Final Project

Puyuan Zhang (pz2334)

2025-11-21

```
library(survival)
library(survminer)
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.5.2
```

```
## Loading required package: ggpubr
```

```
##
## Attaching package: 'survminer'
```

```
## The following object is masked from 'package:survival':
##
##     myeloma
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.5.2
```

```
## Warning: package 'readr' was built under R version 4.5.2
```

```
## Warning: package 'stringr' was built under R version 4.5.2
```

```
## Warning: package 'lubridate' was built under R version 4.5.2
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats   1.0.1     v stringr   1.6.0
## v lubridate 1.9.4     v tibble    3.3.0
## v purrr     1.1.0     v tidyr     1.3.1
## v readr     2.1.6
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(skimr)
```

```
## Warning: package 'skimr' was built under R version 4.5.2
```

```
library(janitor)
```

```
## Warning: package 'janitor' was built under R version 4.5.2
```

```
##
## Attaching package: 'janitor'
##
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```
library(ggplot2)
```

```
cirrhosis <- read.csv("data/cirrhosis.csv") |> clean_names()
```

```
# Status convert to event indicator (1 = death, 0 = censored)
cirrhosis <- cirrhosis |>
  mutate(
    event = case_when(
      status == "D" ~ 1,
      status %in% c("C", "CL") ~ 0,
      TRUE ~ NA_real_
    ),
    sex = factor(sex),
    drug = factor(drug),
    ascites = factor(ascites),
    hepatomegaly = factor(hepatomegaly),
    spiders = factor(spiders),
    edema = factor(edema, levels = c("N", "S", "Y"), ordered = TRUE)
  )

cirrhosis <- cirrhosis |>
  mutate(age_years = age / 365.25)
```

