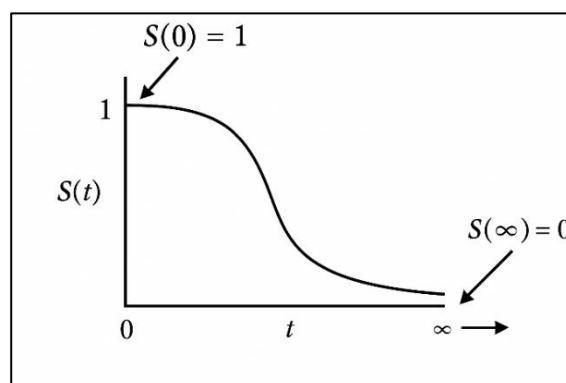
Variabel acak T merepresentasikan waktu hingga suatu peristiwa (*event*) terjadi dalam analisis *survival*. Karena T menunjukkan waktu, maka nilai-nilainya bersifat non-negatif, artinya T harus bernilai lebih besar dari atau sama dengan nol. Sementara itu, t merujuk pada nilai tertentu dari variabel acak T . Fungsi *survival*, dilambangkan dengan $S(t)$, adalah probabilitas bahwa waktu terjadinya suatu peristiwa (*event*) melebihi waktu tertentu t . Dengan kata lain, $S(t)$ menunjukkan peluang bahwa variabel acak T lebih besar dari t . Secara teoritis, rentang nilai t

berada dalam interval dari nol hingga tak hingga. $S(t)$ ini umumnya divisualisasikan dalam bentuk kurva kontinu, dan memiliki beberapa karakteristik khas, yaitu:

1. Kurva tidak meningkat; seiring bertambahnya waktu (t), nilai fungsi $S(t)$ akan menurun atau tetap, karena semakin banyak individu yang mengalami *event* seiring waktu berjalan.
2. Untuk $t = 0, S(t) = 1$; hal ini menunjukkan kondisi awal penelitian, di mana belum ada subjek yang mengalami kejadian, sehingga probabilitas belum mengalami *event* pada waktu ke-0 bernilai 1.
3. Untuk $t = \infty, S(t) = 0$; secara teoritis, jika waktu pengamatan berlangsung sangat lama (mendekati tak hingga), maka seluruh subjek diasumsikan akan mengalami *event*, sehingga probabilitas belum mengalami *event* mendekati nol.

Karakteristik ini mencerminkan sifat dasar dari fungsi *survival*, yaitu semakin lama waktu berlalu, semakin kecil kemungkinan subjek untuk belum mengalami *event*.

Gambar 2.1 menunjukkan ilustrasi kurva fungsi *survival*.



Gambar 2.1 Ilustrasi Kurva Fungsi *Survival* saat $t = 0$ dan $t = \infty$

Fungsi *hazard* menggambarkan laju atau potensi sesaat terjadinya suatu peristiwa (*event*) pada waktu tertentu (t), dengan asumsi bahwa individu belum mengalami

event tersebut hingga waktu (t). Berbeda dengan fungsi *survival* yang menunjukkan probabilitas individu belum mengalami *event* hingga waktu tertentu, fungsi *hazard* menitikberatkan pada kemungkinan terjadinya *event* dalam selang waktu yang sangat kecil setelah waktu (t). Oleh karena itu, fungsi *hazard* dan fungsi *survival* saling berkaitan dalam menggambarkan distribusi waktu terjadinya suatu *event*. Kurva fungsi *hazard* memiliki beberapa karakteristik utama, yaitu (Kleinbaum & Klein, 2012):

1. Selalu bernilai non-negatif, artinya nilai fungsi ini tidak pernah kurang dari nol karena menunjukkan laju terjadinya suatu *event*.
2. Tidak memiliki batas atas, sehingga secara teoritis nilai fungsi *hazard* dapat sangat besar tergantung pada konteks data dan proses terjadinya *event* yang diamati.

Selain itu, fungsi *hazard* juga memiliki tujuan untuk:

1. Menunjukkan laju terjadinya suatu peristiwa (*event rate*),
2. Menentukan bentuk model yang sesuai,
3. Membuat model matematik untuk analisis *survival*.

Misalkan T melambangkan waktu *survival* dari awal pengamatan hingga terjadinya suatu *event*, yang merupakan variabel acak dengan karakteristik fungsi *survival* dan fungsi *hazard*. Fungsi *survival* $S(t)$ didefinisikan sebagai probabilitas suatu individu belum mengalami *event* sampai waktu yang lebih besar atau sama dengan t waktu. Apabila diketahui fungsi distribusi kumulatif, yaitu (As'ari, 2012):

$$F(t) = P(T \leq t) = \int_0^t f(u)du, t \geq 0 \quad (2.8)$$

Maka, dari persamaan (2.8) diperoleh fungsi *survival*:

$$\begin{aligned}
S(t) = P(T \geq t) &= \int_t^{\infty} f(u)du \\
&= 1 - \int_0^t f(u)du \\
&= 1 - F(t), t \geq 0
\end{aligned} \tag{2.9}$$

Fungsi *survival* pada persamaan (2.9) digunakan untuk menggambarkan peluang bahwa suatu individu belum mengalami *event* sejak awal pengamatan hingga waktu t (Collett, 2015).

Fungsi *hazard* $h(t)$ didefinisikan sebagai laju kejadian di mana suatu individu akan mengalami *event* dalam interval waktu antara t dan $t + \Delta_t$, dengan catatan bahwa individu tersebut belum mengalami *event* hingga waktu t . Dalam matematis dapat ditulis :

$$P(t \leq T < t + \Delta_t | T \geq t) \tag{2.10}$$

dengan T sebagai variabel acak waktu terjadinya *event* dan t adalah waktu tertentu yang sedang diamati. Dengan pemanfaatan teorema peluang bersyarat, maka dari persamaan (2.10) dapat diperoleh fungsi *hazard* pada persamaan (2.11) (Collett, 2015):

$$\begin{aligned}
h(t) &= \lim_{\Delta_t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta_t | T \geq t)}{\Delta_t} \right\} \\
&= \lim_{\Delta_t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta_t)}{P(T \geq t)\Delta_t} \right\} \\
&= \lim_{\Delta_t \rightarrow 0} \left\{ \frac{P(T < t + \Delta_t) - P(T < t)}{P(T \geq t)\Delta_t} \right\}
\end{aligned} \tag{2.11}$$

atau dapat ditulis seperti pada persamaan (2.12):

$$h(t) = \lim_{\Delta_t \rightarrow 0} \left\{ \frac{F(t + \Delta_t) - F(t)}{\Delta_t} \right\} \frac{1}{S(t)}$$

$$= \frac{f(t)}{S(t)} \quad (2.12)$$

Karena $S(t) = 1 - F(t)$, maka $S'(t) = -f(t)$, selanjutnya digunakan turunan logaritma alami:

$$\begin{aligned} \frac{S'(t)}{S(t)} &= \frac{d}{dt} \{\log S(t)\} \\ \frac{-f(t)}{S(t)} &= \frac{d}{dt} \{\log S(t)\} \\ h(t) &= -\frac{d}{dt} \{\log S(t)\} \end{aligned} \quad (2.13)$$

Untuk mencari $S(t)$, maka perlu menghilangkan turunan $(\frac{d}{dt})$ pada persamaan (2.13) dengan cara mengintegralkan kedua sisi terhadap t .

$$\begin{aligned} \int_0^t h(u) du &= - \int_0^t \frac{d}{du} \{\log S(u)\} du \\ &= -[\log S(t) - \log S(0)] \end{aligned} \quad (2.14)$$

Karena $\log S(0) = \log 1 = 0$, maka persamaan (2.14) menjadi:

$$\int_0^t h(u) du = -\log S(t) \quad (2.15)$$

$H(t)$ disebut fungsi *hazard* kumulatif. Sehingga, dari persamaan (2.15), dapat ditulis ulang menjadi persamaan (2.16):

$$H(t) = -\log S(t) \leftrightarrow S(t) = \exp[-H(t)] \quad (2.16)$$

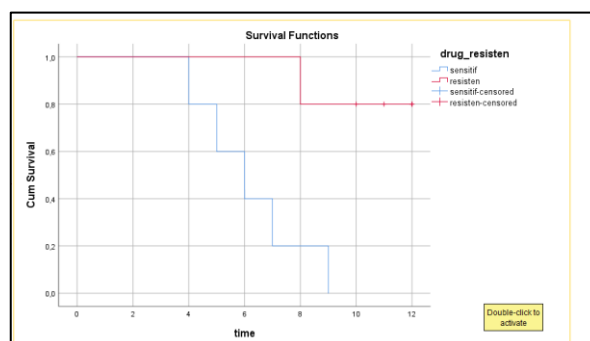
2.1.5 Metode Kaplan-Meier

Metode *Kaplan-Meier* merupakan teknik analisis *survival* yang dibangun berdasarkan informasi waktu terjadinya suatu *event* pada masing-masing individu. Dalam penerapannya, metode ini mengasumsikan bahwa kejadian sensor terjadi secara independen terhadap waktu terjadinya *event*, artinya alasan suatu individu

tersensor tidak berhubungan langsung dengan peluang terjadinya *event* yang diamati. Metode ini mengandalkan dua prinsip utama:

1. Adanya individu-individu yang tersensor dalam data, yang tetap disertakan dalam perhitungan hingga waktu sensor terjadi.
2. Probabilitas *survival* dihitung secara berurutan, yaitu dengan mengalikan peluang belum mengalami *event* pada setiap waktu kejadian.

Keunggulan metode *Kaplan-Meier* terletak pada kemampuannya dalam memberikan estimasi fungsi *survival* yang lebih akurat, karena menggunakan data waktu hingga terjadinya *event* pada setiap individu secara aktual, bukan berdasarkan pengelompokan dalam interval waktu (Kleinbaum & Klein, 2012). Ilustrasi kurva *Kaplan-Meier* dapat dipelajari pada Gambar 2.2.



Gambar 2.2 Ilustrasi Kurva *Kaplan-Meier* terhadap Waktu Kesembuhan

Berdasarkan Gambar 2.2, sumbu horizontal atau sumbu X menampilkan waktu hingga terjadinya suatu *event*, sedangkan sumbu vertikal atau sumbu Y menunjukkan fungsi *survival*, yaitu probabilitas individu belum mengalami *event* hingga waktu tertentu. Garis berwarna biru menunjukkan kategori pasien dengan obat sensitif, sementara garis berwarna merah menunjukkan kategori pasien dengan obat resisten. Terlihat bahwa nilai fungsi *survival* pada kategori sensitif menurun lebih cepat dibandingkan dengan kategori resisten, yang menunjukkan bahwa pasien dengan obat sensitif memiliki laju kesembuhan lebih tinggi selama periode

pengamatan. Titik-titik silang (*censored*) pada kedua garis menunjukkan data pasien yang tidak mengalami kejadian sembuh hingga akhir waktu pengamatan.

Estimasi *Kaplan-Meier* digunakan untuk menghitung fungsi *survival* dan menyajikannya dalam bentuk kurva, sehingga memudahkan dalam mengamati hubungan antara fungsi *survival* dan waktu terjadinya *event* (Khairunnisa dkk., 2022). Fungsi *survival* dilambangkan dengan $S(t)$. Formula Kaplan Meier memberikan probabilitas bahwa suatu individu belum mengalami peristiwa (*event*) hingga melewati waktu kejadian ($t_{(j)}$) dengan mempertimbangkan dua komponen utama:

1. Probabilitas belum mengalami *event* hingga waktu sebelumnya $t_{(j-1)}$, dan
2. Probabilitas bersyarat untuk belum mengalami *event* melewati $t_{(j)}$, dengan syarat individu belum mengalami *event* hingga $t_{(j)}$.

Dalam konteks pengobatan TB, proses estimasi *Kaplan-Meier* dilakukan dengan memperbarui peluang pasien belum sembuh pada setiap waktu kejadian sembuh. Pada setiap titik waktu kejadian, probabilitas belum sembuh pada waktu sebelumnya dikalikan dengan peluang pasien tetap belum sembuh hingga waktu kejadian sembuh saat ini. Dengan cara ini, diperoleh kurva *survival* yang menggambarkan kemungkinan pasien belum mencapai kesembuhan dari waktu ke waktu. Secara matematis, estimasi *Kaplan-Meier* dapat ditulis sebagai berikut (Kleinbaum & Klein, 2012):

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \times \hat{P}(T > t_{(j)} | T \geq t_{(j)}) \quad (2.17)$$

dengan

$$\hat{P}(T > t_{(j)} | T \geq t_{(j)}) = 1 - \frac{d_j}{b_j}.$$

Keterangan:

d_j : Banyaknya individu yang mengalami *event* pada waktu $t_{(j)}$

b_j : Banyaknya individu yang masih dalam pengamatan/belum mengalami *event* tepat sebelum $t_{(j)}$

$t_{(j)}$: Waktu *event* ke- j

Pemberian syarat $T \geq t_{(j)}$ diperlukan untuk memastikan bahwa perhitungan peluang hanya didasarkan pada individu yang masih berisiko mengalami *event* pada waktu tersebut. Dengan demikian, individu yang telah tersensor sebelum $t_{(j)}$ tidak dimasukkan dalam perhitungan, sehingga estimasi fungsi *survival* menjadi tidak bias dan merepresentasikan kondisi sebenarnya dari individu yang masih dalam pengamatan.

2.1.6 Uji *Log-rank*

Uji *log-rank* merupakan salah satu metode yang paling sering digunakan dalam analisis *survival*, khususnya untuk membandingkan kurva *survival* antara dua atau lebih kategori dari suatu variabel kategorik (Kleinbaum & Klein, 2012). Secara teknis, uji *log-rank* termasuk dalam kategori uji *Chi-Square* yang diaplikasikan pada data dengan jumlah sampel besar. Hipotesis yang diuji dalam prosedur ini adalah sebagai berikut (Harlan, 2017):

1. H_0 : Tidak terdapat perbedaan kurva *survival* di antara kategori yang dibandingkan.
2. H_1 : Setidaknya terdapat satu kategori dengan kurva *survival* yang berbeda dari kategori lainnya.

Uji ini bertujuan untuk menilai kesamaan bentuk kurva *survival* serta mendeteksi adanya perbedaan yang signifikan dalam waktu hingga terjadinya *event* berdasarkan variabel-variabel kategorik yang digunakan dalam penelitian. Ketika hasil uji *log-rank* menunjukkan nilai yang signifikan, maka dapat disimpulkan bahwa distribusi *survival* antar kategori berbeda secara nyata (Khairunnisa dkk., 2022). Statistik uji yang digunakan dalam uji *log-rank* yaitu (Kleinbaum & Klein, 2012):

$$\chi^2_{hitung} = \sum_{i=1}^C \frac{(d_i - E_i)^2}{E_i} \quad (2.18)$$

dengan

$$d_i - E_i = \sum_{j=1}^J (d_{ij} - e_{ij})$$

$$e_{ij} = \left(\frac{b_{ij}}{\sum_{i=1}^C b_{ij}} \right) \left(\sum_{i=1}^C d_{ij} \right)$$

Keterangan:

C : Banyaknya kategori

d_i : Banyaknya individu yang mengalami kejadian pada kategori ke- i , $i = 1, 2, \dots, G$

E_i : Banyaknya individu yang diharapkan mengalami kejadian pada kategori ke- i , yaitu nilai ekspektasi jika tidak ada perbedaan antar kategori

J : Banyaknya waktu kejadian

d_{ij} : Banyaknya individu yang mengalami kejadian pada kategori ke- i dan waktu ke- j

e_{ij} : Banyaknya individu yang diharapkan mengalami kejadian pada kategori ke- i dan waktu ke- j

b_{ij} : Banyaknya individu yang belum mengalami *event* pada kategori ke- i dan waktu ke- j

j : Urutan waktu terjadinya *event* $(1, 2, 3, \dots, J)$

Pengambilan keputusan yang digunakan dalam uji *log-rank* ini yaitu tolak H_0 jika $\chi^2 > \chi^2_{(\alpha, C-1)}$ atau $p\text{-value} < \alpha = 0.05$.

2.1.7 Model Cox Proportional Hazard

Model *Cox Proportional Hazard* termasuk dalam kategori model semiparametrik, karena tidak memerlukan asumsi distribusi tertentu untuk waktu *survival*. Estimasi parameter regresi pada model ini dapat dilakukan tanpa harus menentukan bentuk eksplisit dari fungsi *hazard* dasar (*baseline hazard*) (T.Lee & Wang, 2003). Disebut model Cox karena model ini berdasarkan pada asumsi *proportional hazard*, yakni fungsi *hazard* antar individu yang berbeda memiliki proporsi yang konstan sepanjang waktu. Dengan kata lain, rasio *hazard* antara dua individu tidak berubah seiring waktu (konstan). Dengan menggunakan model Cox, maka dapat dilihat hubungan antara variabel independen terhadap variabel dependen, yaitu waktu *survival* melalui fungsi *hazard* yang ada. Rumus umum dari model *Cox Proportional Hazard* dapat dituliskan (Kleinbaum & Klein, 2012):

$$\begin{aligned} h(t, \mathbf{X}) &= h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_q X_q) \\ &= h_0(t) \exp\left(\sum_{p=1}^q \beta_p X_p\right) \end{aligned} \quad (2.19)$$

Keterangan:

$h(t, \mathbf{X})$: Fungsi *hazard* pada waktu t berdasarkan kovariat $\mathbf{X}$

$h_0(t)$: Fungsi dasar *hazard* (*baseline hazard function*) pada waktu t , yaitu nilai fungsi *hazard* pada saat semua variabel kovariat bernilai nol

X_p : Nilai variabel prediktor (kovariat) ke- p , dengan $p = 1, 2, \dots, q$

β_p : Koefisien regresi pada model Cox yang menyatakan pengaruh kovariat ke- p terhadap *hazard*

Estimasi parameter $\beta_1, \beta_2, \dots, \beta_q$ dapat menggunakan metode MPLE atau *Maximum Partial Likelihood Estimation*. Fungsi dari *partial likelihood* dapat dinyatakan pada persamaan (2.20):

$$L(\beta) = \prod_{j=1}^J \left(\frac{\exp(\beta' X_{(j)})}{\sum_{m \in R(t_{(j)})} \exp(\beta' X_m)} \right) \quad (2.20)$$

dengan

$\beta' = (\beta_1, \beta_2, \dots, \beta_q)$: Transpose dari vektor parameter

$X_{(j)} = (X_{1(j)}, X_{2(j)}, \dots, X_{q(j)})'$: Vektor kovariat dari individu yang mengalami *event* pada waktu $t_{(j)}$

$X_m = (X_{1m}, X_{2m}, \dots, X_{qm})'$: Vektor kovariat individu ke- m yang termasuk dalam *risk set* $R(t_{(j)})$

$R(t_{(j)})$ atau *risk set* adalah himpunan individu yang masih berisiko (*at risk*) mengalami kejadian tepat sebelum waktu $t_{(j)}$. Apabila data terdiri dari J waktu kejadian yang berbeda, yang dinotasikan sebagai $t_{(1)}, t_{(2)}, \dots, t_{(J)}$, maka fungsi *log partial likelihood* dapat dinyatakan pada persamaan (2.21):

$$\begin{aligned} \log L(\beta) &= \sum_{j=1}^J \left[\beta' X_{(j)} - \log \left(\sum_{m \in R(t_j)} \exp(\beta' X_m) \right) \right] \\ &= \sum_{j=1}^J \left[\sum_{p=1}^q \beta_p X_{p(j)} - \log \left(\sum_{m \in R(t_j)} \exp \left(\sum_{p=1}^q \beta_p x_{pm} \right) \right) \right] \\ &= \sum_{j=1}^J \sum_{p=1}^q \beta_p X_{p(j)} - \sum_{j=1}^J \log \left(\sum_{m \in R(t_j)} \exp \left(\sum_{p=1}^q \beta_p x_{pm} \right) \right) \end{aligned} \quad (2.21)$$

Dimisalkan $u(\beta)$ adalah turunan pertama dari $\log L(\beta)$ yang dapat dituliskan dalam bentuk matriks $q \times 1$:

$$u(\boldsymbol{\beta}) = \begin{pmatrix} u(\beta_1) \\ u(\beta_2) \\ \vdots \\ u(\beta_q) \end{pmatrix}$$

dengan

$$u(\beta_p) = \frac{\partial \log L(\boldsymbol{\beta})}{\partial (\beta_p)}, p = 1, 2, \dots, q \quad (2.22)$$

Didapatkan persamaan (2.23) dari turunan pertama fungsi log likelihood yaitu

$$\begin{aligned} \frac{\partial \log L(\boldsymbol{\beta})}{\partial \beta_p} &= \sum_{j=1}^J X_{p(j)} - \sum_{j=1}^J \delta_j \frac{\sum_{m \in R(t_{(j)})} X_{pm} \exp(\sum_{p=1}^q \beta_p X_{pm})}{\sum_{m \in R(t_{(j)})} \exp(\sum_{p=1}^q \beta_p X_{pm})} \\ &= \sum_{j=1}^J \left[X_{p(j)} - \left(\frac{\sum_{m \in R(t_{(j)})} X_{pm} \exp(\sum_{p=1}^q \beta_p X_{pm})}{\sum_{m \in R(t_{(j)})} \exp(\sum_{p=1}^q \beta_p X_{pm})} \right) \right] \end{aligned} \quad (2.23)$$

$$\text{Sehingga diperoleh } u(\beta_p) = \sum_{j=1}^J \left[X_{p(j)} - \left(\frac{\sum_{m \in R(t_{(j)})} X_{pm} \exp(\sum_{p=1}^q \beta_p X_{pm})}{\sum_{m \in R(t_{(j)})} \exp(\sum_{p=1}^q \beta_p X_{pm})} \right) \right] = 0$$

Selanjutnya, dimisalkan $l(\boldsymbol{\beta})$ adalah matriks yang berukuran $q \times q$ dari turunan kedua fungsi *log partial likelihood*, dengan

$$l(\beta_p) = - \left(\frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_p \partial \beta_z} \right), p = 1, 2, \dots, q \text{ dan } z = 1, 2, \dots, q \quad (2.24)$$

maka

$$l(\boldsymbol{\beta})_{q \times q} = - \begin{pmatrix} \frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_1 \partial \beta_1} & \dots & \frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_1 \partial \beta_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_q \partial \beta_1} & \dots & \frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_q \partial \beta_q} \end{pmatrix}$$

Digunakan pembagian vektor $\frac{u}{v}$ dalam persamaan (2.24) untuk mencari turunan

kedua dari fungsi $\log L(\boldsymbol{\beta})$, yaitu

$$\begin{aligned} u &= \sum_{m \in R(t_{(j)})} X_{pm} \exp \left(\sum_{p=1}^q \beta_p X_{pm} \right) \\ u' &= \sum_{m \in R(t_{(j)})} X_{pm} X_{zm} \exp \left(\sum_{p=1}^q \beta_p X_{pm} \right) \end{aligned}$$

$$v = \sum_{m \in R(t_{(j)})} \exp \left(\sum_{p=1}^q \beta_p X_{pm} \right)$$

$$v' = \sum_{m \in R(t_{(j)})} X_{zm} \exp \left(\sum_{p=1}^q \beta_p X_{pm} \right)$$

Sehingga diperoleh

$$\frac{u'v - v'u}{v^2} = \sum_{j=1}^J \frac{u'}{v} - \sum_{j=1}^J \left[\frac{v'}{v} \right] \left[\frac{u}{v} \right]$$

$$\frac{\partial^2 \log(L(\boldsymbol{\beta}))}{\partial \beta_p \partial \beta_z} = - \left\{ \sum_{j=1}^J \left\{ \frac{u'}{v} - \left[\frac{v'}{v} \right] \left[\frac{u}{v} \right] \right\} \right\} \quad (2.25)$$

persamaan (2.22) dan (2.25) bukan merupakan persamaan *closed form*, sehingga tidak dapat diselesaikan secara langsung dan membutuhkan metode numerik. Salah satu metode numerik yang efektif untuk menyelesaikan sistem persamaan tersebut adalah metode iterasi Newton–Raphson (Widyastuti dkk., 2024). Oleh karena itu, estimasi parameter pada model *Cox Proportional Hazard* dilakukan melalui prosedur iteratif Newton–Raphson, yang dinyatakan pada persamaan (2.26) (Safitri dkk., 2016):

$$(\widehat{\boldsymbol{\beta}}_{s+1})_{q \times 1} = (\widehat{\boldsymbol{\beta}}_s)_{q \times 1} + (l^{-1}(\widehat{\boldsymbol{\beta}}_s))_{q \times q} (u(\widehat{\boldsymbol{\beta}}_s))_{q \times 1}, s = 1, 2, \dots \quad (2.26)$$

Proses iterasi dimulai dengan menentukan nilai awal $(\widehat{\boldsymbol{\beta}}_1)_{q \times 1} = \mathbf{0}$. Pada setiap iterasi ke- s , vektor parameter $\widehat{\boldsymbol{\beta}}_s$ yang berisi koefisien regresi $(\beta_{1,s}, \beta_{2,s}, \dots, \beta_{q,s})'$ akan diperbarui untuk menghasilkan vektor parameter baru $\widehat{\boldsymbol{\beta}}_{s+1} = (\beta_{1,s+1}, \beta_{2,s+1}, \dots, \beta_{q,s+1})'$. Dengan demikian, $\widehat{\boldsymbol{\beta}}_s$ adalah nilai parameter sebelum pembaruan, sedangkan $\widehat{\boldsymbol{\beta}}_{s+1}$ merupakan nilai parameter setelah pembaruan pada iterasi ke- $s + 1$. Untuk menentukan apakah iterasi sudah mencapai konvergensi, digunakan ukuran perubahan parameter antar-iterasi dengan menghitung jarak Euclidean antara dua vektor parameter berturut-turut, yang dinyatakan dalam persamaan (2.27):

$$\|\hat{\beta}_{s+1} - \hat{\beta}_s\| = \sqrt{\sum_{p=1}^q (\beta_{p,s+1} - \beta_{p,s})^2} \quad (2.27)$$

Iterasi Newton–Raphson dihentikan apabila perubahan parameter sudah sangat kecil, yaitu Ketika persamaan (2.28) terpenuhi.

$$\|\hat{\beta}_{s+1} - \hat{\beta}_s\| < \varepsilon \quad (2.28)$$

dengan $\|\dots\|$ menyatakan norma Euclidean atau panjang vector.

Dalam beberapa penelitian, ambang konvergensi ditetapkan sebesar $\varepsilon = 10^{-6}$, yang memastikan bahwa iterasi berhenti ketika nilai parameter telah stabil dan perubahan antar-iterasi menjadi sangat kecil (Alim & Maulana, 2025).

2.1.8 Uji Asumsi *Cox Proportional Hazard*

Sebuah situasi dikatakan memenuhi asumsi *proportional hazard* apabila rasio *hazard* (perbandingan laju terjadinya *event* antara dua kategori) konstan terhadap waktu. Untuk mengevaluasi apakah asumsi ini terpenuhi, terdapat dua pendekatan yang umum digunakan, yaitu:

1. Uji Visual

Pengujian secara visual terhadap asumsi *proportional hazard* dapat dilakukan menggunakan grafik *Kaplan-Meier*. Apabila kurva *survival* antar kategori cenderung sejajar dan tidak menunjukkan pola perpotongan yang jelas, maka asumsi *proportional hazard* kemungkinan terpenuhi. Sebaliknya, apabila kurva *survival* saling berpotongan, maka terdapat indikasi bahwa asumsi *proportional hazard* mungkin tidak terpenuhi (Dahlan, 2012).

2. Uji Formal

Pengujian asumsi *proportional hazard* secara formal dapat dilakukan menggunakan uji *Schoenfeld residual*. Pengujian ini bertujuan untuk

mengetahui apakah terdapat hubungan antara residual Schoenfeld dengan waktu *survival*. Jika tidak terdapat hubungan, maka pengaruh kovariat terhadap *hazard* bersifat konstan terhadap waktu sehingga asumsi *proportional hazard* terpenuhi. Langkah-langkah pengujian adalah sebagai berikut (Kleinbaum & Klein, 2012):

- a. Mengestimasi model *Cox Proportional Hazard* untuk memperoleh residual Schoenfeld dari masing-masing kovariat.
- b. Mengurutkan waktu *survival* dari yang terkecil hingga terbesar
- c. Menguji hubungan antara residual Schoenfeld dan fungsi waktu

Dalam implementasinya, pengujian dilakukan dengan mengevaluasi hubungan antara residual Schoenfeld dan fungsi waktu menggunakan statistik uji yang mengikuti distribusi χ^2 . Hipotesis yang digunakan adalah sebagai berikut (Widyastuti dkk., 2024):

- i. H_0 : Pengaruh kovariat terhadap *hazard* konstan terhadap waktu (asumsi *proportional hazard* terpenuhi)
- ii. H_1 : Pengaruh kovariat berubah terhadap waktu (asumsi *proportional hazard* tidak terpenuhi)

Statistik uji yang digunakan adalah:

$$\chi^2 = \frac{\hat{\gamma}^2}{Var(\hat{\gamma})} \quad (2.29)$$

Keterangan:

$\hat{\gamma}$: Koefisien hubungan antara residual Schoenfeld dan waktu

$Var(\hat{\gamma})$: Varians dari koefisien

Kriteria pengujiannya adalah jika $p - value > 0.05$ atau $\chi^2_{hitung} \leq \chi^2_{0.05,df}$, maka H_0 gagal ditolak, sehingga tidak terdapat hubungan antara residual Schoenfeld dan waktu *survival*, dan asumsi *proportional hazard* terpenuhi. Jika $p - value \leq 0.05$ atau $\chi^2_{hitung} > \chi^2_{0.05,df}$, maka H_0 ditolak, sehingga terdapat hubungan antara residual Schoenfeld dan waktu *survival*, dan asumsi *proportional hazard* tidak terpenuhi.

