

DIALYSIS SURVIVAL MODEL

Enock Bereka

2024-12-05

Load the required library and the dataset

```
library(survival)
```

```
library(tidyverse)
```

```
dialysis <- read_csv("C:/Users/PC/OneDrive/Desktop/Data  
Science/Datasets/survival/dialysis survival dataset.csv")
```

Explore your data

```
head(dialysis)
```

```
## # A tibble: 6 × 9
```

```
##   event time   age begin center disease_diabetes disease_hypert  
disease_other
```

```
##   <dbl> <dbl> <dbl> <dbl>  <dbl>          <dbl>          <dbl>  
<dbl>
```

```
## 1     0     1    59    35    120              0              1  
0
```

```
## 2     0     3    49    38    120              0              1  
0
```

```
## 3     0    18    49    22    120              0              0  
1
```

```
## 4     0     2    52    21    120              0              1  
0
```

```
## 5     1     1    89    41    120              0              1  
0
```

```
## 6     1     3    72    33    120              0              0  
0
```

```
## # i 1 more variable: disease_renal <dbl>
```

```
names(dialysis)
```

```
## [1] "event"           "time"           "age"           "begin"
```

```
## [5] "center"         "disease_diabetes" "disease_hypert"
```

```
"disease_other"
```

```
## [9] "disease_renal"
```

KAPLAN-MEIER MODEL

Fit the Kaplan-Meier model

```
km_model <- survfit(Surv(time, event)~disease_diabetes, data = dialysis,  
                    type = "kaplan-meier")
```

Get the summary of the model

```
km_model
```

```
## Call: survfit(formula = Surv(time, event) ~ disease_diabetes, data =  
dialysis,
```

```
##   type = "kaplan-meier")
```

```
##
```

```
##               n events median 0.95LCL 0.95UCL
```

```
## disease_diabetes=0 5522  1200    NA      NA      NA
```

```
## disease_diabetes=1 1283   403    37      30      NA
```

```
summary(km_model)
```

```
## Call: survfit(formula = Surv(time, event) ~ disease_diabetes, data =  
dialysis,
```

```
##   type = "kaplan-meier")
```

```
##
```

```
##               disease_diabetes=0
```

```
## time n.risk n.event survival std.err lower 95% CI upper 95% CI
```

```
##    1  5522   181   0.967 0.00240   0.963   0.972
```

```
##    2  5002   192   0.930 0.00349   0.923   0.937
```

```
##    3  4473   123   0.905 0.00409   0.897   0.913
```

```
##    4  4127    84   0.886 0.00447   0.877   0.895
```

```
##    5  3898    65   0.871 0.00476   0.862   0.881
```

```
##    6  3715    49   0.860 0.00497   0.850   0.870
```

```
##    7  3551    32   0.852 0.00511   0.842   0.862
```

```
##    8  3408    49   0.840 0.00533   0.829   0.850
```

```
##    9  3209    34   0.831 0.00549   0.820   0.842
```

```
##   10  3059    34   0.822 0.00565   0.811   0.833
```

```
##   11  2871    22   0.815 0.00576   0.804   0.827
```

```
##   12  2729    32   0.806 0.00594   0.794   0.818
```

```
##   13  2570    26   0.798 0.00609   0.786   0.810
```

```
##   14  2401    27   0.789 0.00626   0.777   0.801
```

```
##   15  2263    11   0.785 0.00634   0.773   0.797
```

```
##   16  2151    23   0.776 0.00651   0.764   0.789
```

```
##   17  2031    21   0.768 0.00667   0.756   0.782
```

```
##   18  1906    15   0.762 0.00680   0.749   0.776
```

```
##   19  1797    20   0.754 0.00698   0.740   0.768
```

```
##   20  1687     9   0.750 0.00707   0.736   0.764
```

```
##   21  1603     4   0.748 0.00712   0.734   0.762
```

```
##   22  1527    13   0.742 0.00727   0.728   0.756
```

```
##   23  1427    14   0.734 0.00746   0.720   0.749
```

```
##   24  1326    11   0.728 0.00762   0.714   0.743
```

```
##   25  1240    10   0.722 0.00778   0.707   0.738
```

```
##   26  1161    13   0.714 0.00801   0.699   0.730
```

```
##   27  1069     8   0.709 0.00817   0.693   0.725
```

```
##   28   996     7   0.704 0.00833   0.688   0.721
```

```
##   29   915    13   0.694 0.00866   0.677   0.711
```

```
##   30   827    11   0.685 0.00898   0.667   0.703
```

```
##   31   756     5   0.680 0.00914   0.663   0.698
```