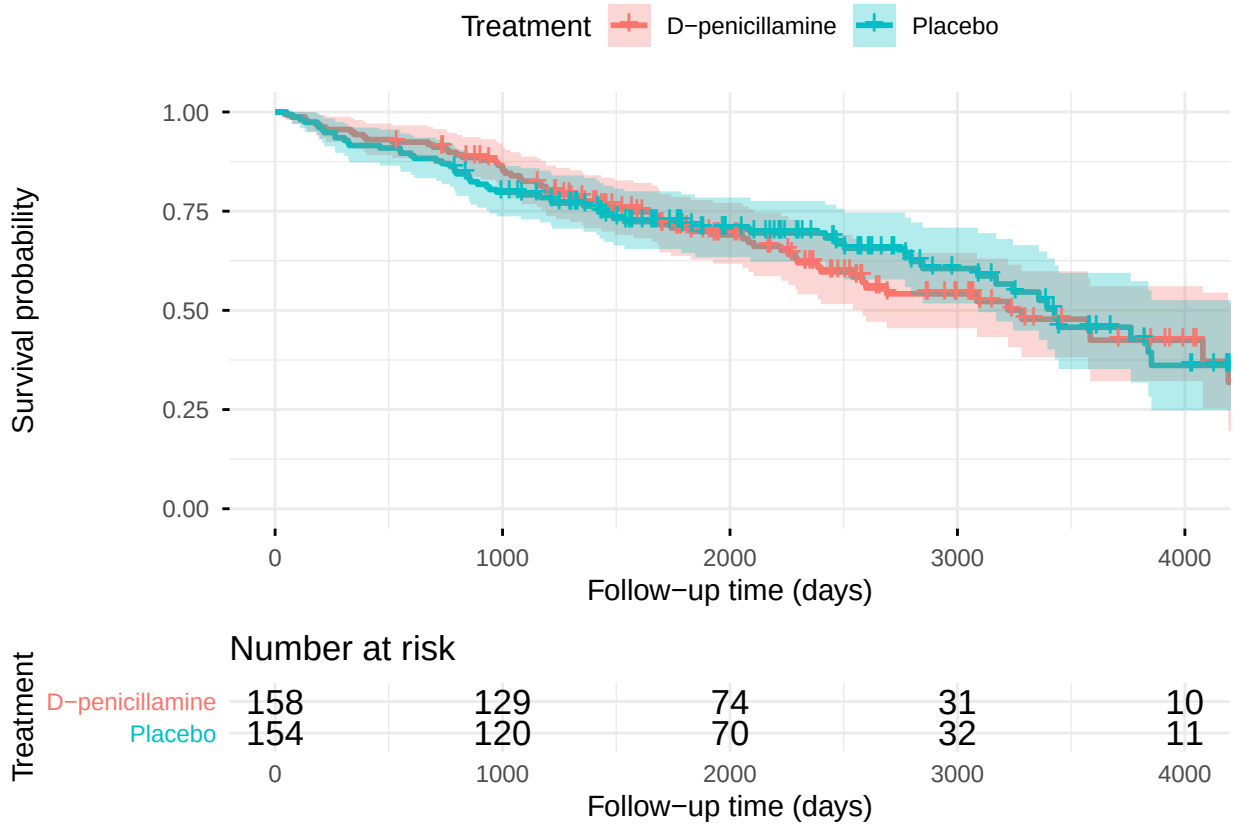
```r
km_dat <- cirrhosis |>
  filter(!is.na(event), !is.na(n_days), !is.na(drug)) %>%
  mutate(
    age_group = ntile(age_years, 3),
    age_group = factor(age_group, labels = c("Low", "Medium", "High"))
  )  %>%
  mutate(
    bili_q = ntile(bilirubin, 4),
    bili_q = factor(bili_q,
                    labels = c("Q1 (lowest)", "Q2", "Q3", "Q4 (highest)"))
  ) %>%
  mutate(
    alb_t = ntile(albumin, 3),
    alb_t = factor(alb_t,
                   labels = c("Low (T1)", "Medium (T2)", "High (T3)"))
  )

fit_km <- survfit(Surv(n_days, event) ~ drug, data = km_dat)

ggsurvplot(
  fit_km,
  data = km_dat,
  risk.table = TRUE,
  conf.int = TRUE,
  xlab = "Follow-up time (days)",
  ylab = "Survival probability",
  legend.title = "Treatment",
  legend.labs = c("D-penicillamine", "Placebo"),
  ggtheme = theme_minimal()
)
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## i The deprecated feature was likely used in the ggpubr package.
##   Please report the issue at <https://github.com/kassambara/ggpubr/issues>.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```

```
## Ignoring unknown labels:
## * colour : "Treatment"
```



The Kaplan–Meier survival curves for the D-penicillamine and placebo groups showed nearly identical trajectories over the 4,000-day follow-up period. The two curves overlapped substantially, with only mild and random fluctuations across time. This pattern indicates that, in the unadjusted analysis, there was no visible survival advantage associated with D-penicillamine treatment.

The number-at-risk tables demonstrated comparable follow-up patterns in both groups, suggesting that differences in censoring did not explain the similarity in survival curves. Together, the visual evidence indicates that D-penicillamine did not meaningfully improve overall survival compared with placebo in this dataset.

```
logrank <- survdiff(Surv(n_days, event) ~ drug, data = km_dat)
logrank
```

```
## Call:
## survdiff(formula = Surv(n_days, event) ~ drug, data = km_dat)
##
##                         N Observed Expected (O-E)^2/E (O-E)^2/V
## drug=D-penicillamine 158       65     63.2    0.0502     0.102
## drug=Placebo         154       60     61.8    0.0513     0.102
##
##  Chisq= 0.1  on 1 degrees of freedom, p= 0.7
```

```
p_val <- 1 - pchisq(logrank$chisq, length(logrank$n) - 1)
p_val
```

```
## [1] 0.7497925
```

According to the MH logrank test, since we have p-value($p = 0.75$) much larger than 0.05, we can conclude that there is not significant difference between different treatment group.

