

Survival Analysis

Asad

2023-05-10

Calling libraries

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

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

```
## 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)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v tibble  3.2.1    v purrr   1.0.1
```

```
## v tidyr   1.3.0    v stringr 1.5.0
```

```
## v readr   2.1.2    v forcats 1.0.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

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

```
library(ggsurvfit)
library(gtsummary)
library(tidycmprsk)
```

```
##
## Attaching package: 'tidycmprsk'

## The following object is masked from 'package:gtsummary':
##
## trial
```

```
library(condsurv)
```

Reading data

```
setwd('E:/R-Programming-Practices/Data Visualization/KM Plot')
data<- read.csv('Lung_surv.csv')
```

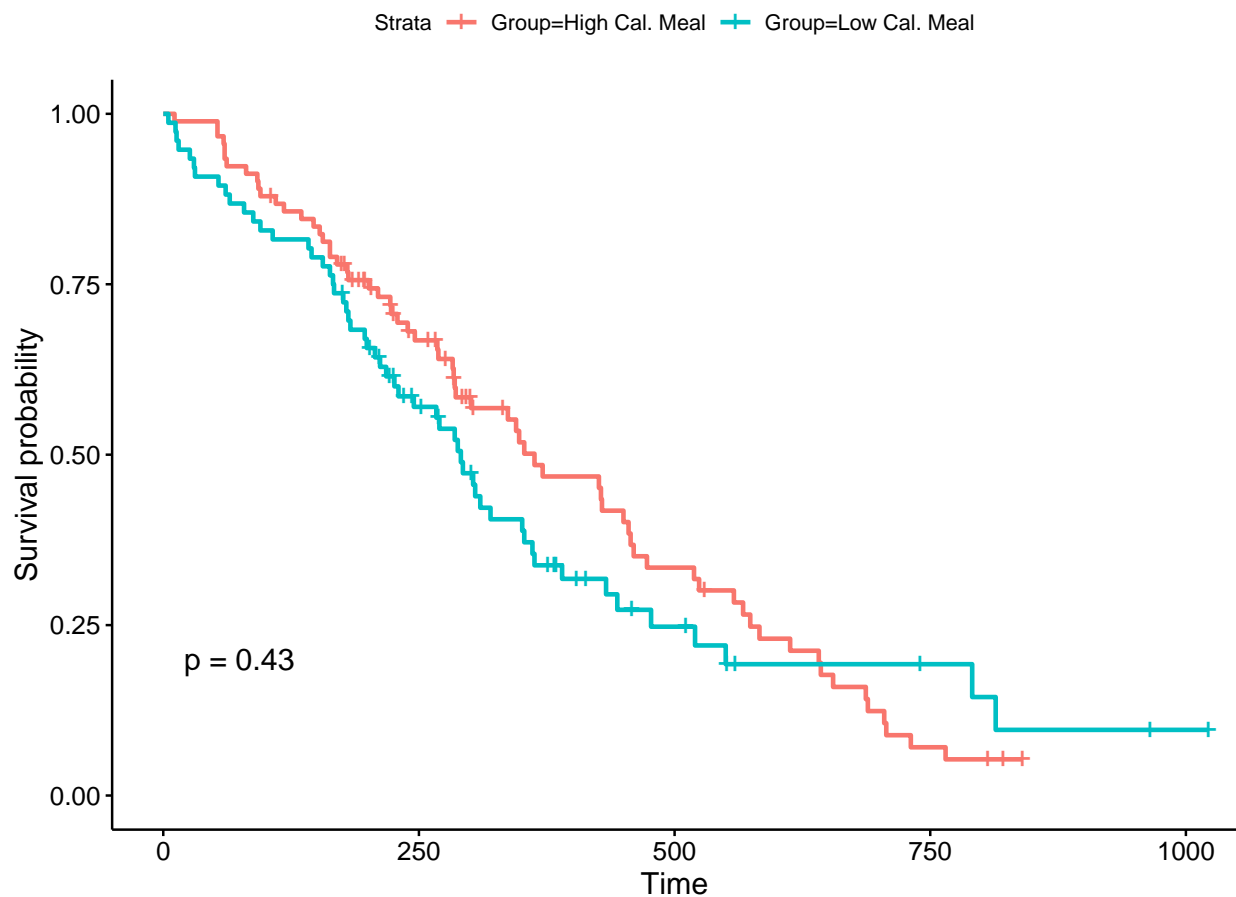
Fit model and prepare basic KM plot

```
fit<- survfit(Surv(Time, Status)~Group, data=data)
fit
```

```
## Call: survfit(formula = Surv(Time, Status) ~ Group, data = data)
##
##              n events median 0.95LCL 0.95UCL
## Group=High Cal. Meal 91      66    363    286    460
## Group=Low Cal. Meal 76      54    291    226    361
```

Plot KM plot

```
ggsurvplot(fit, data=data, pval = T)
```



Summary

#With the summary function we can compare the survival probability of two groups #at any given time. The time variable in this dataset is in months and we'll calculate #the survival probability in one year.