##	32	718	4	0.676	0.00929	0.659	0.695
##	33	640	11	0.665	0.00977	0.646	0.684
##	34	574	3	0.661	0.00992	0.642	0.681
##	35	528	5	0.655	0.01021	0.635	0.675
##	36	476	5	0.648	0.01056	0.628	0.669
##	37	418	6	0.639	0.01107	0.618	0.661
##	38	373	2	0.635	0.01127	0.614	0.658
##	39	314	1	0.633	0.01142	0.611	0.656
##	40	256	2	0.629	0.01185	0.606	0.652
##	41	202	2	0.622	0.01253	0.598	0.647
##	43	90	1	0.615	0.01417	0.588	0.644
##							
##							
##				disease_diabetes=1			
##	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
##	1	1283	48	0.963	0.00530	0.952	0.973
##	2	1166	56	0.916	0.00786	0.901	0.932
##	3	1030	41	0.880	0.00939	0.862	0.898
##	4	935	25	0.856	0.01025	0.837	0.877
##	5	875	24	0.833	0.01103	0.812	0.855
##	6	819	15	0.818	0.01151	0.795	0.840
##	7	775	21	0.795	0.01217	0.772	0.820
##	8	735	20	0.774	0.01277	0.749	0.799
##	9	683	21	0.750	0.01339	0.724	0.777
##	10	641	12	0.736	0.01374	0.710	0.763
##	11	606	14	0.719	0.01415	0.692	0.747
##	12	563	13	0.702	0.01455	0.674	0.731
##	13	523	12	0.686	0.01495	0.658	0.716
##	14	484	8	0.675	0.01523	0.646	0.705
##	15	450	4	0.669	0.01538	0.639	0.700
##	16	426	10	0.653	0.01580	0.623	0.685
##	17	390	7	0.641	0.01613	0.611	0.674
##	18	366	5	0.633	0.01638	0.601	0.666
##	19	342	2	0.629	0.01649	0.598	0.662
##	20	327	3	0.623	0.01667	0.591	0.657
##	21	303	3	0.617	0.01688	0.585	0.651
##	22	282	6	0.604	0.01735	0.571	0.639
##	23	264	3	0.597	0.01760	0.564	0.633
##	24	244	4	0.587	0.01798	0.553	0.624
##	25	226	2	0.582	0.01820	0.548	0.619
##	26	210	5	0.568	0.01879	0.533	0.606
##	27	186	3	0.559	0.01922	0.523	0.598
##	28	168	3	0.549	0.01972	0.512	0.589
##	29	154	3	0.538	0.02028	0.500	0.580
##	30	137	2	0.531	0.02073	0.491	0.573
##	31	121	1	0.526	0.02102	0.487	0.569
##	33	101	2	0.516	0.02185	0.475	0.560
##	34	86	1	0.510	0.02241	0.468	0.556
##	36	63	1	0.502	0.02347	0.458	0.550
##	37	56	1	0.493	0.02470	0.447	0.544

##	39	36	1	0.479	0.02754	0.428	0.536
##	41	20	1	0.455	0.03507	0.391	0.529

Key Findings

Group 1: Patients Without Diabetes (disease_diabetes=0)

Survival Time Summary

Total patients: **5522**.

Events (deaths): **1200**.

Median survival time: **Not reached** during the study period, indicating that more than 50% of the patients survived throughout the observed time.

Daily Risk and Survival

Day 1:

At risk: **5522** patients.

Events (deaths): **181**.

Survival probability: **96.7%** (meaning 96.7% of patients survived past Day 1).

Day 10:

At risk: **3059** patients.

Events (deaths): **34**.

Survival probability: **82.2%** (82.2% of patients survived past Day 10).

Day 20:

At risk: **1687** patients.

Events (deaths): **9**.

Survival probability: **75.0%**.

Day 30:

At risk: **827** patients.

Events (deaths): **11**.

Survival probability: **68.5%**.

Day 43:

At risk: **90** patients.

Events (deaths): **1**.

Survival probability: **61.5%**.

Interpretation

The survival probability steadily decreases over time, indicating that patients without diabetes generally survive longer, but the risk of death accumulates with time

Group 2: Patients With Diabetes (disease_diabetes=1)

Survival Time Summary

Total patients: **1283**.

Events (deaths): **403**.

Median survival time: **37 days** (50% of the patients with diabetes died by Day 37).

Daily Risk and Survival

Day 1:

At risk: **1283** patients.

Events (deaths): **48**.

Survival probability: **96.3%**.

Day 10:

At risk: **641** patients.

Events (deaths): **12**.

Survival probability: **73.6%**.

Day 20:

At risk: **327** patients.

Events (deaths): **3**.

Survival probability: **62.3%**.

Day 30:

At risk: **137** patients.

Events (deaths): **2**.

Survival probability: **53.1%**.

Day 37:

At risk: **56** patients.

Events (deaths): **1**.

Survival probability: **49.3%**.

Interpretation

Patients with diabetes have a significantly lower survival probability compared to those without diabetes.

The median survival time of **37 days** highlights the severity of outcomes for this group.

Comparison Between Groups

Initial Survival (Day 1):

Patients without diabetes: **96.7% survival**.

Patients with diabetes: **96.3% survival**.

Minimal difference in the short term.

Long-Term Survival:

The survival gap widens over time.

By Day 30, survival probability:

Without diabetes: **68.5%**.

With diabetes: **53.1%**.

This indicates that diabetes significantly increases mortality risk over time.

Median Survival Time:

Patients without diabetes: Median not reached (longer survival).

Patients with diabetes: **37 days**.

General Observations

Risk Stratification:

Diabetes significantly impacts survival outcomes in patients undergoing dialysis. Non-diabetic patients survive longer on average.

Clinical Implications:

Early interventions and continuous monitoring for diabetic patients undergoing dialysis are critical to improving outcomes.

Survival Patterns:

The survival probabilities decline more rapidly for diabetic patients, highlighting the need for disease-specific management.

