

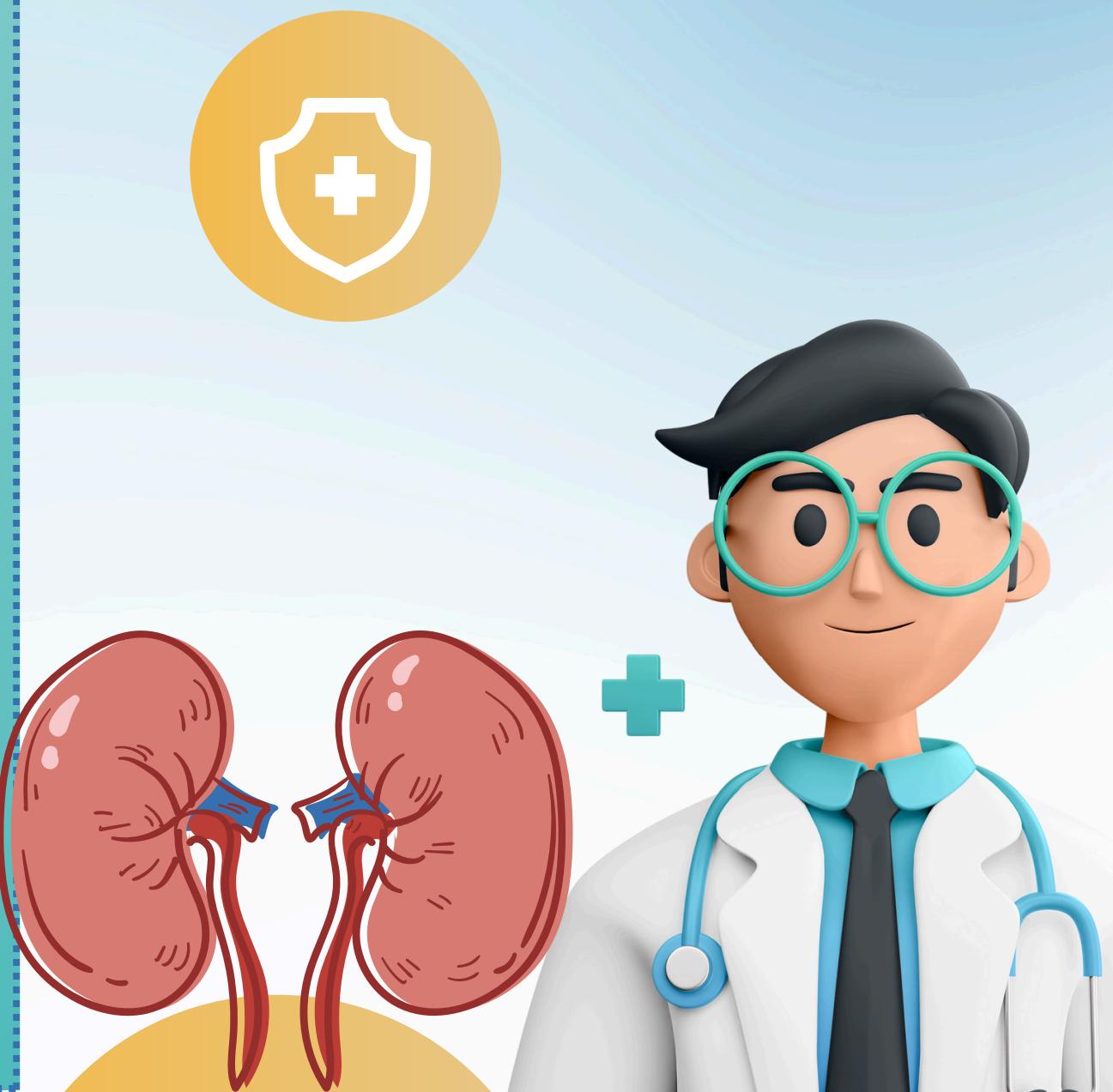


# Analisis Survival dengan Stratified Cox Model: Studi Kasus Kekambuhan Pasien Penyakit Ginjal

Analisis Survival E

R Studio®

STATISTIKA ITS  
2024/2025





# Anggota



**Dini Athirah Yasmin**  
5003211033



**Nabilah Azharudin**  
5003211093



**Dzakia Fathimatul Azmi**  
5003211161



# Latar Belakang



Transplantasi ginjal merupakan prosedur medis yang umum dilakukan untuk menangani gagal ginjal kronis. Namun, setelah transplantasi, pasien berisiko mengalami kejadian berulang seperti infeksi atau kekambuhan penyakit yang memengaruhi keberhasilan jangka panjang prosedur ini. Pemahaman mendalam terhadap faktor-faktor yang berkontribusi terhadap kejadian berulang tersebut menjadi krusial dalam meningkatkan perawatan pasca-transplantasi.