```
summary(fit, times = 365)
```

```
## Call: survfit(formula = Surv(Time, Status) ~ Group, data = data)
##
##           Group=High Cal. Meal
##      time      n.risk      n.event      survival      std.err lower 95% CI
##  365.0000      29.0000      41.0000       0.4847       0.0586       0.3825
## upper 95% CI
##    0.6143
##
##           Group=Low Cal. Meal
##      time      n.risk      n.event      survival      std.err lower 95% CI
##  365.0000      20.0000      46.0000       0.3377       0.0589       0.2399
## upper 95% CI
##    0.4753
```

#Alternatively

```
tbl_survfit(fit, times = 365, label_header = "**1-year survival (95% CI)**")
```

```
## Table printed with 'knitr::kable()', not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Characteristic	1-year survival (95% CI)
Group	
High Cal. Meal	48% (38%, 61%)
Low Cal. Meal	34% (24%, 48%)

Producing the median survival table for two groups

```
tbl_survfit(fit, probs = 0.5, label_header = "**Median survival (95% CI)**")
```

```
## Table printed with 'knitr::kable()', not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Characteristic	Median survival (95% CI)
Group	
High Cal. Meal	363 (286, 460)
Low Cal. Meal	291 (226, 361)

Performing log-rank t test between two groups

```
survdifff(formula = Surv(Time, Status) ~ Group, data = data)
```

```
## Call:
## survdifff(formula = Surv(Time, Status) ~ Group, data = data)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## Group=High Cal. Meal 91      66      70.2      0.256      0.628
## Group=Low Cal. Meal  76      54      49.8      0.362      0.628
##
## Chisq= 0.6  on 1 degrees of freedom, p= 0.4
```

Fit cox-regression model

```
cox<-coxph(Surv(Time, Status) ~ Group, data = data)
tbl_regression(cox, exp = TRUE)
```

```
## Table printed with 'knitr::kable()', not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Characteristic	HR	95% CI	p-value
Group			
High Cal. Meal	—	—	
Low Cal. Meal	1.16	0.81, 1.66	0.4