Dalam analisis menggunakan MICE, uji asumsi formal dilakukan pada masing-masing dataset imputasi. Perbedaan hasil antar dataset merupakan hal yang wajar karena adanya variasi akibat proses imputasi. Oleh karena itu, keputusan tidak didasarkan pada jumlah dataset yang melanggar secara sederhana, melainkan pada konsistensi pola pelanggaran (Rubin, 1987).

2.1.9 Pengujian Parameter

Pengujian terhadap parameter dilakukan untuk menentukan apakah variabel-variabel independen memiliki pengaruh signifikan terhadap variabel dependen, yaitu waktu *survival* (*survival time*) atau waktu hingga terjadinya *event*. Setelah model terbentuk, pengujian dilakukan baik secara serentak maupun parsial.

1. Uji Serentak (*Likelihood Ratio Test*)

Uji ini bertujuan untuk menguji apakah seluruh parameter regresi secara bersama-sama signifikan, atau dengan kata lain, apakah model secara keseluruhan memiliki kekuatan prediksi. Hipotesis yang digunakan dalam pengujian serentak adalah sebagai berikut:

- a. $H_0 : \beta_p = 0, \forall p = 1, 2, \dots, q$, yang berarti seluruh variabel prediktor tidak berpengaruh terhadap waktu kejadian

- b. H_1 : Paling tidak terdapat satu $\beta_p \neq 0$, yang berarti terdapat variabel prediktor yang berpengaruh terhadap waktu kejadian.

Pada uji serentak menggunakan statistik uji pada persamaan (2.30):

$$LR = -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})} \sim \chi^2_{\alpha, df} \quad (2.30)$$

Keterangan:

$L\hat{\omega}$: Nilai likelihood dari model tanpa memasukkan variabel prediktor (model kosong).

$L\hat{\Omega}$: Nilai likelihood dari model yang menyertakan seluruh variabel prediktor (model penuh).

Dalam penelitian ini, uji signifikansi secara serentak menggunakan *likelihood ratio test* dilakukan pada masing-masing dataset imputasi, kemudian hasilnya dibandingkan untuk melihat konsistensi. Pendekatan ini dilakukan karena likelihood tidak dapat digabungkan secara langsung dalam MICE (Rubin, 1987). Pengambilan keputusan dalam uji serentak *likelihood* ini adalah tolak H_0 jika $LR > \chi^2_{\alpha, df}$ atau $p - value < \alpha = 5\% = 0.05$ (Kleinbaum & Klein, 2012).

2. Uji Parsial (Wald Test)

Uji parsial dilakukan untuk menilai apakah masing-masing variabel prediktor secara individu memberikan kontribusi yang signifikan terhadap model. Artinya, setiap parameter regresi diuji satu per satu untuk mengetahui pengaruhnya terhadap waktu hingga terjadinya *event*. Hipotesis yang digunakan dalam uji Wald adalah:

- a. $H_0 : \beta_p = 0$ dengan $p = 1, 2, \dots, q$ (variabel prediktor ke- p tidak memiliki pengaruh yang signifikan).

- b. $H_1 : \beta_p \neq 0$ dengan $p = 1, 2, \dots, q$ (variabel prediktor ke- p memiliki pengaruh yang signifikan).

Statistik uji yang digunakan dalam uji Wald dihitung dalam persamaan (2.31):

$$W^2 = \frac{(\hat{\beta}_p)^2}{(SE(\hat{\beta}_p))^2} \sim \chi^2_{\alpha, df} \quad (2.31)$$

dengan

$$SE(\hat{\beta}_p) = \sqrt{var(\hat{\beta}_p)}$$

Keterangan:

$\hat{\beta}_p$: Koefisien regresi untuk kovariat ke- p

$SE(\hat{\beta}_p)$: *Standard error* dari $\hat{\beta}_p$

$var(\hat{\beta}_p)$: Varians dari $\hat{\beta}_p$

Pengambilan keputusan dalam uji parsial ini adalah tolak H_0 jika $W^2 > \chi^2_{\alpha, 1}$ atau $p - value < \alpha = 5\% = 0.05$ (Kleinbaum & Klein, 2012).

2.1.10 Rasio Hazard

Hazard Ratio (HR) merupakan ukuran yang digunakan untuk membandingkan laju terjadinya suatu peristiwa (*event*) antara dua kategori atau dua individu yang berbeda pada variabel prediktor (Hosmer & Lemeshow, 2000). Estimasi nilai HR diperoleh dengan menghitung eksponensial dari koefisien regresi Cox pada masing-masing variabel prediktor yang berpengaruh signifikan terhadap *hazard*. Misal X_p adalah nilai dari variabel prediktor ke- p , dan variabel tersebut memiliki dua kategori, yaitu 0 dan 1. Hubungan antara variabel X_p dengan *hazard rate* dapat dinyatakan dengan:

$$h(t|X_p) = h_0(t)e^{\beta_p X_p} \quad (2.32)$$

Maka dari persamaan (2.32), individu dengan kategori $X_p = 1$, fungsi *hazard*nya:

$$h(t|X_p = 1) = h_0(t)e^{\beta_p \cdot 1} = h_0(t)e^{\beta_p}$$

Individu dengan $X_p = 0$, fungsi *hazard*nya:

$$h(t|X_p = 0) = h_0(t)e^{\beta_p \cdot 0} = h_0(t)$$

Sehingga nilai HR dapat dihitung dengan rumus:

$$HR = \frac{h(t|X_p = 1)}{h(t|X_p = 0)} = \frac{h_0(t)e^{\beta_p}}{h_0(t)} = e^{\beta_p} \quad (2.33)$$

Nilai *Hazard Ratio* (HR) yang diperoleh dari persamaan (2.33) menunjukkan perbandingan laju terjadinya *event* antara dua kategori dari variabel prediktor. Jika $HR = e^{\beta_p} > 1$, maka interpretasinya adalah bahwa individu dengan kategori $X_p = 1$ memiliki laju terjadinya *event* sebesar e^{β_p} kali lebih tinggi dibandingkan individu dengan kategori $X_p = 0$. Sebaliknya, jika $HR < 1$, maka individu dengan kategori $X_p = 1$ memiliki laju terjadinya *event* yang lebih rendah dibandingkan individu dengan kategori $X_p = 0$. Untuk variabel prediktor kontinu, nilai e^{β_p} menyatakan perbandingan laju terjadinya *event* antara dua individu yang berbeda satu satuan pada nilai variabel X_p . Dengan kata lain, setiap peningkatan satu satuan pada variabel X_p akan mengubah *hazard* sebesar e^{β_p} kali dibandingkan individu dengan nilai X_p yang lebih rendah satu satuan.

2.1.11 Tuberkulosis

Tuberkulosis (TB) merupakan penyakit infeksi menular yang disebabkan oleh *Mycobacterium Tuberculosis*, yaitu bakteri berbentuk batang (basil) yang dikenal dengan sebutan Basil Tahan Asam (BTA). Penularan penyakit ini umumnya terjadi melalui udara ketika penderita TB paru batuk atau bersin, sehingga bakteri

dapat masuk ke paru-paru orang lain dan menginfeksi jaringan parenkim paru. Infeksi ini dapat menyebabkan berbagai komplikasi serius dan berpotensi fatal apabila tidak segera ditangani dengan pengobatan dan pencegahan yang tepat (Avy dkk., 2024).

Tuberkulosis tidak hanya menyerang paru-paru, tetapi juga dapat menginfeksi organ tubuh lainnya seperti sistem saraf pusat, persendian (artikulasi), kelenjar getah bening, dan peritoneum. Berdasarkan organ yang terinfeksi, TB diklasifikasikan menjadi dua jenis utama, yaitu tuberkulosis paru dan tuberkulosis ekstra paru. TB paru merupakan infeksi yang menyerang jaringan parenkim paru, sedangkan TB ekstra paru menyerang organ tubuh lain di luar paru, seperti tulang, otak, atau ginjal. Pengobatan tuberkulosis dilakukan melalui terapi kombinasi beberapa jenis obat dengan dosis yang tepat sekitar 6-12 bulan untuk memastikan bakteri penyebab infeksi dapat terbunuh seluruhnya. Namun, lamanya durasi pengobatan sering kali menjadi kendala karena sebagian penderita menghentikan terapi sebelum waktunya. Faktor lain yang turut memengaruhi ketidaktuntasan pengobatan adalah kurangnya pengetahuan pasien tentang pentingnya kepatuhan terapi dan perawatan selama proses penyembuhan. Seseorang dinyatakan sembuh dari tuberkulosis apabila telah memenuhi kriteria kesembuhan berdasarkan standar nasional, yaitu tingkat keberhasilan pengobatan lebih dari 85%. Setelah dinyatakan sembuh, mantan penderita TB juga perlu mendapatkan edukasi lanjutan mengenai penyebab dan risiko kekambuhan untuk mencegah terjadinya infeksi ulang di kemudian hari (Nurfajilah dkk., 2023). Proses seseorang terinfeksi tuberkulosis (TB) dipengaruhi oleh tiga faktor utama, yaitu:

1. Durasi kontak dengan penderita TB; semakin lama seseorang berinteraksi dengan pasien TB, semakin tinggi risiko penularannya.
2. Jumlah bakteri TB; semakin banyak jumlah bakteri *Mycobacterium Tuberculosis* yang terdapat pada pasien, semakin besar pula tingkat penularannya.
3. Kekuatan daya tahan tubuh; individu dengan imunitas rendah seperti penderita HIV/AIDS, Diabetes Mellitus (DM), lansia, dan anak-anak memiliki risiko lebih tinggi untuk tertular TB.

Pencegahan utama agar seseorang tidak terinfeksi *Mycobacterium Tuberculosis* adalah dengan menemukan kasus TB secara dini dan segera memberikan pengobatan hingga tuntas, sehingga rantai penularan dapat diputus. Jika pasien TB tidak segera diobati, ia berpotensi menularkan penyakitnya kepada 10–15 orang lainnya. Program penanggulangan TB yang berkesinambungan dan tepat sasaran meliputi langkah pencegahan, deteksi dini, serta pengobatan. Pengobatan TB menjadi strategi paling efektif dalam mencegah penyebaran bakteri penyebab penyakit ini. Terapi utama dilakukan dengan pemberian kombinasi beberapa antibiotik yang disebut Obat Anti Tuberkulosis (OAT), yang umumnya terdiri atas empat jenis obat utama: Rifampisin, Isoniazid, Pirazinamid, dan Etambutol. (Kementerian Kesehatan, 2025).

Berikut faktor-faktor penduga laju kesembuhan TB yang digunakan dalam penelitian ini sebagai variabel independen:

1. Usia

Berdasarkan penelitian yang dilakukan di salah satu puskesmas di Lampung Selatan, diketahui bahwa peluang kesembuhan (*survival probability*) pada

pasien tuberkulosis berusia produktif mencapai 4,3%, sedangkan pada pasien dengan usia tidak produktif sebesar 0%. Hasil tersebut menunjukkan bahwa pasien tuberkulosis dengan usia produktif memiliki kemungkinan sembuh yang lebih tinggi dibandingkan dengan pasien yang berada pada usia tidak produktif (Antika dkk., 2025).

2. Jenis Kelamin

Beberapa penelitian menyatakan bahwa perempuan lebih berpeluang untuk mengalami keberhasilan pengobatan tuberkulosis paru. Pasien berjenis kelamin laki-laki cenderung memiliki tingkat kepatuhan yang lebih rendah dibandingkan perempuan, serta kurang memperhatikan aspek kesehatan dan menerapkan gaya hidup yang kurang sehat. Hal ini sejalan dengan beberapa teori yang menyebutkan bahwa perempuan lebih sering melaporkan gejala penyakit dan berkonsultasi dengan tenaga medis karena memiliki sifat yang lebih teliti dan tekun dibandingkan laki-laki. Selain itu, perempuan juga dinilai lebih sabar dalam menjalani proses pengobatan, sehingga tingkat putus pengobatan lebih banyak terjadi pada pasien laki-laki dibandingkan perempuan (Antika dkk., 2025).

3. Status Pengobatan

Status pengobatan yang dimaksud adalah sesuai standar/tidak yang diukur berdasarkan standar nasional, yaitu jika tingkat keberhasilan pengobatan lebih dari 85%. Pengobatan TB yang tidak dilakukan sesuai standar berisiko menurunkan efektivitas terapi dan dapat berujung pada kegagalan pengobatan, yang akhirnya berpotensi menimbulkan TB resisten obat (TB RO). Sebaliknya, pasien dianggap mengikuti standar jika menjalankan

pengobatan dengan paduan OAT (Obat Anti Tuberkulosis) yang tepat, dosis benar, teratur, dan diawasi oleh PMO (Pengawas Menelan Obat) hingga selesai (Kementerian Kesehatan, 2025).

4. Riwayat Pengobatan

Hasil penelitian yang dilakukan oleh Tri Wahyuni menunjukkan bahwa riwayat pengobatan yang paling berpengaruh terhadap lamanya durasi pengobatan adalah kasus putus obat. Dalam penelitiannya, kasus putus obat memiliki nilai *odds ratio* (OR) sebesar 39.2, yang berarti pasien dengan riwayat putus obat memiliki risiko 39.2 kali lebih besar untuk mengalami *multidrug resistance* (MDR), sehingga cenderung memerlukan waktu pengobatan yang lebih lama dibandingkan pasien dengan riwayat kambuh (Adawiyah dkk., 2023).

5. Jenis Fasilitas Layanan Kesehatan (Fasyankes)

Fasilitas kesehatan merupakan salah satu faktor pendukung utama dalam upaya peningkatan derajat kesehatan, karena perubahan perilaku pasien sangat dipengaruhi oleh kualitas sarana kesehatan yang tersedia. Banyak individu terdorong untuk memperbaiki perilaku kesehatannya ketika memperoleh pelayanan yang baik dan fasilitas yang memadai, yang dapat menunjang proses penyembuhan. Hal ini diperkuat oleh penelitian Pratama yang menunjukkan bahwa pasien TB lebih rajin menjalani pengobatan apabila fasilitas kesehatan yang disediakan sesuai dengan harapan mereka. Dengan demikian, ketersediaan fasilitas kesehatan yang memadai dapat meningkatkan tingkat kesembuhan pasien TB (Pratama, 2021).

6. Tipe Diagnosis TB

Pasien TB terkonfirmasi bakteriologis merupakan pasien yang terbukti positif melalui pemeriksaan bakteriologi, seperti uji pada dahak (*sputum*), cairan tubuh, atau jaringan, dengan metode pemeriksaan mikroskopis langsung, TCM TB, maupun biakan. Sementara itu, pasien dengan diagnosis klinis adalah mereka yang tidak memenuhi kriteria diagnosis bakteriologis, namun berdasarkan penilaian dokter didiagnosis sebagai penderita TB aktif dan diputuskan untuk menjalani pengobatan TB (Kementerian Kesehatan, 2025).

7. Hasil Tes HIV (*Human Immunodeficiency Virus*)

Penelitian yang dilakukan oleh Laili (2024) menunjukkan bahwa individu dengan ko-infeksi TB-HIV memiliki risiko 2.3 kali lebih besar mengalami kegagalan pengobatan TB resisten obat (TB RO) di Indonesia. Temuan ini konsisten dengan hasil penelitian Koroma (2024), yang menyebutkan bahwa penderita HIV positif berisiko 2.2 kali lebih tinggi mengalami kegagalan pengobatan *Multidrug Resistant Tuberculosis* (MDR-TB) dibandingkan dengan individu yang negatif HIV (Laili dkk., 2024).

8. Hasil Tes DM (*Diabetes Mellitus*)

Hubungan antara TB dan diabetes penting karena keduanya saling memengaruhi dalam hal pengobatan dan respons tubuh. Diabetes dapat melemahkan sistem imun, sehingga pengobatan TB menjadi kurang efektif dan risiko kekambuhan atau kematian meningkat. Sebaliknya, infeksi TB dapat menaikkan kadar gula darah, yang menyulitkan pengendalian diabetes. Selain itu, pengobatan jangka panjang untuk kedua penyakit ini

berpotensi menimbulkan interaksi obat yang mengurangi efektivitas terapi. Penyakit diabetes melitus dapat meningkatkan risiko terjadinya TB sebesar dua hingga tiga kali lipat, serta memperbesar kemungkinan terjadinya kegagalan pengobatan, kekambuhan, dan kematian pada penderita TB (Nugraha dkk., 2021).

2.2 Kajian Integrasi Topik dengan Al-Quran/Hadits

Dalam pembahasan kitab tauhid karya Dr. Shalih bin Fauzan bin Abdullah Al Fauzan, ibadah secara etimologis berarti merendahkan diri serta tunduk, sedangkan secara syar'i memiliki beberapa definisi namun bermakna sama, yaitu ketaatan total kepada Allah dengan melaksanakan perintah-Nya sebagaimana yang disampaikan melalui para rasul. Ibadah juga dimaknai sebagai bentuk ketundukan dan kerendahan diri yang disertai rasa cinta yang mendalam kepada Allah Subhanahu wa Ta'ala. Secara lebih lengkap, ibadah mencakup seluruh perbuatan dan ucapan baik yang tampak maupun tersembunyi yang dicintai dan diridhai oleh Allah. Ibadah memiliki berbagai bentuk, di antaranya ibadah hati seperti rasa takut (*khauf*), harapan (*raja'*), cinta (*mahabbah*), dan tawakkal. Ibadah lisan seperti dzikir, tasbih, dan membaca Al-Qur'an; serta ibadah anggota badan seperti shalat, zakat, puasa, haji, dan jihad. Ciri khas ibadah adalah pelaksanaannya semata-mata karena Allah, sesuai dengan syariat yang telah ditetapkan-Nya, disertai keikhlasan dan niat untuk mendekatkan diri kepada-Nya. Bahkan perbuatan yang bersifat duniawi seperti bekerja, makan, atau tidur dapat bernilai ibadah apabila diniatkan sebagai sarana untuk menaati dan mendekatkan diri kepada Allah. Dengan demikian, ibadah tidak hanya terbatas pada ritual keagamaan, tetapi mencakup seluruh aspek kehidupan

seorang mukmin yang dilandasi niat tulus dan ketaatan kepada Allah Subhanahu wa Ta'ala (Fauzan, 2011). Hal ini ditegaskan dalam firman Allah SWT pada QS. Al-An'am ayat 162:

قُلْ إِنَّ صَلَاتِي وَنُسُكِي وَمَحْيَايَ وَمَمَاتِي لِلَّهِ رَبِّ الْعَالَمِينَ

“Katakanlah (Nabi Muhammad), sesungguhnya salatku, ibadahku, hidupku, dan matiku hanyalah untuk Allah, Tuhan semesta alam.” (QS. Al-An'am, 6:162).

Prof. Quraish Shihab dalam *Tafsir Al-Misbah*, ayat ini mengandung pesan ketulusan dalam beramal sebagaimana ajaran Nabi Ibrahim AS, dan menjadi teladan bagi umat Islam untuk menyerahkan seluruh hidupnya kepada Allah semata tanpa menyekutukan-Nya. Penempatan shalat di awal ayat menunjukkan kedudukannya sebagai ibadah yang paling utama dan pondasi bagi amal lainnya (Shihab, 2002). Selanjutnya, dalam QS. Al-Kahfi ayat 110, Allah SWT berfirman:

قُلْ إِنَّمَا أَنَا بَشَرٌ مِّثْلُكُمْ يُوحَىٰ إِلَيَّ أَنَّمَا إِلَهُكُمُ إِلَهٌ وَاحِدٌ فَمَنْ كَانَ يَرْجُوا لِقَاءَ رَبِّهِ فَلْيَعْمَلْ عَمَلًا صَالِحًا وَلَا يُشْرِكْ بِعِبَادَةِ رَبِّهِ أَحَدًا

Artinya: “Katakanlah, sesungguhnya aku ini hanya seorang manusia seperti kalian, yang diwahyukan kepadaku bahwa sesungguhnya Tuhan kalian adalah Tuhan Yang Maha Esa. Barang siapa mengharapakan perjumpaan dengan Tuhannya, maka hendaklah ia mengerjakan amal saleh dan janganlah ia mempersekutukan seorang pun dalam beribadat kepada Tuhannya.” (QS. Al-Kahfi, 18:110).

Asbabun nuzul (sebab turunnya) Al-Kahfi ayat 110 adalah ketika seorang sahabat bertanya kepada Rasulullah SAW tentang amal saleh yang ia lakukan demi meraih ridha Allah, namun di saat yang sama ia juga ingin amal tersebut dilihat dan diakui oleh orang lain. Ayat ini kemudian diturunkan untuk menegaskan pentingnya keikhlasan dan menjauhi riya (pamer), karena Allah hanya menerima amal yang tulus karena-Nya. Dari ayat tersebut dapat dipahami bahwa setiap amal, khususnya ibadah, akan menjadi sia-sia apabila dicampuri dengan unsur syirik. Dengan demikian, keikhlasan menjadi wujud penegasan terhadap keyakinan akan keesaan

Allah, sebagaimana terkandung dalam syahadat “*Tiada Tuhan selain Allah*”. Makna utama dari syahadat ini menegaskan penolakan terhadap segala bentuk kesyirikan dalam hati seorang muslim sebagai syarat diterimanya suatu ibadah (Abdullah, 2007).

Keikhlasan dan ketulusan ini juga tercermin dalam ibadah shalat yang menjadi pusat pengendalian moral dan spiritual seorang muslim. Sebagaimana firman Allah dalam QS. Al-‘Ankabut ayat 45:

أَتْلُ مَا أُوحِيَ إِلَيْكَ مِنَ الْكِتَابِ وَأَقِمِ الصَّلَاةَ إِنَّ الصَّلَاةَ تَنْهَى عَنِ الْفَحْشَاءِ وَالْمُنْكَرِ وَلَذِكْرُ اللَّهِ أَكْبَرُ وَاللَّهُ يَعْلَمُ مَا تَصْنَعُونَ

Artinya: “*Bacalah (Nabi Muhammad) Kitab (Al-Qur’an) yang telah diwahyukan kepadamu dan tegakkanlah salat. Sesungguhnya salat itu mencegah dari (perbuatan) keji dan mungkar. Sungguh, mengingat Allah (salat) itu lebih besar (keutamaannya daripada ibadah yang lain). Allah mengetahui apa yang kamu kerjakan.*” (QS. Al-Ankabut, 29:45).

Sesuai tafsir *Tafsir Al-Misbah* surah Al-‘Ankabut ayat 45 ini, setelah memerintahkan membaca, mempelajari, dan melaksanakan ajaran-ajaran Al-Qur’an, maka Allah memerintahkan agar kaum muslimin mengerjakan salat wajib, yaitu salat lima waktu. Shalat hendaknya dilaksanakan dengan memenuhi seluruh rukun dan syaratnya serta dilakukan dengan penuh kekhayusan. Disarankan pula untuk menyempurnakannya dengan menjalankan berbagai sunah salat. Apabila salat dilakukan dengan sempurna, maka ia akan mampu mencegah pelakunya dari perbuatan keji dan mungkar. Pelaksanaan salat mencerminkan keyakinan yang tertanam dalam hati seorang hamba bahwa dirinya bergantung sepenuhnya kepada Allah. Karena itu, ia berusaha sekuat tenaga untuk menaati perintah-Nya dan menjauhi larangan-Nya, sebagaimana terkandung dalam doa surat Al-Fatihah yang dibaca setiap salat: “*Tunjukkanlah kepada kami jalan yang lurus, yaitu jalan*

orang-orang yang telah Engkau beri nikmat, bukan jalan mereka yang dimurkai, dan bukan pula jalan mereka yang sesat.” Doa tersebut senantiasa tertanam dalam ingatannya, sehingga menumbuhkan tekad untuk menjauhi segala bentuk perbuatan keji dan mungkar (Abdullah, 2007).

Rasulullah saw menjelaskan tentang keutamaan dan manfaat yang diperoleh bagi mereka yang melaksanakan salat, serta kerugian dan azab yang akan dialami oleh orang yang meninggalkannya, sebagaimana dijelaskan dalam sebuah hadis yang artinya:

Dari Nabi SAW, bahwasanya ia pada suatu hari menyebut tentang salat, maka ia berkata, “Barang siapa yang memelihara salat, ia akan memperoleh cahaya, petunjuk, dan keselamatan pada hari Kiamat, dan barang siapa yang tidak memeliharanya, ia tidak akan memperoleh cahaya, petunjuk, dan keselamatan. Dan ia pada hari Kiamat bersama Karun, Fir’aun, Haman, dan Ubai bin Khalaf. (Riwayat Ahmad dan ath-Thabrani dari ‘Abdullah bin ‘Umar).

Nabi saw juga menjelaskan tentang keadaan seseorang yang melaksanakan salat lima waktu dengan penuh kesungguhan, memenuhi seluruh rukun dan syaratnya, serta menunaikannya tepat pada waktunya. Menurut beliau, orang yang demikian seolah-olah membersihkan dirinya dari dosa lima kali dalam sehari, hingga tidak ada dosa yang tersisa. Rasulullah SAW bersabda yang artinya:

“Bagaimanakah pendapatmu, andaikata ada sebuah sungai dekat pintu rumah salah seorang dari kamu, ia mandi di sungai itu lima kali setiap hari. Adakah masih ada dakinya yang tinggal barang sedikit pun?” Para sahabat menjawab, “Tidak ada daki yang tertinggal barang sedikit pun.” Rasulullah bersabda, “Maka demikianlah perumpamaan salat yang lima waktu, dengan salat itu Allah akan menghapus semua kesalahannya.” (Riwayat at-Tirmidzi dari Abu Hurairah).

Demikianlah perumpamaan yang diberikan Rasulullah saw tentang keadaan orang yang mengerjakan salat lima waktu dengan sungguh-sungguh hanya karena Allah. Dari ayat dan hadis Rasulullah yang telah disebutkan sebelumnya, dapat

disimpulkan bahwa terdapat tiga tujuan utama bagi seseorang yang melaksanakan salat, yaitu:

1. Munculnya keikhlasan
2. Tumbuhnya sikap takwa kepada Allah
3. Senantiasa mengingat Allah

Salat seharusnya menumbuhkan keikhlasan dalam diri orang yang melakukannya, karena ia dilaksanakan semata-mata untuk Allah sebagai bentuk pemurnian ketaatan hanya kepada-Nya. Wujud dari keikhlasan ini tercermin dalam dorongan hati untuk melakukan segala sesuatu yang diridai oleh Allah. Rasa takwa kepada Allah berarti munculnya keinginan untuk melaksanakan seluruh perintah-Nya dan menjauhi segala larangan-Nya. Melalui salat pula seseorang akan senantiasa mengingat Allah, sebab di dalamnya terkandung bacaan tasbih, tahmid, dan takbir yang menumbuhkan kesadaran akan keagungan serta kebesaran-Nya (Departemen Agama RI, 2009).

2.3 Kajian Topik dengan Teori Pendukung

Pada penelitian ini, analisis *survival* dilakukan menggunakan dua pendekatan utama, yaitu metode *Kaplan-Meier* dan *Cox Proportional Hazard*. Kedua metode ini digunakan secara berurutan untuk menggambarkan pola waktu menuju kesembuhan pasien dan mengidentifikasi faktor-faktor yang berpengaruh terhadap waktu kesembuhan. Selain itu, analisis *survival* juga dapat mengakomodasi adanya data tersensor, yaitu pasien yang belum sembuh hingga akhir masa observasi atau hilang dalam pemantauan (*lost to follow up*).

Langkah pertama dalam penelitian ini adalah melakukan *preprocessing* data. Tahapan ini mencakup beberapa proses penting untuk memastikan kualitas dan kelayakan data sebelum dilakukan analisis. Pertama, dilakukan pengkodean terhadap variabel kategorik dengan mengubah data yang semula berbentuk teks menjadi kode numerik agar dapat diolah menggunakan metode analisis statistik. Kedua, dilakukan perhitungan waktu hingga terjadinya *event* (*survival time*) dalam satuan hari, karena pada data mentah hanya tersedia informasi berupa tanggal mulai pengobatan dan tanggal selesai pengobatan. Oleh karena itu, variabel waktu dihitung berdasarkan selisih antara kedua tanggal tersebut. Ketiga, dilakukan identifikasi terhadap nilai hilang (*missing value*), yang meliputi pendeteksian serta analisis pola dan mekanisme *missing data*. Berdasarkan hasil analisis tersebut, penanganan data hilang dilakukan menggunakan metode imputasi, yaitu *Multiple Imputation by Chained Equations* (MICE), sehingga data tetap dapat dimanfaatkan secara optimal tanpa mengurangi jumlah sampel secara signifikan. Proses MICE menghasilkan beberapa dataset lengkap yang kemudian digunakan pada tahap analisis selanjutnya. Setelah seluruh proses *pre-processing* selesai dilakukan, data dinyatakan siap untuk digunakan dalam analisis *survival*.

Langkah kedua dilakukan analisis deskriptif untuk menggambarkan karakteristik pasien berdasarkan masing-masing variabel independen. Analisis ini bertujuan memberikan gambaran umum mengenai profil pasien. Distribusi frekuensi dan persentase digunakan untuk menyajikan data, yang kemudian divisualisasikan dalam bentuk diagram agar lebih mudah dipahami.