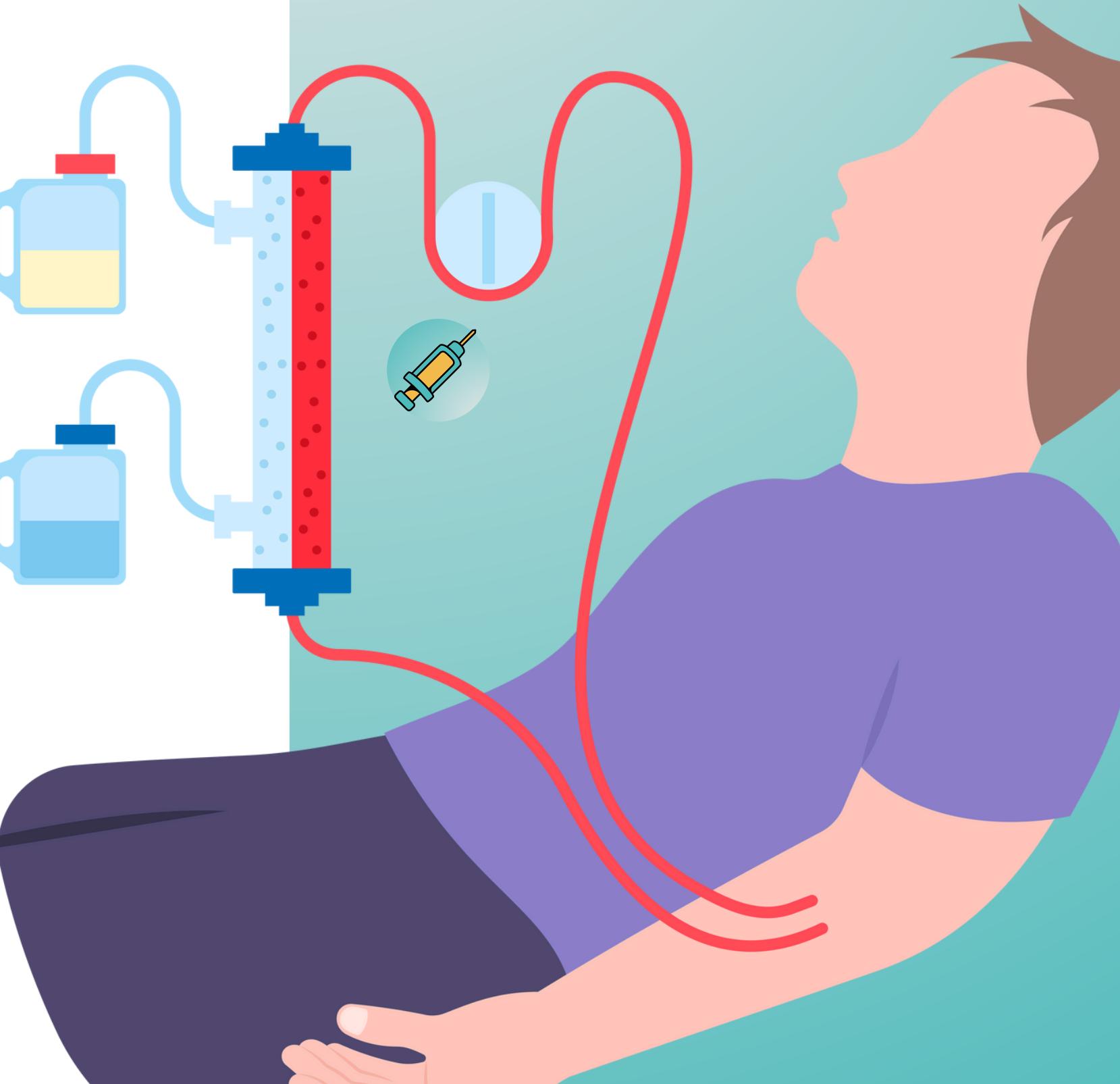


# Sumber Data



Dataset **kidrecurr** yang berhubungan dengan data pasien gagal ginjal, disediakan dalam **package KMSurv** di **RStudio**. Dataset ini mencakup informasi tentang **waktu** hingga kejadian berulang terjadi,,

**status infeksi, usia, jenis kelamin**, serta jenis penyakit ginjal, seperti glomerulonephritis (**gn**), amyloidosis (**an**), dan polycystic kidney disease (**pkd**).