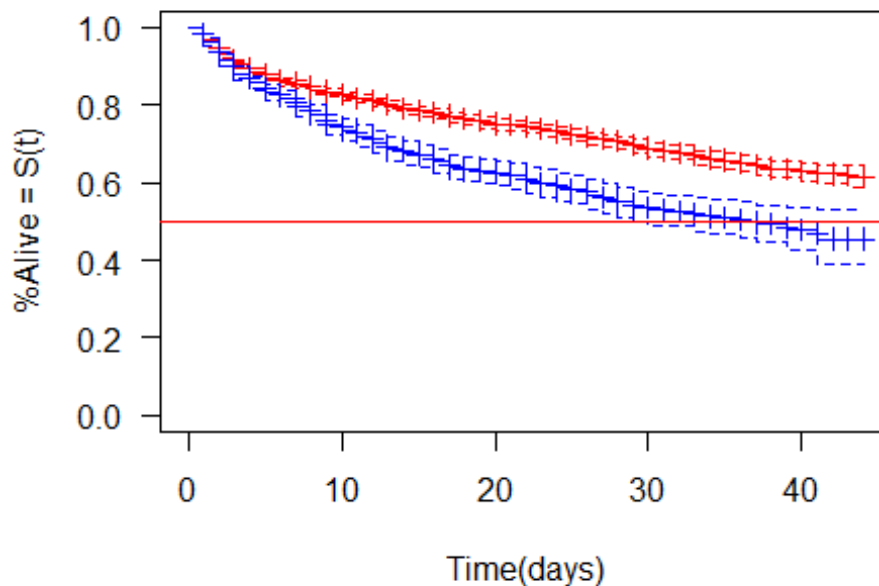
Event Rate:

The number of deaths per day is higher among diabetic patients compared to non-diabetic ones, even when adjusted for initial group size.

Visualize Survival curves

```
plot(km_model, conf.int = T, xlab = "Time(days)",  
     ylab = "%Alive = S(t)", main = "KM.Model",  
     mark.time = T, col = c("red", "blue"), las = 1)  
abline(h = 0.5, col = "red")
```

KM.Model



The Kaplan-Meier survival curve shows clear stratification based on diabetes status, with diabetic patients having significantly shorter survival times

#Log-rank test

```
log_rank <- survdiff(Surv(time, event) ~ disease_diabetes , data = dialysis)
log_rank
```

Call:

```
## survdiff(formula = Surv(time, event) ~ disease_diabetes, data = dialysis)
##
```

	N	Observed	Expected	(O-E) ² /E	(O-E) ² /V
## disease_diabetes=0	5522	1200	1320	10.9	63.1
## disease_diabetes=1	1283	403	283	50.7	63.1

##

Chisq= 63.1 on 1 degrees of freedom, p= 2e-15

There was a significant difference in survival between those people with diabetes and the ones without diabetes.

COX-PROPORTIONAL HAZARD MODEL

Fit Cox Model

```
cox_model <- coxph(Surv(time, event)~disease_hypert+
                    disease_renal+begin+center, data = dialysis)
```


Summarize the model

```
summary(cox_model)
```

```
## Call:
## coxph(formula = Surv(time, event) ~ disease_hypert + disease_renal +
##       begin + center, data = dialysis)
##
##      n= 6805, number of events= 1603
##
##              coef exp(coef)    se(coef)      z Pr(>|z|)
## disease_hypert -1.905e-01  8.265e-01  5.556e-02 -3.429 0.000606 ***
## disease_renal  -2.489e-01  7.796e-01  6.988e-02 -3.562 0.000368 ***
## begin          7.592e-03  1.008e+00  2.366e-03  3.209 0.001333 **
## center         -2.524e-05  1.000e+00  1.321e-05 -1.910 0.056130 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## disease_hypert    0.8265    1.2099    0.7412    0.9216
## disease_renal     0.7796    1.2826    0.6798    0.8941
## begin             1.0076    0.9924    1.0030    1.0123
## center            1.0000    1.0000    0.9999    1.0000
##
## Concordance= 0.544 (se = 0.008 )
## Likelihood ratio test= 32.98 on 4 df,   p=1e-06
## Wald test               = 33.2 on 4 df,   p=1e-06
## Score (logrank) test = 33.3 on 4 df,   p=1e-06
```

Model Fit and Statistics

Sample Size and Events:

Total participants in the study were 6805.

Number of participants who experienced the event (death) were 1603.

Model Statistics:

Likelihood Ratio Test, Wald Test, and Score Test:

These tests assess whether the included predictors are collectively significant.

All tests yield a p value of 1×10^{-6} , indicating the predictors significantly explain variations in hazard rates.

Concordance Index:

Measures the model's discriminative ability (higher values are better, up to 1).

A value of 0.544 suggests modest predictive power.

Coefficient and Hazard Ratios

Disease_Hypert (Hypertension):

Coefficient: -0.1905: Decreases hazard, increasing survival time.

HR: 0.82650.82650.8265 (CI: [0.7412,0.9216]):

Patients with hypertension have 17.35% ($1 - 0.82651 - 0.82651 - 0.8265$) lower risk of death compared to those without.

p-value=0.000606: Highly significant effect.

Disease_Renal (Renal Disease):

Coefficient: -0.2489: Decreases hazard, increasing survival time.

HR: 0.77960.77960.7796 (CI: [0.6798,0.8941]):

Patients with renal disease have 22.04% lower risk of death compared to those without.

p-value=0.00036: Highly significant effect.

Begin (Start Time of Dialysis):

Coefficient: 0.007590: Increases hazard, reducing survival time.

HR: 1.00761.00761.0076 (CI: [1.0030,1.0123]):

For each unit increase in start time, the hazard of death increases by 0.76%.

p-value=0.001333: Statistically significant.

Center (Dialysis Center):

Coefficient: -0.00002524: Decreases hazard, but the effect is minimal.

HR: 1.00001.00001.0000 (CI: [0.9999,1.0000]):

The impact of the dialysis center on the hazard is negligible.

p-value=0.05613: Marginally insignificant.