```
cox_unadj <- coxph(Surv(n_days, event) ~ drug, data = km_dat)
summary(cox_unadj)
```

```
## Call:
## coxph(formula = Surv(n_days, event) ~ drug, data = km_dat)
##
##   n= 312, number of events= 125
##
##                 coef exp(coef) se(coef)      z Pr(>|z|)
## drugPlacebo -0.05722   0.94438  0.17916 -0.319    0.749
##
##             exp(coef) exp(-coef) lower .95 upper .95
## drugPlacebo    0.9444      1.059    0.6647     1.342
##
## Concordance= 0.499  (se = 0.025 )
## Likelihood ratio test= 0.1  on 1 df,   p=0.7
## Wald test            = 0.1  on 1 df,   p=0.7
## Score (logrank) test = 0.1  on 1 df,   p=0.7
```

```
# HR and CI
HR <- exp(coef(cox_unadj))
CI <- exp(confint(cox_unadj))
p_wald <- summary(cox_unadj)$coefficients[,"Pr(>|z|)"]

# combine
result <- data.frame(
  HR = HR,
  CI_lower = CI[1],
  CI_upper = CI[2],
  logrank_p = p_val,
  wald_p = p_wald
)
result
```

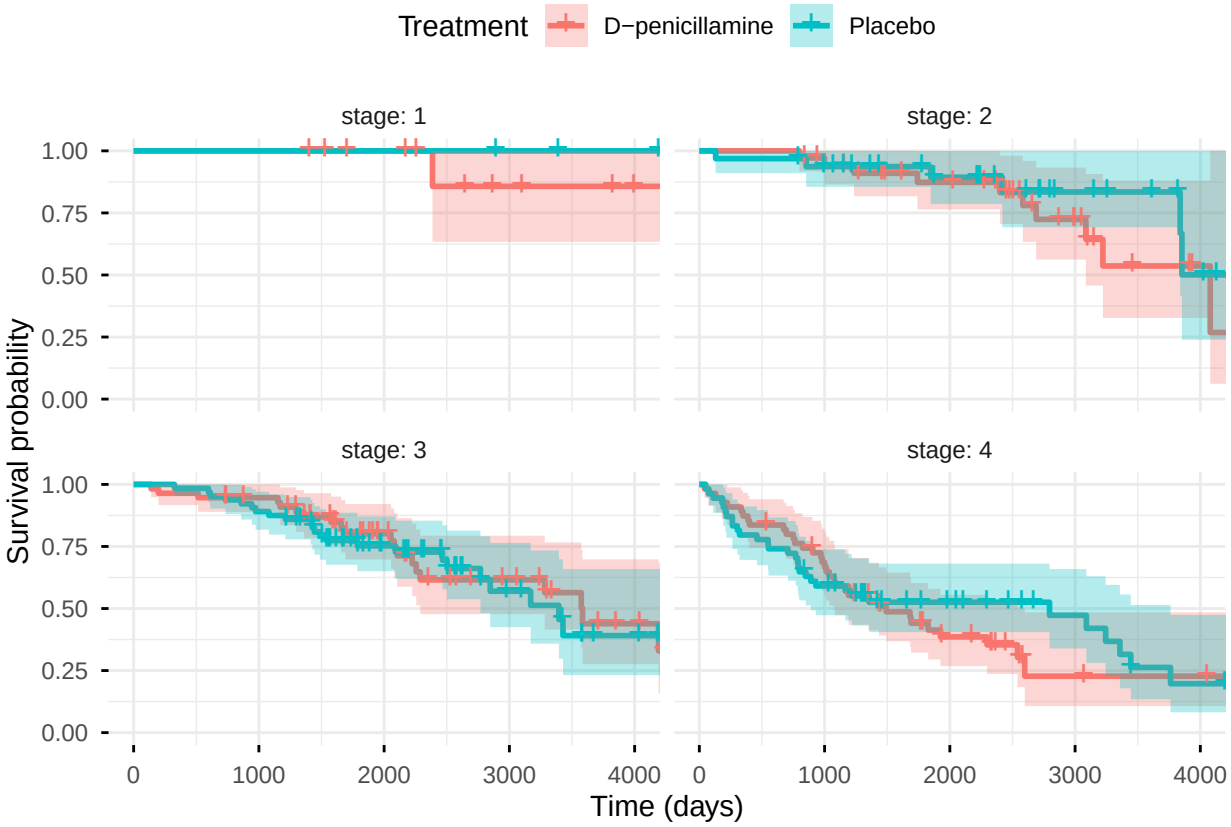
```
##                     HR CI_lower CI_upper logrank_p    wald_p
## drugPlacebo 0.9443827 0.664726 1.341694 0.7497925 0.7494294
```

The unadjusted Cox proportional hazards model yielded a hazard ratio of 0.94 (95% CI: 0.66–1.35, $p = 0.75$), indicating no evidence that D-penicillamine improved survival compared with placebo. The confidence interval includes 1, and the effect size is small, suggesting that any potential treatment effect is likely negligible in magnitude.