# Struktur Data

Variabel	Struktur
Time	Rasio
Infect	Nominal (0: sensor, 1: tidak sensor)
Age	Rasio
Gender	Nominal (0: laki-laki, 1: perempuan)
Glomerulonephritis (gn)	Nominal (0: tidak ada, 1: ada riwayat)
Amyloidosis (an)	Nominal (0: tidak ada, 1: ada riwayat)
Polycystic kidney disease (pkd)	Nominal (0: tidak ada, 1: ada riwayat)



patient	time	infect	age	gender	gn	an	pkd	disease
1	1	16	1	28.0	0	0	0	0
2	2	13	0	48.0	1	1	0	1
3	3	22	1	32.0	0	0	0	0
4	4	318	1	31.5	1	0	0	0
5	5	30	1	10.0	0	0	0	0
6	6	24	1	16.5	1	0	0	0
7	7	9	1	51.0	0	1	0	1
8	8	30	1	55.5	1	1	0	1
9	9	53	1	69.0	1	0	1	2
10	10	154	1	51.5	0	1	0	1

# Processing Data

Pemrosesan data dilakukan untuk memastikan data terstruktur dengan baik, akurat, dan efisien untuk

dianalisis.

```
data("kidrecur")
attach(kidrecur)
patient = c(patient, patient)
time = c(time1, time2)
infect = c(infect1, infect2)
age = c(age, age)
gender = c(gender, gender)
gn = c(gn, gn)
an = c(an, an)
pkd = c(pkd, pkd)

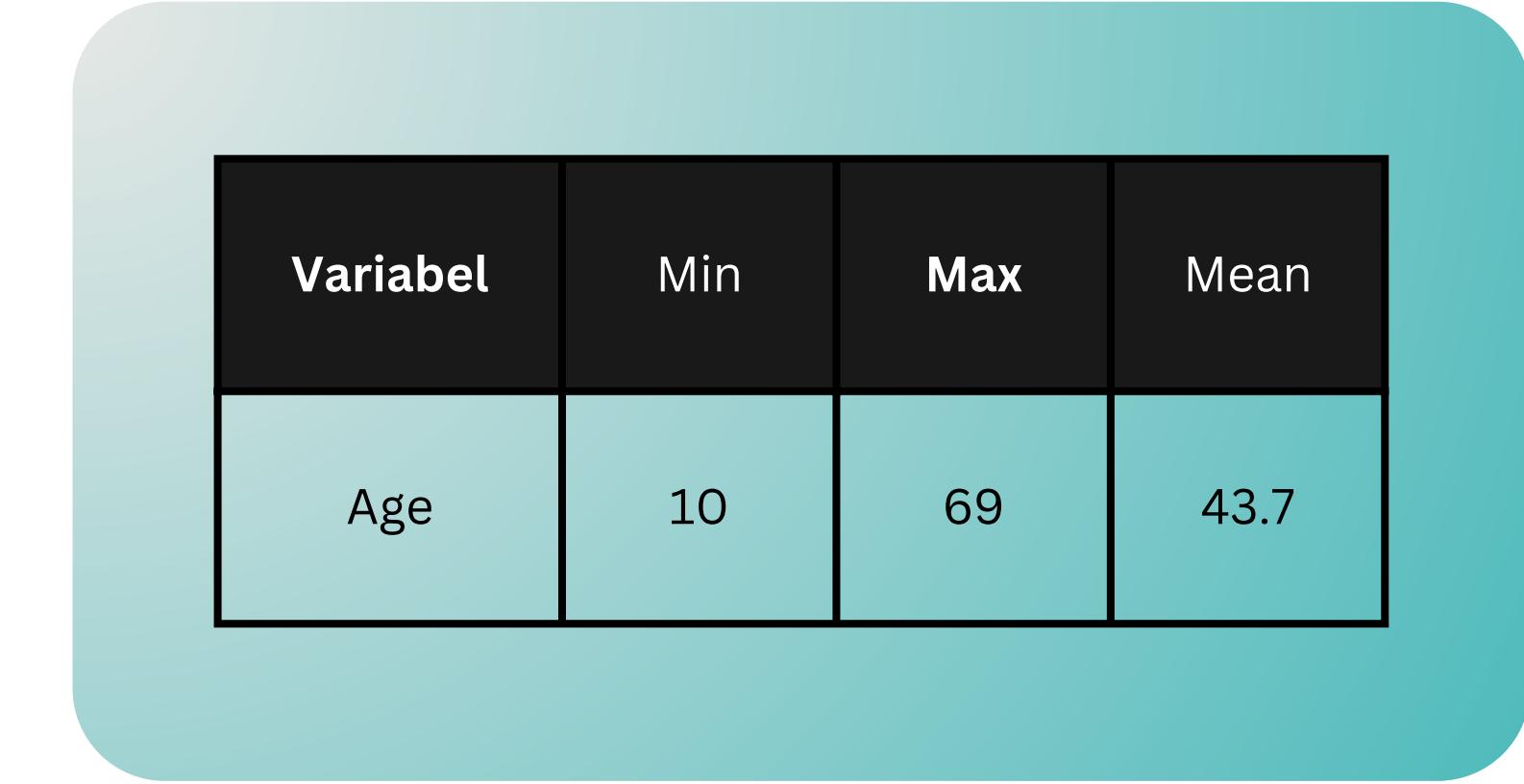
# Menambahkan kolom disease
df$disease <- with(df, ifelse(gn == 1, 1, # gn
                               ifelse(an == 1, 2, # an
                                      ifelse(pkd == 1, 3, # pkd
                                          0)))) # no disease)
```



