CH1

2023-06-03

패키지

```
library(tidyverse)
library(survival)
library(survminer)
library(ggsci)
library(ggsignif)
library(gtsummary)
library(forestmodel)
```

테이블 만들기

```
dt1<-read_csv("C:\\Users\\phl02\\Desktop\\P\\bio\\Ch1_table1.csv")

dt1 %>%
  select(-id, -hcc_yr, -m6_alb) %>%
  tbl_summary(by=LC,missing='no') %>%
  add_p() %>%
  add_overall() %>%
  modify_spanning_header( c('stat_1','stat_2')~'**Liver Function**')
```

Characteristic	Overall, N = 1,000	Cirrhosis, N = 526	No cirrhosis, N = 474	p-value
Sex				0.038
F	317 (32%)	182 (35%)	135 (28%)	
M	683 (68%)	344 (65%)	339 (72%)	
Age	47 (40, 53)	50 (44, 55)	43 (35, 49)	<0.001
ALT	107 (61, 203)	86 (49, 155)	134 (86, 258)	<0.001
Bilirubin	1.20 (0.90, 1.60)	1.30 (1.10, 1.80)	1.10 (0.90, 1.38)	<0.001
PT	1.09 (1.03, 1.18)	1.14 (1.06, 1.27)	1.05 (1.01, 1.10)	<0.001
Creatinine	0.90 (0.70, 1.00)	0.90 (0.70, 0.90)	0.90 (0.80, 1.00)	0.002
Platelet	149 (104, 186)	109 (81, 145)	182 (157, 210)	<0.001
Albumin	3.80 (3.50, 4.10)	3.70 (3.20, 4.10)	4.00 (3.70, 4.13)	<0.001
HBeAg				<0.001
Negative	351 (35%)	214 (41%)	137 (29%)	
Positivie	641 (65%)	307 (59%)	334 (71%)	
HCC	151 (15%)	135 (26%)	16 (3.4%)	<0.001