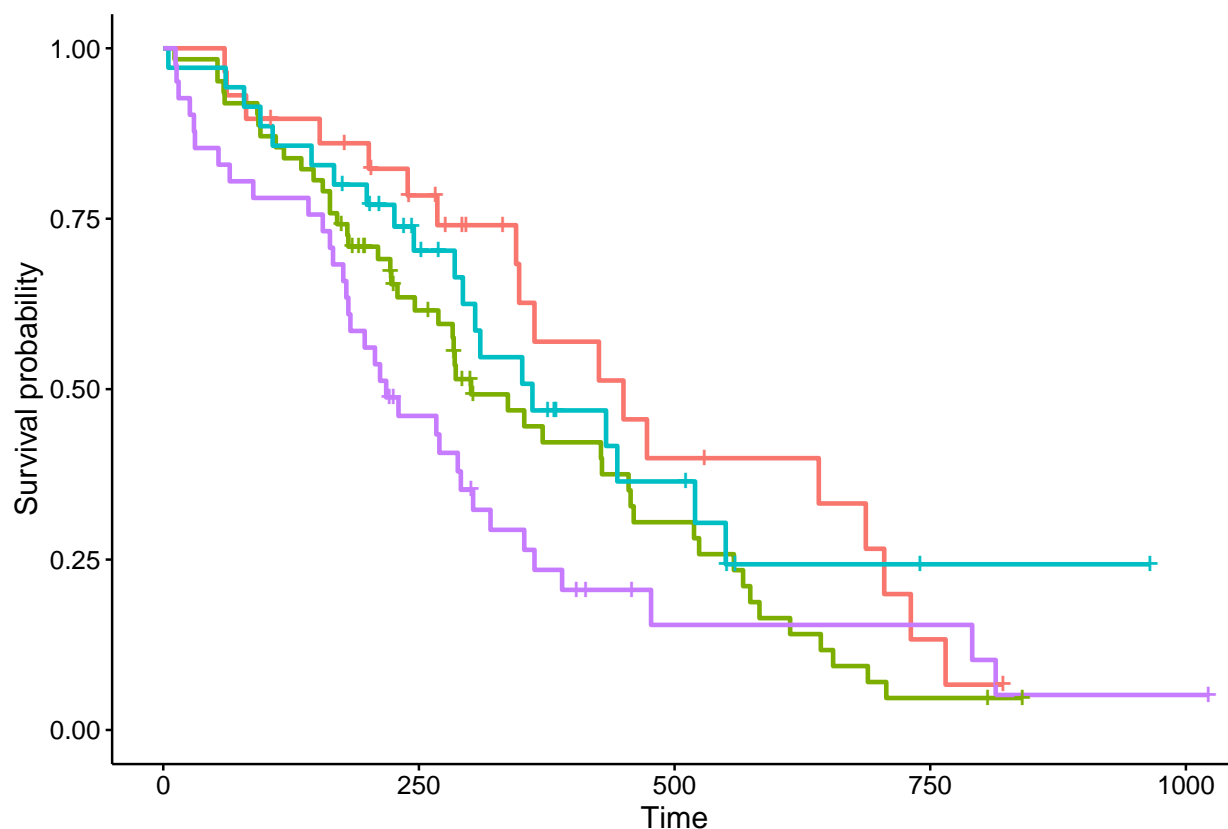
Checking survival trend among multiple groups

```
fit2<- survfit(Surv(Time, Status)~Group+Sex, data=data)
fit2
```

```
## Call: survfit(formula = Surv(Time, Status) ~ Group + Sex, data = data)
##
##
##           n events median 0.95LCL 0.95UCL
## Group=High Cal. Meal, Sex=Female 29      18    450      348      731
## Group=High Cal. Meal, Sex=Male   62      48    301      269      457
## Group=Low Cal. Meal, Sex=Female  35      20    361      293      NA
## Group=Low Cal. Meal, Sex=Male   41      34    218      181      320
```

```
ggsurvplot(fit2, data = data)
```

up=High Cal. Meal, Sex=Female + Group=High Cal. Meal, Sex=Male + Group=Low Cal. Meal, Sex=Female + Group=L



Cox model on multiple groups

```
cox2<-coxph(Surv(Time, Status) ~ Group+Sex+Race+Stage+
             Smoking.Habit+Metastasis.Status, data = data)
cox2
```

```
## Call:
## coxph(formula = Surv(Time, Status) ~ Group + Sex + Race + Stage +
##       Smoking.Habit + Metastasis.Status, data = data)
##
##               coef exp(coef) se(coef)      z      p
## GroupLow Cal. Meal      0.2399   1.2712  0.1995  1.203 0.22909
## SexMale                 0.5989   1.8200  0.2151  2.784 0.00536
## RaceAsian              -0.3212   0.7253  0.2670 -1.203 0.22894
## RaceCaucasian          -0.2804   0.7555  0.2739 -1.024 0.30602
## RaceLatino             -0.4419   0.6428  0.3242 -1.363 0.17281
## StageStage II          -0.1132   0.8930  0.2759 -0.410 0.68172
## StageStage III          0.2424   1.2743  0.3244  0.747 0.45501
## StageStage IV          -0.1132   0.8930  0.2967 -0.382 0.70282
## Smoking.HabitReformed Smoker -0.1702  0.8435  0.2471 -0.689 0.49107
## Smoking.HabitSmoker     -0.1436  0.8663  0.2516 -0.571 0.56831
```

```
## Metastasis.StatusN1      -0.3414    0.7108    0.2725 -1.253 0.21022
## Metastasis.StatusN2      -0.1216    0.8855    0.2778 -0.438 0.66158
## Metastasis.StatusN3      -0.3387    0.7127    0.2439 -1.389 0.16482
##
## Likelihood ratio test=15.23 on 13 df, p=0.2931
## n= 167, number of events= 120
```