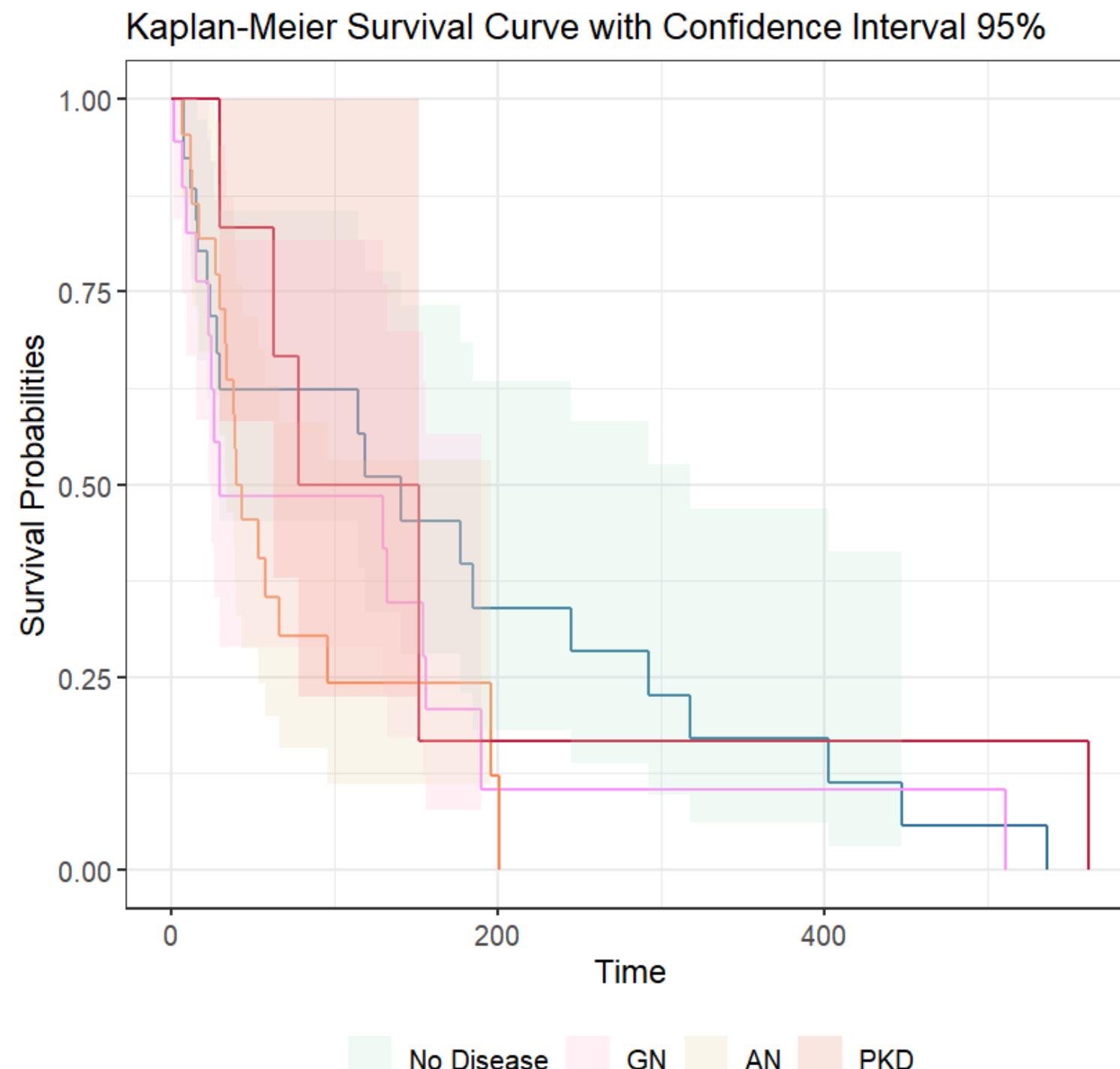


# Statistika Deskriptif

Variabel	Kategori	Frekuensi
Gender	Pria	20
	Wanita	56
GN	Ya	18
	Tidak	58
AN	Ya	24
	Tidak	52
PKD	Ya	8
	Tidak	68