```
fit_stage <- survfit(Surv(n_days, event) ~ drug + stage, data = km_dat)
ggsurvplot(
  fit_stage,
  data = km_dat,
  facet.by = "stage",
  conf.int = T,
  risk.table = TRUE,
```

```
  legend.title = "Treatment",
  legend.labs = c("D-penicillamine", "Placebo"),
  xlab = "Time (days)",
  ylab = "Survival probability",
  ggtheme = theme_minimal()
)
```

```
## Warning in (function (survsummary, times, survtable = c("cumevents",
## "risk.table", : The length of legend.labs should be 8
```



The stratified Kaplan–Meier curves demonstrate clear differences in overall survival across disease stages, confirming that stage is a major prognostic factor in this cohort. As expected, patients in stages 3 and 4 experienced markedly poorer survival compared with stages 1 and 2.

When comparing treatment groups within each stage, the survival trajectories for D-penicillamine and placebo remained largely overlapping.

Stage 1: Very few events occurred, and both curves remained close to 1.0 throughout follow-up, making treatment differences difficult to assess.

Stage 2: Both groups showed similar gradual declines, with no consistent separation between curves.

Stage 3: The curves nearly overlapped for the entire follow-up period, indicating no observable treatment benefit in this intermediate-risk subgroup.

Stage 4: Although mortality was highest in this group, the D-penicillamine and placebo curves again followed similar patterns without evidence of divergence.

Overall, stage-stratified KM curves suggest that disease severity strongly predicts survival, but the treatment effect does not vary meaningfully across stages. No stage subgroup demonstrated a survival advantage for D-penicillamine.

```
logrank_stage <- survdiff(Surv(n_days, event) ~ drug + strata(stage), data = km_dat)
logrank_stage
```