Langkah ketiga adalah melakukan estimasi fungsi *survival* menggunakan metode *Kaplan-Meier*. Kurva *Kaplan-Meier* digunakan untuk menggambarkan

peluang pasien belum mengalami kejadian kesembuhan hingga waktu tertentu. Metode ini menghitung probabilitas belum sembuh secara bertahap pada setiap waktu kejadian, sehingga menghasilkan kurva empiris yang merepresentasikan fungsi *survival*. Dalam konteks pengobatan tuberkulosis, setiap kali seorang pasien dinyatakan sembuh, probabilitas belum sembuh akan diperbarui. Penurunan kurva menunjukkan semakin banyak pasien yang mengalami kejadian, sedangkan bagian datar menunjukkan tidak adanya kejadian baru dalam interval waktu tertentu. Langkah berikutnya adalah melakukan uji *log-rank* untuk mengetahui apakah terdapat perbedaan yang signifikan antara kurva *survival* dari dua atau lebih kategori berdasarkan variabel independen.

Langkah keempat adalah analisis regresi *Cox Proportional Hazard*, yang bertujuan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap waktu kesembuhan pasien. Analisis ini dilakukan pada setiap dataset hasil imputasi MICE. Model Cox bersifat semiparametrik, karena tidak memerlukan asumsi bentuk tertentu dari fungsi *hazard* dasar, namun mengasumsikan bahwa rasio *hazard* antar individu bersifat konstan terhadap waktu. Secara matematis, model Cox dinyatakan sebagai:

$$h(t, \mathbf{X}) = h_0(t) \exp \left(\sum_{p=1}^q \beta_p X_p \right)$$

di mana $h(t, \mathbf{X})$ adalah fungsi *hazard* pada waktu t , $h_0(t)$ merupakan *baseline hazard*, dan β_p menunjukkan koefisien regresi sampai variabel ke- q yang mengukur pengaruh faktor tersebut terhadap kesembuhan pasien.

Setelah model Cox dibangun, dilakukan estimasi parameter menggunakan metode *Maximum Partial Likelihood*. Metode ini merupakan pengembangan dari *Maximum Likelihood Estimation* (MLE) yang hanya memanfaatkan informasi

urutan kejadian (*event*) tanpa memerlukan pendefinisian bentuk fungsi *hazard* dasar. Dengan demikian, estimasi parameter dapat dilakukan tanpa harus menentukan bentuk distribusi tertentu dari waktu *survival*. Proses estimasi dilakukan secara iteratif menggunakan metode numerik Newton-Raphson hingga mencapai kondisi konvergen.

Selanjutnya, dilakukan pengujian asumsi *Proportional Hazard* (PH) untuk memastikan bahwa rasio *hazard* antar kategori bersifat konstan sepanjang waktu. Pemeriksaan asumsi ini dapat dilakukan secara visual maupun formal. Secara visual, asumsi PH dinilai melalui grafik *Kaplan-Meier* dengan melihat pola kurva *survival* antar kategori. Jika kurva *survival* antar kategori menunjukkan pola yang relatif sejajar dan tidak saling berpotongan secara signifikan, maka secara visual terdapat indikasi bahwa asumsi *proportional hazard* terpenuhi, meskipun hal ini tidak dapat dijadikan sebagai dasar keputusan utama. Secara formal, pengujian dilakukan menggunakan residual Schoenfeld. Apabila tidak terdapat hubungan yang signifikan antara residual dengan waktu, maka tidak terdapat bukti pelanggaran terhadap asumsi PH.

Karena analisis dilakukan pada beberapa dataset hasil imputasi, maka hasil estimasi parameter dari model Cox selanjutnya digabungkan (*pooling*) menggunakan aturan Rubin (*Rubin's Rules*). Variansi dalam imputasi (*within-imputation variance*) merupakan rata-rata variansi dari estimasi parameter pada masing-masing dataset hasil imputasi, yang mencerminkan ketidakpastian akibat variasi sampel dalam model. Sementara itu, variansi antar imputasi (*between-imputation variance*) merupakan variasi antar estimasi parameter dari beberapa dataset imputasi, yang mencerminkan ketidakpastian akibat adanya data yang

hilang. Kedua komponen ini kemudian digabungkan menggunakan *Rubin's Rules* untuk memperoleh estimasi variansi total yang lebih akurat. Selanjutnya, untuk menguji signifikansi model, digunakan dua jenis uji statistik, yaitu uji serentak (*Likelihood Ratio Test*) yang dilakukan pada masing-masing dataset imputasi untuk menilai pengaruh seluruh variabel independen secara bersama-sama, serta uji parsial (*Wald Test*) yang diperoleh dari hasil *pooling* untuk menguji pengaruh masing-masing variabel secara individual terhadap waktu kesembuhan pasien. Tahap akhir adalah interpretasi hasil regresi Cox berdasarkan nilai *Hazard Ratio* (HR), yang menunjukkan besar risiko relatif suatu kejadian terjadi pada kategori tertentu dibandingkan dengan kategori referensi.

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian ini menggunakan pendekatan deskriptif kuantitatif yang didukung oleh studi literatur. Studi literatur dilakukan dengan menghimpun berbagai sumber pustaka dan teori-teori yang relevan sebagai dasar dalam mengkaji serta memecahkan permasalahan penelitian. Sementara itu, pendekatan deskriptif kuantitatif digunakan untuk menganalisis data numerik secara statistik guna menggambarkan suatu kondisi atau fenomena, serta menarik kesimpulan berdasarkan pola atau hubungan yang muncul dalam data tersebut.

3.2 Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data kuantitatif sekunder yang diperoleh dari Dinas Kesehatan (Dinkes) Kota Probolinggo pada rentang waktu tahun 2021 hingga 2023. Data penelitian mencakup 1.365 pasien tuberkulosis yang memenuhi kriteria penelitian, data dapat dilihat pada Lampiran 1. Data yang digunakan telah melalui proses seleksi dan pembersihan (*data cleaning*) untuk memastikan kelengkapan, konsistensi, dan keakuratan informasi yang dianalisis. Data tersebut mencakup rekam medis pasien tuberkulosis yang menjalani pengobatan menggunakan Obat Anti Tuberkulosis (OAT), mulai dari awal pengobatan hingga status akhir pengobatan (sembuh, putus obat, meninggal, atau masih dalam pengobatan). Selain itu, data juga memuat variabel-variabel pendukung yang berpotensi memengaruhi kesembuhan TB.

3.3 Variabel Penelitian dan Struktur Data

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Deskripsi
T (Waktu hingga terjadinya <i>event</i> atau <i>survival time</i>)	Variabel dependen	Lama waktu (dalam hari) dari awal pengobatan tuberkulosis hingga pasien dinyatakan sembuh atau hingga akhir observasi (<i>censored</i>).
δ Status kejadian	Variabel dependen (biner)	1 = Sembuh (<i>event</i> terjadi) 0 = Belum sembuh / putus obat / hilang <i>follow-up</i> / meninggal (<i>censored</i>).
X_1 Usia	Variabel independen (biner)	Usia pasien saat memulai pengobatan 1 = usia produktif (15 – 64 tahun) 0 = usia tidak produktif (< 15 tahun dan > 64 tahun)
X_2 Jenis kelamin	Variabel independen (biner)	1 = Laki-laki 0 = Perempuan
X_3 Status pengobatan	Variabel independen (biner)	1 = Sesuai standar 0 = Tidak sesuai standar
X_4 Riwayat pengobatan	Variabel independen (biner)	1 = Kasus baru 0 = Pernah diobati sebelumnya
X_5 Jenis fasilitas layanan kesehatan	Variabel independen (kategori)	1 = Puskesmas 2 = Rumah sakit 3 = Fasilitas kesehatan lainnya
X_6 Tipe diagnosis TB	Variabel independen (biner)	1 = Terkonfirmasi bakteriologis 0 = Terdiagnosis klinis
X_7 Hasil Tes HIV	Variabel independen (biner)	1 = Negatif 0 = Positif
X_8 Hasil Tes DM	Variabel independen (biner)	1 = Negatif 0 = Positif

Dalam penelitian ini, terdapat sejumlah variabel yang akan dianalisis, di mana masing-masing variabel memiliki klasifikasi atau deskripsinya sendiri yang sudah dijelaskan pada Tabel 3.1

Variabel kategorik pada penelitian ini dikonversi menjadi variabel faktor (*factor*) dalam perangkat lunak R, sehingga pada pemodelan *cox proportional hazard* secara otomatis dibentuk variabel dummy berdasarkan kategori referensi. Kategori referensi ditentukan oleh level pertama (*first level*) dari masing-masing variabel faktor. Untuk variabel dengan dua kategori, yaitu $X_1, X_2, X_3, X_4, X_6, X_7$, dan X_8 , dibentuk satu variabel dummy dengan kategori 0 sebagai referensi dan kategori 1 sebagai pembanding. Sementara itu, X_5 yang memiliki tiga kategori (puskesmas, rumah sakit, dan fasilitas kesehatan lainnya) direpresentasikan menjadi dua variabel dummy ($3 - 1$ dummy), dengan puskesmas sebagai kategori referensi. Tabel 3.2 menunjukkan ilustrasi pembentukan variabel dummy untuk X_5 .

Tabel 3. 2 Pembentukan Variabel Dummy untuk X_5

Kategori X_5	$D_{5,1}$ (Rumah Sakit)	$D_{5,2}$ (Fasyankes Lain)
Puskesmas	0	0
Rumah Sakit	1	0
Fasyankes lain	0	1

Berdasarkan Tabel 3.2, kategori puskesmas berperan sebagai kategori referensi yang direpresentasikan dengan nilai $D_{5,1}$ dan $D_{5,2} = 0$. Rumah sakit direpresentasikan oleh $D_{5,1} = 1$ dan $D_{5,2} = 0$, sedangkan fasilitas kesehatan lainnya direpresentasikan oleh $D_{5,2} = 1$ dan $D_{5,1} = 0$. Dengan demikian, interpretasi model *cox proportional hazard* dilakukan dengan membandingkan masing-masing kategori terhadap puskesmas sebagai referensi. Selanjutnya, struktur data penelitian yang telah dikategorikan dan siap digunakan dalam proses analisis ditampilkan pada Tabel 3.3.

Tabel 3.3 Struktur Data Penelitian

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Status	Tanggal Mulai	Tanggal Selesai
1	1	1	0	1	0	1	1	0	02/01/2021	20/06/2021
1	1	1	1	1	0	1	0	0	02/01/2021	17/06/2021
1	1	1	1	1	0	1	1	0	17/06/2021	30/11/2021
1	1	1	1	1	0	1	1	0	04/01/2021	19/06/2021
1	1	1	1	1	1	1	1	1	28/04/2021	12/10/2021
1	1	1	1	1	0	1	1	0	04/01/2021	19/06/2021
1	1	1	1	1	0	1	1	0	29/04/2021	12/10/2021
1	0	1	1	1	1	1	0	0	03/03/2021	19/05/2021
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

3.4 Lokasi Penelitian

Penelitian ini dilakukan di Kota Probolinggo dengan fokus pada data pasien tuberkulosis yang terdaftar di fasilitas pelayanan kesehatan yang berada di bawah koordinasi dinas kesehatan Kota Probolinggo. Pemilihan lokasi ini didasarkan pada tingginya angka kasus tuberkulosis di Kota Probolinggo dalam beberapa tahun terakhir, sehingga penelitian diharapkan dapat memberikan kontribusi nyata dalam mendukung upaya pencegahan dan penanggulangan tuberkulosis di daerah tersebut.

3.5 Teknik Analisis Data

Berikut langkah-langkah analisis yang akan dilakukan pada penelitian ini:

1. Melakukan *preprocessing* data yang meliputi:
 - a. Mengatur kode untuk variabel kategori seperti pada deskripsi Tabel 3.1.
 - b. Melakukan perhitungan *survival time* pada Tabel 3.3 dalam satuan hari, yaitu dengan menghitung selisih tanggal selesai dan tanggal mulai pengobatan.
 - c. Data diperiksa untuk mengidentifikasi adanya *missing value*. Variabel-variabel yang mengandung *missing value* diuji hubungannya dengan

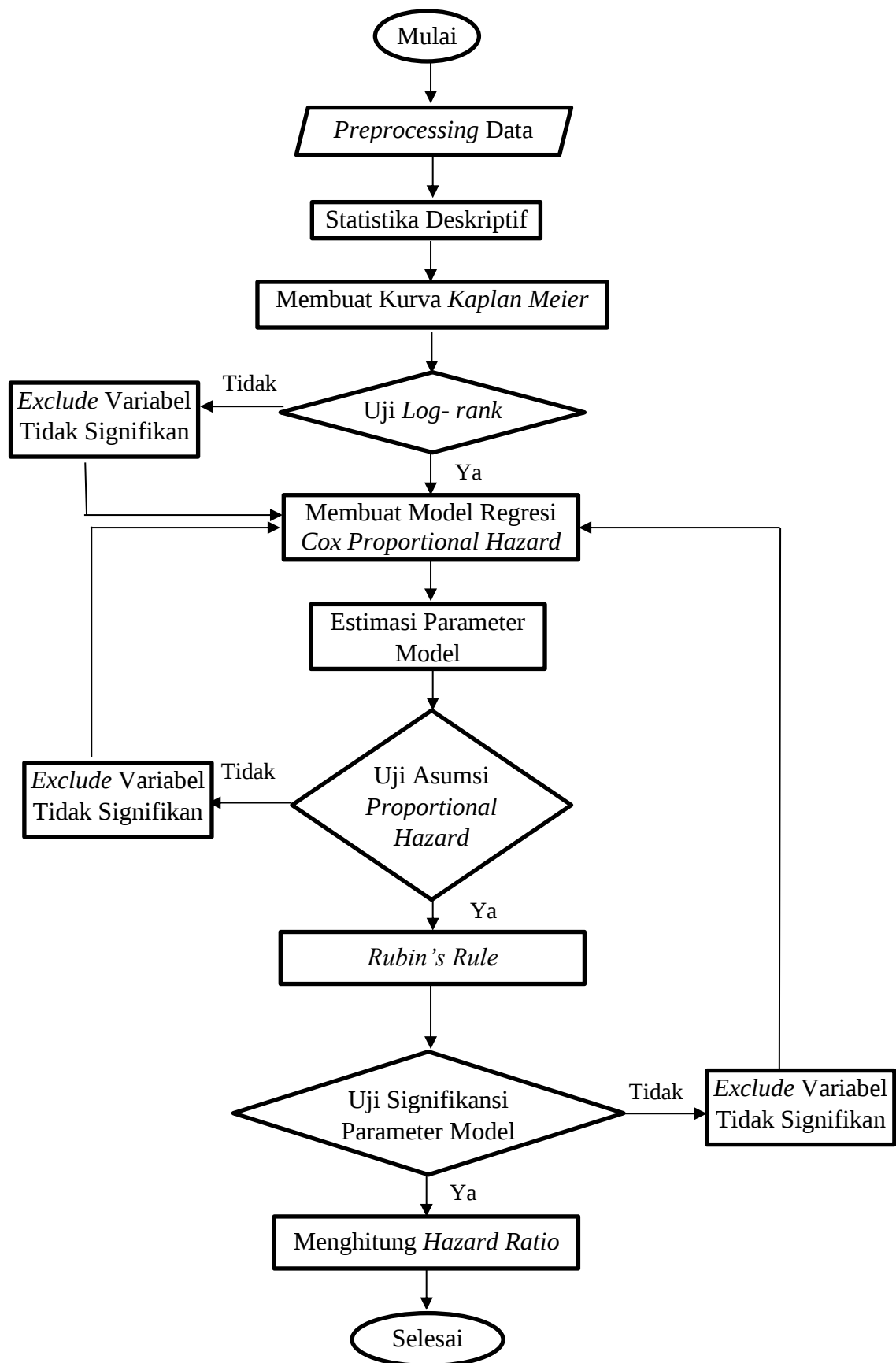
variabel lain menggunakan uji *Chi-Square* atau *Fisher Exact*. Berdasarkan hasil pengujian tersebut, ditentukan mekanisme *missing data*, apakah termasuk *Missing Completely at Random* (MCAR), *Missing at Random* (MAR), atau *Missing Not at Random* (MNAR). Setelah mekanisme *missing data* teridentifikasi, ditentukan metode penanganan yang sesuai. Dalam penelitian ini, penanganan *missing value* dilakukan menggunakan metode *Multiple Imputation by Chained Equations* (MICE), karena metode ini mampu memanfaatkan hubungan antar variabel dan menghasilkan estimasi yang lebih akurat pada data dengan mekanisme MAR.

2. Melakukan analisis deskriptif untuk menggambarkan karakteristik pasien berdasarkan masing-masing variabel independen menggunakan distribusi frekuensi dan persentase yang kemudian disajikan dalam bentuk *pie chart*.
3. Analisis *Kaplan-Meier*, yang meliputi:
 - a. Membuat kurva *Kaplan-Meier* dari masing-masing kategori dari variabel independen untuk memvisualisasikan pola waktu menuju kesembuhan, sebagaimana dirumuskan pada Persamaan (2.17).
 - b. Melakukan uji *log-rank* untuk mengetahui apakah terdapat perbedaan yang signifikan antara kurva *survival* dari dua atau lebih kategori berdasarkan variabel independen, sebagaimana dirumuskan pada Persamaan (2.18).
4. Analisis regresi *Cox Proportional Hazard* meliputi:
 - a. Membangun model regresi *Cox Proportional Hazard*, seperti pada Persamaan (2.19).
 - b. Melakukan estimasi parameter model pada masing-masing dataset hasil imputasi menggunakan metode *Maximum Partial Likelihood Estimation*

yang diselesaikan secara numerik menggunakan metode Newton-Raphson seperti pada Persamaan (2.26) hingga mencapai konvergensi seperti pada Pertaksamaan (2.28).

- c. Melakukan uji asumsi *Cox Proportional Hazard*:
 - i. Secara visual menggunakan grafik *Kaplan-Meier*
 - ii. Secara formal menggunakan uji residual Schoenfeld dengan meregresikan residual Schoenfeld terhadap waktu *survival* dan menguji signifikansinya menggunakan statistik uji *Chi-Square* seperti pada Persamaan (2.29).
- d. Menggabungkan hasil estimasi parameter dari seluruh dataset imputasi menggunakan *Rubin's Rule* melalui proses *pooling* untuk memperoleh estimasi parameter akhir model yang lebih akurat.
- e. Melakukan uji signifikansi parameter model:
 - i. Uji serentak menggunakan *Likelihood Ratio Test* sebagaimana pada Persamaan (2.30) untuk melihat pengaruh variabel secara simultan.
 - ii. Uji parsial menggunakan *Wald Test* sebagaimana pada Persamaan (2.31) untuk melihat pengaruh variabel secara individual.
- f. Menghitung *hazard ratio* dari masing-masing variabel independen seperti pada Persamaan (2.33) dan menginterpretasikannya untuk mengetahui seberapa besar laju kesembuhan pasien tuberkulosis berdasarkan masing-masing faktor.

Dari langkah teknik analisis data di atas, dapat dibuat *flowchart* pada Gambar 3.1.



Gambar 3.1 Flowchart Teknik Analisis Data

BAB IV

HASIL DAN PEMBAHASAN

4.1 *Preprocessing Data*

Tahap awal dalam penelitian ini adalah melakukan *preprocessing data* untuk memastikan kualitas data sebelum dilakukan analisis lebih lanjut. Tahapan ini meliputi identifikasi *missing value*, pengolahan tipe data, serta pengujian pola *missing data*.

4.1.1 Identifikasi *Missing Value*

Data yang digunakan dalam penelitian ini terdiri dari 1365 observasi dengan 10 variabel, yaitu delapan variabel independen (X_1 hingga X_8), satu variabel dependen status kejadian (Status), dan waktu hingga terjadinya *event* (T). Berdasarkan hasil eksplorasi data, ditemukan adanya *missing value* pada beberapa variabel. Jumlah dan persentase *missing value* pada masing-masing variabel disajikan pada Tabel 4.1.

Tabel 4.1 Jumlah *Missing Value* pada Setiap Variabel

Variabel	Jumlah <i>Missing</i>	Persentase (%)
X_1	0	0.00
X_2	0	0.00
X_3	0	0.00
X_4	62	4.54
X_5	0	0.00
X_6	0	0.00
X_7	170	12.45
X_8	141	10.33
Status	0	0.00
T	0	0.00

Berdasarkan Tabel 4.1, dapat diketahui bahwa variabel yang memiliki *missing value* adalah X_4 , X_7 , dan X_8 , dengan persentase terbesar terdapat pada variabel X_7 .

Selanjutnya, seluruh variabel kategorik diubah ke dalam bentuk *factor* agar dapat diperlakukan sesuai dengan sifatnya sebagai data kategorik dalam analisis statistik, khususnya untuk uji χ^2 , *Fisher Exact*, serta proses imputasi menggunakan metode MICE.

4.1.2 Identifikasi Mekanisme *Missing Data*

Untuk menentukan mekanisme *missing data*, dilakukan pengujian hubungan antara indikator *missing* pada masing-masing variabel dengan variabel lainnya menggunakan uji χ^2 atau *Fisher Exact*. Uji *Fisher Exact* digunakan apabila asumsi χ^2 tidak terpenuhi, yaitu ketika lebih dari 20% sel memiliki frekuensi harapan kurang dari 5 atau terdapat sel dengan frekuensi harapan kurang dari 1. Hipotesis yang digunakan dalam uji χ^2 maupun *Fisher Exact* yaitu:

H_0 : Tidak terdapat hubungan antara *missing data* suatu variabel dengan variabel lain

H_1 : Terdapat hubungan antara *missing data* suatu variabel dengan variabel lain

Kriteria pengujian ditentukan berdasarkan nilai *p – value* sebagai berikut:

1. Jika $p - value < \alpha = 0.05$, maka H_0 ditolak
2. Jika $p - value \geq \alpha = 0.05$, maka H_0 gagal ditolak

Berdasarkan hasil pengujian pada Tabel 4.2, diketahui bahwa:

1. Variabel X_4 memiliki minimal satu hubungan signifikan dengan variabel lain, yaitu X_3 , sehingga termasuk dalam mekanisme *Missing At Random* (MAR).
2. Variabel X_7 menunjukkan beberapa hubungan signifikan dengan variabel lain, yaitu X_1, X_3, X_5, X_6 , dan status, sehingga termasuk dalam mekanisme MAR.
3. Variabel X_8 juga menunjukkan hubungan signifikan dengan beberapa variabel lain, yaitu X_1, X_3, X_6 , dan status sehingga termasuk dalam mekanisme MAR.

Dengan demikian, seluruh variabel yang memiliki *missing value*, yaitu X_4, X_7 , dan X_8 dikategorikan dalam mekanisme *Missing At Random* (MAR).

Tabel 4.2 Hasil Uji Mekanisme *Missing Data*

Variabel <i>Missing</i>	Variabel Pembanding	Metode	<i>p-value</i>	Keputusan
X_4	X_1	χ^2	0.5002	Tidak ada hubungan
X_4	X_2	χ^2	0.4872	Tidak ada hubungan
X_4	X_3	Fisher	0.0484	Ada hubungan
X_4	X_5	Fisher	1.0000	Tidak ada hubungan
X_4	X_6	χ^2	0.1981	Tidak ada hubungan
X_4	Status	χ^2	0.7101	Tidak ada hubungan
X_7	X_1	χ^2	0.0000	Ada hubungan
X_7	X_2	χ^2	0.3009	Tidak ada hubungan
X_7	X_3	Fisher	0.0029	Ada hubungan
X_7	X_5	χ^2	0.0002	Ada hubungan
X_7	X_6	χ^2	0.0007	Ada hubungan
X_7	Status	χ^2	0.0000	Ada hubungan
X_8	X_1	χ^2	0.0024	Ada hubungan
X_8	X_2	χ^2	0.0912	Tidak ada hubungan
X_8	X_3	Fisher	0.0339	Ada hubungan
X_8	X_5	χ^2	0.4675	Tidak ada hubungan
X_8	X_6	χ^2	0.0002	Ada hubungan
X_8	Status	χ^2	0.0000	Ada hubungan

4.1.3 Penanganan *Missing Value* dengan Metode *Multiple Imputation*

Penanganan *missing value* dalam penelitian ini dilakukan menggunakan metode *Multiple Imputation by Chained Equations* (MICE) dengan bantuan *package mice*. Metode ini dipilih karena mampu menangani data hilang dengan mempertimbangkan hubungan antar variabel serta sesuai untuk mekanisme *Missing At Random* (MAR) yang telah diidentifikasi pada tahap sebelumnya.

Langkah awal dalam proses imputasi adalah mengidentifikasi pola *missing data*. Hasil identifikasi disajikan dalam Tabel 4.3.

Tabel 4.3 Pola *Missing* Data

Jumlah Observasi	Pola <i>Missing</i>	Jumlah <i>Missing Value</i>
1075	Tidak ada <i>missing</i>	0
93	<i>Missing</i> pada X_7 saja	93
70	<i>Missing</i> pada X_8 saja	70
65	<i>Missing</i> pada X_7 dan X_8	130
48	<i>Missing</i> pada X_4 saja	48
14	Kombinasi <i>missing</i> lainnya	32
1365	Total Observasi	373

Berdasarkan Tabel 4.3, sebagian besar observasi (1075 data) tidak memiliki *missing value*. *Missing* data terjadi pada variabel X_4 , X_7 , dan X_8 , baik secara tunggal maupun dalam kombinasi. Jumlah *missing value* total sebesar 373 tidak hanya berasal dari observasi dengan satu *missing*, tetapi juga dari observasi yang memiliki lebih dari satu *missing value*. Sebagai contoh, terdapat 65 observasi yang mengalami *missing* pada dua variabel sekaligus, yaitu X_7 dan X_8 , sehingga berkontribusi sebesar 130 *missing value*. Selain itu, terdapat 14 observasi dengan pola *missing* yang merupakan kombinasi dari dua atau tiga variabel, yang secara keseluruhan menyumbang 32 *missing value*. Hal ini menunjukkan bahwa pola *missing data* bersifat tidak teratur (*non-monotone*), sehingga metode imputasi sederhana tidak memadai dan diperlukan metode yang lebih kompleks seperti MICE.

Metode MICE bekerja dengan pendekatan iteratif yang dikenal sebagai *chained equations*. Pada tahap awal, nilai yang hilang diisi dengan nilai awal sementara. Selanjutnya, setiap variabel yang memiliki *missing value* dimodelkan sebagai fungsi dari variabel lain yang tersedia. Dalam penelitian ini, variabel X_4 , X_7 , dan X_8 diimputasi menggunakan metode regresi logistik karena ketiga variabel tersebut bersifat kategorik biner. Proses imputasi dilakukan secara bergantian untuk masing-masing variabel dalam beberapa iterasi hingga mencapai kondisi stabil

(konvergen). Model regresi logistik yang digunakan dalam proses imputasi dinyatakan dalam persamaan (2.3), dimana $P(Y = 1)$ merupakan probabilitas suatu observasi berada pada kategori 1 dibandingkan dengan kategori referensi, yaitu kategori 0, dan nilai logit merupakan transformasi log odds dari probabilitas tersebut. Berdasarkan Lampiran 2 sampai Lampiran 4, diperoleh variabel X_4, X_7 , dan X_8 untuk model logistik dataset imputasi 1 yaitu:

Untuk variabel X_4 , model logistik dibentuk dengan menjadikan X_4 sebagai variabel respon dan variabel $X_1, X_2, X_3, X_5, X_6, X_7, X_8$ sebagai prediktor:

$$\begin{aligned} \text{logit}(P(X_4 = 1)) = & 19.42 - 1.31X_1 - 0.12X_2 - 14.27X_3 - 0.15D_{5,1} + 14.36D_{5,2} \\ & - 0.11X_6 - 0.76X_7 - 0.08X_8 \end{aligned} \quad (4.1)$$

Untuk variabel X_7 , model logistik dibentuk dengan menjadikan X_7 sebagai variabel respon dan variabel $X_1, X_2, X_3, X_4, X_5, X_6, X_8$ sebagai prediktor:

$$\begin{aligned} \text{logit}(P(X_7 = 1)) = & 4.78 - 0.76X_1 - 0.66X_2 + 0.65X_3 - 0.55X_4 - 1.54D_{5,1} \\ & - 3.00D_{5,2} + 1.25X_6 - 1.15X_8 \end{aligned} \quad (4.2)$$

Untuk variabel X_8 , model logistik dibentuk dengan menjadikan X_8 sebagai variabel respon dan variabel $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ sebagai prediktor:

$$\begin{aligned} \text{logit}(P(X_8 = 1)) = & 2.66 - 0.33X_1 + 0.23X_2 + 0.07X_3 - 0.07X_4 + 0.21D_{5,1} \\ & + 0.65D_{5,2} - 0.31X_6 - 1.13X_7 \end{aligned} \quad (4.3)$$

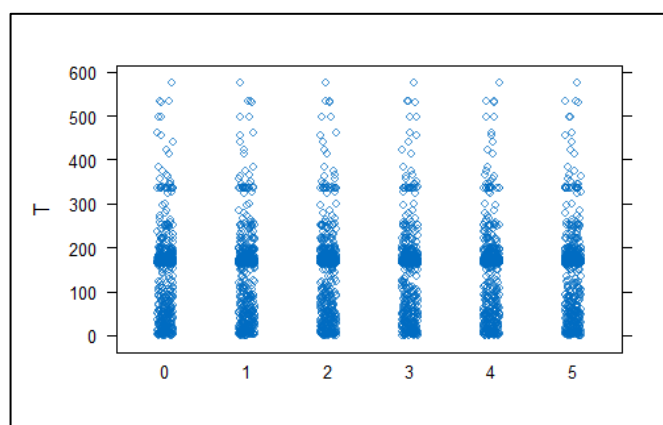
Dalam persamaan (4.1) sampai (4.3), variabel $D_{5,1}$ dan $D_{5,2}$ merupakan representasi dummy dari variabel jenis fasyankes (X_5), dimana $D_{5,1} = 1$ untuk rumah sakit dan 0 selainnya, serta $D_{5,2} = 1$ untuk fasyankes lainnya dan 0 selainnya. Kategori puskesmas digunakan sebagai kategori referensi yang direpresentasikan oleh kondisi $D_{5,1} = 0$ dan $D_{5,2} = 0$. Selanjutnya, untuk model logistik untuk dataset imputasi 2 sampai imputasi 5 disajikan pada Lampiran 2 sampai Lampiran 4.