# Kurva Kaplan-Meier

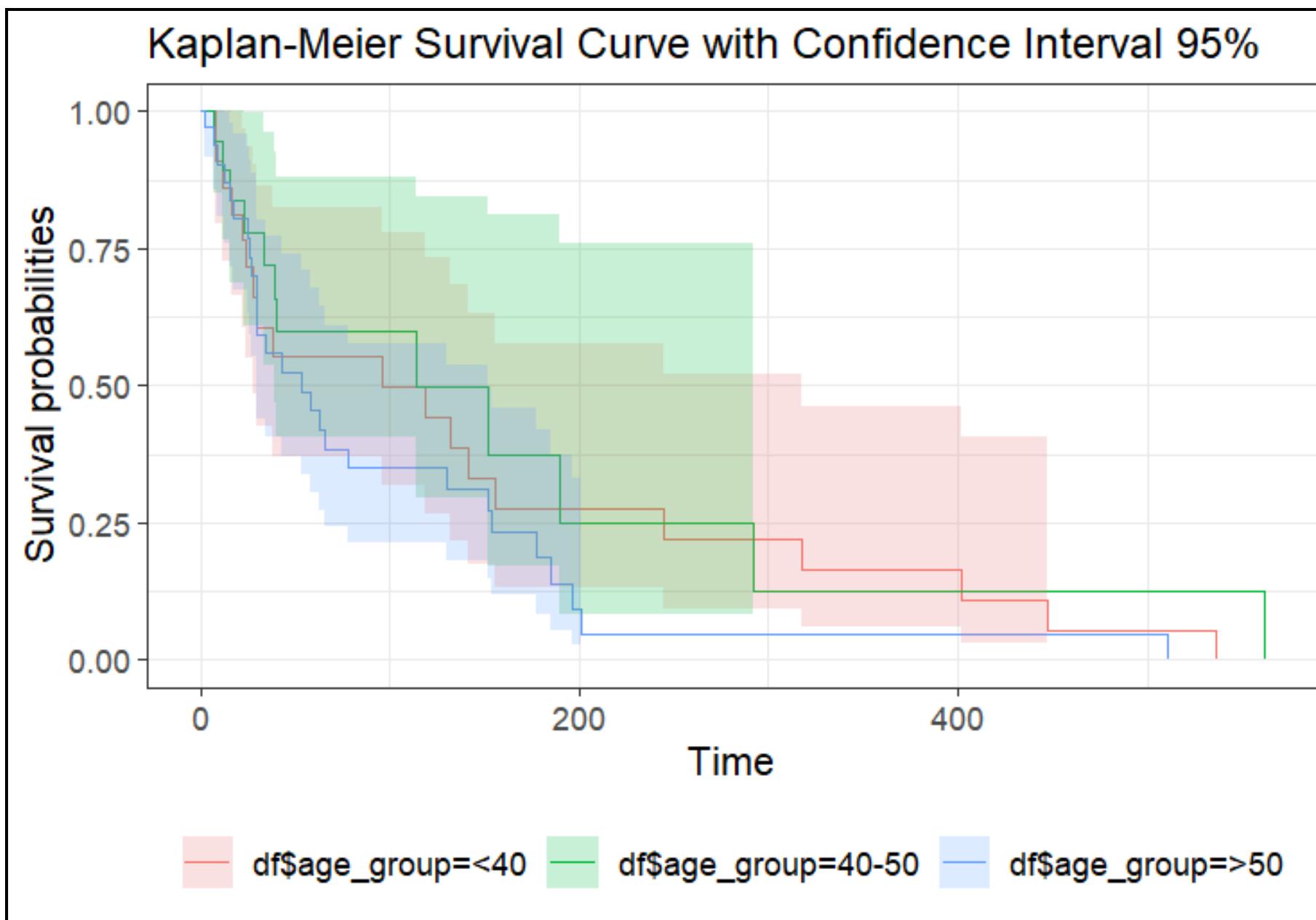


## Type of Disease

Kurva Kaplan-Meier untuk variabel Type of Disease menunjukkan bahwa :

- Pasien dalam kelompok Kelompok No Disease (warna biru muda) memiliki angka survival tertinggi dibanding kelompok penyakit lain. Kurva menunjukkan penurunan survival lebih lambat, artinya pasien cenderung bertahan hidup lebih lama.
- Kurva GN menunjukkan penurunan survival yang tajam pada awal periode waktu. Hal ini mengindikasikan bahwa pasien dengan tipe penyakit Glomerulonephritis (GN) memiliki risiko kematian yang lebih cepat dibanding kelompok lain.
- Kurva AN berada di antara kelompok GN dan No Disease. Risiko penurunan survival terjadi relatif cepat tetapi lebih baik dibanding GN.
- Kelompok PKD juga menunjukkan penurunan survival yang signifikan di awal periode. Kurva ini hampir mirip dengan GN, tetapi dengan tingkat risiko kematian yang sedikit lebih baik.