```
## Call:
## survdiff(formula = Surv(n_days, event) ~ drug + strata(stage),
##     data = km_dat)
##
##                         N Observed Expected (O-E)^2/E (O-E)^2/V
## drug=D-penicillamine 158       65     61.8     0.169     0.343
## drug=Placebo         154       60     63.2     0.165     0.343
##
##  Chisq= 0.3  on 1 degrees of freedom, p= 0.6
```

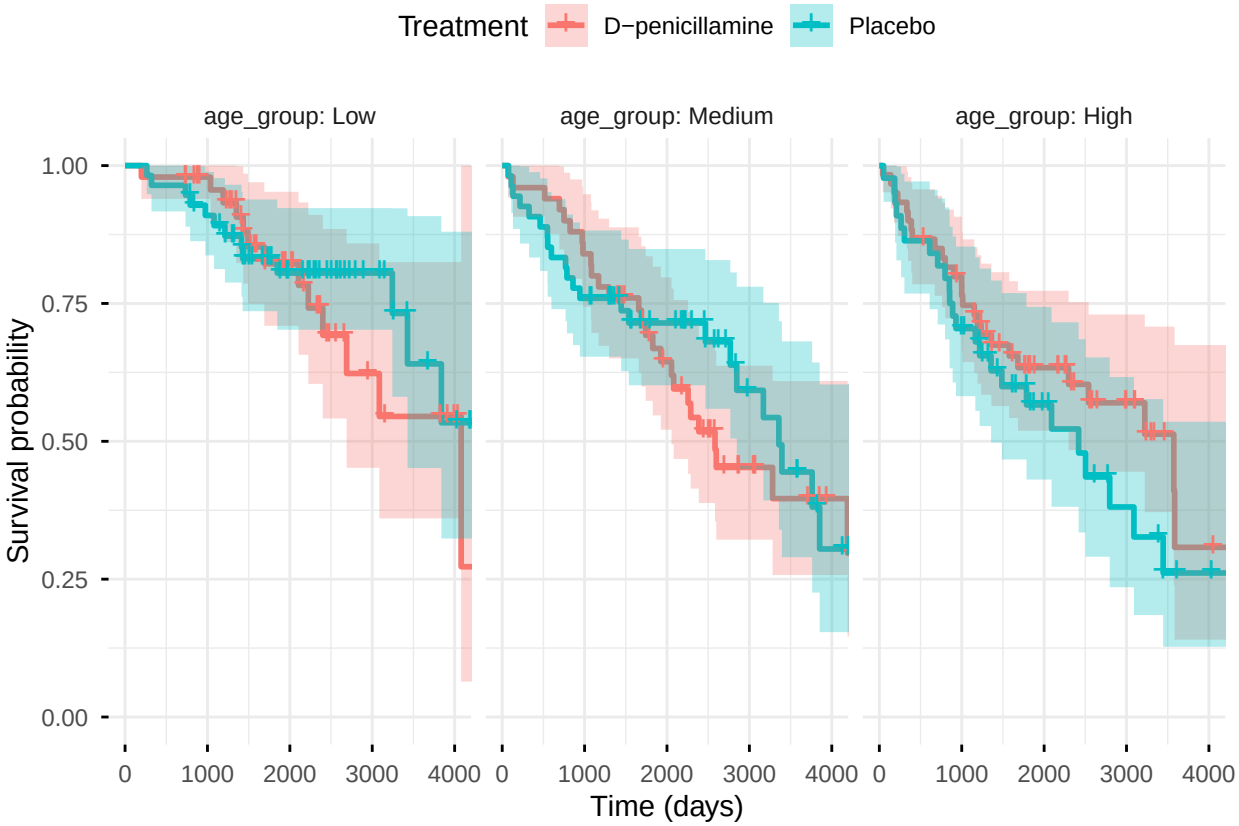
```
p_val_stage <- 1 - pchisq(logrank_stage$chisq, length(logrank_stage$n) - 1)
p_val_stage
```

```
## [1] 0.5578978
```

Because disease stage is a strong predictor of survival, we performed a stratified log-rank test to compare treatment groups while controlling for stage. The p-value for this test is about 0.56, which is also much higher than 0.05, indicating no significant difference in survival probability between different treatment among stages.

```
fit_age <- survfit(Surv(n_days, event) ~ drug + age_group, data = km_dat)
ggsurvplot(
  fit_age,
  data = km_dat,
  facet.by = "age_group",
  conf.int = T,
  risk.table = TRUE,
  legend.title = "Treatment",
  legend.labs = c("D-penicillamine", "Placebo"),
  xlab = "Time (days)",
  ylab = "Survival probability",
  ggtheme = theme_minimal()
)
```

```
## Warning in (function (survsummary, times, survtable = c("cumevents",
## "risk.table", : The length of legend.labs should be 6
```



When we divided patients into three age groups (Low, Medium, High), we saw that age is strongly related to survival. Younger patients lived longer, and older patients had the highest risk of death.

However, within each age group, the D-penicillamine and placebo survival curves look almost the same.

In the Low age group, both treatments showed very similar survival patterns.

In the Medium age group, the curves almost completely overlap.

In the High age group, survival is poorer overall, but again the two treatments look very similar.

This means that age does not change the treatment effect. No matter if patients are young, middle-aged, or older, D-penicillamine does not show a survival benefit.