Proses imputasi dilakukan secara iteratif sebanyak 20 iterasi ($maxit = 20$) hingga parameter model menunjukkan kestabilan (konvergen). Pemilihan jumlah

iterasi ini mengacu pada literatur yang menyarankan bahwa 20–30 iterasi umumnya cukup untuk mencapai konvergensi dalam metode MICE dan akan dievaluasi pada Gambar 4.2. Selain itu, jumlah dataset imputasi yang digunakan adalah sebanyak 5 ($m = 5$), yang mengacu pada rekomendasi umum bahwa 5 imputasi sudah memadai untuk menghasilkan estimasi parameter yang stabil dengan efisiensi yang baik. Sehingga, hasil dari proses MICE adalah lima dataset lengkap tanpa *missing value*.

4.1.4 Evaluasi Hasil Imputasi

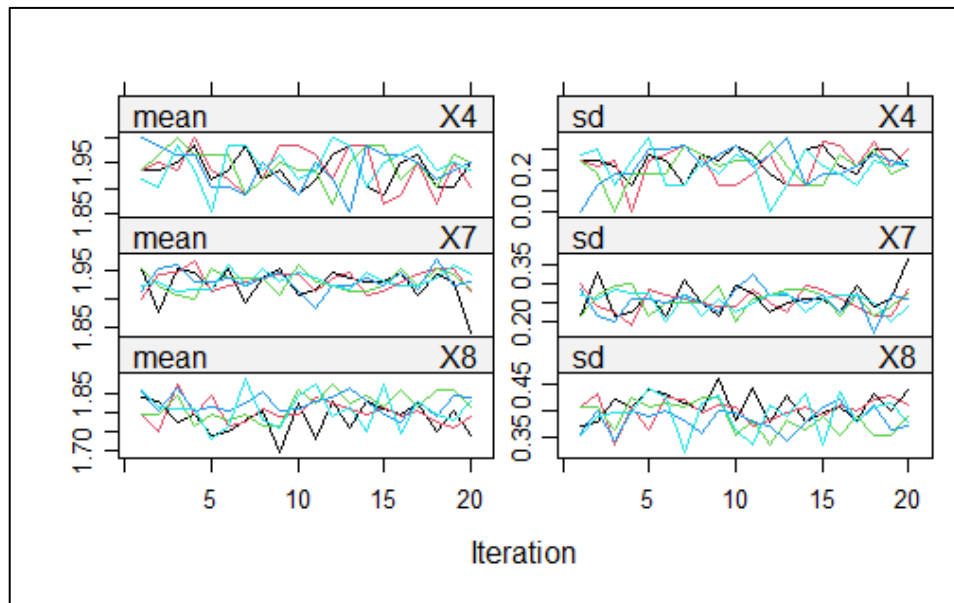
Evaluasi awal dilakukan menggunakan stripplot yang ditampilkan pada Gambar 4.1. Plot ini digunakan untuk membandingkan distribusi data asli dan data hasil imputasi pada variabel yang mengalami *missing value*, yaitu X_4 , X_7 , dan X_8 . Pada grafik tersebut, sumbu X menunjukkan data asli (0) dan lima dataset hasil imputasi (1 - 5), sedangkan sumbu Y menunjukkan nilai waktu menuju kesembuhan pasien (T) dalam satuan hari. Stripplot menampilkan sebaran nilai observasi pada masing-masing dataset sehingga dapat diamati apakah hasil imputasi menghasilkan distribusi yang masih konsisten dengan data asli.



Gambar 4.1 Stripplot Sebaran Nilai pada Lima Dataset Hasil Imputasi

Berdasarkan Gambar 4.1, terlihat bahwa sebaran nilai pada lima dataset hasil imputasi berada dalam rentang yang relatif sama dengan data asli. Selain itu, pola

penyebaran data pada setiap dataset imputasi tampak serupa dan tidak menunjukkan adanya nilai yang sangat menyimpang dari distribusi data observasi. Kondisi ini mengindikasikan bahwa proses imputasi menghasilkan nilai yang realistis dan mampu mempertahankan karakteristik distribusi data tanpa menimbulkan distorsi yang berarti.



Gambar 4.2 Trace Plot Konvergensi Mean dan Standar Deviasi Hasil Imputasi MICE pada Variabel X_4 , X_7 , dan X_8

Selanjutnya, evaluasi konvergensi dilakukan menggunakan *trace plot* yang ditampilkan pada Gambar 4.2. *Trace plot* menunjukkan perkembangan nilai rata-rata (*mean*) dan standar deviasi (*standard deviation*) dari variabel yang diimputasi pada setiap iterasi untuk masing-masing dataset imputasi. Pada grafik tersebut, sumbu X merepresentasikan jumlah iterasi dalam proses imputasi MICE, sedangkan sumbu Y menunjukkan nilai statistik ringkasan (*mean* dan *standard deviation*) dari data hasil imputasi. Setiap garis berwarna merepresentasikan satu dari lima dataset imputasi ($m = 5$), yaitu imputasi ke-1 hingga imputasi ke-5. Perbedaan warna digunakan untuk membedakan masing-masing dataset imputasi sehingga variasi antar imputasi dapat diamati selama proses iterasi.

Berdasarkan Gambar 4.2, seluruh garis yang merepresentasikan lima dataset imputasi menunjukkan pola yang relatif stabil sepanjang iterasi. Tidak terlihat adanya tren yang sistematis, baik meningkat maupun menurun, pada nilai *mean* maupun *standard deviation*. Sebagian besar garis berada dalam rentang yang relatif berdekatan dan saling tumpang tindih. Pada variabel X_7 terlihat adanya satu garis yang mengalami penurunan nilai *mean* dan peningkatan nilai *standard deviation* pada iterasi terakhir dibandingkan garis lainnya. Namun, pola tersebut hanya muncul pada satu titik iterasi dan tidak terjadi secara konsisten sepanjang proses imputasi. Selain itu, pada iterasi-iterasi sebelumnya garis tersebut masih berada dalam rentang yang relatif sama dengan dataset imputasi lainnya. Kondisi ini menunjukkan bahwa perbedaan tersebut merupakan fluktuasi acak yang dapat terjadi akibat proses simulasi pada metode MICE, bukan indikasi ketidakstabilan atau kegagalan konvergensi. Dengan demikian, proses imputasi dapat dianggap telah mencapai kondisi konvergen dan hasil imputasi yang diperoleh layak digunakan untuk analisis selanjutnya.

Tabel 4.4 Perbandingan Proporsi Kategori Sebelum dan Sesudah Imputasi

Variabel	Sebelum imputasi	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5	%Selisih terbesar
$X_4 = 0$	0.047	0.047	0.049	0.047	0.047	0.048	0.2%
$X_4 = 1$	0.953	0.953	0.951	0.953	0.953	0.952	0.2%
$X_7 = 0$	0.051	0.064	0.056	0.055	0.053	0.052	1.3%
$X_7 = 1$	0.949	0.936	0.944	0.945	0.947	0.948	1.3%
$X_8 = 0$	0.214	0.219	0.214	0.211	0.209	0.210	0.5%
$X_8 = 1$	0.786	0.781	0.786	0.789	0.791	0.790	0.5%

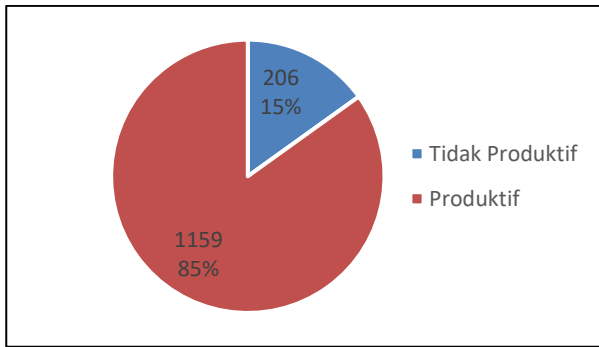
Selain itu, evaluasi hasil imputasi juga diperkuat melalui perbandingan distribusi data (*distribution comparison*) sebelum dan sesudah imputasi, yang disajikan dalam bentuk proporsi kategori pada Tabel 4.4. Berdasarkan Tabel 4.4, menunjukkan bahwa proporsi kategori pada data hasil imputasi relatif mendekati

data asli. Pada variabel X_4 dan X_8 , perbedaan proporsi antara data sebelum dan sesudah imputasi tergolong sangat kecil, yaitu kurang dari 1% (selisih terbesar berada di 0.2%). Sementara itu, pada variabel X_7 terdapat perbedaan proporsi yang sedikit lebih besar, yaitu selisih terbesarnya 1,3%, namun perbedaan tersebut masih dalam batas wajar dan tidak menunjukkan adanya pergeseran pola distribusi yang signifikan.

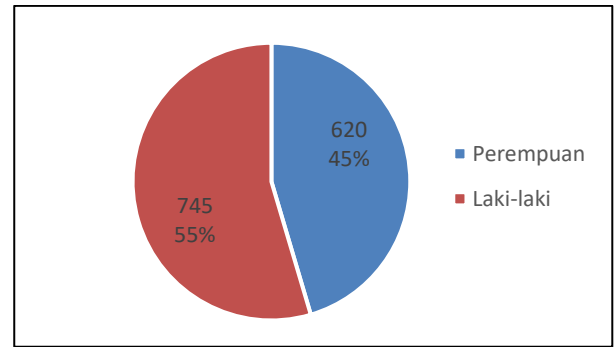
Secara keseluruhan, tidak terlihat adanya perubahan distribusi kategori antara data sebelum dan sesudah imputasi, sehingga dapat disimpulkan bahwa proses imputasi tidak mengubah karakteristik data secara substantif. Selain itu, nilai imputasi pada setiap observasi yang sebelumnya missing dapat berbeda antar dataset imputasi ke-1 hingga ke-5, yang menunjukkan adanya variasi acak sebagai bagian dari pendekatan *multiple imputation*. Rincian nilai imputasi pada masing-masing observasi disajikan pada Lampiran 5 sampai Lampiran 7. Oleh karena itu, metode *Multiple Imputation by Chained Equations* (MICE) dinilai mampu menangani *missing value* pada variabel X_4 , X_7 , dan X_8 dengan baik, serta menghasilkan data yang stabil, realistis, dan layak digunakan untuk analisis selanjutnya.

4.2 Analisis Deskriptif

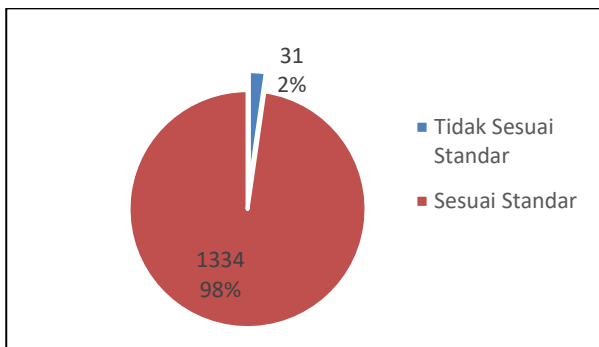
Analisis deskriptif dilakukan untuk menggambarkan karakteristik pasien TB berdasarkan faktor-faktor yang diduga memengaruhi kesembuhan. Hasil disajikan dalam bentuk distribusi frekuensi, persentase, dan *pie chart*.



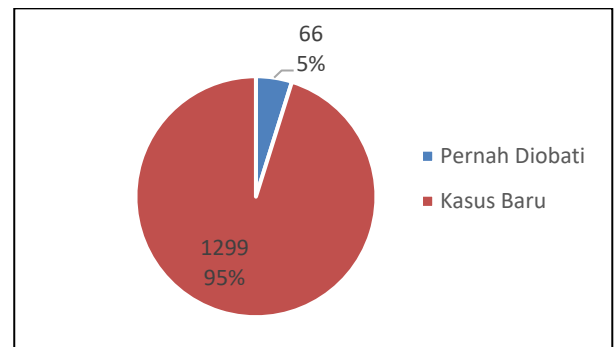
(a) Usia



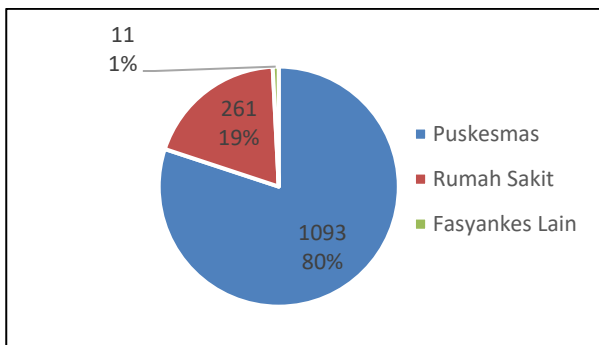
(b) Jenis Kelamin



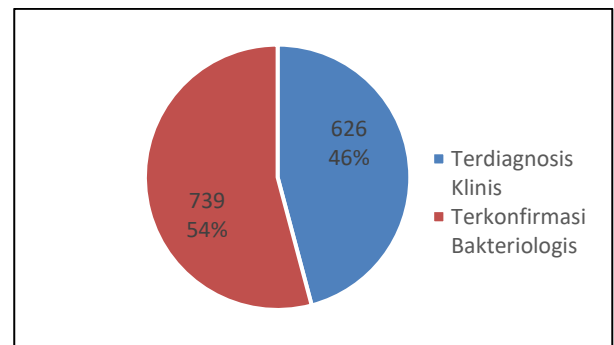
(c) Status Pengobatan



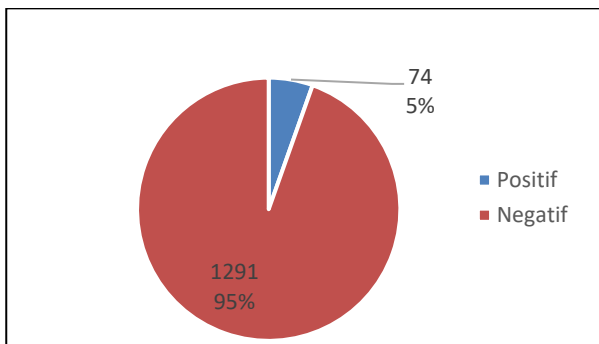
(d) Riwayat Pengobatan



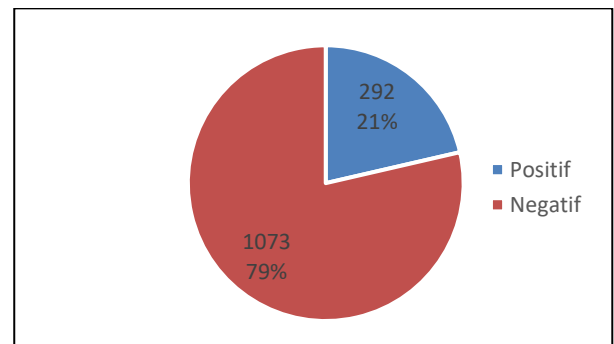
(e) Jenis Fasyankes



(f) Tipe Diagnosis TB



(g) Hasil Tes HIV



(h) Hasil Tes DM

Gambar 4.3 Pie Chart dari Masing-Masing Variabel Independen

Analisis dilakukan dengan menggabungkan seluruh dataset hasil *multiple imputation* ($m = 5$) untuk memperoleh estimasi yang lebih stabil. Frekuensi dihitung sebagai rata-rata dari seluruh imputasi, sedangkan persentase diperoleh dengan membandingkan frekuensi terhadap total pengamatan dan dikalikan 100%.

1. Usia (X_1)

Berdasarkan Gambar 4.3 (a), sebagian besar pasien berada pada kelompok usia produktif (15–64 tahun). Hal ini menunjukkan bahwa kasus tuberkulosis dalam penelitian ini lebih banyak terjadi pada kategori usia yang aktif secara sosial dan ekonomi. Kondisi ini dapat meningkatkan risiko penularan karena kategori usia produktif memiliki mobilitas yang tinggi serta interaksi sosial yang lebih luas.

2. Jenis Kelamin (X_2)

Berdasarkan Gambar 4.3 (b), distribusi pasien berdasarkan jenis kelamin menunjukkan bahwa pasien laki-laki dan perempuan memiliki proporsi yang relatif seimbang meskipun laki-laki sedikit lebih dominan. Hal ini mengindikasikan bahwa tuberkulosis dapat menyerang kedua kategori jenis kelamin, namun laki-laki cenderung memiliki risiko sedikit lebih tinggi, yang dapat dikaitkan dengan faktor gaya hidup atau paparan lingkungan.

3. Status Pengobatan (X_3)

Berdasarkan Gambar 4.3 (c), hampir seluruh pasien menjalani pengobatan sesuai standar yang ditetapkan. Hal ini menunjukkan bahwa sistem pelayanan kesehatan telah berjalan dengan baik dalam memberikan terapi yang sesuai pedoman. Kepatuhan terhadap standar pengobatan merupakan faktor penting dalam meningkatkan peluang kesembuhan pasien.

4. Riwayat Pengobatan (X_4)

Berdasarkan Gambar 4.3 (d), sebagian besar pasien merupakan kasus baru yang belum pernah menjalani pengobatan sebelumnya. Hal ini menunjukkan bahwa mayoritas kasus tuberkulosis dalam penelitian ini merupakan infeksi baru, bukan kasus kambuhan/putus obat.

5. Jenis Fasilitas Layanan Kesehatan (X_5)

Berdasarkan Gambar 4.3 (e), sebagian besar pasien mendapatkan layanan pengobatan di puskesmas dibandingkan rumah sakit maupun fasilitas kesehatan lainnya. Hal ini menunjukkan bahwa puskesmas memiliki peran utama sebagai fasilitas pelayanan kesehatan tingkat pertama dalam penanganan tuberkulosis, terutama dalam program pengendalian TB di masyarakat.

6. Tipe Diagnosis TB (X_6)

Berdasarkan Gambar 4.3 (f), distribusi tipe diagnosis menunjukkan bahwa pasien dengan diagnosis terkonfirmasi bakteriologis dan diagnosis klinis memiliki proporsi yang relatif seimbang. Hal ini menunjukkan bahwa metode diagnosis TB dalam penelitian ini menggunakan kombinasi pemeriksaan klinis dan laboratorium untuk memastikan kondisi pasien.

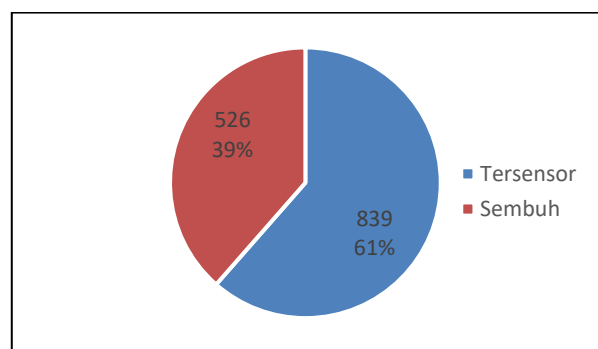
7. Hasil Tes HIV (X_7)

Berdasarkan Gambar 4.3 (g), sebagian besar pasien memiliki hasil tes HIV negatif. Hal ini menunjukkan bahwa sebagian besar pasien TB dalam penelitian ini tidak disertai dengan infeksi HIV. Namun demikian, keberadaan pasien dengan HIV tetap perlu diperhatikan karena koinfeksi TB-HIV dapat memperburuk kondisi pasien dan menurunkan peluang kesembuhan.

8. Hasil Tes DM (X_8)

Berdasarkan Gambar 4.3 (h), mayoritas pasien memiliki hasil tes diabetes mellitus negatif. Hal ini menunjukkan bahwa sebagian besar pasien TB tidak memiliki penyakit penyerta diabetes. Meskipun demikian, pasien dengan diabetes tetap menjadi perhatian karena kondisi tersebut dapat memengaruhi respons terhadap pengobatan TB.

9. Status Kejadian



Gambar 4.4 Pie Chart dari Variabel Status Kejadian

Berdasarkan Gambar 4.4, sebagian besar pasien belum mencapai status sembuh atau termasuk dalam kategori tersensor (*censored*), seperti putus obat, hilang *follow-up*, atau meninggal. Hal ini menunjukkan bahwa masih terdapat tantangan dalam keberhasilan pengobatan tuberkulosis, sehingga perlu dilakukan analisis lebih lanjut terhadap faktor-faktor yang memengaruhi kesembuhan pasien.

4.3 Analisis Kaplan-Meier

Analisis *Kaplan-Meier* pada penelitian ini digunakan untuk menggambarkan probabilitas belum sembuh pasien tuberkulosis berdasarkan waktu pengamatan. Analisis dilakukan pada seluruh dataset hasil imputasi, namun pembahasan pada sub-bab ini difokuskan pada data imputasi 1, sedangkan hasil analisis pada dataset

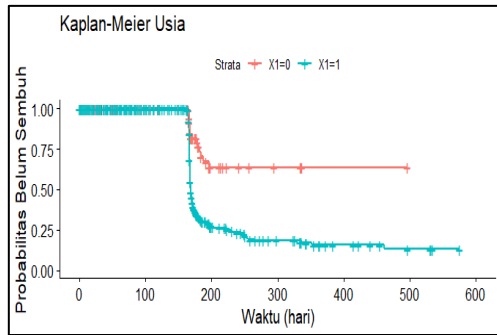
imputasi lainnya disajikan pada Lampiran 8 sampai Lampiran 13. T didefinisikan sebagai lama waktu (dalam hari) sejak pasien memulai pengobatan hingga pasien dinyatakan sembuh ($event = 1$) atau hingga akhir pengamatan jika pasien belum sembuh, putus obat, hilang *follow-up*, atau meninggal ($censored = 0$).

4.3.1 Kurva Kaplan-Meier

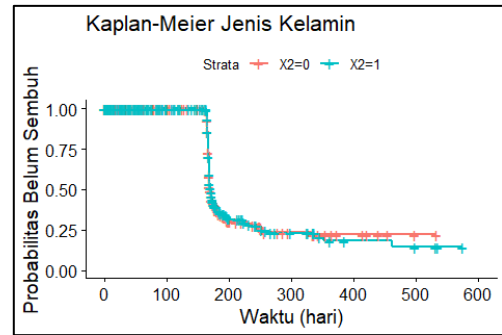
Kurva *Kaplan-Meier* yang dihasilkan menampilkan hubungan antara waktu dan probabilitas pasien belum sembuh. Sumbu horizontal (X) menunjukkan waktu (dalam hari), sedangkan sumbu vertikal (Y) menunjukkan probabilitas pasien untuk belum mengalami kejadian sembuh pada waktu tertentu. Kurva berbentuk fungsi tangga (*step function*), di mana setiap penurunan kurva menunjukkan terjadinya *event* (pasien sembuh), sedangkan data tersensor ditandai tanpa menyebabkan penurunan kurva. Setiap kurva merepresentasikan kelompok (*strata*) berdasarkan kategori variabel independen.

Kurva *Kaplan-Meier* selalu dimulai dari waktu $t = 0$ dengan nilai *survival* sebesar $S(0) = 1$. Hal ini dikarenakan pada awal pengamatan seluruh pasien belum mengalami kejadian sembuh, sehingga probabilitas belum sembuh adalah 1 atau 100%. Meskipun perbedaan antar kelompok baru mulai terlihat pada waktu tertentu, kurva tetap ditampilkan sejak awal pengamatan untuk menggambarkan keseluruhan proses hingga *event* terjadi secara utuh.

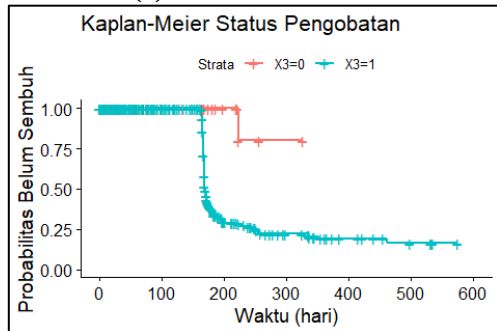
Panjang kurva *Kaplan-Meier* pada masing-masing kelompok dapat berbeda karena kurva akan berakhir pada waktu terakhir ketika masih terdapat pasien yang berada dalam kelompok berisiko (*at risk*), yaitu pasien yang belum mengalami *event* sembuh dan belum tersensor.



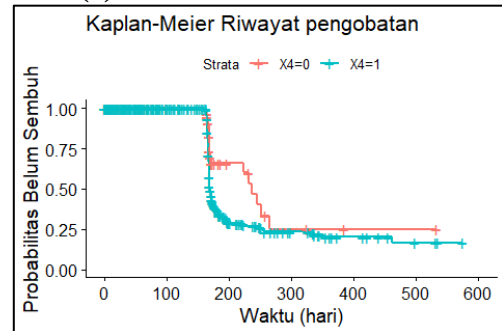
(a) Berdasarkan Usia



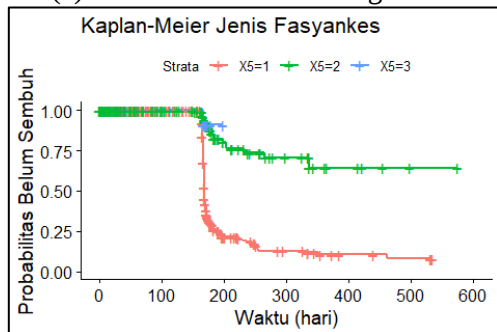
(b) Berdasarkan Jenis Kelamin



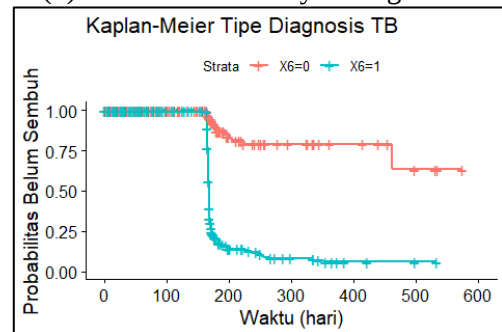
(c) Berdasarkan Status Pengobatan



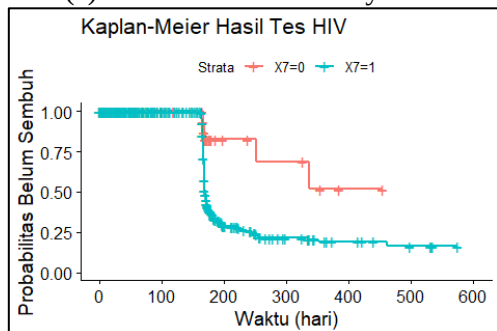
(d) Berdasarkan Riwayat Pengobatan



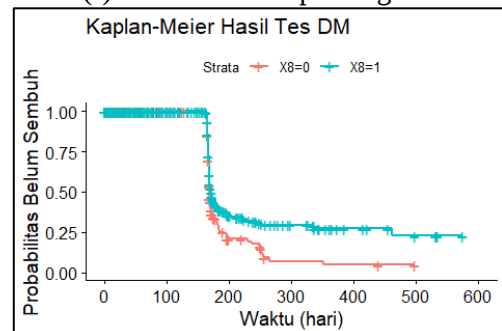
(e) Berdasarkan Jenis Fasyankes



(f) Berdasarkan Tipe Diagnosis TB



(g) Berdasarkan Hasil Tes HIV



(h) Berdasarkan Hasil Tes DM

Gambar 4.5 Kurva Kaplan-Meier pada Dataset Imputasi 1

Apabila pada suatu kelompok tidak terdapat lagi individu yang tersisa dalam pengamatan, baik karena seluruhnya telah mengalami *event* (sembuh) maupun tersensor, maka kurva pada kelompok tersebut akan berhenti lebih awal. Sebaliknya, kelompok yang masih memiliki individu dalam pengamatan hingga waktu yang lebih panjang akan menghasilkan kurva yang lebih panjang.

Estimasi fungsi *survival* untuk masing-masing kategori pada variabel independen berdasarkan Persamaan (2.17) disajikan pada bagian Lampiran 8 sampai Lampiran 15, sedangkan untuk kurva *Kaplan-Meier* dataset imputasi 2 sampai imputasi 5 disajikan pada Lampiran 16 sampai Lampiran 19.