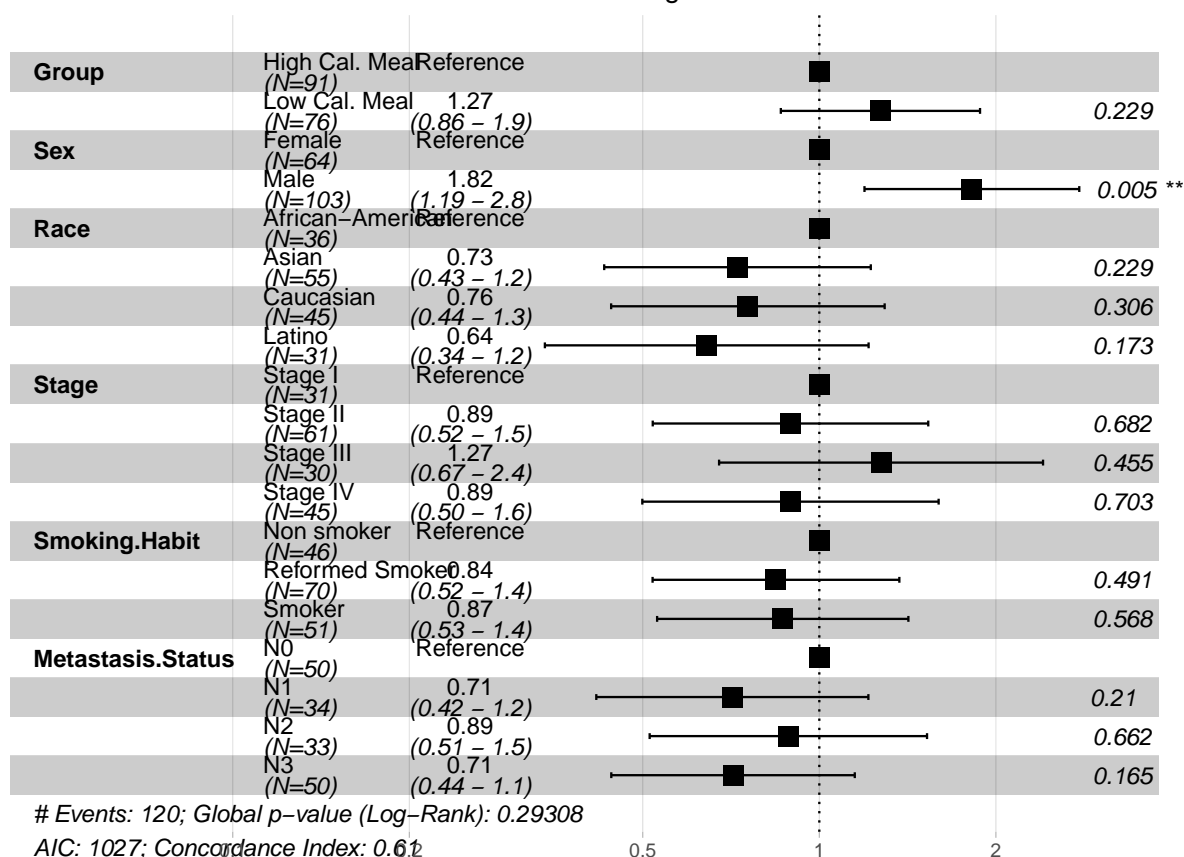
```
tbl_regression(cox2, exp = TRUE)
```

```
## Table printed with 'knitr::kable()', not {gt}. Learn why at
## https://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.
```

Characteristic	HR	95% CI	p-value
Group			
High Cal. Meal	—	—	
Low Cal. Meal	1.27	0.86, 1.88	0.2
Sex			
Female	—	—	
Male	1.82	1.19, 2.77	0.005
Race			
African-American	—	—	
Asian	0.73	0.43, 1.22	0.2
Caucasian	0.76	0.44, 1.29	0.3
Latino	0.64	0.34, 1.21	0.2
Stage			
Stage I	—	—	
Stage II	0.89	0.52, 1.53	0.7
Stage III	1.27	0.67, 2.41	0.5
Stage IV	0.89	0.50, 1.60	0.7
Smoking.Habit			
Non smoker	—	—	
Reformed Smoker	0.84	0.52, 1.37	0.5
Smoker	0.87	0.53, 1.42	0.6
Metastasis.Status			
N0	—	—	
N1	0.71	0.42, 1.21	0.2
N2	0.89	0.51, 1.53	0.7
N3	0.71	0.44, 1.15	0.2

```
ggforest(cox2, data = data, fontsize = 1, refLabel = 'Reference'
, main = 'Forest Plot of Cox-regression')
```