```
logrank_age <- survdiff(Surv(n_days, event) ~ drug + strata(age_group), data = km_dat)
logrank_age
```

```
## Call:
## survdiff(formula = Surv(n_days, event) ~ drug + strata(age_group),
##     data = km_dat)
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## drug=D-penicillamine 158       65     65.2  0.000509   0.00108
## drug=Placebo         154       60     59.8  0.000554   0.00108
##
##  Chisq= 0  on 1 degrees of freedom, p= 1
```

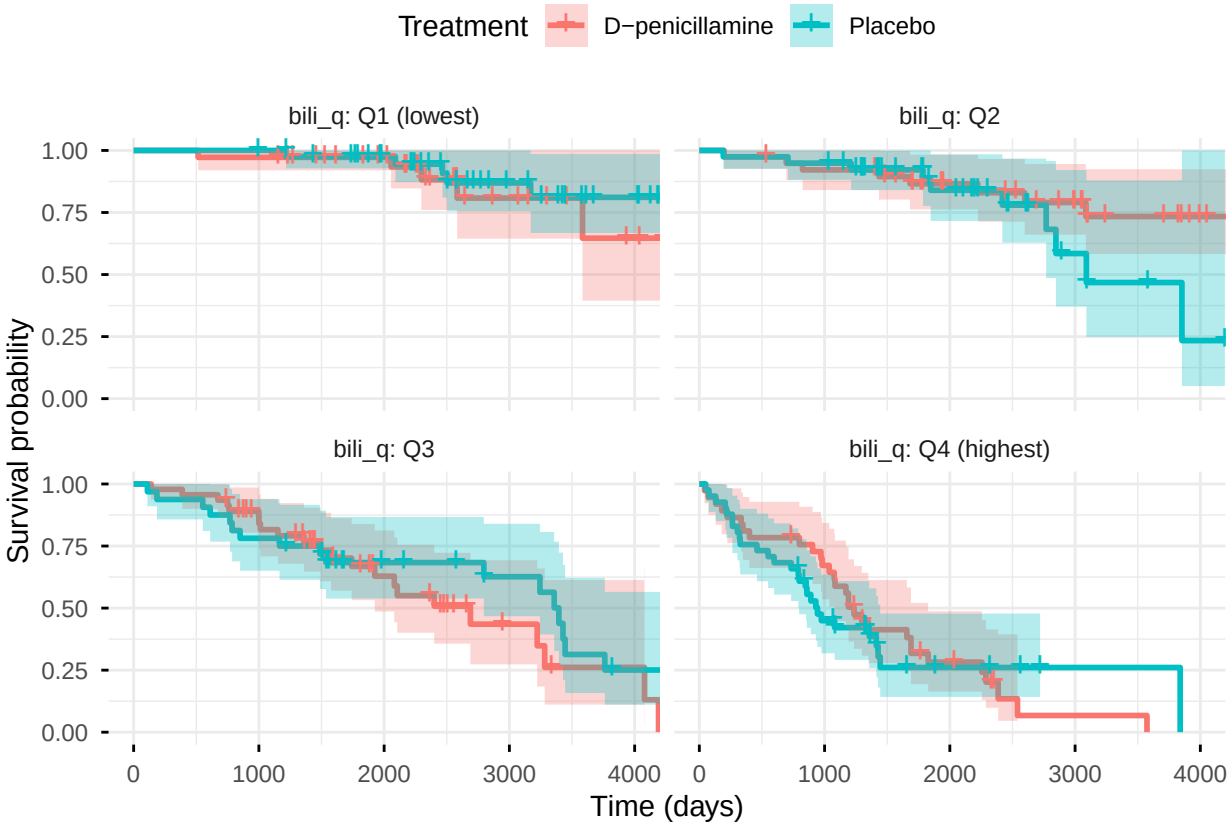
```r
p_val_age <- 1 - pchisq(logrank_age$chisq, length(logrank_age$n) - 1)
p_val_age
```