1. *Kaplan-Meier* Berdasarkan Usia (X_1)

Berdasarkan Gambar 4.5 (a), kurva *Kaplan-Meier* berdasarkan variabel usia (X_1) menunjukkan perbedaan pola probabilitas belum sembuh antara dua kelompok pasien, yaitu usia tidak produktif ($X_1 = 0$; <15 tahun dan >64 tahun) dan usia produktif ($X_1 = 1$; 15–64 tahun). Selama periode pengamatan, kurva *Kaplan-Meier* untuk kelompok usia produktif ($X_1 = 1$) berada lebih rendah dibandingkan kelompok usia tidak produktif ($X_1 = 0$). Hal ini menunjukkan bahwa probabilitas belum sembuh pada pasien usia produktif lebih cepat menurun seiring waktu, yang berarti pasien dalam kelompok ini cenderung lebih cepat mengalami kejadian sembuh. Sebaliknya, kurva pada kelompok usia tidak produktif berada pada posisi yang lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih relatif besar pada berbagai titik waktu pengamatan. Dengan kata lain, pasien pada kelompok usia tidak produktif cenderung membutuhkan waktu yang lebih lama untuk mencapai kesembuhan dibandingkan pasien usia produktif.

2. *Kaplan-Meier* Berdasarkan Jenis Kelamin (X_2)

Berdasarkan Gambar 4.5 (b), kurva *Kaplan-Meier* berdasarkan variabel jenis kelamin (X_2) menunjukkan bahwa pasien laki-laki ($X_2 = 1$) dan perempuan ($X_2 = 0$) memiliki pola probabilitas belum sembuh yang relatif serupa selama periode pengamatan. Hal ini terlihat dari posisi kedua kurva yang cenderung berimpit atau tidak menunjukkan pemisahan yang jelas sepanjang waktu. Kesamaan pola tersebut mengindikasikan bahwa penurunan probabilitas belum sembuh pada kedua kelompok berlangsung dengan laju yang hampir sama, sehingga tidak terdapat perbedaan yang berarti dalam waktu yang dibutuhkan untuk mencapai kesembuhan antara pasien laki-laki dan perempuan. Dengan kata lain, baik pasien laki-laki maupun perempuan memiliki kecenderungan yang relatif sama dalam mengalami kejadian sembuh selama masa pengamatan. Namun, interpretasi berdasarkan grafik bersifat deskriptif, sehingga diperlukan pengujian statistik lebih lanjut menggunakan uji *log-rank* untuk mengevaluasi signifikansi perbedaan antar kelompok yang akan dijelaskan pada sub-bab tersendiri.

3. *Kaplan-Meier* Berdasarkan Status Pengobatan (X_3)

Berdasarkan Gambar 4.5 (c), kurva *Kaplan-Meier* berdasarkan variabel status pengobatan (X_3) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien yang menjalani pengobatan sesuai standar ($X_3 = 1$) dan tidak sesuai standar ($X_3 = 0$). Selama periode pengamatan, kurva *Kaplan-Meier* untuk pasien dengan pengobatan sesuai standar berada lebih rendah dibandingkan pasien dengan pengobatan tidak sesuai standar. Hal ini menunjukkan bahwa probabilitas belum sembuh pada kelompok pasien

dengan pengobatan sesuai standar menurun lebih cepat seiring waktu, yang berarti pasien dalam kelompok ini cenderung lebih cepat mengalami kejadian sembuh. Sebaliknya, pasien yang menjalani pengobatan tidak sesuai standar memiliki kurva yang berada lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih relatif besar pada berbagai titik waktu pengamatan.

4. *Kaplan-Meier* Berdasarkan Riwayat Pengobatan (X_4)

Berdasarkan Gambar 4.5 (d), kurva *Kaplan-Meier* berdasarkan variabel riwayat pengobatan (X_4) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien dengan kasus baru ($X_4 = 1$; belum pernah diobati sebelumnya) dan pasien yang pernah diobati sebelumnya ($X_4 = 0$). Selama periode pengamatan, kurva *Kaplan-Meier* untuk pasien dengan kasus baru berada lebih rendah dibandingkan pasien yang memiliki riwayat pengobatan sebelumnya. Hal ini menunjukkan bahwa probabilitas belum sembuh pada kelompok pasien kasus baru menurun lebih cepat seiring waktu, yang berarti pasien dalam kelompok ini cenderung lebih cepat mengalami kejadian sembuh. Sebaliknya, pasien yang pernah diobati sebelumnya memiliki kurva yang berada lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih relatif besar pada berbagai titik waktu pengamatan.

5. *Kaplan-Meier* Berdasarkan Jenis Fasyankes (X_5)

Berdasarkan gambar 4.5 (e), kurva *Kaplan-Meier* berdasarkan variabel jenis fasilitas layanan kesehatan (X_5) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien yang menjalani pengobatan di puskesmas ($X_5 = 1$), rumah sakit ($X_5 = 2$), dan fasilitas kesehatan lainnya (X_5

= 3). Selama periode pengamatan, terlihat bahwa kurva *Kaplan-Meier* pada masing-masing kategori tidak sepenuhnya berimpit, yang mengindikasikan adanya perbedaan waktu kesembuhan antar kelompok. Berdasarkan grafik, pasien yang menjalani pengobatan di puskesmas memiliki kurva yang berada paling rendah dibandingkan kategori lainnya. Hal ini menunjukkan bahwa probabilitas belum sembuh pada kelompok ini menurun lebih cepat seiring waktu, sehingga pasien yang berobat di puskesmas cenderung lebih cepat mengalami kejadian sembuh. Sementara itu, pasien yang menjalani pengobatan di rumah sakit maupun fasilitas kesehatan lainnya memiliki kurva yang relatif lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh pada kelompok tersebut masih lebih besar pada berbagai titik waktu pengamatan.

6. *Kaplan-Meier* Berdasarkan Tipe Diagnosis TB (X_6)

Berdasarkan Gambar 4.5 (f), kurva *Kaplan-Meier* berdasarkan variabel tipe diagnosis TB (X_6) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien dengan terkonfirmasi bakteriologis ($X_6 = 1$) dan diagnosis klinis ($X_6 = 0$). Selama periode pengamatan, kurva *Kaplan-Meier* untuk pasien dengan terkonfirmasi bakteriologis berada lebih rendah dibandingkan pasien dengan diagnosis klinis. Hal ini terlihat dari penurunan kurva pada kelompok terkonfirmasi bakteriologis yang terjadi lebih cepat seiring waktu, yang menunjukkan bahwa probabilitas belum sembuh pada kelompok ini menurun lebih cepat. Dengan demikian, pasien dengan terkonfirmasi bakteriologis cenderung lebih cepat mengalami kejadian sembuh dibandingkan pasien dengan diagnosis klinis. Sebaliknya, kurva pada

kelompok diagnosis klinis berada pada posisi yang lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih relatif besar pada berbagai titik waktu pengamatan.

7. *Kaplan-Meier* Berdasarkan Hasil Tes HIV (X_7)

Berdasarkan Gambar 4.5 (g), kurva *Kaplan-Meier* berdasarkan variabel hasil tes HIV (X_7) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien dengan HIV negatif ($X_7 = 1$) dan HIV positif ($X_7 = 0$). Selama periode pengamatan, kurva *Kaplan-Meier* untuk pasien dengan HIV negatif berada lebih rendah dibandingkan pasien dengan HIV positif. Hal ini terlihat dari penurunan kurva pada kelompok HIV negatif yang terjadi lebih cepat seiring waktu, yang menunjukkan bahwa probabilitas belum sembuh pada kelompok ini menurun lebih cepat. Dengan demikian, pasien TB dengan HIV negatif cenderung lebih cepat mengalami kejadian sembuh. Sebaliknya, pada kelompok HIV positif berada pada posisi yang lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih relatif besar pada berbagai titik waktu pengamatan.

8. *Kaplan-Meier* Berdasarkan Hasil Tes DM (X_8)

Berdasarkan Gambar 4.5 (h), kurva *Kaplan-Meier* berdasarkan variabel hasil tes Diabetes Mellitus (DM) (X_8) menunjukkan adanya perbedaan pola probabilitas belum sembuh antara pasien dengan DM positif ($X_8 = 0$) dan DM negatif ($X_8 = 1$). Selama periode pengamatan, kurva *Kaplan-Meier* untuk pasien dengan DM positif berada lebih rendah dibandingkan pasien dengan DM negatif, yang terlihat dari penurunan kurva yang lebih cepat pada kelompok DM positif. Hal ini menunjukkan bahwa probabilitas belum

sembuh pada kelompok pasien dengan DM positif menurun lebih cepat seiring waktu, sehingga secara deskriptif pasien pada kelompok ini tampak lebih cepat mengalami kejadian sembuh dibandingkan pasien tanpa DM. Sebaliknya, kurva pada kelompok DM negatif berada pada posisi yang relatif lebih tinggi, yang mengindikasikan bahwa probabilitas belum sembuh masih lebih besar pada berbagai titik waktu pengamatan. Namun demikian, hasil ini perlu diinterpretasikan secara hati-hati. Secara teoritis, kondisi diabetes melitus dapat memperlambat proses penyembuhan tuberkulosis akibat gangguan sistem imun. Ketidakesesuaian antara hasil analisis dan teori ini kemungkinan disebabkan oleh adanya faktor lain yang belum dikendalikan dalam analisis, seperti tingkat kepatuhan pengobatan, intensitas pengawasan medis, maupun karakteristik klinis pasien lainnya. Selain itu, analisis *Kaplan-Meier* merupakan analisis univariat yang belum mempertimbangkan pengaruh variabel lain secara simultan. Oleh karena itu, diperlukan analisis lanjutan menggunakan model *Cox Proportional Hazards* model untuk memperoleh hasil yang lebih komprehensif dan mampu mengontrol variabel-variabel perancu yang mungkin memengaruhi waktu kesembuhan pasien.

4.3.2 Uji *Log-rank*

Uji *log-rank* digunakan untuk mengetahui apakah terdapat perbedaan kurva *survival* (probabilitas belum sembuh) antar kategori pada masing-masing variabel independen. Dalam penelitian ini, uji *log-rank* dilakukan untuk membandingkan kurva *Kaplan-Meier* berdasarkan variabel X_1 hingga X_8 . Analisis dilakukan pada seluruh dataset hasil imputasi, namun pembahasan pada subbab ini difokuskan pada dataset imputasi pertama (imputasi ke-1), sedangkan hasil uji *log-rank* pada dataset

imputasi ke-2 hingga ke-5 disajikan pada Lampiran 20 sampai Lampiran 23. Uji ini digunakan untuk mengidentifikasi variabel-variabel yang memiliki perbedaan signifikan dalam waktu kesembuhan sebelum dilakukan analisis lanjutan menggunakan model *Cox Proportional Hazards*. Hipotesis yang digunakan dalam uji *log-rank* adalah sebagai berikut:

H_0 : Tidak terdapat perbedaan kurva *survival* di antara kategori yang dibandingkan.

H_1 : Setidaknya terdapat satu kategori dengan kurva *survival* yang berbeda dari kategori lainnya.

Pengambilan keputusan yang digunakan dalam uji *log-rank* ini yaitu tolak H_0 jika

$\chi^2 > \chi^2_{(\alpha, C-1)}$ atau $p - value < \alpha = 0.05$.

Tabel 4.5 Hasil Uji *Log-rank*

Variabel	χ^2 hitung	χ^2 tabel	df	$p - value$	Keputusan
X_1	53.3405	3.8415	1	0.0000	Tolak H_0
X_2	0.0068	3.8415	1	0.9342	Gagal Tolak H_0
X_3	18.5661	3.8415	1	0.0000	Tolak H_0
X_4	8.7354	3.8415	1	0.0031	Tolak H_0
X_5	115.7570	5.9915	2	0.0000	Tolak H_0
X_6	329.0837	3.8415	1	0.0000	Tolak H_0
X_7	20.3408	3.8415	1	0.0000	Tolak H_0
X_8	9.4392	3.8415	1	0.0021	Tolak H_0

Berdasarkan Tabel 4.5, interpretasi masing-masing variabel sebagai berikut:

1. Variabel Usia (X_1)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien usia tidak produktif dan usia produktif.

H_1 : Terdapat perbedaan kurva *survival* antara pasien usia tidak produktif dan usia produktif.

Hasil uji menunjukkan $p - value = 0.0000 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel} (53.3405 > 3.8415)$, sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan antar kategori pada pasien usia produktif (15-64 tahun) dan tidak produktif (< 15 dan > 64 tahun)

2. Variabel Jenis Kelamin (X_2)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien laki-laki dan perempuan.

H_1 : Terdapat perbedaan kurva *survival* antara pasien laki-laki dan perempuan.

Hasil uji menunjukkan $p - value = 0.9342 (> 0.05)$ dan $\chi^2_{hitung} \leq \chi^2_{tabel} (0.0068 \leq 3.8415)$, sehingga H_0 gagal ditolak. Artinya, tidak terdapat perbedaan signifikan waktu kesembuhan antar kategori pada pasien laki-laki dan perempuan.

3. Variabel Status Pengobatan (X_3)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien dengan pengobatan secara standar dan tidak sesuai standar.

H_1 : Terdapat perbedaan kurva *survival* antara pasien dengan pengobatan standar dan tidak sesuai standar.

Hasil uji menunjukkan $p - value = 0.0000 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel} (18.5661 > 3.8415)$, sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan antar kategori pada pasien dengan pengobatan standar dan tidak sesuai standar.

4. Variabel Riwayat Pengobatan (X_4)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien dengan kasus baru dan pasien yang pernah diobati sebelumnya.

H_1 : Terdapat perbedaan kurva *survival* antara pasien dengan kasus baru dan pasien yang pernah diobati sebelumnya.

Hasil uji menunjukkan $p - value = 0.0031 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel} (8.7354 > 3.8415)$, sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan antar kategori pada pasien dengan kasus baru dan pernah diobati sebelumnya.

5. Variabel Jenis Fasyankes (X_5)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien yang berobat di puskesmas, rumah sakit, dan fasilitas kesehatan lainnya.

H_1 : Terdapat perbedaan kurva *survival* antara pasien yang berobat di puskesmas, rumah sakit, dan fasyankes lain.

Hasil uji menunjukkan $p - value = 0.0000 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel} (115.7570 > 5.9915)$, sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan antar ketiga kategori berdasarkan jenis fasyankes.

6. Variabel Tipe Diagnosis TB (X_6)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien dengan diagnosis klinis dan bakteriologis.

H_1 : Terdapat perbedaan kurva *survival* antara pasien dengan diagnosis klinis dan bakteriologis.

Hasil uji menunjukkan $p - value = 0.0000 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel}$ ($329.0837 > 3.8415$), sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan antar kategori pada pasien dengan diagnosis klinis dan bakteriologis.

7. Variabel Hasil Tes HIV (X_7)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien HIV positif dan HIV negatif.

H_1 : Terdapat perbedaan kurva *survival* antara pasien dengan HIV positif dan HIV negatif.

Hasil uji menunjukkan $p - value = 0.0000 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel}$ ($20.3408 > 3.8415$), sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan berdasarkan status HIV.

8. Variabel Hasil Tes DM (X_8)

Hipotesis yang diuji:

H_0 : Tidak terdapat perbedaan kurva *survival* antara pasien dengan DM dan tanpa DM.

H_1 : Terdapat perbedaan kurva *survival* antara pasien dengan DM dan tanpa DM.

Hasil uji menunjukkan $p - value = 0.0021 (< 0.05)$ dan $\chi^2_{hitung} > \chi^2_{tabel}$ ($9.4392 > 3.8415$), sehingga H_0 ditolak. Artinya, terdapat perbedaan signifikan waktu kesembuhan berdasarkan status Diabetes Mellitus.

Berdasarkan hasil uji *log-rank*, variabel yang menunjukkan perbedaan signifikan kurva *survival* (probabilitas belum sembuh) adalah $X_1, X_3, X_4, X_5, X_6, X_7, X_8$. Sementara itu, variabel X_2 tidak menunjukkan perbedaan yang signifikan. Hasil ini juga konsisten pada seluruh dataset hasil imputasi (imputasi ke-1 hingga ke-5), di mana variabel X_2 secara konsisten menghasilkan keputusan gagal menolak H_0 . Dengan demikian, variabel-variabel yang signifikan tersebut selanjutnya akan dilanjutkan ke analisis multivariat menggunakan model *Cox Proportional Hazards* untuk mengetahui pengaruhnya secara simultan terhadap waktu kesembuhan pasien.

4.4 Analisis Regresi Cox Proportional Hazard

Pada penelitian ini, analisis *Cox Proportional Hazard* digunakan untuk mengidentifikasi faktor-faktor yang memengaruhi waktu kesembuhan pasien tuberkulosis, di mana kejadian (*event*) yang diamati adalah kesembuhan pasien. Waktu *survival* (T) didefinisikan sebagai lama waktu (dalam hari) sejak awal pengobatan hingga pasien dinyatakan sembuh.

4.4.1 Pembentukan Model Cox Proportional Hazard

Model ini bersifat semiparametrik karena tidak memerlukan asumsi bentuk tertentu dari fungsi *hazard* dasar, namun mengasumsikan bahwa rasio *hazard* antar individu bersifat konstan terhadap waktu. Berdasarkan hasil analisis *Kaplan-Meier* dan uji *log-rank* pada sub-bab sebelumnya, diperoleh bahwa variabel jenis kelamin (X_2) tidak berpengaruh signifikan terhadap waktu kesembuhan, sehingga variabel tersebut tidak dimasukkan dalam model Cox. Dengan demikian, model *Cox Proportional Hazard* yang terbentuk berdasarkan persamaan (2.19) yaitu:

$$h(t, \mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_3 + \beta_3 X_4 + \beta_4 D_{5,1} + \beta_5 D_{5,2} + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8) \quad (4.4)$$

Model ini digunakan untuk melihat pengaruh masing-masing variabel terhadap laju kesembuhan pasien TB.

4.4.2 Estimasi Parameter Model

Estimasi parameter pada model *Cox Proportional Hazard* dilakukan menggunakan metode *Maximum Partial Likelihood Estimation* (MPLE) yang dioptimalkan melalui algoritma Newton–Raphson. Proses estimasi dilakukan secara iteratif hingga parameter model mencapai kondisi konvergen. Dalam penelitian ini, proses iterasi diimplementasikan menggunakan fungsi `coxph()`, di mana nilai parameter diperbarui secara otomatis hingga perubahan antar iterasi menjadi sangat kecil. Fungsi `coxph()` memiliki kriteria konvergensi internal, sehingga nilai jarak Euclidean tidak ditampilkan secara eksplisit oleh output.

Tabel 4.6 Jumlah Iterasi *Cox Proportional Hazard*

Imputasi	Iterasi	Keterangan
1	7	Konvergen
2	7	Konvergen
3	7	Konvergen
4	7	Konvergen
5	7	Konvergen

Berdasarkan Tabel 4.6, seluruh model pada masing-masing dataset imputasi telah mencapai konvergensi setelah 7 iterasi. Jumlah iterasi yang konsisten pada setiap imputasi menunjukkan bahwa proses estimasi parameter berjalan stabil. Sehingga, didapatkan model *Cox Proportional Hazard* pada masing-masing imputasi berdasarkan Persamaan 4.4 sebagai berikut:

1. Model *Cox Proportional Hazard* Dataset Imputasi 1

$$h(t, \mathbf{X}) = h_0(t) \exp(0.4245X_1 + 2.0400X_3 + 0.8776X_4 - 1.2692D_{5,1} - 1.4686D_{5,2} + 2.3918X_6 + 0.6740X_7 - 0.0145X_8) \quad (4.5)$$

2. Model *Cox Proportional Hazard* Dataset Imputasi 2

$$h(t, \mathbf{X}) = h_0(t) \exp(0.4397X_1 + 2.0542X_3 + 0.8867X_4 - 1.2245D_{5,1} - 2.0453D_{5,2} + 2.3788X_6 + 0.7262X_7 - 0.0432X_8) \quad (4.6)$$

3. Model *Cox Proportional Hazard* Dataset Imputasi 3

$$h(t, \mathbf{X}) = h_0(t) \exp(0.4247X_1 + 2.0529X_3 + 0.8613X_4 - 1.2663D_{5,1} - 2.0411D_{5,2} + 2.3743X_6 + 0.5329X_7 - 0.0543X_8) \quad (4.7)$$

4. Model *Cox Proportional Hazard* Dataset Imputasi 4:

$$h(t, \mathbf{X}) = h_0(t) \exp(0.4196X_1 + 2.0424X_3 + 0.8109X_4 - 1.2724D_{5,1} - 2.0346D_{5,2} + 2.3920X_6 + 0.5591X_7 - 0.0006X_8) \quad (4.8)$$

5. Model *Cox Proportional Hazard* Dataset Imputasi 5

$$h(t, \mathbf{X}) = h_0(t) \exp(0.4205X_1 + 2.0326X_3 + 0.8010X_4 - 1.2819D_{5,1} - 2.0242D_{5,2} + 2.3978X_6 + 0.5617X_7 + 0.4150X_8) \quad (4.9)$$

Berdasarkan kelima model yang diperoleh, terlihat bahwa nilai koefisien regresi pada masing-masing variabel relatif konsisten antar imputasi. Hal ini menunjukkan bahwa proses imputasi yang dilakukan menghasilkan estimasi parameter yang stabil. Secara umum, variabel X_1 , X_3 , X_4 , X_6 , dan X_7 memiliki koefisien positif, yang menunjukkan bahwa peningkatan nilai variabel tersebut cenderung meningkatkan *hazard* kesembuhan pasien. Sebaliknya, variabel X_5 memiliki koefisien negatif, yang menunjukkan bahwa kategori tertentu pada variabel tersebut cenderung menurunkan *hazard* kesembuhan. Dalam model Persamaan (4.5) sampai (4.9), variabel $D_{5,1}$ adalah kategori rumah sakit dibandingkan dengan kategori puskesmas, dan $D_{5,2}$ adalah kategori fasyankes lainnya dibandingkan dengan kategori puskesmas.

4.4.3 Uji Asumsi Cox *Proportional Hazard*

Pengujian asumsi *proportional hazard* pada penelitian ini dilakukan secara visual dan formal. Secara visual, grafik *Kaplan-Meier* digunakan sebagai pemeriksaan awal, di mana pola kurva yang relatif sejajar dapat mengindikasikan bahwa rasio *hazard* cenderung konstan terhadap waktu, meskipun hasil ini tidak bersifat konklusif. Oleh karena itu, pengujian dilanjutkan secara formal menggunakan uji Schoenfeld residual, yaitu dengan menguji hubungan antara residual Schoenfeld dan waktu *survival*. Dalam implementasinya, perhitungan dilakukan menggunakan Persamaan (2.29). Hipotesis yang digunakan dalam pengujian ini adalah:

H_0 : Pengaruh kovariat terhadap *hazard* bersifat konstan terhadap waktu

H_1 : Pengaruh kovariat terhadap *hazard* berubah terhadap waktu

Pengambilan keputusan didasarkan pada $p - value$ yang dihitung dari distribusi χ^2 berdasarkan statistik uji yang diperoleh, yaitu:

1. Jika $p - value > 0.05$ atau $\chi^2_{hitung} \leq \chi^2_{0.05,df}$ maka H_0 diterima (asumsi terpenuhi)
2. Jika $p - value \leq 0.05$ atau $\chi^2_{hitung} > \chi^2_{0.05,df}$, maka H_0 ditolak (asumsi tidak terpenuhi)

Berdasarkan Tabel 4.7, keseluruhan hasil pengujian asumsi *proportional hazard* secara formal, dapat disimpulkan bahwa variabel X_6 secara konsisten melanggar asumsi *proportional hazard* pada seluruh dataset hasil imputasi karena $\chi^2_{hitung} > \chi^2_{0.05,df}$ serta $p - value \leq 0.05$. Untuk X_8 tetap dipertahankan dalam model karena tidak melanggar secara konsisten pada seluruh dataset imputasi.

Tabel 4.7 Hasil Uji Asumsi *Cox Proportional Hazard Model* Pertama

Variabel	df	χ^2 tabel	Imputasi 1		Imputasi 2		Imputasi 3		Imputasi 4		Imputasi 5	
			χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
X_1	1	3.841	0.144	0.704	0.103	0.748	0.151	0.698	0.156	0.693	0.171	0.680
X_3	1	3.841	1.950	0.163	1.945	0.163	1.946	0.163	1.954	0.162	1.938	0.164
X_4	1	3.841	1.881	0.170	0.784	0.376	1.992	0.158	1.650	0.199	1.013	0.314
X_5	2	5.991	0.841	0.657	0.979	0.613	1.006	0.605	0.988	0.610	0.977	0.613
X_6	1	3.841	4.809	0.028	5.186	0.023	5.089	0.024	4.885	0.027	4.890	0.027
X_7	1	3.841	0.012	0.913	0.009	0.924	0.004	0.947	0.003	0.956	0.037	0.847
X_8	1	3.841	3.615	0.057	2.949	0.086	3.295	0.069	5.776	0.016	2.776	0.096
Simultan	8	15.507	14.415	0.072	12.655	0.124	14.726	0.065	17.070	0.029	12.792	0.119

Keterangan: Nilai yang ditebali adalah nilai yang melanggar asumsi

Untuk memperoleh model yang lebih valid dan sesuai dengan asumsi *Cox Proportional Hazard*, variabel X_6 dikeluarkan dari model, dan selanjutnya akan dibentuk model kedua *Cox Proportional Hazard* tanpa melibatkan variabel tersebut. Secara umum, model *Cox Proportional Hazard* berdasarkan persamaan (2.19):

$$h(t, \mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_3 + \beta_3 X_4 + \beta_4 D_{5,1} + \beta_5 D_{5,2} + \beta_6 X_7 + \beta_7 X_8) \quad (4.10)$$

Untuk hasil estimasi model kedua Cox tanpa X_6 disajikan pada Lampiran 24.

Selanjutnya, dilakukan uji asumsi *Cox Proportional Hazard* kembali menggunakan uji Schoenfeld residual. Nilai *p-value* pada Tabel 4.8 diperoleh dari statistik uji Schoenfeld yang mengikuti distribusi χ^2 , di mana setiap kolom (Imputasi 1 hingga Imputasi 5) merepresentasikan hasil pengujian pada masing-masing dataset imputasi. Berdasarkan Tabel 4.8, diketahui bahwa sebagian besar variabel telah memenuhi asumsi *proportional hazard* pada sebagian besar dataset imputasi. Namun demikian, terdapat beberapa variabel yang menunjukkan pelanggaran asumsi pada imputasi tertentu, yaitu variabel X_4 dan X_8 . Variabel X_4 teridentifikasi melanggar asumsi *proportional hazard* pada imputasi ke-1, 3, dan 4, sedangkan pada imputasi lainnya variabel ini memenuhi asumsi. Sementara itu, variabel X_8 menunjukkan pelanggaran pada imputasi ke-3, 4, dan 5, namun tetap

memenuhi asumsi pada imputasi lainnya. Selain itu, uji simultan juga menunjukkan bahwa pada imputasi ke-1, 3, dan 4, model secara keseluruhan tidak memenuhi asumsi *proportional hazard*, sedangkan pada imputasi lainnya asumsi terpenuhi.

Tabel 4.8 Hasil Uji Asumsi *Cox Proportional Hazard* Model Kedua

Variabel	df	χ^2 tabel	Imputasi 1		Imputasi 2		Imputasi 3		Imputasi 4		Imputasi 5	
			χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
X_1	1	3.841	0.382	0.537	0.342	0.560	0.448	0.503	0.374	0.541	0.454	0.500
X_3	1	3.841	2.193	0.139	2.192	0.140	2.179	0.140	2.179	0.139	2.111	0.146
X_4	1	3.841	4.660	0.031	2.312	0.130	5.136	0.023	4.275	0.038	3.337	0.068
X_5	2	5.991	2.102	0.350	2.160	0.340	2.148	0.342	2.216	0.330	2.287	0.319
X_7	1	3.841	0.948	0.330	0.410	0.520	0.292	0.589	0.882	0.347	0.324	0.569
X_8	1	3.841	3.520	0.061	2.483	0.120	4.012	0.045	6.801	0.009	3.988	0.046
Simultan	7	14.067	14.585	0.042	10.476	0.160	15.036	0.036	17.655	0.013	13.234	0.067

Keterangan: Nilai yang ditebali adalah nilai yang melanggar asumsi

Meskipun pelanggaran asumsi tidak terjadi pada seluruh dataset imputasi, variabel X_4 dan X_8 menunjukkan pelanggaran pada sebagian besar dataset imputasi. Hal ini mengindikasikan adanya ketidak-konsistenan terhadap asumsi *proportional hazard* yang dapat memengaruhi kestabilan model *Cox Proportional Hazard*. Oleh karena itu, variabel X_4 dan X_8 dikeluarkan dari model dan dilakukan pembentukan model ulang tanpa menyertakan kedua variabel tersebut. Keputusan ini dilakukan agar model akhir lebih memenuhi asumsi *proportional hazard* serta menghasilkan estimasi parameter yang lebih valid dan stabil pada seluruh dataset imputasi.