Hypertension and Renal Disease: Both conditions significantly reduce the hazard of death, suggesting they are associated with better survival outcomes during dialysis. This may reflect effective management or other underlying factors.

Start Time: Later start times slightly increase the risk of death, potentially indicating delayed intervention.

Dialysis Center: Minimal impact on survival, with borderline significance.

Check proportional hazard assumption.

```
ph_tet <- cox.zph(cox_model)
print(ph_tet)

##               chisq df    p
## disease_hypert 0.00675  1 0.93
## disease_renal  0.69199  1 0.41
## begin          1.31572  1 0.25
## center         0.05019  1 0.82
## GLOBAL         2.23212  4 0.69
```

Individual Predictors:

{disease_hypert, disease_renal, begin, center}:

All have $p > 0.05$, indicating no significant time-dependent effects.

Global Test:

p-value=0.69: The PH assumption holds globally, suggesting the model is valid

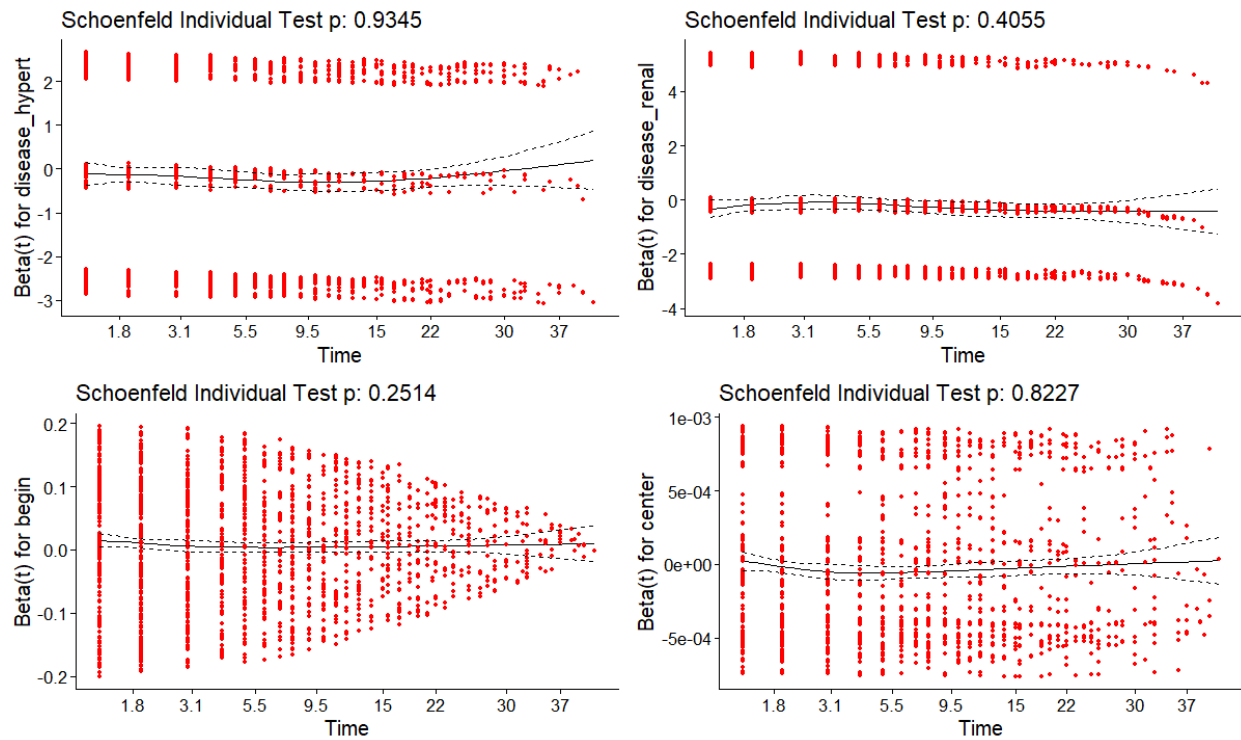
Model Fit: The PH assumption is valid, and the predictors explain hazard variation modestly well.

Visualize model assumptions

```
library(survminer)

ggcoxzph(cox.zph(cox_model))
```