Forest Plot of Cox-regression

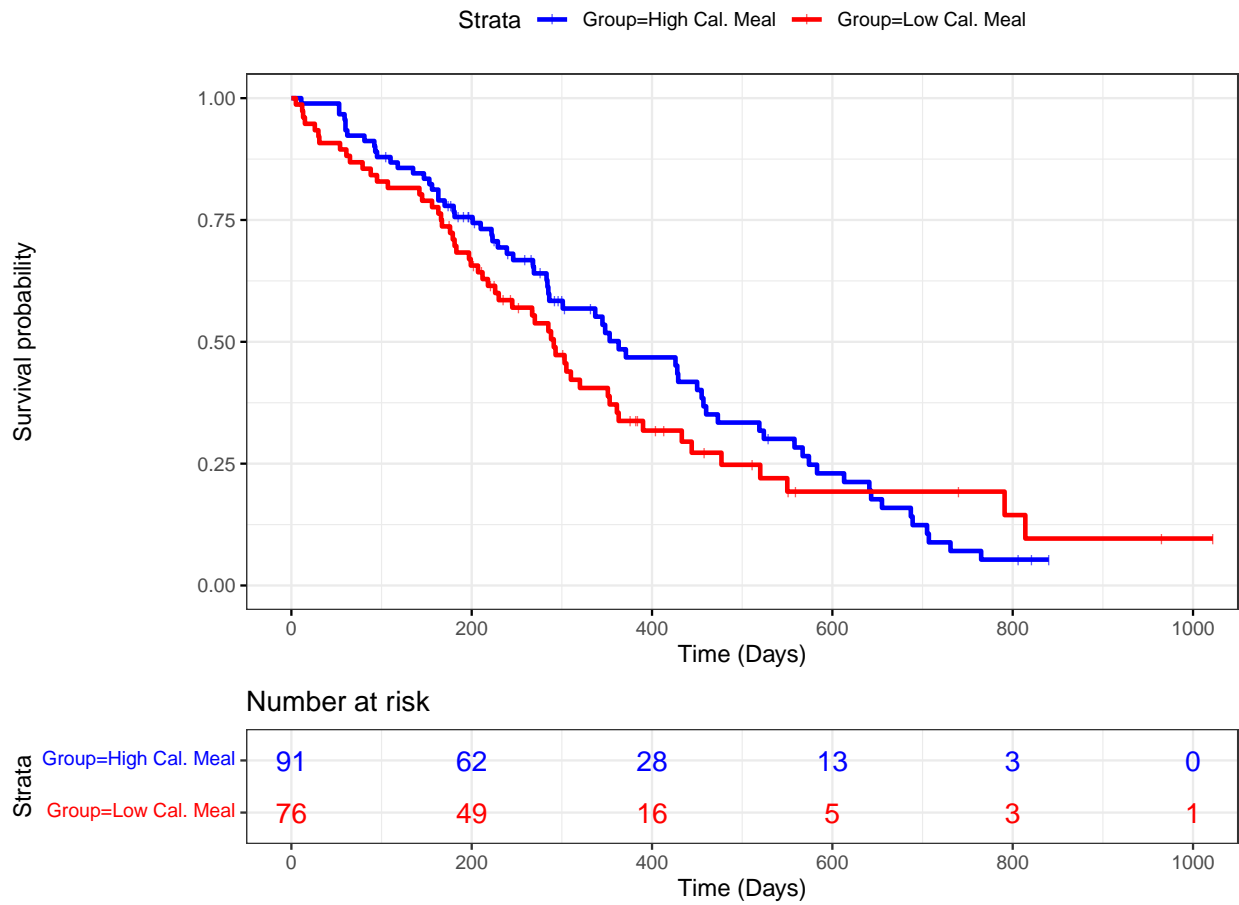


Customize plot

```

KM_plot<-ggsurvplot(
  fit,
  data = data,
  size = 1,
  xlim=c(0, 1000),
  break.x.by = 200,
  xlab="Time (Days)",
  censor.shape="|", censor.size = 1.5, # change plot line size
  palette =c("blue", "red"),# custom color palettes
  conf.int = F,                # Add confidence interval
  pval = F,                    # Add p-value which will perform log-rank T test
  risk.table = TRUE,           # Add risk table below
  risk.table.col = "strata",# Risk table color by groups
  #legend.labs = c("Low Expression", "High Expression"),
  risk.table.height = 0.25, # Useful to change when you have multiple groups
  ggtheme = theme_bw()        # Change ggplot2 theme
)
KM_plot

```

Add HR and p value

```
KM_plot$plot <- KM_plot$plot +
  ggplot2::annotate(
    "text",
    x = Inf, y = Inf,
    vjust = 1, hjust = 1,
    label = "HR = 1.16 \n p=0.43",
    size = 4
  )
KM_plot
```