Multivariable analysis table

```r
dt2<-read_csv("C:\\Users\\phl02\\Desktop\\P\\bio\\Ch1_multi.csv")

fit.multi <- glm(Group~CurrentUser+Age+RaceGroup+Gender+
                 HBV+HCV+Cirrhosis+IBD+Diabetes+Obesity+
                 NAFLD+Smoking,
              family=binomial, data=dt2)

tbl_regression(fit.multi, exponentiate = T) %>%
  bold_labels() %>%
  bold_p()
```
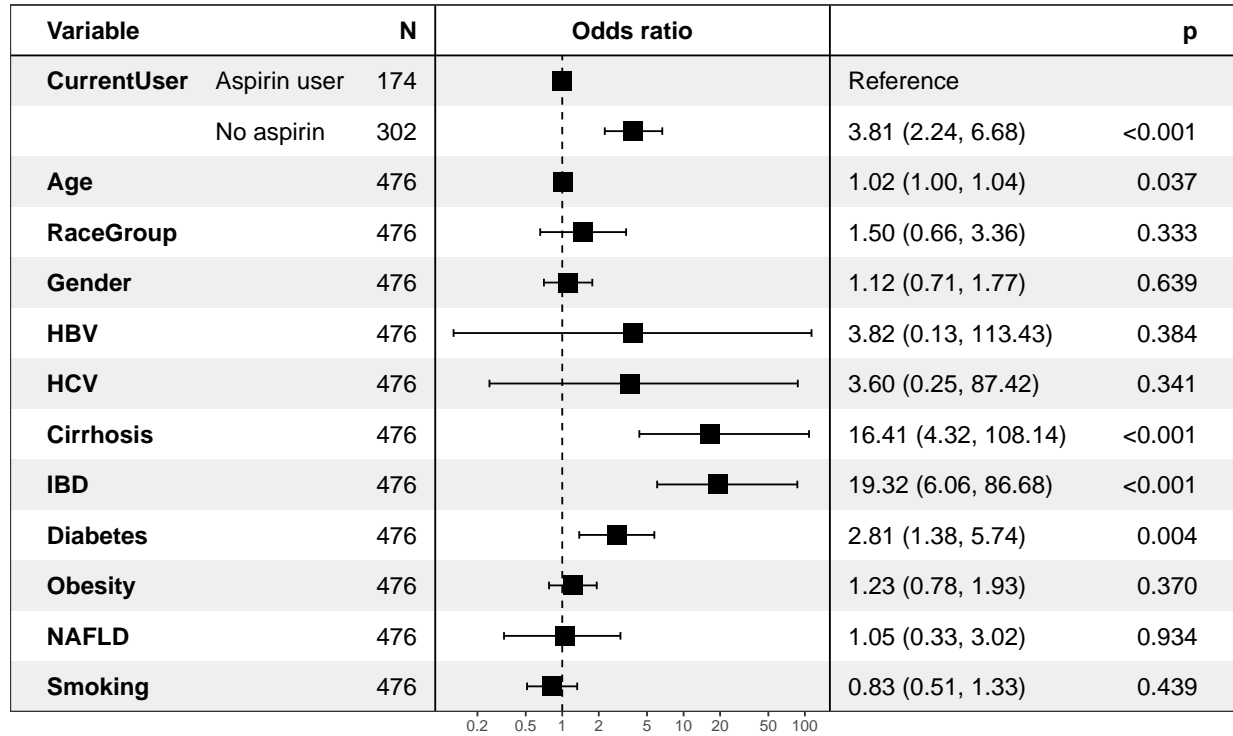
Characteristic	OR	95% CI	p-value
CurrentUser			
Aspirin user	—	—	
No aspirin	3.81	2.24, 6.68	**<0.001**
Age	1.02	1.00, 1.04	**0.037**
RaceGroup	1.50	0.66, 3.36	0.3
Gender	1.12	0.71, 1.77	0.6
HBV	3.82	0.13, 113	0.4
HCV	3.60	0.25, 87.4	0.3
Cirrhosis	16.4	4.32, 108	**<0.001**
IBD	19.3	6.06, 86.7	**<0.001**
Diabetes	2.81	1.38, 5.74	**0.004**
Obesity	1.23	0.78, 1.93	0.4
NAFLD	1.05	0.33, 3.02	>0.9
Smoking	0.83	0.51, 1.33	0.4

Forest plot

```
forest_model(fit.multi)
```

Variable		N	Odds ratio	p
CurrentUser	Aspirin user	174		Reference
	No aspirin	302		3.81 (2.24, 6.68)	<0.001
Age		476		1.02 (1.00, 1.04)	0.037
RaceGroup		476		1.50 (0.66, 3.36)	0.333
Gender		476		1.12 (0.71, 1.77)	0.639
HBV		476		3.82 (0.13, 113.43)	0.384
HCV		476		3.60 (0.25, 87.42)	0.341
Cirrhosis		476		16.41 (4.32, 108.14)	<0.001
IBD		476		19.32 (6.06, 86.68)	<0.001
Diabetes		476		2.81 (1.38, 5.74)	0.004
Obesity		476		1.23 (0.78, 1.93)	0.370
NAFLD		476		1.05 (0.33, 3.02)	0.934
Smoking		476		0.83 (0.51, 1.33)	0.439