# Kurva Kaplan-Meier



Age

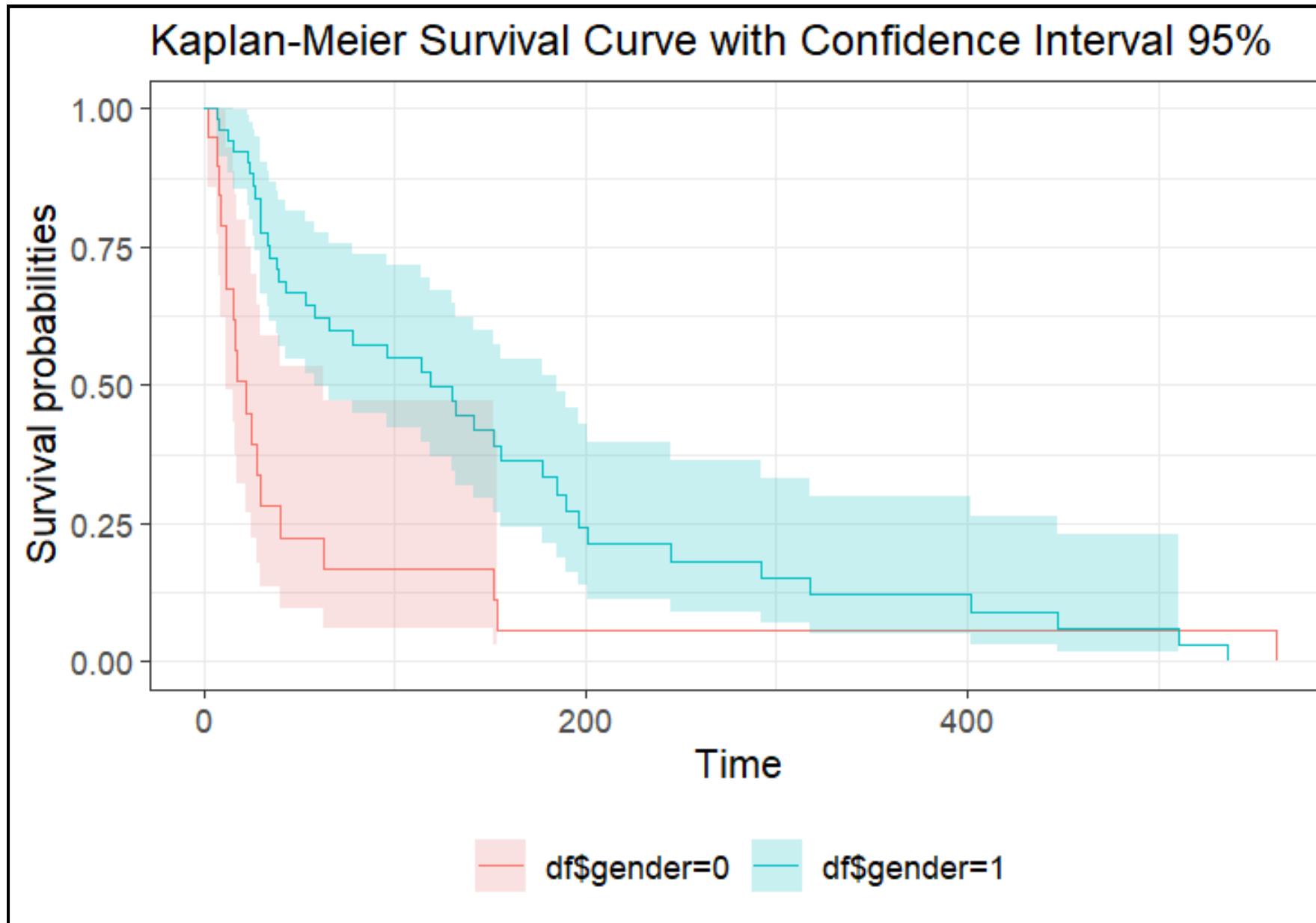
Kurva Kaplan-Meier untuk variabel usia (age) menunjukkan bahwa :

- Kelangsungan hidup pasien usia < 40 tahun cenderung lebih rendah dibandingkan kelompok lainnya dalam jangka panjang, karena garis survival menurun lebih cepat. Ini mengindikasikan risiko kematian lebih tinggi untuk kelompok usia ini.
- Kelangsungan hidup pasien usia 40-50 tahun, memiliki probabilitas kelangsungan hidup yang lebih baik dibanding kelompok ≤40, tetapi masih menurun secara signifikan seiring waktu.
- Kelompok pasien usia > 50 tahun menunjukkan probabilitas kelangsungan hidup tertinggi, karena kurva menurun lebih lambat dibandingkan kedua kelompok lainnya.

# Kurva Kaplan-Meier



## Gender



Kurva Kaplan-Meier untuk variabel jenis kelamin (gender) menunjukkan bahwa :

- Kelompok wanita (gender=1) memiliki survival probability yang lebih tinggi dibandingkan kelompok pria (gender=0). Hasil ini menunjukkan bahwa gender mungkin memiliki pengaruh terhadap kelangsungan hidup, atau dalam kata lain bahwa pasien wanita memiliki kelangsungan hidup lebih baik dibandingkan dengan pasien pria .

# Log Rank Test +

Uji log-rank digunakan untuk mengevaluasi apakah kurva kelangsungan hidup untuk dua atau lebih kelompok setara secara statistik. Hipotesis yang digunakan sebagai berikut.

- **H<sub>0</sub> : the two survival curves are equivalent**
- **H<sub>1</sub> : the two survival curves are different**

## Age

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
df\$age_group=<40	22	19	20.1	0.057	0.0926
df\$age_group=40-50	20	12	16.0	1.013	1.4970
df\$age_group=>50	34	27	21.9	1.188	2.0261

Chisq= 2.4 on 2 degrees of freedom, p= 0.3

- Dengan nilai p = 0.3, sehingga keputusannya adalah gagal tolak hipotesis nol karena p > 0.05.
- Ini berarti **tidak ada perbedaan signifikan** dalam kelangsungan hidup antar kelompok usia pasien.