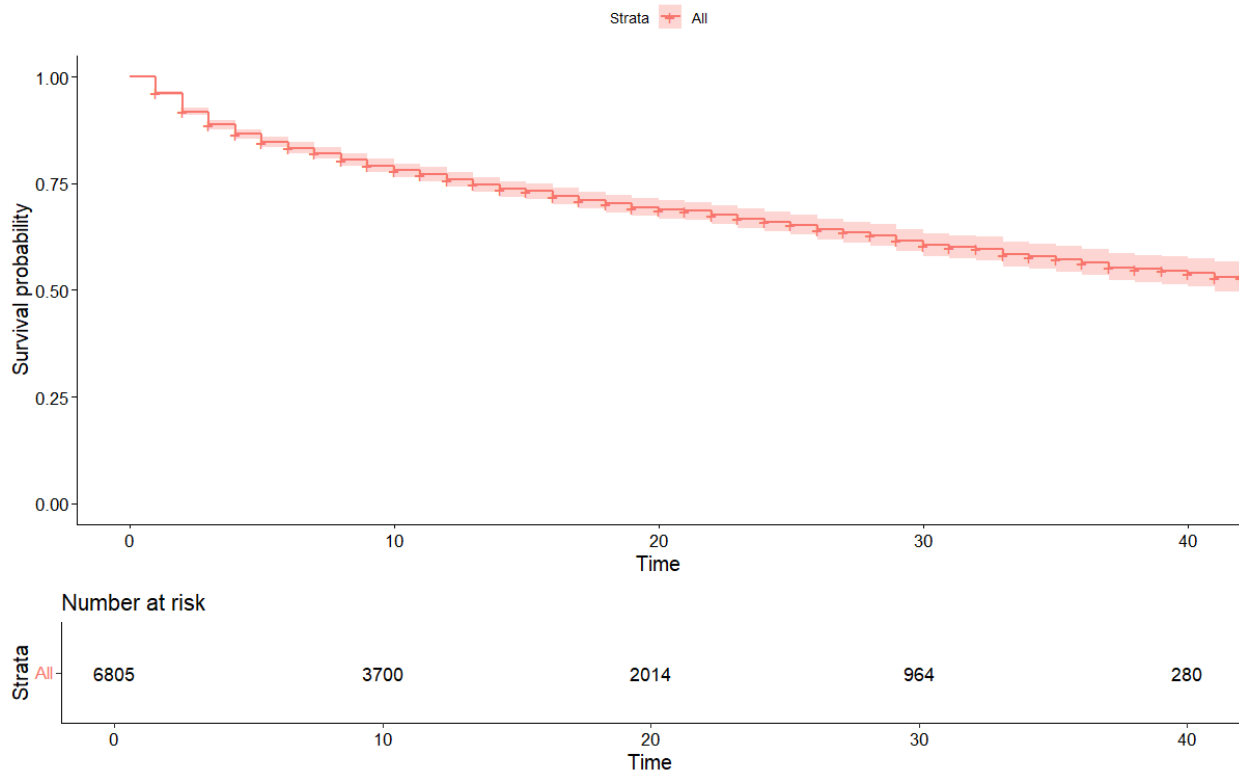
Global Schoenfeld Test p: 0.6932



Visual inspection of Schoenfeld residual plots (output from `ggcoxzph`) confirms that the PH assumption is not violated.

Visualize the survival curves.

```
ggsurvplot(survfit(cox_model), data = dialysis, risk.table = TRUE)
```



Survival curves visualize how the survival probabilities change over time across different groups.

The plot shows higher survival probabilities for patients with hypertension or renal disease compared to those without.

EXPONENTIAL MODEL

The exponential survival model assumes a constant hazard rate over time. It estimates the impact of covariates on the hazard rate and ultimately on survival probabilities. Key findings are derived from the coefficients (est) and their exponentiation ($\exp(est)$), which represent the hazard ratio (HR).

Fit the exponential model

```
library(flexsurv)
```

```
exp_model <- flexsurvreg(Surv(time, event)~disease_hypert+disease_renal
                        +begin+center,
                        data = dialysis,
                        dist = "exponential")
```

Get the summary of the model

exp_model

Call:

```
## flexsurvreg(formula = Surv(time, event) ~ disease_hypert + disease_renal +  
##     begin + center, data = dialysis, dist = "exponential")
```

##

Estimates:

	data	mean	est	L95%	U95%	se
## rate		NA	1.35e-02	1.19e-02	1.54e-02	9.06e-04
## disease_hypert	4.17e-01		-2.08e-01	-3.17e-01	-9.95e-02	5.56e-02
## disease_renal	2.08e-01		-2.51e-01	-3.88e-01	-1.14e-01	6.99e-02
## begin	2.28e+01		2.34e-02	1.91e-02	2.78e-02	2.23e-03
## center	2.55e+03		-3.03e-05	-5.10e-05	-9.70e-06	1.05e-05

	exp(est)	L95%	U95%
## rate	NA	NA	NA
## disease_hypert	8.12e-01	7.28e-01	9.05e-01
## disease_renal	7.78e-01	6.78e-01	8.92e-01
## begin	1.02e+00	1.02e+00	1.03e+00
## center	1.00e+00	1.00e+00	1.00e+00

##

N = 6805, Events: 1603, Censored: 5202

Total time at risk: 96345

Log-likelihood = -8099.732, df = 5

AIC = 16209.46

summary(exp_model)

##

disease_hypert=0.416752387950037,disease_renal=0.207788390889052,begin=22.777
5165319618,center=2553.10639235856

	time	est	lcl	ucl
## 1	1	0.9816002	0.9806185	0.9825365
## 2	2	0.9635390	0.9616127	0.9653780
## 3	3	0.9458101	0.9429753	0.9485192
## 4	4	0.9284075	0.9246990	0.9319548
## 5	5	0.9113250	0.9067770	0.9156796
## 6	6	0.8945568	0.8892023	0.8996887
## 7	7	0.8780972	0.8719683	0.8839770
## 8	8	0.8619404	0.8550683	0.8685397
## 9	9	0.8460809	0.8384958	0.8533720
## 10	10	0.8305132	0.8222445	0.8384692
## 11	11	0.8152320	0.8063082	0.8238266
## 12	12	0.8002319	0.7906808	0.8094397
## 13	13	0.7855078	0.7753563	0.7953041
## 14	14	0.7710547	0.7603287	0.7814153
## 15	15	0.7568675	0.7455924	0.7677691
## 16	16	0.7429413	0.7311418	0.7543612
## 17	17	0.7292713	0.7169712	0.7411874
## 18	18	0.7158529	0.7030752	0.7282437