0.2 0.5 1 2 5 10 20 50 100
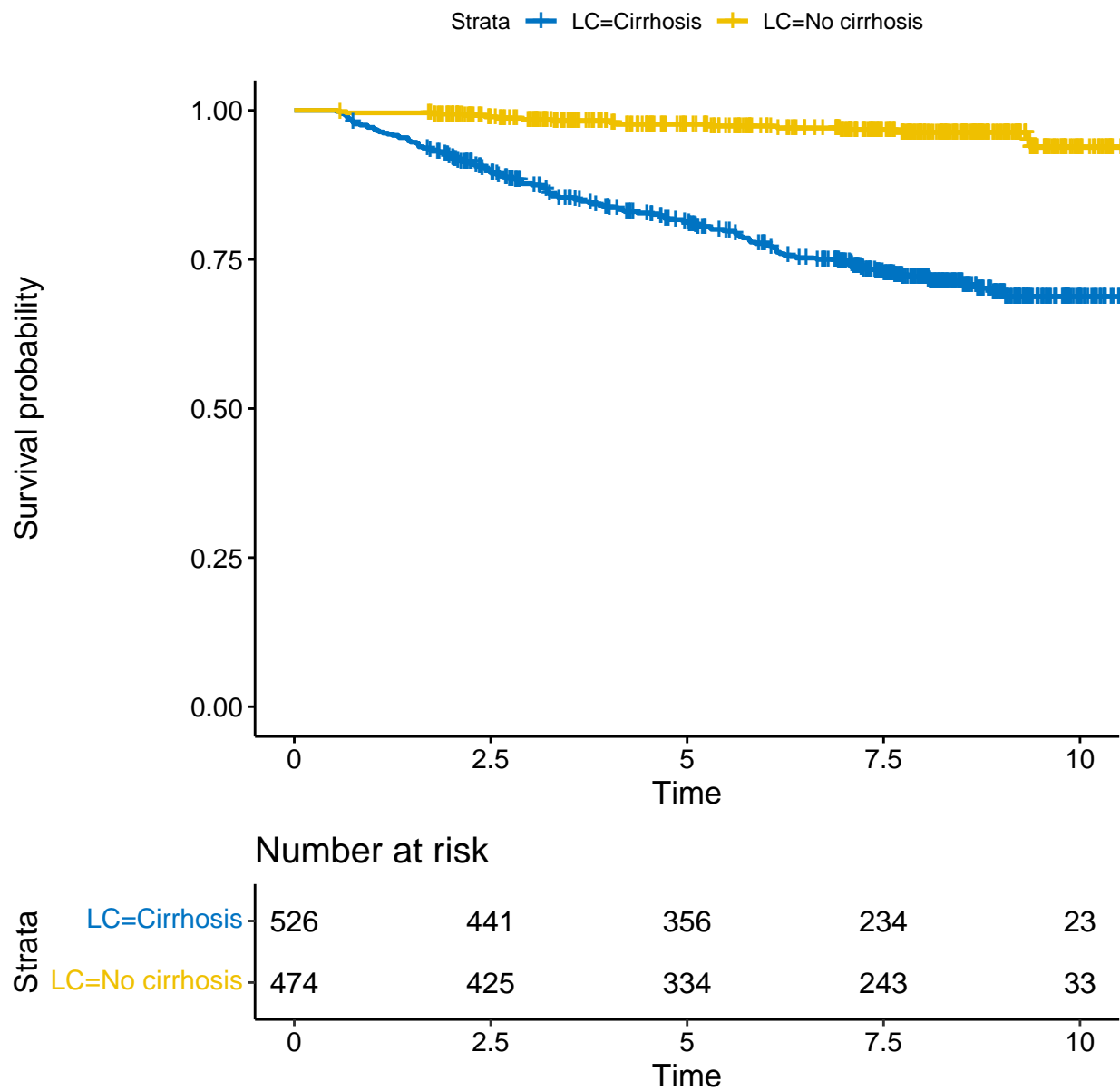
NEJM bar 그래프

```
ggplot(dt1,aes(x= Sex, y=Platelet, fill=factor(HCC)))+
  geom_bar(stat='identity', position='dodge')+
  theme_bw()+
  scale_fill_nejm()+
  geom_signif(comparisons = list(c('F','M')))
```
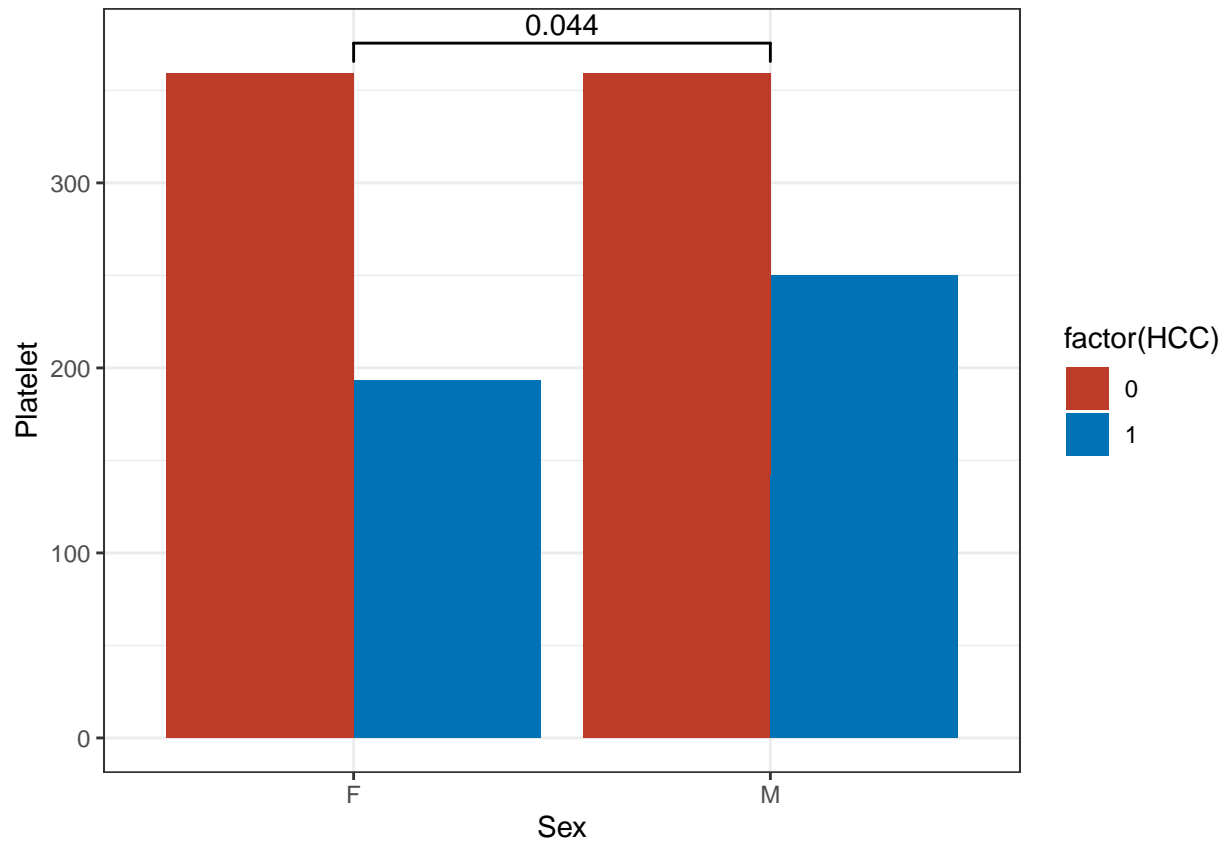
JCO KM 그래프

```
km1<-survfit(Surv(hcc_yr, HCC)~LC, data=dt1)

ggsurvplot(km1,palette = 'jco',risk.table = T)
```

Strata ⊢ LC=Cirrhosis ⊢ LC=No cirrhosis

Number at risk

Strata	0	2.5	5	7.5	10
LC=Cirrhosis	526	441	356	234	23
LC=No cirrhosis	474	425	334	243	33

Waterfall 그래프

```
dt1 %>%
  filter(id<=30) %>%
  select(id,Sex,Albumin,m6_alb) %>%
  mutate(delta_alb=Albumin-m6_alb) %>%
  ggbarplot(x='id', y='delta_alb',
            fill='Sex',
            sort.val='desc',
            sort.by.groups = FALSE)
```