## Gender

	N	Observed	Expected	(O-E) <sup>2</sup> /E	(O-E) <sup>2</sup> /V
df\$gender=0	20	18	10.3	5.69	7.88
df\$gender=1	56	40	47.7	1.23	7.88

Chisq= 7.9 on 1 degrees of freedom, p= 0.005

- Dengan nilai p = 0.005, sehingga keputusannya adalah tolak hipotesis nol karena p < 0.05.
- Ini berarti **ada perbedaan signifikan** dalam kelangsungan hidup antar jenis kelamin pasien.



# Log Rank Test +



## Type of Disease

	N	Observed	Expected	$(O-E)^2/E$	$(O-E)^2/V$
df\$disease=0	26	20	23.87	0.627	1.147
df\$disease=1	18	14	11.93	0.358	0.466
df\$disease=2	24	18	13.57	1.446	2.099
df\$disease=3	8	6	8.63	0.800	1.093

chisq= 3.7 on 3 degrees of freedom, p= 0.3

- Dengan nilai  $p = 0.3$ , sehingga keputusannya adalah gagal tolak hipotesis nol karena  $p > 0.05$ .
- Ini berarti **tidak ada perbedaan signifikan** dalam kelangsungan hidup antar tipe penyakit ginjal yang di derita pasien.



# The Cox Proportional Hazard Model

Pemilihan model terbaik menggunakan pendekatan Likelihood

Ratio Test. Hipotesis yang digunakan adalah sebagai berikut.

- **H0 : a term has no significant effect on survival times**
- **H1 : a term has a significant effect on survival times**

1  $\hat{h}_1(t|x) = \exp(\beta_1 Age)$

2  $\hat{h}_2(t|x) = \exp(\beta_1 Age + \beta_2 Gender)$

3  $\hat{h}_3(t|x) = \exp(\beta_1 Age + \beta_2 Gender + \beta_3 GN)$

4  $\hat{h}_4(t|x) = \exp(\beta_1 Age + \beta_2 Gender + \beta_4 AN)$

5  $\hat{h}_5(t|x) = \exp(\beta_1 Age + \beta_2 Gender + \beta_5 PKD)$

$$\hat{h}(t | \mathbf{X}) = h_0(t) \exp(\hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p)$$

Model	LRT	Keputusan	Model Terbaik
Model 1 vs Model 2	6.5341	Tolak H0	Model 2
Model 2 vs Model 3	0.4837	Gagal Tolak H0	Model 2
Model 2 vs Model 4	3.7563	Gagal Tolak H0	Model 2
Model 2 vs Model 5	9.6277	Tolak H0	Model 5



# Pengujian Asumsi Proportional hazard

Berdasarkan uji log likelihood ratio diperoleh model proportional hazard yang memiliki efek signfikan terhadap kekambuhan pasien penderita penyakit ginjal adalah model 5. Selanjutnya dilakukan pengujian asumsi proportional hazard.

$$\hat{h}(t|x) = h_0(t) \exp(\beta_1 Age + \beta_2 Gender + \beta_5 PKD)$$

Hasil uji goodness of fit menunjukkan bahwa variabel **gender tidak memenuhi asumsi proportional hazard**, dengan nilai p-value sebesar 0.022 sehingga keputusannya adalah Tolak H0 karena p-value < 0.05.

## Goodness of Fit Test

Pengujian asumsi proportional hazard menggunakan pendekatan statistik yaitu **Goodness of Fit Test**. Metode ini menghasilkan p-value untuk setiap faktor yang berpengaruh terhadap kekambuhan pasien penyakit ginjal. Hipotesis yang digunakan sebagai berikut.

- $H_0 : \rho = 0$
- $H_1 : \rho \neq 0$

Variabel	Chisq	P-value	Keputusan
Age	0.1137	0.736	Gagal Tolak H0
Gender	5.2319	0.022	Tolak H0
PKD	0.0991	0.753	Gagal Tolak H0



# Stratified Cox Model

Model regresi stratified Cox adalah salah satu metode yang dapat digunakan untuk **pemodelan data survival jika terdapat satu atau lebih variabel yang tidak memenuhi asumsi proportional hazard**. Model regresi stratified Cox didapatkan dengan memodifikasi model Cox proportional hazard. Modifikasi dilakukan dengan mengontrol variabel yang tidak memenuhi asumsi proportional hazard yaitu variabel gender dengan cara menstratifikasi kan variabel tersebut.

$$\hat{h}_g(t|x) = h_{0g}(t) \exp(\beta_1 Age + \beta_5 PKD)$$

dengan  $g = 1$  untuk female dan  $g = 0$  untuk male



$$Female (g = 1): \hat{h}_1(t|x) = h_{01}(t) \exp(\beta_1 Age + \beta_5 PKD)$$

$$Male (g = 0): \hat{h}_2(t|x) = h_{02}(t) \exp(\beta_1 Age + \beta_5 PKD)$$



# Stratified Cox Model

## Startified vs Non Stratified

Pengujian ini dilakukan untuk mengetahui apakah model dengan stratifikasi lebih baik dibandingkan dengan tanpa stratifikasi. Pengujian menggunakan uji log likelihood ratio dengan hipotesis sebagai berikut.