## 19	19	0.7026814	0.6894486	0.7155261
## 20	20	0.6897522	0.6760861	0.7030305
## 21	21	0.6770610	0.6629825	0.6907532
## 22	22	0.6646032	0.6501330	0.6786902
## 23	23	0.6523747	0.6375324	0.6668379
## 24	24	0.6403711	0.6251761	0.6551926
## 25	25	0.6285884	0.6130593	0.6437507
## 26	26	0.6170226	0.6011773	0.6325086
## 27	27	0.6056695	0.5895256	0.6214628
## 28	28	0.5945253	0.5780998	0.6106099
## 29	29	0.5835862	0.5668953	0.5999465
## 30	30	0.5728484	0.5559081	0.5894694
## 31	31	0.5623081	0.5451338	0.5791752
## 32	32	0.5519617	0.5345683	0.5690608
## 33	33	0.5418058	0.5242076	0.5591230
## 34	34	0.5318367	0.5140477	0.5493588
## 35	35	0.5220510	0.5040847	0.5397651
## 36	36	0.5124454	0.4943148	0.5303389
## 37	37	0.5030165	0.4847342	0.5210774
## 38	38	0.4937611	0.4753394	0.5119776
## 39	39	0.4846761	0.4661266	0.5030367
## 40	40	0.4757581	0.4570924	0.4942519
## 41	41	0.4670043	0.4482333	0.4856206
## 42	42	0.4584115	0.4395458	0.4771399
## 43	43	0.4499769	0.4310268	0.4688074
## 44	44	0.4416974	0.4226729	0.4606204

Key Results

Baseline Hazard Rate:

The model's baseline hazard rate is $\text{rate}=0.0135$, meaning the average constant hazard rate is approximately 1.35% per unit of time.

Covariate Effects on Hazard:

disease_hypert ($\text{est}=-0.208$, $\text{exp}(\text{est})=0.812$):

Patients with hypertension (disease_hypert) have a hazard rate that is **18.8% lower** compared to those without hypertension.

This reduction in hazard translates to improved survival probability.

disease_renal ($\text{est}=-0.251$, $\text{exp}(\text{est})=0.778$):

Patients with renal disease (disease_renal) have a hazard rate that is **22.2% lower** compared to those without renal disease.

This also implies better survival for patients with renal disease.

begin (est=0.0234, exp(est)=1.023):

Each additional unit of "begin" increases the hazard rate by **2.3%**

This negatively impacts survival.

center (est=-0.00003, exp(est)=1.000)

The hazard rate decreases slightly with "center," but the effect is very small and may not be clinically significant.

Survival Probability Over Time

The survival probabilities decrease over time, as shown in the model's output:

Time Survival Probability (est) 95% Confidence Interval

1	0.9816	[0.9806, 0.9826]
5	0.9113	[0.9068, 0.9160]
10	0.8305	[0.8224, 0.8390]
20	0.6898	[0.6763, 0.7040]
40	0.4758	[0.4573, 0.4956]

At **1 unit of time**, the survival probability is **98.16%**, indicating most patients survive early on.

At **10 units of time**, survival probability decreases to **83.05%**.

At **40 units of time**, survival probability further declines to **47.58%**, showing that almost half of the population survives at this point.

Model Performance:

Log-likelihood: -8099.732, it is small indicating a better model fit.

AIC: 16209.4616209, used for comparison with other models

The exponential model is suitable for this data but did not capture time-varying hazard rates hence I went on to employ the Weibull model.

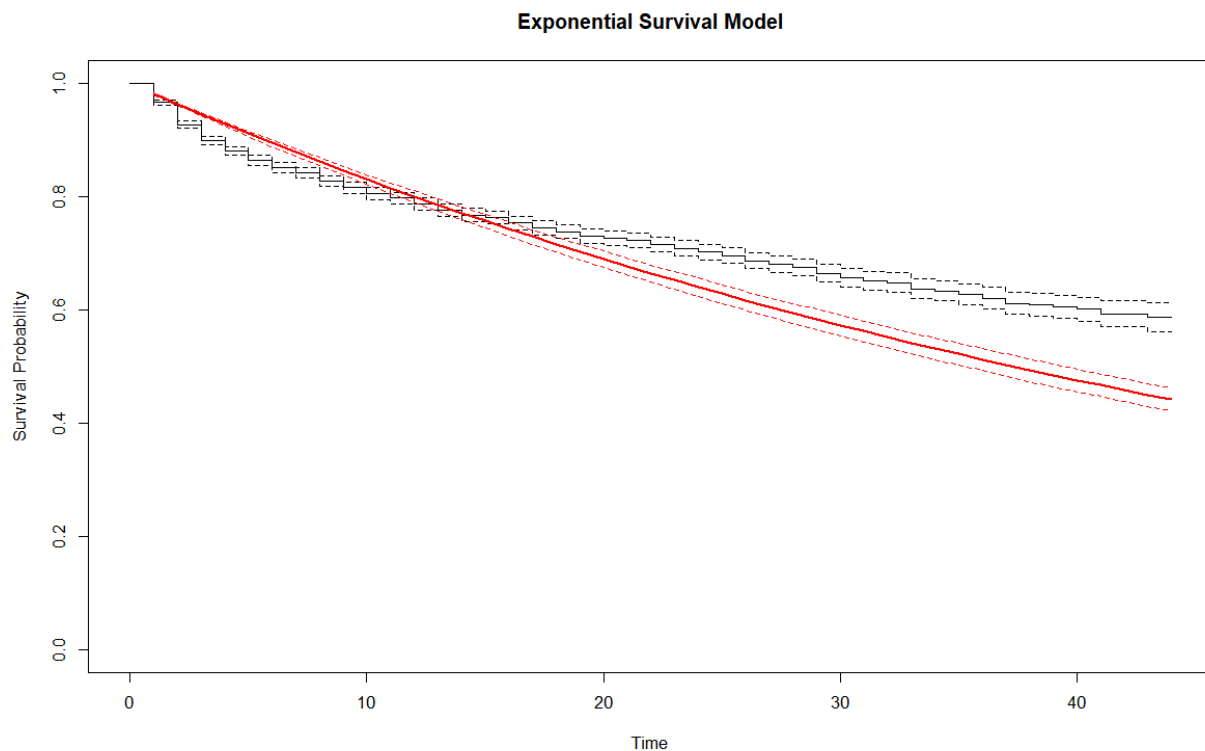
Hypertension and **renal disease** improve survival by reducing the hazard rate.