Pada analisis *Cox Proportional Hazard* model final, dilakukan pembentukan model ulang dengan mengeluarkan variabel X_4 dan X_8 karena pada sebagian besar dataset imputasi melanggar asumsi PH. Model regresi *Cox Proportional Hazard* baru yang dibentuk pada masing-masing dataset imputasi disajikan pada persamaan berikut:

$$h(t, \mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_3 + \beta_3 D_{5,1} + \beta_4 D_{5,2} + \beta_5 X_7) \quad (4.11)$$

Proses estimasi dilakukan secara iteratif hingga parameter model mencapai kondisi konvergen pada iterasi ke-7, yang telah disajikan pada Tabel 4.6 sebelumnya. Hasil estimasi parameter untuk setiap imputasi disajikan pada tabel berikut:

Tabel 4.9 Estimasi Parameter *Cox Proportional Hazard* Model Final

Parameter	Variabel	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
β_1	X_1	0.9602	0.9675	0.9492	0.9517	0.9668
β_2	X_3	2.7372	2.7417	2.7543	2.7545	2.7305
β_3	$D_{5,1}$	-1.7148	-1.7358	-1.7486	-1.7332	-1.7396
β_4	$D_{5,2}$	-1.6822	-2.2040	-2.2049	-2.1951	-2.1166
β_5	X_7	1.0153	1.1340	1.0666	0.9055	0.8269

Berdasarkan Tabel 4.9, dapat dibentuk model pada Persamaan (4.11) pada masing-masing dataset imputasi sebagai berikut:

Dataset imputasi 1:

$$h(t, X) = h_0(t) \exp(0.9602X_1 + 2.7372X_3 - 1.7148D_{5,1} - 1.6822D_{5,2} + 1.0153X_7) \quad (4.12)$$

Dataset imputasi 2:

$$h(t, X) = h_0(t) \exp(0.9675X_1 + 2.7417X_3 - 1.7098D_{5,1} - 1.7358D_{5,2} + 1.1340X_7) \quad (4.13)$$

Dataset imputasi 3:

$$h(t, X) = h_0(t) \exp(0.9492X_1 + 2.7543X_3 - 1.7486D_{5,1} - 2.2049D_{5,2} + 1.0666X_7) \quad (4.14)$$

Dataset imputasi 4:

$$h(t, X) = h_0(t) \exp(0.9517X_1 + 2.7545X_3 - 1.7332D_{5,1} - 2.1951D_{5,2} + 0.9055X_7) \quad (4.15)$$

Dataset imputasi 5:

$$h(t, X) = h_0(t) \exp(0.9668X_1 + 2.7305X_3 - 1.7396D_{5,1} - 2.1166D_{5,2} + 0.8269X_7) \quad (4.16)$$

Pada Persamaan (4.12) hingga (4.16), variabel X_1 , X_3 , dan X_7 memiliki koefisien positif, yang menunjukkan bahwa variabel tersebut meningkatkan *hazard* kesembuhan (mempercepat waktu sembuh). Sementara itu, variabel X_5 yang direpresentasikan melalui variabel dummy $D_{5,1}$ dan $D_{5,2}$ memiliki koefisien negatif. Variabel $D_{5,1}$ menunjukkan kategori rumah sakit dibandingkan dengan kategori

referensi, yaitu puskesmas, sedangkan $D_{5,2}$ menunjukkan kategori fasilitas pelayanan kesehatan lainnya dibandingkan dengan puskesmas. Koefisien negatif pada kedua dummy tersebut menunjukkan bahwa pasien yang menjalani pengobatan di rumah sakit maupun fasilitas kesehatan lainnya memiliki *hazard* kesembuhan yang lebih rendah dibandingkan pasien yang menjalani pengobatan di puskesmas, sehingga cenderung membutuhkan waktu kesembuhan yang lebih lama.

Setelah model dibentuk, dilakukan kembali pengujian asumsi *Proportional Hazard* (PH) secara formal menggunakan residual Schoenfeld untuk setiap variabel prediktor. Nilai $p - value$ pada Tabel 4.10 diperoleh dari statistik uji Schoenfeld yang mengikuti distribusi χ^2 .

Tabel 4.10 Hasil Uji Asumsi Cox *Proportional Hazard* Model Final

Variabel	df	χ^2 tabel	Imputasi 1		Imputasi 2		Imputasi 3		Imputasi 4		Imputasi 5	
			χ^2	p -value	χ^2	p -value	χ^2	p -value	χ^2	p -value	χ^2	p -value
X_1	1	3.841	0.459	0.50	0.442	0.51	0.563	0.45	0.450	0.50	0.524	0.47
X_3	1	3.841	2.250	0.13	2.268	0.13	2.244	0.13	2.246	0.13	2.157	0.14
X_5	2	5.991	1.978	0.37	2.003	0.37	2.064	0.36	2.126	0.35	2.189	0.33
X_7	1	3.841	0.473	0.49	0.064	0.80	0.108	0.74	0.429	0.51	0.078	0.78
Simultan	5	11.070	4.940	0.42	4.569	0.47	4.719	0.45	4.960	0.42	4.688	0.46

Berdasarkan Tabel 4.10, diketahui bahwa pada seluruh dataset imputasi, variabel X_1 , X_3 , X_5 , dan X_7 memiliki nilai $\chi^2_{hitung} \leq \chi^2_{0.05,df}$ serta p -value > 0.05 , sehingga H_0 diterima. Hal ini menunjukkan bahwa variabel-variabel tersebut memenuhi asumsi PH, yaitu pengaruhnya terhadap *hazard* bersifat konstan sepanjang waktu. Secara keseluruhan, hasil uji simultan pada setiap imputasi menunjukkan nilai p -value > 0.05 , yang berarti secara simultan model memenuhi asumsi PH.

4.4.4 Pooling Rubin's Rule

Selanjutnya, karena analisis dilakukan pada data hasil *multiple imputation*, maka langkah berikutnya adalah melakukan proses *pooling* terhadap hasil estimasi parameter dari masing-masing dataset imputasi menggunakan aturan Rubin (*Rubin's Rules*). Proses ini menggabungkan rata-rata estimasi parameter dari setiap imputasi serta memperhitungkan variansi dalam imputasi (*within-imputation*), variansi antar imputasi (*between-imputation*), dan variansi total. Hasil *pooling* dan komponen variansi tersebut disajikan pada Tabel 4.11.

Tabel 4.11 Hasil *Pooling Rubin's Rule* Model Final

Parameter	Variabel	Estimate	$\bar{U}$ (Within)	B (Between)	T (Total)
β_1	X_1	0.9576	0.0415	0.00033	0.0419
β_2	X_3	2.7533	1.0035	0.00016	1.0037
β_3	$D_{5,1}$	-1.7367	0.0493	0.00006	0.0494
β_4	$D_{5,2}$	-2.1347	1.0159	0.01026	1.0283
β_5	X_7	0.9005	0.1630	0.03908	0.2099

Berdasarkan tabel 4.11, model regresi *Cox Proportional Hazard* yang terbentuk adalah:

$$h(t, X) = h_0(t) \exp (0.9576X_1 + 2.7533X_3 - 1.7367D_{5,1} - 2.1347D_{5,2} + 0.9005X_7) \quad (4.17)$$

Model persamaan (4.17) merupakan hasil penggabungan dari lima dataset imputasi, sehingga estimasi parameter yang diperoleh telah mempertimbangkan ketidakpastian akibat data yang hilang dan menghasilkan estimasi yang lebih stabil. Berdasarkan Tabel 4.11 juga diketahui bahwa nilai variansi dalam imputasi ($\bar{U}$) lebih besar dibandingkan variansi antar imputasi (B), pada seluruh variabel. Hal ini menunjukkan bahwa ketidakpastian estimasi parameter lebih banyak berasal dari variasi dalam masing-masing model pada setiap dataset imputasi, sedangkan perbedaan estimasi antar dataset imputasi relatif kecil. Dengan demikian, hasil imputasi yang diperoleh cenderung konsisten antar dataset, sehingga estimasi

parameter yang dihasilkan dapat dianggap stabil. Variansi total (VT) yang merupakan kombinasi dari $\bar{U}$ dan B digunakan untuk menghitung *standard error* pada masing-masing parameter.

4.4.5 Uji Signifikansi Parameter Model

Uji signifikansi parameter dilakukan untuk mengetahui apakah variabel prediktor/independen dalam model regresi *Cox Proportional Hazard* berpengaruh terhadap waktu kesembuhan pasien. Pengujian dilakukan secara serentak menggunakan *Likelihood Ratio Test* (LRT) berdasarkan Persamaan (2.30) dan secara parsial menggunakan *Wald Test* berdasarkan Persamaan (2.31).

1. Uji Serentak (*Likelihood Ratio Test*)

Hipotesis :

$H_0 : \beta_p = 0, \forall p = 1, 2, \dots, 5$, yang berarti seluruh variabel prediktor tidak berpengaruh terhadap waktu kejadian.

H_1 : Paling tidak terdapat satu $\beta_p \neq 0$, yang berarti terdapat variabel prediktor yang berpengaruh terhadap waktu kejadian.

Pengambilan keputusan dalam uji serentak *likelihood* ini adalah tolak H_0 jika $G^2 > \chi^2_{\alpha, df}$ atau $p - value < \alpha = 5\% = 0.05$. Dalam analisis dengan *multiple imputation*, nilai *Likelihood Ratio Test* tidak dapat digabungkan menggunakan aturan Rubin karena berbasis *likelihood*. Oleh karena itu, pengujian dilakukan pada masing-masing dataset imputasi, dan hasilnya disajikan pada Tabel 4.12.

Berdasarkan Tabel 4.12, diperoleh bahwa pada seluruh dataset imputasi, nilai statistik uji G^2 jauh lebih besar dibandingkan nilai $\chi^2_{0.05, 7} = 11.0705$, dengan seluruh $p - value < 0.05$. Dengan demikian, H_0 ditolak pada

seluruh imputasi. Hal ini menunjukkan bahwa secara simultan terdapat paling tidak satu variabel prediktor yang berpengaruh signifikan terhadap waktu kesembuhan pasien tuberkulosis. Konsistensi hasil pada seluruh dataset imputasi juga menunjukkan bahwa model yang terbentuk signifikan secara keseluruhan dan stabil terhadap variasi akibat proses imputasi.

Tabel 4.12 Hasil Uji Serentak Model Final pada Tiap Dataset Imputasi

Imputasi	G^2	$\chi^2_{0.05,5}$	df	$p - value$	Keputusan
1	218.4021	11.0705	5	0.000	Tolak H_0
2	217.6463	11.0705	5	0.000	Tolak H_0
3	224.2117	11.0705	5	0.000	Tolak H_0
4	220.5851	11.0705	5	0.000	Tolak H_0
5	214.2559	11.0705	5	0.000	Tolak H_0

2. Uji Parsial (Wald Test)

Hipotesis:

$H_0 : \beta_p = 0$ dengan $p = 1, 2, \dots, 5$ (variabel prediktor ke- p tidak memiliki pengaruh yang signifikan).

$H_1 : \beta_p \neq 0$ dengan $p = 1, 2, \dots, 5$ (variabel prediktor ke- p memiliki pengaruh yang signifikan).

Pengambilan keputusan dalam uji parsial ini adalah tolak H_0 jika $W^2 > \chi^2_{\alpha,1}$ atau $p - value < \alpha = 5\% = 0.05$.

Tabel 4.13 Hasil Uji Wald Model Final

Parameter	Variabel	W^2 hitung	χ^2 tabel	$p-value$	Keputusan
β_1	X_1	22.0672	3.84	0.0000	Tolak H_0
β_2	X_3	7.5543	3.84	0.0062	Tolak H_0
β_3	$D_{5,1}$	61.0819	3.84	0.0000	Tolak H_0
β_4	$D_{5,2}$	4.4854	3.84	0.0347	Tolak H_0
β_5	X_7	4.9738	3.84	0.0310	Tolak H_0

Berdasarkan Tabel 4.13, diketahui bahwa variabel X_1 (usia), X_3 (status pengobatan), dan X_7 (hasil tes HIV) memiliki nilai $W^2 > \chi^2_{0.05,1} = 3.84$

dengan $p - value < 0.05$, sehingga H_0 ditolak. Hal ini menunjukkan bahwa variabel-variabel tersebut berpengaruh signifikan terhadap waktu kesembuhan pasien tuberkulosis dalam model. Untuk variabel X_5 , pengujian dilakukan melalui variabel dummy yang terbentuk, yaitu D_1 dan D_2 . Hasil pengujian menunjukkan bahwa kedua parameter tersebut signifikan, sehingga dapat disimpulkan bahwa variabel X_5 secara keseluruhan berpengaruh terhadap waktu kesembuhan pasien.

4.4.6 Hazard Ratio

Hazard Ratio (HR) merupakan ukuran yang digunakan dalam model regresi *Cox Proportional Hazard* untuk menggambarkan besarnya pengaruh variabel prediktor terhadap laju kejadian (*hazard*), dalam hal ini adalah kesembuhan pasien TB. Nilai HR berdasarkan Persamaan (2.33) diperoleh dari eksponensial koefisien regresi atau $\exp(\beta)$. Interpretasi *hazard ratio* adalah sebagai berikut:

1. Jika $HR > 1$, maka variabel tersebut meningkatkan laju kejadian (mempercepat kesembuhan).
2. Jika $HR < 1$, maka variabel tersebut menurunkan laju kejadian (memperlambat kesembuhan).
3. Jika $HR = 1$, maka variabel tidak berpengaruh terhadap *hazard*.

Tabel 4.14 Nilai *Hazard Ratio* Tiap Variabel pada Model Final

Variabel	Koefisien	HR
X_1	0.958	2.606
X_3	2.753	15.695
$D_{5,1}$	-1.737	0.176
$D_{5,2}$	-2.135	0.118
X_7	0.901	2.461

Berdasarkan Tabel 4.14, interpretasi masing-masing variabel adalah sebagai berikut:

1. Variabel X_1 memiliki nilai HR sebesar 2.606. Hal ini menunjukkan bahwa pasien pada kategori $X_1 = 1$ memiliki laju kesembuhan 2.60 kali lebih tinggi dibandingkan kategori referensi ($X_1 = 0$). Atau dengan kata lain, pasien berusia produktif (15-64 tahun) memiliki laju kesembuhan 2.60 kali lebih tinggi dibandingkan pasien dengan usia tidak produktif. Temuan ini sejalan dengan penelitian Antika dkk. (2025) yang menyatakan bahwa pasien usia produktif memiliki laju kesembuhan yang lebih tinggi dibandingkan usia tidak produktif, yang dipengaruhi oleh kondisi fisik yang lebih baik serta kemampuan menjalani pengobatan secara optimal.
2. Variabel X_3 memiliki nilai HR sebesar 15.695. Artinya, pasien pada kategori $X_3 = 1$ memiliki laju kesembuhan 15.69 kali lebih tinggi dibandingkan kategori referensi. Atau dengan kata lain, pasien dengan pengobatan sesuai standar memiliki laju kesembuhan 15.69 kali lebih tinggi dibandingkan dengan pasien dengan pengobatan yang tidak sesuai standar. Nilai ini menunjukkan pengaruh yang sangat kuat terhadap percepatan kesembuhan. Pengobatan sesuai standar yang dimaksud dalam penelitian ini adalah pengobatan yang mengikuti pedoman nasional, yaitu penggunaan paduan Obat Anti Tuberkulosis (OAT) yang tepat, dosis yang benar, diberikan secara teratur, diawasi oleh Pengawas Menelan Obat (PMO), serta dijalankan hingga tuntas dengan tingkat keberhasilan pengobatan lebih dari 85%. Kondisi ini menunjukkan kepatuhan terhadap standar terapi yang telah ditetapkan oleh Kementerian Kesehatan. Secara klinis, pengobatan yang sesuai standar

memiliki efektivitas yang lebih tinggi dalam membunuh bakteri penyebab tuberkulosis, sehingga mempercepat proses penyembuhan dan menurunkan risiko kegagalan pengobatan. Sebaliknya, pengobatan yang tidak sesuai standar dapat menyebabkan proses penyembuhan menjadi lebih lama dan berisiko menimbulkan resistensi obat.

3. Variabel $D_{5,1}$ dan $D_{5,2}$ merupakan variabel dummy yang merepresentasikan jenis fasilitas pelayanan kesehatan dengan kategori referensi puskesmas. Variabel $D_{5,1}$ memiliki nilai HR sebesar 0,176, yang menunjukkan bahwa pasien yang menjalani pengobatan di rumah sakit memiliki laju kesembuhan sebesar 0,176 kali dibandingkan pasien yang berobat di puskesmas. Dengan kata lain, laju kesembuhan pasien di rumah sakit sekitar 82,4% lebih rendah dibandingkan pasien di puskesmas. Sementara itu, variabel $D_{5,2}$ memiliki nilai HR sebesar 0.118, yang menunjukkan bahwa pasien yang menjalani pengobatan selain puskesmas dan rumah sakit memiliki laju kesembuhan sebesar 0.118 kali dibandingkan pasien yang berobat di puskesmas. Atau dengan kata lain, laju kesembuhan pasien di fasilitas layanan kesehatan selain rumah sakit dan puskesmas sekitar 82,4% lebih rendah dibandingkan pasien di puskesmas. Hasil ini menunjukkan bahwa puskesmas memiliki laju kesembuhan yang lebih tinggi dibandingkan fasilitas kesehatan lainnya. Secara teoritis, puskesmas merupakan fasilitas pelayanan kesehatan tingkat pertama yang memiliki akses lebih dekat dengan masyarakat. Temuan ini juga sejalan dengan penelitian Pratama (2021) yang menyatakan bahwa kepatuhan pasien TB meningkat ketika pelayanan kesehatan mudah diakses.

4. Variabel X_7 memiliki nilai HR sebesar 2.461. Hal ini menunjukkan bahwa pasien pada kategori $X_7 = 1$ memiliki laju kesembuhan 2.461 kali lebih tinggi dibandingkan kategori referensi. Atau dengan kata lain, pasien tanpa HIV memiliki laju kesembuhan 2.461 kali lebih tinggi dibandingkan pasien dengan HIV. Hasil ini sejalan dengan penelitian Laili dkk. (2024) yang menyatakan bahwa ko-infeksi TB-HIV dapat memperlambat proses penyembuhan serta meningkatkan risiko kegagalan pengobatan karena sistem imun pasien yang lebih lemah.

Secara keseluruhan, variabel X_1 (usia), X_3 (status pengobatan), dan X_7 (hasil tes HIV) berpengaruh terhadap peningkatan laju kesembuhan pasien TB, sedangkan variabel $D_{5,1}$ dan $D_{5,2}$ berpengaruh terhadap penurunan laju kesembuhan. Hasil ini menunjukkan bahwa setiap variabel prediktor memiliki arah pengaruh yang berbeda terhadap laju kesembuhan pasien, sehingga faktor klinis, demografis, maupun karakteristik pelayanan kesehatan berperan dalam menentukan peningkatan atau penurunan proses kesembuhan pasien tuberkulosis. Temuan ini dapat menjadi dasar dalam penyusunan strategi penanganan pasien TB yang lebih terarah berdasarkan karakteristik masing-masing pasien.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah dan pembahasan yang telah dijabarkan pada bab sebelumnya, maka diperoleh kesimpulan dari penelitian ini sebagai berikut:

1. Hasil analisis *Kaplan-Meier* menunjukkan bahwa terdapat perbedaan pola kurva *survival* pada beberapa variabel atau faktor-faktor yang berpengaruh, yang mengindikasikan adanya perbedaan waktu kesembuhan antar kategori pasien TB. Hasil dikuatkan berdasarkan uji *log-rank*, yang menunjukkan bahwa variabel usia (X_1), status pengobatan (X_3), riwayat pengobatan (X_4), jenis fasyankes (X_5), tipe diagnosis TB (X_6), hasil tes HIV (X_7), dan hasil tes DM (X_8) memiliki perbedaan kurva *survival* yang signifikan antar kategori. Hal ini menunjukkan bahwa variabel-variabel tersebut berpengaruh terhadap perbedaan waktu kesembuhan pasien TB. Sedangkan variabel jenis kelamin (X_2) tidak menunjukkan perbedaan yang signifikan antar kategori.
2. Model regresi *Cox Proportional Hazard* dari faktor-faktor yang diduga memengaruhi laju kesembuhan tuberkulosis di Kota Probolinggo pada tahun 2021-2023 adalah:

$$h(t, X) = h_0(t) \exp (0.9576X_1 + 2.7533X_3 - 1.7367D_{5,1} - 2.1347D_{5,2} + 0.9005X_7)$$

dan berdasarkan *Hazard Ratio* (HR), diperoleh kesimpulan bahwa:

- a. Pasien berusia produktif (15-64 tahun) memiliki laju kesembuhan 2.60 kali lebih tinggi dibandingkan pasien dengan usia tidak produktif.
- b. Pasien dengan pengobatan sesuai standar memiliki laju kesembuhan 15.69 kali lebih tinggi dibandingkan pasien dengan pengobatan yang tidak sesuai

standar. Pengobatan sesuai standar yang dimaksud sesuai pedoman nasional tuberculosis berupa penggunaan Obat Anti Tuberkulosis (OAT) yang tepat, dosis yang benar, diminum secara teratur, serta diawasi oleh Pengawas Menelan Obat (PMO) hingga tuntas.

- c. Pasien dengan jenis fasilitas layanan kesehatan di rumah sakit memiliki laju kesembuhan sebesar 0.176 kali dibandingkan di puskesmas. Selain itu juga, pasien dengan jenis fasyankes selain puskesmas dan rumah sakit memiliki laju kesembuhan sebesar 0.118 kali dibandingkan di puskesmas.
- d. Pasien tanpa HIV memiliki laju kesembuhan 2.46 kali lebih tinggi dibandingkan pasien dengan HIV.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, disarankan kepada peneliti selanjutnya untuk dapat mengembangkan analisis lebih lanjut, khususnya dalam mengeksplorasi pendekatan *survival* yang lebih fleksibel seperti model Cox stratifikasi atau model *Cox extended*, guna memperoleh pemahaman yang lebih mendalam terhadap kemungkinan pengaruh kovariat yang berubah terhadap waktu. Selain itu, penelitian selanjutnya juga diharapkan dapat mengeksplorasi pendekatan metodologis lain serta mempertimbangkan variabel tambahan yang relevan, sehingga model yang dihasilkan menjadi lebih komprehensif dan mampu memberikan gambaran yang lebih luas terkait faktor-faktor yang memengaruhi waktu kesembuhan pasien TB.

DAFTAR PUSTAKA

- Abdullah, M. (2007). *Tafsir Ibnu Katsir Jilid 7*. Bogor: Pustaka Imam Asy-Syafi'i.
- Adawiyah, R. N., Akaputra, R., W, M. R., & Fachri, M. (2023). Faktor – Faktor yang Mempengaruhi Lama Waktu Pengobatan Tuberkulosis Paru Klinis di Rumah Sakit Umum Daerah Pasar Minggu Tahun 2019 - 2023. *Prosiding Seminar Nasional Penelitian LPPM UMJ*, 1–13.
- Alim, M. K., & Maulana, D. A. (2025). *Prediksi Kekeringan Ekstrem di Jawa Timur Menggunakan Generalized Pareto Distribution (GPD)*. 13(02), 597–606.
- Antika, B. R., Perdana, A. A., & Muhani, N. (2025). Analisis Survival Lama Keberhasilan Pengobatan Pasien Tuberkulosis di Puskesmas Hajimena Lampung Selatan 2021-2023. *Jurnal Dunia Kesmas*, 14(1), 81–91. <https://doi.org/10.33024/jdk.v14i1.16817>
- As'ari, A. W. (2012). *Pendekatan Regresi Cox Proporsional Hazard dalam Penentuan Faktor – Faktor yang Berpengaruh terhadap Lama Studi Mahasiswa S-1 Matematika di Universitas Airlangga*. Universitas Airlangga.
- Avy, A. H., Hutami, B. P., Alfalah, M. Z., & Febriyanti, S. (2024). Faktor Risiko Kejadian Tuberkulosis Paru di Berbagai Wilayah Indonesia. *Indonesia Journal Chest*, 11(1), 61–65.
- Collett, D. (2015). *Modelling Survival Data in Medical Research* (3rd ed.). Chapman and Hall/CRC. <https://www.taylorfrancis.com/books/9781135441449>
- Dahlan, M. S. (2012). *Analisis Survival: Dasar-Dasar Teori & Aplikasi*. Epidemiologi Indonesia.
- Departemen Agama RI. (2009). *Al-Qur'an dan Tafsirnya: Jilid 7 (Juz 19, 20, 21)*. Departemen Agama Republik Indonesia
- Faisal, A. R., Bustan, M. N., & Annas, S. (2020). Analisis Survival dengan Pemodelan Regresi Cox Proportional Hazard Menggunakan Pendekatan Bayesian (Studi Kasus: Pasien Rawat Inap Penderita Demam Tifoid di RSUD Haji Makassar). *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, 2(2), 62. <https://doi.org/10.35580/variansiunm14629>
- Fauzan, D. S. bin F. bin A. Al. (2011). *Pembahasan Kitab Tauhid*. Al-Qur'an dan Sunnah. <https://alquran-sunnah.com/artikel/pilihan/kitab-tauhid/kitab-tauhid-bab-1/668-12-ibadah-pengertian-macam-dan-keluasan-cakupannya.html>
- Fitri, A., Rahim, R., Nurhayati, Natsir, I., Pagiling, A. S. L., Simanjuntak, A. M. D. N., Hutagaol, K., & Anugrah, N. E. (2023). *Dasar-Dasar Statistika untuk Penelitian*. Yayasan Kita Menulis.
- Hari, N. C., Komalig, H., & Langi, Y. (2018). Analisis Survival Dalam Menentukan Faktor-faktor Yang Mempengaruhi Lama Studi Mahasiswa Matematika Di

- Jurusan Matematika FMIPA Universitas Sam Ratulangi Manado. *d'CARTESIAN*, 7(2). <https://doi.org/10.35799/dc.7.2.2018.21455>
- Harlan, J. (2017). Analisis Survival. In *Gunadarma*. Penerbit Gunadarma.
- Heymans, M. W., & Twisk, J. W. R. (2022). Handling Missing Data in Clinical Research. *Journal of Clinical Epidemiology*, 151, 185–188. <https://doi.org/10.1016/j.jclinepi.2022.08.016>
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). John Wiley and Sons, Inc.
- Kementerian Agama RI. (2019). *Al-Qur'an dan Terjemahannya*. Lajnah Pentashihan Mushaf Al-Qur'an.
- Kementerian Kesehatan. (2025). *Buku Panduan Tenaga Medis dan Tenaga Kesehatan Tuberkulosis*.
- Khairunnisa, F., Saumi, F., & Amelia, A. (2022). Survival Analysis of Dengue Hemorrhagic Fever Patients (Dhf). *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 16(3), 897–908. <https://doi.org/10.30598/barekengvol16iss3pp897-908>
- Klein, J. P., & Moeschberger, M. L. (2003). *SURVIVAL ANALYSIS Techniques for Censored and Truncated Data* (2nd ed.). Springer. <https://dl.acm.org/doi/10.1145/390011.808243>
- Kleinbaum, D. G., & Klein, M. (2012). Survival Analysis A Self-Learning Text. In *Statistics for Biology and Health* (3rd ed.). Springer Science. <https://actaojs.org.ar/ojs/index.php/acta/article/view/164>
- Laili, F., Ronoatmodjo, S., & Murtiani, F. (2024). Ko-Infeksi TB-HIV terhadap Kegagalan Pengobatan Pasien Tuberkulosis Resistan Obat di Indonesia. *The Indonesian Journal of Infectious Disease*, 10(2), 153–165.
- Little, R. J. A., & Rubin, D. B. (2020). *Statistical Analysis with Missing Data* (3rd ed.). Wiley.
- Nugraha, I. B. A., Gotera, W., & Yustin, W. E. F. (2021). Diabetes Melitus sebagai Faktor Risiko Tuberkulosis. *Jurnal Kedokteran Meditek*, 27(3), 273–281.
- Nurfajilah, Saumi, F., Sari, R. P., & Amelia. (2023). Analisis Ketahanan Hidup Penderita Tuberkulosis di RSUD Dr. Zainoel Abidin. *Gamma-Pi : Jurnal Matematika dan Terapan*, 5(2), 1–11.
- Pramudita, C., Aiswarya, N., Wicaksono, A., & Kharisma, A. (2025). Menangani Missing Value dengan Metode Imputasi Menggunakan Package MICE di R. *PRISMA, Prosiding Seminar Nasional Matematika*, 8, 399–406.
- Pratama, F. A. V. R. (2021). Faktor Sarana/Fasilitas Kesehatan Mempengaruhi Perubahan Perilaku Pasien TBC. In *OSF Preprints*. <https://doi.org/https://doi.org/10.31219/osf.io/nm69y>
- Rahem, A., Utami, W., Hermansyah, A., Sukorini, A. I., Rahayu, T. P., Athiyah,

- U., Yuda, A., Sulistyarini, A., Wijaya, I. N., Noorrizka V, G., Priyandani, Y., Budiadin, A. S., Aryani, T., & Zulkamain, B. S. (2024). Peningkatan Pemahaman Kelompok Tani Terhadap Pengobatan Tuberculosis Dan Pencegahannya Melalui Focus Group Discussion. *Jiis (Jurnal Ilmiah Ibnu Sina): Ilmu Farmasi dan Kesehatan*, 9(1), 170–178. <https://doi.org/10.36387/jiis.v9i1.1727>
- Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Canada: Library of Congress Cataloging in Publication.
- Safitri, W., Wuryandari, T., & Suparti. (2016). Analisis Ketahanan Hidup Penderita Tuberculosis dengan Menggunakan Metode Regresi Cox Kegagalan Proporsional (Studi Kasus di Puskesmas Kecamatan Kembangan Jakarta Barat). *Jurnal Gaussian*, 5(4), 781–790.
- Shihab, M. Q. (2002). *Tafsir Al-Mishbah: Pesan, Kesan dan Keserasian Al-Qur'an*. Lentera Hati.
- T.Lee, E., & Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis* (3rd ed.). John Wiley & Sons, Inc., Hoboken, New Jersey. <https://doi.org/10.1002/0471458546>
- Utami, R. S., & Danardono. (2019). Metode Multiple Imputation Untuk Mengatasi Kovariat Tak lengkap pada Data Kejadian Berulang. *Journal Of Fundamental Mathematics*, 2(2), 47–57.
- Widyastuti, R., Wulandari, D., & Prasetyowati, D. (2024). Penerapan Regresi Cox Untuk Menganalisis Variabel yang Berpengaruh Terhadap Durasi Studi Mahasiswa. *Jurnal Gaussian*, 13(1), 88–98. <https://doi.org/10.14710/j.gauss.13.1.88-98>
- Zukhaila, N. (2016). *Analisis Model Persamaan Regresi Cox Proportional Hazard pada Data Status Gizi Balita untuk Mengetahui Faktor-Faktor Penyebab Terjadinya Kekurangan Gizi*. Universitas Negeri Semarang.