- **H<sub>0</sub>: Penambahan variabel ke dalam Stratified Cox model tidak meningkatkan performa model.**
- **H<sub>1</sub>: Penambahan variabel ke dalam Stratified Cox model meningkatkan performa model.**

Model	-2ln L	df	Pvalue
Non Stratified	-147.9494		
Stratified	-179.9819	3	0.000

Hasil pengujian menunjukkan bahwa **model dengan stratifikasi lebih baik** dibandingkan tanpa stratifikasi dengan nilai p-value sebesar 0.000 sehingga keputusannya adalah Tolak H<sub>0</sub> karena p-value < 0.05.





# Stratified Cox Model

## Startified with Interaction vs Non Interaction Model

Untuk menguji ada tidaknya interaksi pada model stratified Cox digunakan uji likelihood ratio (LR) yaitu dengan membandingkan statistik log likelihood untuk model interaksi dan model tanpa interaksi. Hipotesis yang digunakan sebagai berikut.

- **H<sub>0</sub>: tidak ada interaksi antara variabel stratifikasi dengan variabel independen yang masuk dalam model**
- **H<sub>1</sub>: terdapat interaksi antara variabel stratifikasi dengan variabel independen yang masuk dalam model**

Model	-2ln L	df	Pvalue
Statified No Interaction	300.7875		
Stratified With Interaction	295.2354	3	0.1355

Hasil pengujian memberikan keputusan **Gagal Tolak H<sub>0</sub>** karena nilai p-value sebesar 0.1355 dimana > 0.05. Dapat disimpulkan bahwa **tidak ada interaksi** antara variabel gender dengan variabel age dan PKD.



# Estimasi Parameter



Variabel	Estimasi Parameter
Age (40-50) (X1(1))	0.04273
Age (>50) (X1(2))	0.0396
PKD (X2)	-0.68474

**Model gender = 1 (Wanita)**

$$\hat{h}_1(t|x) = h_{01}(t) \exp (0.04273 X_1(1) + \exp (0.04273 X_1(1) - 0.68474) X_2$$

**Model gender = 0 (Pria)**

$$\hat{h}_2(t|x) = h_{02}(t) \exp (0.04273 X_1(1) + \exp (0.04273 X_1(1) - 0.68474) X_2$$



# Hazard Ratio



Variabel	Hazard Ratio
Age (40-50) (X1(1))	1.0436
Age (>50) (X1(2))	1.4859
PKD (X2)	0.5042

- Kelompok **Usia 40-50 tahun** memiliki resiko kekambuhan **4.36% lebih tinggi** dibandingkan dengan kelompok usia < 40 tahun, namun peningkatan resiko ini sangat kecil.
- Kelompok **Usia > 50 tahun** memiliki resiko kekambuhan **48.59% lebih tinggi** dibandingkan dengan kelompok usia < 40 tahun. Hal ini menunjukkan peningkatan risiko yang lebih besar dibandingkan kelompok usia 40-50.
- Individu dengan **PKD (Penyakit Ginjal Polikistik)** memiliki risiko kekambuhan **49.58% lebih rendah** dibandingkan individu tanpa PKD. HR < 1 menunjukkan efek protektif atau penurunan risiko kejadian.



# Kesimpulan



1. Berdasarkan pengujian asumsi proportional hazard diperoleh bahwa variabel gender tidak memenuhi asumsi PH. Dilakukan pemodelan menggunakan stratified cox model untuk mengatasi hal tersebut.
2. Diperoleh hasil estimasi parameter model regresi stratified cox sebagai berikut.

$$\hat{h}_1(t|x) = h_{01}(t) \exp(0.04273 X_1(1) + \exp(0.04273 X_1(1) - 0.68474)X_2)$$

$$\hat{h}_2(t|x) = h_{02}(t) \exp(0.04273 X_1(1) + \exp(0.04273 X_1(1) - 0.68474)X_2)$$





# Rekomendasi



- 1. Prioritaskan perhatian pada individu usia > 50 tahun karena risiko kekambuhan signifikan lebih tinggi.**
  - Hal ini dapat dilakukan dengan cara meningkatkan frekuensi pemantauan kesehatan dan pemeriksaan berkala.
  
- 2. Tingkatkan pengawasan dan perlindungan bagi individu tanpa PKD karena memiliki risiko kekambuhan yang lebih tinggi.**
  - Menerapkan program pencegahan yang lebih intensif.
  - Memberikan edukasi kesehatan terkait manajemen risiko kekambuhan.





<https://bit.ly/FinalProjectAnsur>



# Thank You