Begin increases hazard, reducing survival, while **center** has negligible impact.

Survival probabilities decrease progressively over time, with nearly half of the population surviving at 40 time units.

Visualize survival curves

```
plot(exp_model, main = "Exponential Survival Model",  
     xlab = "Time", ylab = "Survival Probability")
```



The survival curve from the plot shows a **monotonically decreasing trend**, consistent with the exponential model's assumption of constant hazard. Survival probabilities decline steadily over time, reflecting the risk factors' influence.

WEIBULL MODEL

Fit the weibull model

```
weibull_model <- flexsurvreg(Surv(time, event)~disease_hypert+disease_renal+  
                             begin+center, data = dialysis, dist =  
                             "weibull")
```

Get the summary of the model

weibull_model

```
## Call:
## flexsurvreg(formula = Surv(time, event) ~ disease_hypert + disease_renal +
##     begin + center, data = dialysis, dist = "weibull")
##
## Estimates:
##           data mean  est      L95%      U95%      se
## shape              NA  8.42e-01  8.06e-01  8.80e-01  1.89e-02
## scale              NA  8.25e+01  7.04e+01  9.66e+01  6.66e+00
## disease_hypert    4.17e-01  2.38e-01  1.09e-01  3.68e-01  6.61e-02
## disease_renal     2.08e-01  2.96e-01  1.33e-01  4.59e-01  8.32e-02
## begin             2.28e+01 -1.90e-02 -2.43e-02 -1.38e-02  2.68e-03
## center            2.55e+03  3.37e-05  9.15e-06  5.82e-05  1.25e-05
##           exp(est)  L95%      U95%
## shape              NA          NA
## scale              NA          NA
## disease_hypert    1.27e+00  1.11e+00  1.44e+00
## disease_renal     1.34e+00  1.14e+00  1.58e+00
## begin             9.81e-01  9.76e-01  9.86e-01
## center            1.00e+00  1.00e+00  1.00e+00
##
## N = 6805,  Events: 1603,  Censored: 5202
## Total time at risk: 96345
## Log-likelihood = -8068.356, df = 6
## AIC = 16148.71
```

summary(weibull_model)

```
##
disease_hypert=0.416752387950037,disease_renal=0.207788390889052,begin=22.777
5165319618,center=2553.10639235856
##   time      est      lcl      ucl
## 1    1 0.9719588 0.9685425 0.9747557
## 2    2 0.9502827 0.9455289 0.9542650
## 3    3 0.9307567 0.9251311 0.9356304
## 4    4 0.9126210 0.9062777 0.9183111
## 5    5 0.8955272 0.8886314 0.9018732
## 6    6 0.8792730 0.8718584 0.8862553
## 7    7 0.8637261 0.8558733 0.8713155
## 8    8 0.8487924 0.8403590 0.8569517
## 9    9 0.8344017 0.8256430 0.8429893
## 10   10 0.8204990 0.8113640 0.8296891
## 11   11 0.8070404 0.7975408 0.8167504
## 12   12 0.7939897 0.7839834 0.8042125
## 13   13 0.7813164 0.7704904 0.7919708
## 14   14 0.7689947 0.7574054 0.7800833
## 15   15 0.7570022 0.7448736 0.7685946
## 16   16 0.7453192 0.7332317 0.7573021
```

```

## 17    17 0.7339286 0.7211177 0.7463046
## 18    18 0.7228148 0.7094280 0.7356643
## 19    19 0.7119641 0.6977850 0.7251923
## 20    20 0.7013640 0.6866623 0.7149717
## 21    21 0.6910031 0.6761916 0.7052904
## 22    22 0.6808710 0.6653352 0.6958202
## 23    23 0.6709583 0.6547327 0.6863950
## 24    24 0.6612562 0.6446098 0.6773251
## 25    25 0.6517566 0.6347153 0.6682027
## 26    26 0.6424518 0.6250318 0.6592231
## 27    27 0.6333350 0.6155511 0.6503920
## 28    28 0.6243995 0.6062344 0.6419932
## 29    29 0.6156392 0.5969505 0.6335974
## 30    30 0.6070483 0.5878494 0.6253554
## 31    31 0.5986214 0.5789263 0.6172713
## 32    32 0.5903533 0.5701727 0.6096121
## 33    33 0.5822392 0.5615906 0.6016099
## 34    34 0.5742746 0.5531740 0.5938468
## 35    35 0.5664549 0.5449168 0.5864412
## 36    36 0.5587762 0.5368139 0.5792494
## 37    37 0.5512345 0.5288630 0.5721797
## 38    38 0.5438259 0.5211825 0.5650773
## 39    39 0.5365470 0.5136708 0.5579716
## 40    40 0.5293943 0.5062967 0.5509830
## 41    41 0.5223645 0.4988965 0.5443762
## 42    42 0.5154545 0.4914538 0.5379210
## 43    43 0.5086613 0.4843862 0.5315717
## 44    44 0.5019820 0.4774491 0.5252779

```

Model Overview

Dependent Variable: Time to an event (e.g., survival time in days).