LAMPIRAN

Lampiran 1. Data Penelitian

No	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	Status	Tanggal Mulai	Tanggal Selesai
1	37	L	Sesuai standar	Diobati setelah putus berobat	Puskesmas	Terdiagnosis klinis	Bukan ODHIV	Negatif	Meninggal	02/01/2021	20/06/2021
2	25	L	Sesuai standar	Baru	Puskesmas	Terdiagnosis klinis	Bukan ODHIV	Positif	Pengobatan Lengkap	02/01/2021	17/06/2021
3	17	L	Sesuai standar	Baru	Puskesmas	Terdiagnosis klinis	Bukan ODHIV	Negatif	Pengobatan Lengkap	17/06/2021	30/11/2021
4	37	L	Sesuai standar	Baru	Puskesmas	Terdiagnosis klinis	Bukan ODHIV	Negatif	Pengobatan Lengkap	04/01/2021	19/06/2021
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
16	75	L	Sesuai standar	Baru	Puskesmas	Terkonfirmasi bakteriologis	NA	NA	Meninggal	23/03/2021	09/04/2021
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
55	39	L	Sesuai standar	Baru	Puskesmas	Terdiagnosis klinis	NA	NA	Putus berobat (lost to follow up)	03/03/2021	19/05/2021
56	30	L	Sesuai standar	Baru	Puskesmas	Terdiagnosis klinis	Bukan ODHIV	NA	Pengobatan Lengkap	02/01/2021	18/06/2021
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
154	38	P	Sesuai standar	NA	Puskesmas	Terkonfirmasi bakteriologis	Bukan ODHIV	Positif	Sembuh	10/07/2021	23/12/2021
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1365	11	P	Sesuai standar	Baru	Rumah Sakit	Terdiagnosis klinis	Bukan ODHIV	Negatif	Putus berobat (lost to follow up)	01/11/2023	31/01/2024

Lampiran 2. Model Logistik X_4 pada Setiap Dataset Imputasi

Variabel	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
Intercept	19.42281	19.3045	19.23122	19.4394	19.13524
X_1	-1.31334	-1.3689	-1.31038	-1.3049	-1.32153
X_2	-0.13455	-0.2010	-0.05740	-0.1255	-0.16441
X_3	-14.27782	14.3256	-14.3310	-14.3175	-14.3003
D_1 (Rumah Sakit)	-0.15026	-0.1597	-0.08733	-0.1191	-0.07496
D_2 (Faskes Lainnya)	14.36266	14.6779	14.64149	14.6433	14.59837
X_6	-0.11374	-0.1844	-0.04296	-0.1374	-0.07340
X_7	-0.76818	-0.3167	-0.64004	-0.6168	-0.57721
X_8	-0.08934	-0.3118	-0.05662	-0.2429	0.04126

Lampiran 3. Model Logistik X_7 pada Setiap Dataset Imputasi

Variabel	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
Intercept	4.78050	5.02991	5.1445	5.5999	5.1292
X_1	-0.76020	-0.83569	-0.7662	-0.6233	-0.9239
X_2	-0.66700	-0.76252	-0.6592	-0.8478	-1.0468
X_3	0.65750	0.73167	0.2764	-0.1699	0.8551
X_4	-0.55850	-0.09818	-0.4609	-0.4062	-0.4478
D_1 (Rumah Sakit)	-1.54190	-1.42869	-1.0798	-1.2349	-1.2029
D_2 (Faskes Lainnya)	-3.00400	13.62763	13.7769	13.7115	-0.5352
X_6	1.25320	1.15894	1.3647	1.4674	1.0392
X_8	-1.15300	-1.75044	-1.3828	-1.3642	-1.3025

Lampiran 4. Model Logistik X_8 pada Setiap Dataset Imputasi

Variabel	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
Intercept	2.66088	3.5062	3.55274	3.6648	3.54574
X_1	-0.33205	-0.3614	-0.58722	-0.4465	-0.38544
X_2	0.23079	0.1747	0.21173	0.1138	0.13994
X_3	0.07992	0.1775	-0.29128	-0.3209	-0.62053
X_4	-0.07326	-0.2975	-0.04477	-0.2306	0.05188
D_1 (Rumah Sakit)	0.21394	0.2124	0.31002	0.2316	0.35745
D_2 (Faskes Lainnya)	0.65451	1.0760	1.07023	1.0486	1.00243
X_6	-0.31499	-0.4562	-0.44478	-0.4405	-0.39315
X_7	-1.13733	-1.6975	-1.33792	-1.2958	-1.26886

Lampiran 5. Hasil Imputasi Variabel X_4

	Baris	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
1	153	1	1	1	1	1
2	163	1	1	1	1	1
3	170	1	1	1	1	1
4	175	1	1	1	1	0
5	203	1	1	1	1	1
6	285	1	1	1	1	1
7	321	1	1	1	1	1
8	335	1	0	1	1	0
:	:	:	:	:	:	:
62	1364	1	1	1	1	1

Lampiran 6. Hasil Imputasi Variabel X_7

	Baris	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
1	14	1	0	1	1	1
2	15	1	1	1	1	1
3	49	1	1	1	1	1
4	54	1	0	0	1	1
5	81	1	1	1	1	1
6	86	1	1	1	1	1
7	89	1	1	1	1	1
8	95	1	1	1	1	1
:	:	:	:	:	:	:
170	1363	0	1	1	1	1

Lampiran 7. Hasil Imputasi Variabel X_8

	Baris	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
1	15	1	0	1	1	0
2	18	1	1	1	1	1
3	47	1	1	0	1	1
4	48	1	1	0	1	1
5	54	1	1	0	1	1
6	55	1	0	1	1	1
7	65	1	1	1	0	1
8	81	0	0	1	0	1
:	:	:	:	:	:	:
141	1352	0	1	1	1	1

Lampiran 8. Estimasi Kaplan-Meier Berdasarkan Usia
 (a) Usia Produktif (b) Usia Tidak Produktif

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1159	0	1.0000
150	988	1	0.9989
161	981	2	0.9959
162	979	1	0.9949
163	975	2	0.9928
164	972	2	0.9908
165	969	67	0.9223
166	817	14	0.8433
⋮	⋮	⋮	⋮
575	1	0	0.1311

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	206	0	1.0000
165	156	1	0.9935
166	142	8	0.9376
167	126	3	0.9152
168	81	6	0.8474
169	54	2	0.8161
179	32	1	0.7905
180	29	1	0.7633
⋮	⋮	⋮	⋮
497	1	0	0.6368

Lampiran 9. Estimasi Kaplan-Meier Berdasarkan Jenis Kelamin
 (a) Perempuan (b) Laki-laki

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	620	0	1.0000
161	530	2	0.9962
162	528	1	0.9943
163	525	1	0.9924
164	524	1	0.9905
165	523	33	0.9280
166	443	33	0.8589
167	382	58	0.7285
⋮	⋮	⋮	⋮
532	1	0	0.2211

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	745	0	1.0000
150	614	1	0.9983
154	613	1	0.9967
163	607	1	0.9951
164	605	1	0.9934
165	602	35	0.9356
166	516	45	0.8540
167	422	74	0.7043
⋮	⋮	⋮	⋮
575	1	0	0.1450

Lampiran 10. Estimasi Kaplan-Meier Berdasarkan Status Pengobatan
 (a) Tidak Sesuai Standar (b) Sesuai Standar

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	31	0	1.0000
223	5	1	0.8000
326	1	0	0.8000

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1334	0	1.0000
150	1121	1	0.9991
154	1119	1	0.9982
161	1114	2	0.9964
162	1112	1	0.9955
163	1108	2	0.9937
164	1105	2	0.9919
165	1101	68	0.9306
⋮	⋮	⋮	⋮
575	1	0	0.1635

Lampiran 11. Estimasi Kaplan-Meier Berdasarkan Riwayat Pengobatan
 (a) Pernah Diobati Sebelumnya (b) Kasus Baru

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	64	0	1.0000
161	56	1	0.9821
162	55	1	0.9642
166	52	1	0.9457
167	48	2	0.9063
168	34	3	0.8263
169	26	3	0.7310
170	21	1	0.6962
⋮	⋮	⋮	⋮
532	1	0	0.2498

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1301	0	1.0000
161	1082	1	0.9972
163	1078	2	0.9953
164	1075	2	0.9935
165	1071	68	0.9304
166	907	77	0.8514
167	756	130	0.7050
168	482	90	0.5734
⋮	⋮	⋮	⋮
575	1	0	0.1678

Lampiran 12. Estimasi Kaplan-Meier Berdasarkan Jenis Fasyankes
 (a) Puskesmas (b) Rumah Sakit

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1093	0	1.0000
150	981	1	0.9989
161	978	2	0.9969
162	976	1	0.9959
163	973	2	0.9938
164	970	2	0.9918
165	967	68	0.9920
166	808	74	0.8376
⋮	⋮	⋮	⋮
534	1	0	0.7967

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	261	0	1.0000
154	152	1	0.9934
166	140	4	0.9650
167	131	1	0.9576
168	105	2	0.9394
169	85	1	0.9283
172	71	1	0.9153
173	70	1	0.9022
⋮	⋮	⋮	⋮
575	1	0	0.6429

(c) Fasyankes Lainnya

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
167	11	1	0.9090
168	6	0	0.9090
170	5	0	0.9090
173	4	0	0.9090
174	3	0	0.9090
179	2	0	0.9090
199	1	0	0.9090

Lampiran 13. Estimasi *Kaplan-Meier* Berdasarkan Tipe Diagnosis
(a) Terdiagnosis Klinis (b) Terkonfirmasi Bakteriologis

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	626	0	1.0000
154	470	1	0.9978
165	462	2	0.9935
166	401	2	0.9885
167	340	8	0.9653
168	206	1	0.9606
169	137	1	0.9536
170	114	1	0.9452
⋮	⋮	⋮	⋮
575	1	0	0.6314

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	739	0	1.0000
150	674	1	0.9985
161	671	2	0.9955
162	669	1	0.9940
163	668	2	0.9910
164	666	2	0.9880
165	663	66	0.8897
166	558	76	0.7685
⋮	⋮	⋮	⋮
532	1	0	0.0604

Lampiran 14. Estimasi *Kaplan-Meier* Berdasarkan Hasil Tes HIV
(a) Positif (b) Negatif

Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	88	0	1.0000
167	44	3	0.9318
168	30	2	0.8696
170	20	1	0.8262
251	6	1	0.6885
337	4	1	0.5163
⋮	⋮	⋮	⋮
454	1	0	0.5163

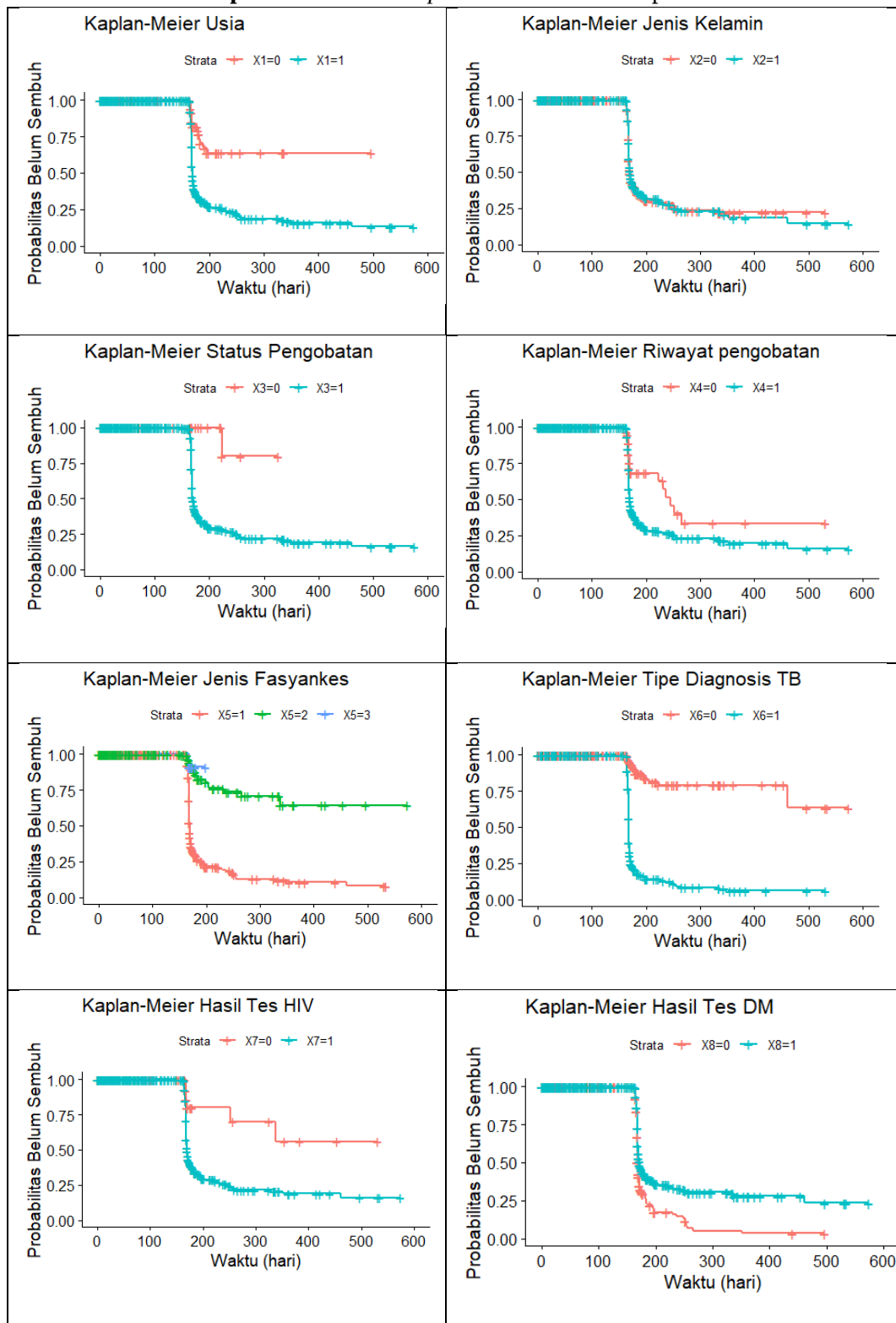
Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1277	0	1.0000
150	1095	1	0.9990
154	1093	1	0.9981
161	1089	2	0.9963
162	1087	1	0.9954
163	1083	2	0.9935
164	1080	2	0.9917
165	1076	68	0.9290
⋮	⋮	⋮	⋮
575	1	0	0.1640

Lampiran 15. Estimasi *Kaplan-Meier* Berdasarkan Hasil Tes DM
(a) Positif (b) Negatif

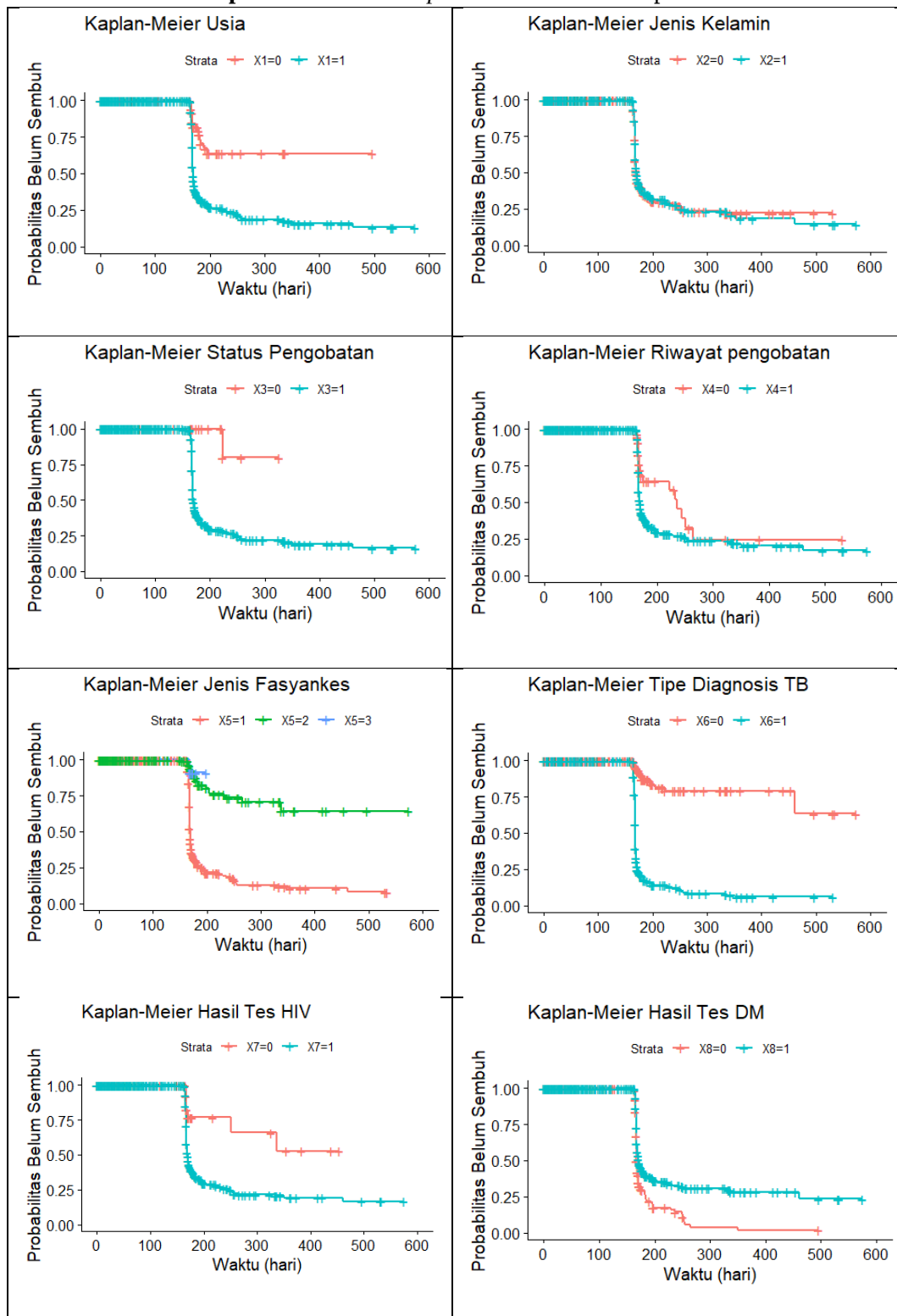
Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	299	0	1.0000
150	258	1	0.9961
163	257	1	0.9922
165	255	15	0.9338
166	225	20	0.8508
167	188	34	0.6969
168	129	30	0.5348
169	89	13	0.4567
⋮	⋮	⋮	⋮
497	1	0	0.0458

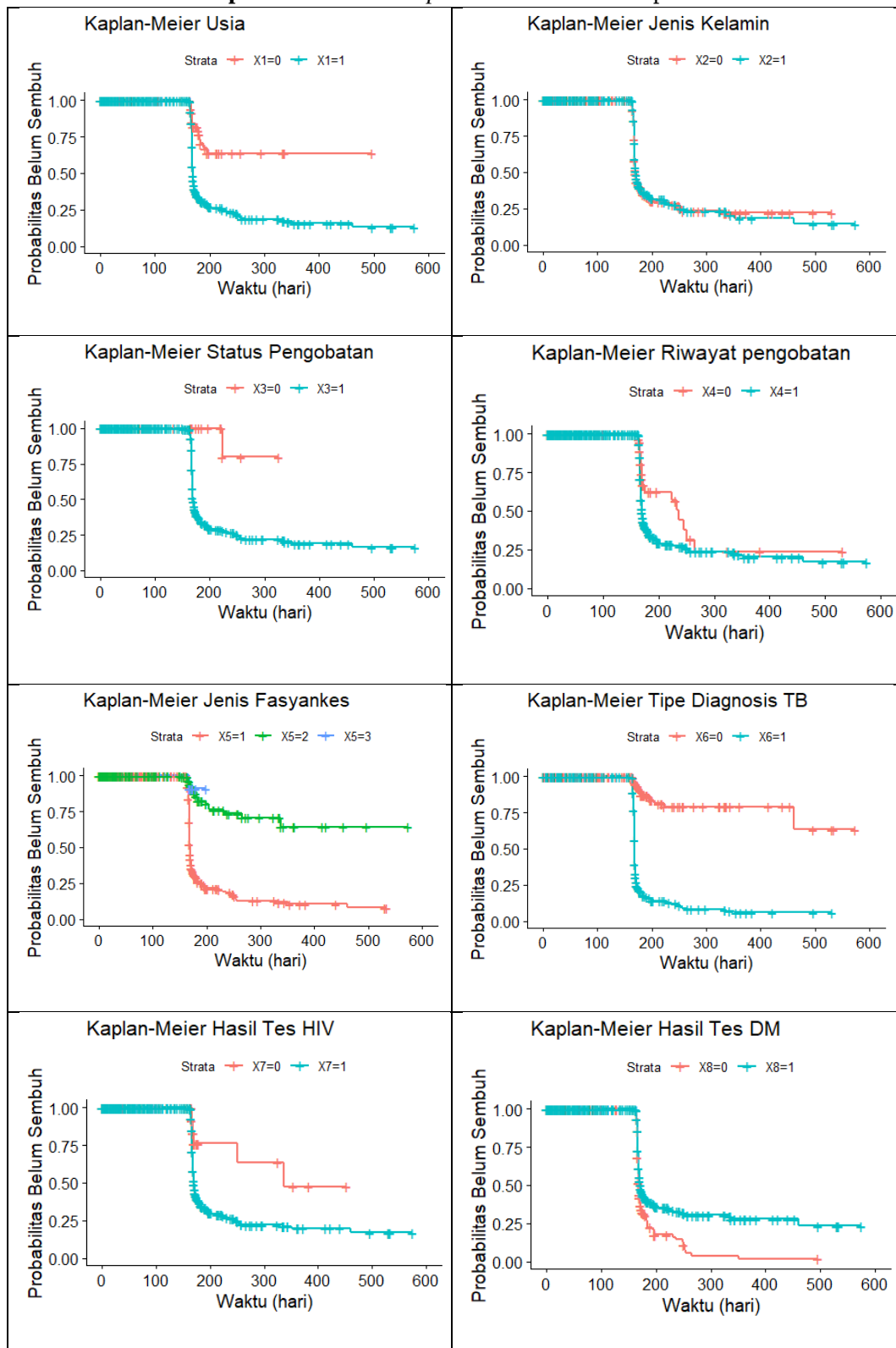
Waktu (t)	Jumlah Berisiko (b_i)	Jumlah Event (d_i)	Survival $S(t)$
0	1066	0	1.0000
154	886	1	0.9988
161	881	2	0.9966
162	879	1	0.9954
163	875	1	0.9943
164	873	2	0.9920
165	870	53	0.9316
⋮	⋮	⋮	⋮
575	1	0	0.2258

Lampiran 16. Kurva *Kaplan Meier* Dataset Imputasi 2

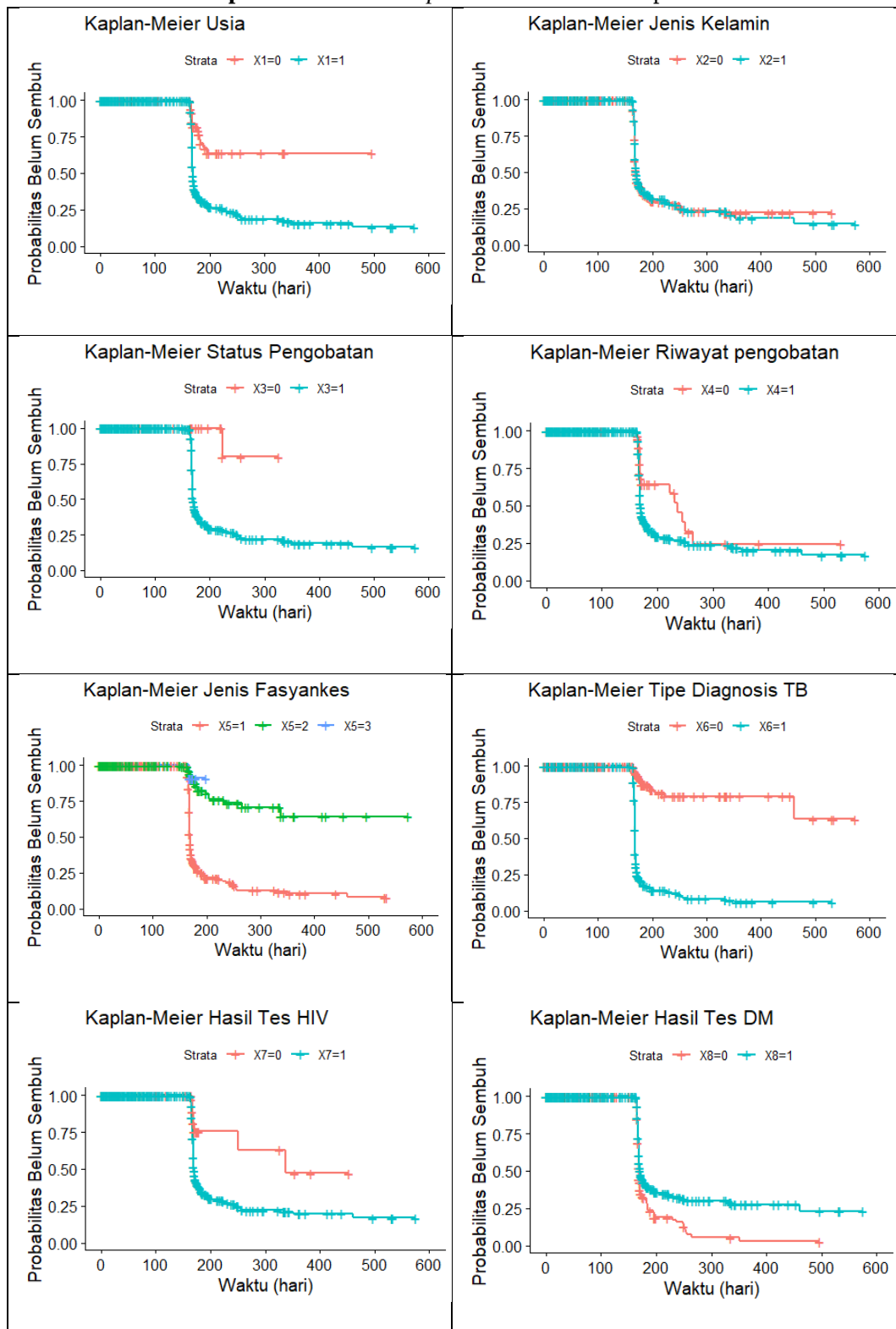


Lampiran 17. Kurva Kaplan-Meier Dataset Imputasi 3



Lampiran 18. Kurva *Kaplan-Meier* Dataset Imputasi 4

Lampiran 19. Kurva *Kaplan-Meier* Dataset Imputasi 5



Lampiran 20. Hasil Uji *Log-rank* Dataset Imputasi 2

Variabel	χ^2 hitung	χ^2 tabel	df	<i>p-value</i>	Keputusan
X_1	53.3405	3.8415	1	0.0000	Tolak H_0
X_2	0.0068	3.8415	1	0.9342	Gagal Tolak H_0
X_3	18.5661	3.8415	1	0.0000	Tolak H_0
X_4	11.1751	3.8415	1	0.0008	Tolak H_0
X_5	115.7570	5.9915	2	0.0000	Tolak H_0
X_6	329.0837	3.8415	1	0.0000	Tolak H_0
X_7	16.5696	3.8415	1	0.0000	Tolak H_0
X_8	19.2735	3.8415	1	0.0000	Tolak H_0

Lampiran 21. Hasil Uji *Log-rank* Dataset Imputasi 3

Variabel	χ^2 hitung	χ^2 tabel	df	<i>p-value</i>	Keputusan
X_1	53.3405	3.8415	1	0.0000	Tolak H_0
X_2	0.0068	3.8415	1	0.9342	Gagal Tolak H_0
X_3	18.5661	3.8415	1	0.0000	Tolak H_0
X_4	8.6129	3.8415	1	0.0033	Tolak H_0
X_5	115.7570	5.9915	2	0.0000	Tolak H_0
X_6	329.0837	3.8415	1	0.0000	Tolak H_0
X_7	14.3315	3.8415	1	0.0002	Tolak H_0
X_8	20.5609	3.8415	1	0.0000	Tolak H_0

Lampiran 22. Hasil Uji *Log-rank* Dataset Imputasi 4

Variabel	χ^2 hitung	χ^2 tabel	df	<i>p-value</i>	Keputusan
X_1	53.3405	3.8415	1	0.0000	Tolak H_0
X_2	0.0068	3.8415	1	0.9342	Gagal Tolak H_0
X_3	18.5661	3.8415	1	0.0000	Tolak H_0
X_4	7.4359	3.8415	1	0.0064	Tolak H_0
X_5	115.7570	5.9915	2	0.0000	Tolak H_0
X_6	329.0837	3.8415	1	0.0000	Tolak H_0
X_7	11.8217	3.8415	1	0.0006	Tolak H_0
X_8	14.7705	3.8415	1	0.0001	Tolak H_0