```
## [1] 0.9737294
```

After adjusting for age, there is still no difference in survival between D-penicillamine and placebo. The stratified log-rank test gives p = 1, which means the groups are almost identical.

```r
fit_bili <- survfit(Surv(n_days, event) ~ drug + bili_q, data = km_dat)
ggsurvplot(
  fit_bili,
  data = km_dat,
  facet.by = "bili_q",
  conf.int = T,
  risk.table = TRUE,
  legend.title = "Treatment",
  legend.labs = c("D-penicillamine", "Placebo"),
  xlab = "Time (days)",
  ylab = "Survival probability",
  ggtheme = theme_minimal()
)
```

```
## Warning in (function (survsummary, times, survtable = c("cumevents",
## "risk.table", : The length of legend.labs should be 8
```



When we divided patients into four bilirubin groups (Q1 = lowest, Q4 = highest), we observed clear differences in baseline survival. Patients with higher bilirubin levels (Q3 and especially Q4) had much poorer survival.

However, within each bilirubin group, the survival curves of the D-penicillamine and placebo groups remained very similar. These patterns show that bilirubin strongly predicts prognosis, but it does not modify or influence the treatment effect.

```
logrank_bili <- survdiff(Surv(n_days, event) ~ drug + strata(bili_q), data = km_dat)
logrank_bili
```