Independent Variables:

disease_hypert: Presence of hypertension-related disease.

disease_renal: Presence of renal-related disease.

begin: The beginning value (context-dependent, perhaps starting dialysis date).

center: Likely a numeric variable related to the treatment center.

Distribution: Weibull, characterized by two parameters:

Shape (0.842): Determines whether the hazard rate increases, decreases, or stays constant over time. A value less than 1 indicates a **decreasing hazard** over time (subjects are less likely to experience the event as time progresses).

Scale (82.5): Represents a time measure related to the average survival time.

Key Parameter Estimates

disease_hypert (Estimate: 0.238, Exp(Estimate): 1.27):

Patients with hypertension-related diseases have a **27% higher risk of experiencing the event** (e.g., death or failure) compared to those without this condition, holding other factors constant.

Confidence Interval: The effect is statistically significant as the 95% CI (1.11–1.44) does not include 1.

disease_renal (Estimate: 0.296, Exp(Estimate): 1.34):

Interpretation: Patients with renal-related diseases have a **34% higher risk of experiencing the event**, compared to those without this condition, holding other factors constant.

Statistically significant (95% CI: 1.14–1.58).

begin (Estimate: -0.019, Exp(Estimate): 0.981):

Interpretation: A one-unit increase in the `begin` variable is associated with a **1.9% decrease in the hazard of the event**.

Statistically significant (95% CI: 0.976–0.986).

center (Estimate: 3.37e-05, Exp(Estimate): ~1):

Interpretation: The effect of `center` on the hazard rate is negligible, as its exponent is effectively

May not be practically significant (small effect).

Model Fit

Events: 1603 out of 6805 observations experienced the event.

Censored Data: 5202 observations did not experience the event within the observation period.

Log-likelihood: -8068.36

AIC: 16148.71 (used for model comparison; lower AIC is better).

Survival Probabilities

The survival probabilities ($S(t)$) for time points from 1 to 42 indicate the likelihood of surviving beyond each time point. For instance:

At **Time = 1**, the survival probability is 97.2% (high likelihood of survival early on).

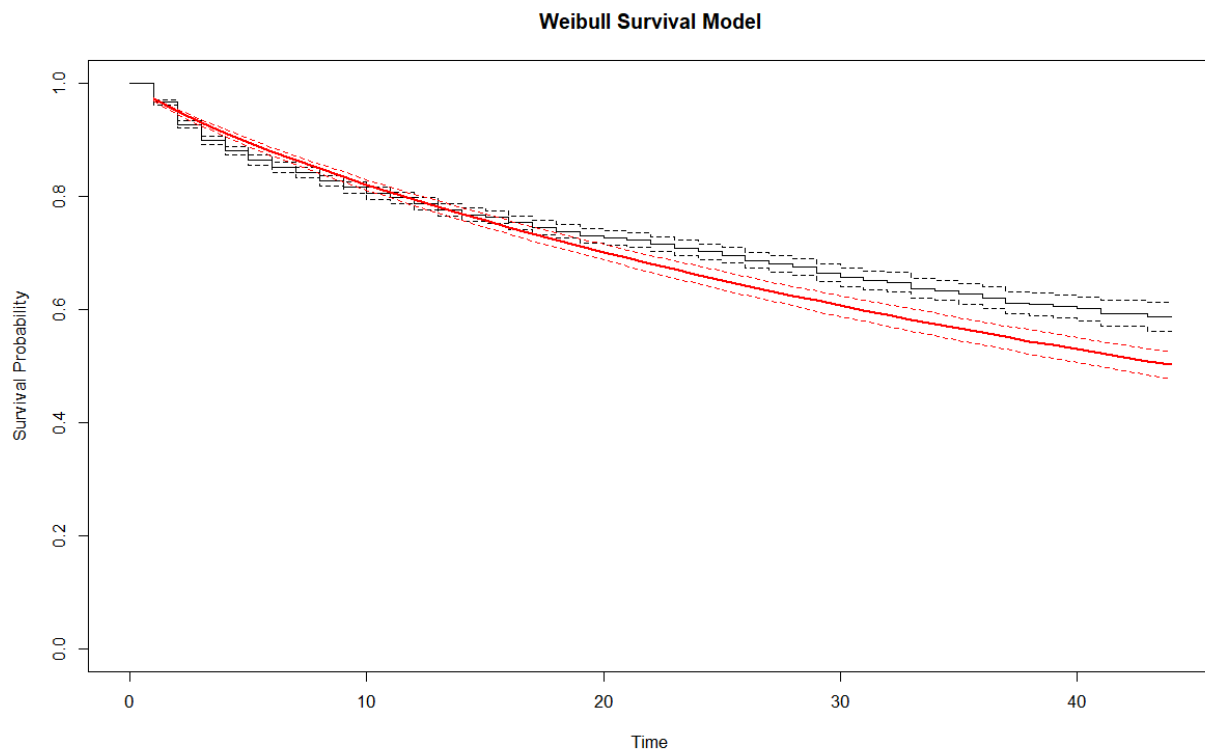
At **Time = 20**, the survival probability drops to 70.1%.

At **Time = 40**, the survival probability drops further to 52.9%.

This trend shows declining survival probabilities over time.

Visualize survival curves

```
plot(weibull_model, main = "Weibull Survival Model",  
     xlab = "Time", ylab = "Survival Probability")
```



The survival curve provides a graphical representation of the declining survival probabilities over time.

Interpretation: There is a steeper decline in the survival curve indicating faster reduction in survival probability over time, confirming the decreasing hazard suggested by the Weibull shape parameter (<1).