Lampiran 23. Hasil Uji *Log-rank* Dataset Imputasi 5

Variabel	χ^2 hitung	χ^2 tabel	df	<i>p-value</i>	Keputusan
X_1	53.3405	3.8415	1	0.0000	Tolak H_0
X_2	0.0068	3.8415	1	0.9342	Gagal Tolak H_0
X_3	18.5661	3.8415	1	0.0000	Tolak H_0
X_4	7.5256	3.8415	1	0.0061	Tolak H_0
X_5	115.7570	5.9915	2	0.0000	Tolak H_0
X_6	329.0837	3.8415	1	0.0000	Tolak H_0
X_7	10.7723	3.8415	1	0.0010	Tolak H_0
X_8	12.8263	3.8415	1	0.0003	Tolak H_0

Lampiran 24. Hasil Estimasi Parameter Model Kedua

Variabel	Imputasi 1	Imputasi 2	Imputasi 3	Imputasi 4	Imputasi 5
X_1	0.986	0.987	0.966	0.974	0.988
X_3	2.771	2.775	2.773	2.779	2.762
X_4	0.748	0.773	0.740	0.675	0.693
$D_{5,1}$	-1.684	-1.688	-1.709	-1.698	-1.707
$D_{5,2}$	-1.744	-2.230	-2.229	-2.219	-2.153
X_7	0.911	0.947	0.964	0.782	0.714
X_8	-0.108	-0.192	-0.214	-0.169	-0.124

Lampiran 25. Script R

```
#LANGKAH 1: PREPOCESSING DATA#
#1.IDENTIFIKASI MISSING VALUE#
data <- read.csv("C:/Users/USER/Downloads/skripsi tbc/SIAP OLAH (ADA MISSING).csv", sep = ";")
str(data)
summary(data)
data$X1 <- as.factor(data$X1)
data$X2 <- as.factor(data$X2)
data$X3 <- as.factor(data$X3)
data$X4 <- as.factor(data$X4)
data$X5 <- as.factor(data$X5)
data$X6 <- as.factor(data$X6)
data$X7 <- as.factor(data$X7)
data$X8 <- as.factor(data$X8)
data$Status <- as.factor(data$Status)
# Jumlah missing
colSums(is.na(data))
# Persentase missing
round(colSums(is.na(data)) / nrow(data) * 100, 2)
#ambil variabel kategorik
var_kat <- c("X1","X2","X3","X4","X5","X6","X7","X8","Status")
var_missing <- var_kat[colSums(is.na(data[var_kat])) > 0]
var_missing
#fungsi uji chi square/fisher otomatis
uji_missing <- function(data, var_miss, var_lain) {
  miss_indicator <- as.factor(is.na(data[[var_miss]]))
  pembanding <- as.factor(data[[var_lain]])
  tab <- table(miss_indicator, pembanding)
  if (nrow(tab) < 2 || ncol(tab) < 2) {
    return(NULL)}
  chi <- suppressWarnings(chisq.test(tab))
  expected <- chi$expected
#cek asumsi
  prop_small <- sum(expected < 5) / length(expected)
  if (prop_small > 0.2 || any(expected < 1)) {
    test <- fisher.test(tab)
    metode <- "Fisher"
    pval <- test$p.value}
  else {
    metode <- "Chi-Square"
    pval <- chi$p.value}
  return(list(
    hasil = data.frame(
      Variabel_Missing = var_miss,
      Variabel_Pembanding = var_lain,
      Metode = metode,
      P_value = round(pval, 4)),
    expected = expected)))
var_kat_pembanding <- setdiff(var_kat, c("X4", "X7", "X8"))
#loop semua variabel
hasil <- data.frame()
expected_list <- list()
for (vm in var_missing) {
  for (vl in var_kat_pembanding) {
    if (vm != vl) {
      res <- uji_missing(data, vm, vl)
      if (!is.null(res)) {
        hasil <- rbind(hasil, res$hasil)
        expected_list[[paste(vm, "vs", vl)]] <- res$expected}}}}
hasil
```

```

#interpretasi otomatis
hasil$Keputusan <- ifelse(hasil$P_value < 0.05,
                          "Ada hubungan",
                          "Tidak ada hubungan")

hasil
expected_list
# =====
# MICE (MULTIPLE IMPUTATION)
# =====
library(mice)
method <- make.method(data)
# Variabel biner → logistic regression
method["X4"] <- "logreg"
method["X7"] <- "logreg"
method["X8"] <- "logreg"
predictorMatrix <- make.predictorMatrix(data)
imp <- mice(data,
            m = 5,          # jumlah dataset imputasi
            maxit = 20,     # jumlah iterasi (sesuai teori 20–30)
            method = method,
            predictorMatrix = predictorMatrix,
            seed = 123)
summary(imp)
#CEK KONVERGENSI
plot(imp) # trace plot (WAJIB untuk bukti konvergen)

data1 <- complete(imp, 1)
data2 <- complete(imp, 2)
data3 <- complete(imp, 3)
data4 <- complete(imp, 4)
data5 <- complete(imp, 5)
write.csv(data1, "C:/Users/USER/Downloads/data_imputasi_1.csv", row.names = FALSE)
write.csv(data2, "C:/Users/USER/Downloads/data_imputasi_2.csv", row.names = FALSE)
write.csv(data3, "C:/Users/USER/Downloads/data_imputasi_3.csv", row.names = FALSE)
write.csv(data4, "C:/Users/USER/Downloads/data_imputasi_4.csv", row.names = FALSE)
write.csv(data5, "C:/Users/USER/Downloads/data_imputasi_5.csv", row.names = FALSE)
#TABEL PERBANDINGAN IMPUTASI
proporsi <- function(x) {
  round(prop.table(table(x)), 3)}
var_imputasi <- c("X4", "X7", "X8")
hasil_all <- list()
for (v in var_imputasi) {
  # proporsi data asli
  asli <- proporsi(data[[v]])
  # proporsi tiap imputasi
  imp_list <- list(
    Imp1 = proporsi(data1[[v]]),
    Imp2 = proporsi(data2[[v]]),
    Imp3 = proporsi(data3[[v]]),
    Imp4 = proporsi(data4[[v]]),
    Imp5 = proporsi(data5[[v]]))
  # gabungkan semua kategori
  semua_level <- unique(unlist(lapply(c(list(asli), imp_list), names)))
  df <- data.frame(Kategori = semua_level)
  # isi data asli
  df$Asli <- asli[semua_level]
  df$Asli[is.na(df$Asli)] <- 0
  # isi semua imputasi
  for (i in 1:5) {
    temp <- imp_list[[i]][semua_level]
    temp[is.na(temp)] <- 0
  }
}

```

```

    df[[paste0("Imputasi_", i)]] <- temp}
df$Variabel <- v
hasil_all[[v]] <- df}
# Gabungkan semua variabel
tabel_imputasi <- do.call(rbind, hasil_all)
tabel_imputasi
# Simpan tabel imputasi
write.csv(tabel_imputasi,
          "C:/Users/USER/Downloads/tabel_perbandingan_imputasi.csv",
          row.names = FALSE)
# ambil index NA
idx_na_X4 <- which(is.na(data$X4))
# ambil nilai imputasi
hasil_X4 <- data.frame(
  ID = idx_na_X4,
  Imp1 = data1$X4[idx_na_X4],
  Imp2 = data2$X4[idx_na_X4],
  Imp3 = data3$X4[idx_na_X4],
  Imp4 = data4$X4[idx_na_X4],
  Imp5 = data5$X4[idx_na_X4])
hasil_X4
idx_na_X7 <- which(is.na(data$X7))
hasil_X7 <- data.frame(
  ID = idx_na_X7,
  Imp1 = data1$X7[idx_na_X7],
  Imp2 = data2$X7[idx_na_X7],
  Imp3 = data3$X7[idx_na_X7],
  Imp4 = data4$X7[idx_na_X7],
  Imp5 = data5$X7[idx_na_X7])
hasil_X7
tail(hasil_X7)
idx_na_X8 <- which(is.na(data$X8))
hasil_X8 <- data.frame(
  ID = idx_na_X8,
  Imp1 = data1$X8[idx_na_X8],
  Imp2 = data2$X8[idx_na_X8],
  Imp3 = data3$X8[idx_na_X8],
  Imp4 = data4$X8[idx_na_X8],
  Imp5 = data5$X8[idx_na_X8])
hasil_X8
#model logreg
data_i <- complete(imp, 1)
model_X4 <- glm(X4 ~ X1 + X2 + X3 + X5 + X6 + X7 + X8,
               data = data_i,
               family = binomial)
summary(model_X4)
model_X7 <- glm(X7 ~ X1 + X2 + X3 + X4 + X5 + X6 + X8,
               data = data_i,
               family = binomial)
summary(model_X7)
model_X8 <- glm(X8 ~ X1 + X2 + X3 + X4 + X5 + X6 + X7,
               data = data_i,
               family = binomial)
summary(model_X8)
stripplot(imp)
densityplot(imp, ~X4, groups = .imp)
# =====
# LANGKAH 2: ANALISIS DESKRIPTIF
# =====
#Gabungkan 5 imputasi
data1$imp <- 1

```

```

data2$imp <- 2
data3$imp <- 3
data4$imp <- 4
data5$imp <- 5
data_deskriptif <- rbind(data1, data2, data3, data4, data5)
# Variabel kategorik
var_kat <- c("X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8", "Status")
#Fungsi Frekuensi & Persentase
frek_persen <- function(x) {
  freq <- table(x) / 5 # ←
  persen <- prop.table(freq) * 100
  data.frame(
    Kategori = names(freq),
    Frekuensi = round(as.vector(freq), 0),
    Persentase = round(as.vector(persen), 2))}
#Hitung Semua Variabel
hasil_deskriptif <- list()
for (v in var_kat) {
  hasil_deskriptif[[v]] <- frek_persen(data_deskriptif[[v]])}
#Gabungkan Jadi 1 Tabel
library(dplyr)
tabel_deskriptif <- bind_rows(
  lapply(names(hasil_deskriptif), function(v) {
    df <- hasil_deskriptif[[v]]
    df$Variabel <- v
    return(df)}))
tabel_deskriptif
#PIE CHART
nama_variabel <- c(
  X1 = "Usia",
  X2 = "Jenis Kelamin",
  X3 = "Status Pengobatan",
  X4 = "Riwayat Pengobatan",
  X5 = "Jenis Fasilitas Kesehatan",
  X6 = "Tipe Diagnosis TB",
  X7 = "Hasil Tes HIV",
  X8 = "Hasil Tes DM",
  Status = "Status Kejadian")
label_kategori <- list(
  X1 = c("0"="Tidak Produktif", "1"="Produktif"),
  X2 = c("0"="Perempuan", "1"="Laki-laki"),
  X3 = c("0"="Tidak Standar", "1"="Sesuai Standar"),
  X4 = c("0"="Pernah Diobati", "1"="Kasus Baru"),
  X5 = c("0"="Puskesmas", "1"="Rumah Sakit", "2"="Lainnya"),
  X6 = c("0"="Klinis", "1"="Bakteriologis"),
  X7 = c("0"="HIV Positif", "1"="HIV Negatif"),
  X8 = c("0"="DM Positif", "1"="DM Negatif"),
  Status = c("0"="Censored", "1"="Sembuh"))
# Hitung frekuensi (dibagi 5)
freq <- table(data_deskriptif[[v]]) / 5
persen <- prop.table(freq) * 100
# Label kategori
kategori_asli <- names(freq)
kategori_label <- label_kategori[[v]][kategori_asli]
# Label di dalam pie (frekuensi + persen)
label_pie <- paste0(round(freq,0), "\n(", round(persen,1), "%)")
# Warna
warna <- rainbow(length(freq))
# Plot pie
pie(freq,
  labels = label_pie,

```

```

    col = warna,
    main = paste("Distribusi", nama_variabel[v]))
# Tambah legend
legend("topright",
      legend = kategori_label,
      fill = warna,
      cex = 0.8,
      bty = "n")
dev.off()
# =====
# LANGKAH 3: Kaplan-Meier & Uji Log-Rank
library(survival)
library(survminer)
data_survival <- data5 #ganti sesuai dataset berapa
data_survival$T <- as.numeric(data_survival$T)
data_survival$Status <- as.numeric(as.character(data_survival$Status))
data_survival <- na.omit(data_survival)
#Membentuk Objek Survival
surv_object <- Surv(time = data_survival$T,
                    event = data_survival$Status)
#Kaplan-Meier per Variabel + Log-Rank
km_X1 <- survfit(Surv(T, Status) ~ X1, data = data_survival)
ggsurvplot(km_X1,
            data = data_survival,
            title = "Kaplan-Meier Usia",
            xlab = "Waktu (hari)",
            ylab = "Probabilitas Belum Sembuh")
tabel_km_X1 <- data.frame(
  time = km_X1$time,
  n_risk = km_X1$n.risk,
  n_event = km_X1$n.event,
  survival = km_X1$surv,
  std_error = km_X1$std.err)
tabel_km_X1
tail(tabel_km_X1)
km_X2 <- survfit(Surv(T, Status) ~ X2, data = data_survival)
ggsurvplot(km_X2,
            data = data_survival,
            title = "Kaplan-Meier Jenis Kelamin",
            xlab = "Waktu (hari)",
            ylab = "Probabilitas Belum Sembuh")
tabel_km_X2 <- data.frame(
  time = km_X2$time,
  n_risk = km_X2$n.risk,
  n_event = km_X2$n.event,
  survival = km_X2$surv,
  std_error = km_X2$std.err)
tabel_km_X2
tail(tabel_km_X2, n=69)
km_X3 <- survfit(Surv(T, Status) ~ X3, data = data_survival)
ggsurvplot(km_X3,
            data = data_survival,
            title = "Kaplan-Meier Status Pengobatan",
            xlab = "Waktu (hari)",
            ylab = "Probabilitas Belum Sembuh")
tabel_km_X3 <- data.frame(
  time = km_X3$time,
  n_risk = km_X3$n.risk,
  n_event = km_X3$n.event,
  survival = km_X3$surv,
  std_error = km_X3$std.err)

```

```

tabel_km_X3
tail(tabel_km_X3,n=7)
km_X4 <- survfit(Surv(T, Status) ~ X4, data = data_survival)
ggsurvplot(km_X4,
  data = data_survival,
  title = "Kaplan-Meier Riwayat pengobatan",
  xlab = "Waktu (hari)",
  ylab = "Probabilitas Belum Sembuh")
tabel_km_X4 <- data.frame(
  time = km_X4$time,
  n_risk = km_X4$n.risk,
  n_event = km_X4$n.event,
  survival = km_X4$surv,
  std_error = km_X4$std.err)
tabel_km_X4
tail(tabel_km_X4,n=15)
km_X5 <- survfit(Surv(T, Status) ~ X5, data = data_survival)
ggsurvplot(km_X5,
  data = data_survival,
  title = "Kaplan-Meier Jenis Fasyankes",
  xlab = "Waktu (hari)",
  ylab = "Probabilitas Belum Sembuh")
tabel_km_X5 <- data.frame(
  time = km_X5$time,
  n_risk = km_X5$n.risk,
  n_event = km_X5$n.event,
  survival = km_X5$surv,
  std_error = km_X5$std.err)
tabel_km_X5
tail(tabel_km_X5,n=64)
km_X6 <- survfit(Surv(T, Status) ~ X6, data = data_survival)
ggsurvplot(km_X6,
  data = data_survival,
  title = "Kaplan-Meier Tipe Diagnosis TB",
  xlab = "Waktu (hari)",
  ylab = "Probabilitas Belum Sembuh")
tabel_km_X6 <- data.frame(
  time = km_X6$time,
  n_risk = km_X6$n.risk,
  n_event = km_X6$n.event,
  survival = km_X6$surv,
  std_error = km_X6$std.err)
tabel_km_X6
tail(tabel_km_X6,n=65)
km_X7 <- survfit(Surv(T, Status) ~ X7, data = data_survival)
ggsurvplot(km_X7,
  data = data_survival,
  title = "Kaplan-Meier Hasil Tes HIV",
  xlab = "Waktu (hari)",
  ylab = "Probabilitas Belum Sembuh")
tabel_km_X7 <- data.frame(
  time = km_X7$time,
  n_risk = km_X7$n.risk,
  n_event = km_X7$n.event,
  survival = km_X7$surv,
  std_error = km_X7$std.err)
tabel_km_X7
tail(tabel_km_X7,n=39)
km_X8 <- survfit(Surv(T, Status) ~ X8, data = data_survival)
ggsurvplot(km_X8,
  data = data_survival,

```

```

    title = "Kaplan-Meier Hasil Tes DM",
    xlab = "Waktu (hari)",
    ylab = "Probabilitas Belum Sembuh")
tabel_km_X8 <- data.frame(
  time = km_X8$time,
  n_risk = km_X8$n.risk,
  n_event = km_X8$n.event,
  survival = km_X8$surv,
  std_error = km_X8$std.err)
tabel_km_X8
tail(tabel_km_X8,n=49)

hasil_logrank_all <- list()
for (i in 1:5) {
  cat("\n=====\\n")
  cat("IMPUTASI KE-", i, "\\n")
  cat("=====\\n")
  data_i <- complete(imp, i)
  data_i$Status <- as.numeric(as.character(data_i$Status))
  data_i$T <- as.numeric(data_i$T)
  cat("Distribusi Status:\\n")
  print(table(data_i$Status))
  hasil_logrank <- data.frame()
  for (v in c("X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8")) {
    lr <- survdiff(as.formula(paste("Surv(T, Status) ~", v)),
      data = data_i)
    df <- length(lr$n) - 1
    p_val <- 1 - pchisq(lr$chisq, df = df)
    O <- lr$obs
    E <- lr$exp
    chi_manual <- sum((O - E)^2 / E)
    chi_tabel <- qchisq(0.95, df)
    keputusan <- ifelse(chi_manual > chi_tabel | p_val < 0.05,
      "Tolak H0",
      "Gagal Tolak H0")
    hasil_logrank <- rbind(hasil_logrank, data.frame(
      Imputasi = i,
      Variabel = v,
      Chi_R = round(lr$chisq, 4),
      Chi_manual = round(chi_manual, 4),
      Chi_tabel = round(chi_tabel, 4),
      df = df,
      p_value = round(p_val, 4),
      Keputusan = keputusan)))
  print(hasil_logrank)
  hasil_logrank_all[[i]] <- hasil_logrank}
# =====
#LANGKAH 4: COX PROPORTIONAL HAZARD
# =====
#Loop untuk Imputasi 1–5
library(mice)
library(survival)
iterasi <- numeric(5)
for (i in 1:5) {
  cat("\n=====\\n")
  cat("IMPUTASI KE-", i, "\\n")
  cat("=====\\n")
  # Ambil dataset imputasi ke-i
  data_i <- complete(imp, i)
  # Perbaiki tipe data
  data_i$Status <- as.numeric(as.character(data_i$Status))

```

```

data_i$T <- as.numeric(data_i$T)
#Model Cox + Estimasi Parameter (MPLE)
cox_model <- coxph(Surv(T, Status) ~
                    X1 + X3 + X4 + X5 + X6 + X7 + X8,
                    data = data_i)
cat("\nModel Cox (Imputasi", i, "):\n")
cat("\nKoefisien Model:\n")
print(coef(cox_model))# beta
# Ambil jumlah iterasi
iterasi[i] <- cox_model$iter
cat("\nJumlah Iterasi sampai Konvergen:", iterasi[i], "\n")
tabel_iterasi <- data.frame(
  Imputasi = 1:5,
  Iterasi = iterasi,
  Status = rep("Konvergen", 5))
print(tabel_iterasi)
# Uji Asumsi PH (Schoenfeld)
for (i in 1:5) {
  cat("\n===== \n")
  cat("IMPUTASI KE-", i, "\n")
  cat("===== \n")
  data_i <- complete(imp, i)
  data_i$Status <- ifelse(data_i$Status == 1, 1, 0)
  data_i$Status <- as.numeric(data_i$Status)
  data_i$T <- as.numeric(data_i$T)
  cox_model <- coxph(Surv(T, Status) ~
                    X1 + X3 + X4 + X5 + X6 + X7 + X8,
                    data = data_i)
  uji_ph <- cox.zph(cox_model)
  print(uji_ph)}
# =====
# MODEL COX TANPA X6
# =====
library(survival)
library(mice)
for (i in 1:5) {
  cat("\n===== \n")
  cat("IMPUTASI KE-", i, "\n")
  cat("===== \n")
  data_i <- complete(imp, i)
  data_i$Status <- as.numeric(as.character(data_i$Status))
  data_i$T <- as.numeric(data_i$T)
  cox_model_new <- coxph(Surv(T, Status) ~
                        X1 + X3 + X4 + X5 + X7 + X8,
                        data = data_i)
  cat("\nKoefisien Model\n")
  print(coef(cox_model_new))
  cat("\nJumlah Iterasi sampai Konvergen:\n")
  print(cox_model_new$iter)
  # Uji asumsi PH ulang tanpa X6
  uji_ph_new <- cox.zph(cox_model_new)
  cat("\nUji PH Model Baru:\n")
  print(uji_ph_new)}
# =====
# MODEL FINAL
# TANPA X2, X4, X6, DAN X8
# =====
library(survival)
library(mice)
library(mitools)
library(dplyr)

```

```

model_list <- list()
for (i in 1:5) {
  data_i <- complete(imp, i)
  data_i$Status <- as.numeric(as.character(data_i$Status))
  data_i$T <- as.numeric(data_i$T)
  cox_model_final <- coxph(
    Surv(T, Status) ~ X1 + X3 + X5 + X7,
    data = data_i)
  model_list[[i]] <- cox_model_final
  cat("\nKoefisien Model:\n")
  print(coef(cox_model_final))
  # Jumlah iterasi konvergen
  cat("\nJumlah Iterasi sampai Konvergen:",
    cox_model_final$iter, "\n")
  # Uji ASSUMSI PH
  uji_ph <- cox.zph(cox_model_final)
  cat("\nUji PH Model Final:\n")
  print(uji_ph)}
# =====
# POOLING RUBIN'S RULE
# =====
fit_final <- as.mira(model_list)
pool_final <- pool(fit_final)
summary_pool <- summary(pool_final)
print(summary_pool)
# Variansi dalam imputasi (Ubar)
Ubar <- summary_pool$std.error^2
coef_list <- lapply(model_list, function(x) coef(x))
coef_matrix <- do.call(rbind, coef_list)
coef_matrix <- as.data.frame(coef_matrix)
# Between-imputation variance
B <- sapply(coef_matrix, stats::var)
B
# Banyak imputasi
m <- 5
# Variansi total (T)
Tvar <- Ubar + (1 + 1/m) * B
# Tabel Rubin
tabel_rubin <- data.frame(
  Variabel = summary_pool$term,
  Estimate = summary_pool$estimate,
  Ubar = Ubar,
  B = B,
  T = Tvar)
print(tabel_rubin)
# =====
# LIKELIHOOD RATIO TEST (LRT)
# =====
hasil_LRT <- data.frame()
for (i in 1:5) {
  data_i <- complete(imp, i)
  data_i$Status <- as.numeric(as.character(data_i$Status))
  data_i$T <- as.numeric(data_i$T)
  # Model full
  model_full <- coxph(
    Surv(T, Status) ~ X1 + X3 + X5 + X7,
    data = data_i)
  # Model null
  model_null <- coxph(
    Surv(T, Status) ~ 1,
    data = data_i)

```

```

# Statistik G2
G2 <- as.numeric(
  2 * (logLik(model_full) - logLik(model_null)))
# df
# X1, X3, X7 = 3 parameter
# X5 = 2 dummy
# Total = 5 parameter
df <- 5
# Chi-square tabel
chi_tabel <- qchisq(0.95, df)
# p-value
p_value <- 1 - pchisq(G2, df)
# Keputusan
keputusan <- ifelse(
  G2 > chi_tabel | p_value < 0.05,
  "Tolak H0",
  "Gagal Tolak H0")
# Simpan hasil
hasil_LRT <- rbind(
  hasil_LRT,
  data.frame(
    Imputasi = i,
    G2 = round(G2, 4),
    Chi_tabel = round(chi_tabel, 4),
    df = df,
    p_value = round(p_value, 4),
    Keputusan = keputusan))}
print(hasil_LRT)
# =====
# UJI WALD
# =====
summary_pool$W2 <- summary_pool$statistic^2
summary_pool$Chi_tabel <- qchisq(0.95, df = 1)
summary_pool$Keputusan <- ifelse(
  summary_pool$W2 > summary_pool$Chi_tabel |
  summary_pool$p.value < 0.05,
  "Tolak H0",
  "Gagal Tolak H0")
hasil_wald <- data.frame(
  Variabel = summary_pool$term,
  W2 = round(summary_pool$W2, 4),
  Chi_tabel = round(summary_pool$Chi_tabel, 4),
  p_value = round(summary_pool$p.value, 4),
  Keputusan = summary_pool$Keputusan)
print(hasil_wald)

# Hitung Hazard Ratio
HR <- exp(summary_pool$estimate)
# Confidence Interval 95%
Lower_CI <- exp(summary_pool$estimate - 1.96 * summary_pool$std.error)
Upper_CI <- exp(summary_pool$estimate + 1.96 * summary_pool$std.error)
# Tabel hasil
tabel_HR <- data.frame(
  Variabel = summary_pool$term,
  Koefisien = round(summary_pool$estimate, 3),
  HR = round(HR, 3),
  CI_95_Bawah = round(Lower_CI, 3),
  CI_95_Atas = round(Upper_CI, 3),
  p_value = round(summary_pool$p.value, 4))
# Tampilkan hasil
tabel_HR

```

RIWAYAT HIDUP



Aini Intan Nur Lailiyah, lahir di Kota Probolinggo pada tanggal 24 Januari 2004. Penulis merupakan anak kedua dari ketiga bersaudara dari Bapak Indung Setya dan Ibu Dewi Camelia. Penulis telah menempuh pendidikan mulai dari TK di TK ABA III Kota Probolinggo yang lulus pada tahun 2010. Kemudian menempuh Pendidikan sekolah dasar di MI Muhammadiyah 1 Kota Probolinggo (kelas 1-3), dilanjutkan di SD Al-Ikhlas Lumajang, yang kemudian lulus di tahun 2016. Setelah itu, penulis melanjutkan pendidikan sekolah menengah pertama di SMP Ar-Rohmah Putri IIBS Malang dan lulus pada tahun 2019. Selanjutnya, penulis melanjutkan pendidikan sekolah menengah atas di MAN 2 Kota Probolinggo dan lulus pada tahun 2022. Pada tahun yang sama, penulis melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika, Fakultas Sains dan Teknologi.



KEMENTERIAN AGAMA RI

UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM

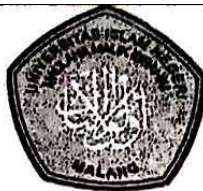
MALANG FAKULTAS SAINS DAN TEKNOLOGI

Jl. Gajayana No. 50 Malang 65144 Telp. / Fax. (0341) 558933

BUKTI KONSULTASI SKRIPSI

Nama : Aini Intan Nur Lailiyah
NIM : 2206011110106
Fakultas / Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Analisis *Survival* Menggunakan Metode *Kaplan-Meier* dan *Cox Proportional Hazard* pada Waktu Kesembuhan Penderita Tuberkulosis Paru (Studi Kasus di Kota Probolinggo)
Pembimbing I : Dr. Fachrur Rozi, M.Si.
Pembimbing II : Ari Kusumastuti, M.Pd, M.Si. .

No	Tanggal	Hal	Tanda Tangan
1.	16 Mei 2025	Pembahasan Topik	1.
2.	26 Mei 2025	Konsultasi Latar Belakang	2.
3.	28 Mei 2025	Konsultasi Bab I, II, III	3.
4.	12 September 2025	Konsultasi Kajian Agama Bab I dan II	4.
5.	18 Juni 2025	Revisi Bab I, II, III	5.
6.	27 Oktober 2025	ACC Kajian Agama Bab I dan II	6.
7.	10 November 2025	ACC Bab I, II, dan III	7.
8.	1 Desember 2025	ACC Seminar Proposal	8.
9.	9 Februari 2026	Konsultasi Revisi Seminar Proposal	9.
10.	13 Februari 2026	Konsultasi Bab IV dan V	10.
11.	13 April 2026	Kosultasi Kajian Agama Bab IV	11.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No. 50 Malang 65144 Telp. / Fax. (0341) 558933

12.	16 April 2026	ACC Kajian Agama Bab IV	12.
13.	19 Februari 2026	Revisi Bab IV dan V	13.
14.	24 Februari 2026	Konsultasi Bab IV dan V	14.
15.	13 April 2026	ACC Bab IV dan V	15.
16.	27 April 2026	ACC Seminar Hasil	16.
17.	11 Mei 2026	Konsultasi Revisi Seminar Hasil	17.
18.	12 Mei 2026	Revisi Seminar Hasil	18.
19.	20 Mei 2026	ACC Sidang Skripsi	19.
20.	2 Juni 2026	ACC Keseluruhan	20.

Malang, 2 Juni 2026
Mengetahui,
Ketua Program Studi Matematika



Dr. Fachrur Rozi, M.Si.

NIP. 19800527 200801 1 012