```
## Call:
## survdiff(formula = Surv(n_days, event) ~ drug + strata(bili_q),
##     data = km_dat)
##
##                          N Observed Expected (O-E)^2/E (O-E)^2/V
## drug=D-penicillamine 158       65     64.4   0.00627    0.0137
## drug=Placebo         154       60     60.6   0.00665    0.0137
##
##  Chisq= 0  on 1 degrees of freedom, p= 0.9
```

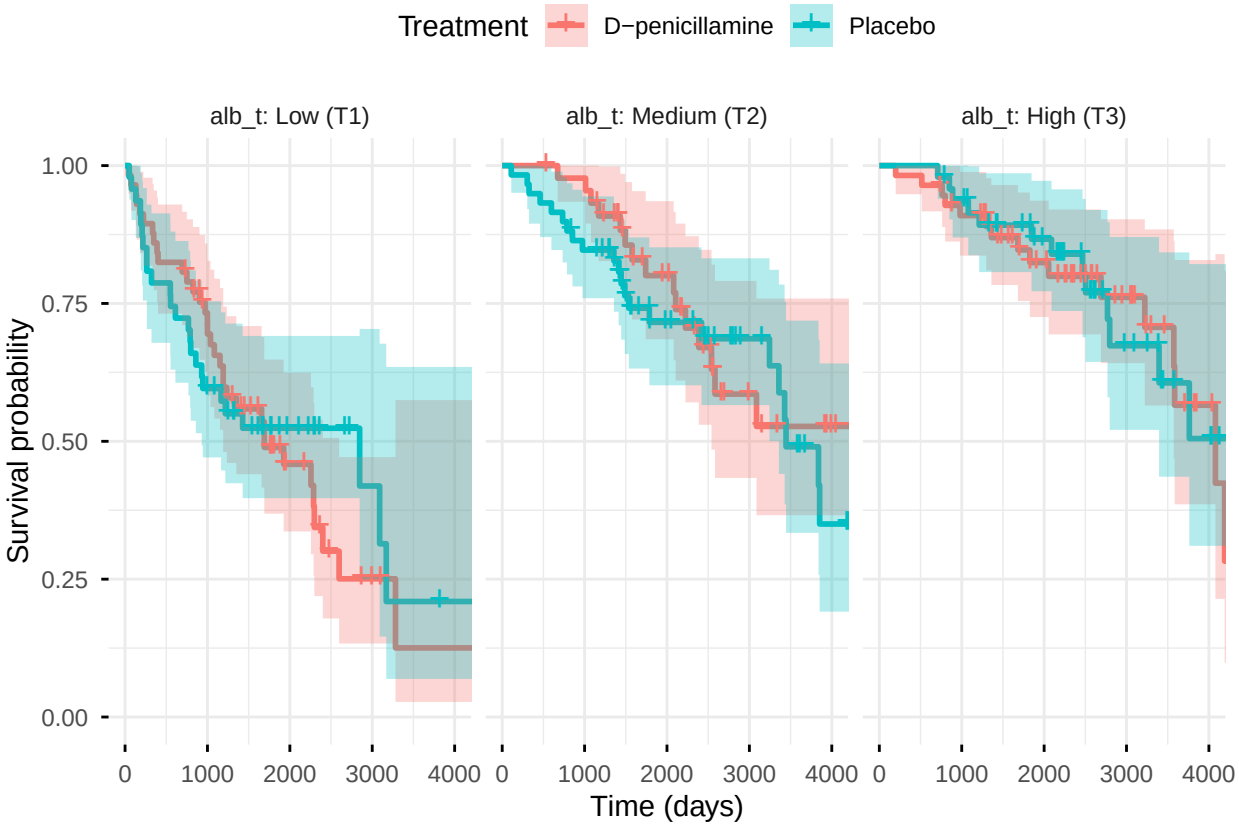
```
p_val_bili <- 1 - pchisq(logrank_bili$chisq, length(logrank_bili$n) - 1)
p_val_bili
```

```
## [1] 0.9067018
```

The stratified logrank test support the conclusion above which indicate that there is no significant difference in survival between D-penicillamine and placebo among different bilirubin group.

```
fit_alb <- survfit(Surv(n_days, event) ~ drug + alb_t, data = km_dat)
ggsurvplot(
  fit_alb,
  data = km_dat,
  facet.by = "alb_t",
  conf.int = T,
  risk.table = TRUE,
  legend.title = "Treatment",
  legend.labs = c("D-penicillamine", "Placebo"),
  xlab = "Time (days)",
  ylab = "Survival probability",
  ggtheme = theme_minimal()
)
```

```
## Warning in (function (survsummary, times, survtable = c("cumevents",
## "risk.table", : The length of legend.labs should be 6
```



Patients with lower albumin have worse survival, and patients with higher albumin live longer. But within each albumin group, the D-penicillamine and placebo curves are almost the same. This means albumin affects prognosis, but it does not change the treatment effect — the drug still shows no benefit in any albumin level.

```
logrank_alb <- survdiff(Surv(n_days, event) ~ drug + strata(alb_t), data = km_dat)
logrank_alb
```

```
## Call:
## survdiff(formula = Surv(n_days, event) ~ drug + strata(alb_t),
##     data = km_dat)
##
##                         N Observed Expected (O-E)^2/E (O-E)^2/V
## drug=D-penicillamine 158       65     65.1  0.000260   0.00055
## drug=Placebo         154       60     59.9  0.000283   0.00055
##
##  Chisq= 0  on 1 degrees of freedom, p= 1
```

```
p_val_alb <- 1 - pchisq(logrank_alb$chisq, length(logrank_alb$n) - 1)
p_val_alb
```

```
## [1] 0.9812976
```

The stratified logrank test support the conclusion above which indicate that there is no significant difference in survival between D-penicillamine and placebo among different albumin group.