

Code ▾

EORTC HNSCC Honore et al_07232024 Final Analysis

```
library(swimplot) library(grid) library(gtable) library(readr) library(mosaic) library(dplyr) library(survival)
library(survminer) library(ggplot2) library(scales) library(coxphf) library(ggthemes) library(tidyverse)
library(gtsummary) library(flextable) library(parameters) library(car) library(ComplexHeatmap) library(tidyverse)
library(readxl) library(janitor) library(DT) library(rms)
```

#ctDNA Detection Rates by Window and Stages

Hide

```
#ctDNA at Baseline
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_data$ctDNA.Base <- factor(circ_data$ctDNA.Base, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.Base %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("III","IVA","IVB"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.Base == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.Base, by=list(circ_data$Stage), FUN=length)
combined_data <- data.frame(
  Stage = total_counts_by_stage$Group.1,
  Total_Count = total_counts_by_stage$x,
  Positive_Count = positive_counts_by_stage$x,
  Rate = (positive_counts_by_stage$x / total_counts_by_stage$x) * 100 # Convert to percentage
)
combined_data$Rate <- sprintf("%.2f%%", combined_data$Rate)
overall_total_count <- nrow(circ_data)
overall_positive_count <- nrow(circ_data[circ_data$ctDNA.Base == "POSITIVE",])
overall_positivity_rate <- (overall_positive_count / overall_total_count) * 100 # Convert to percentage
overall_row <- data.frame(
  Stage = "Overall",
  Total_Count = overall_total_count,
  Positive_Count = overall_positive_count,
  Rate = sprintf("%.2f%%", overall_positivity_rate)
)
combined_data <- rbind(combined_data, overall_row)
print(combined_data)
```

Stage <fctr>	Total_Count <int>	Positive_Count <int>	Rate <chr>
III	16	16	100.00%
IVA	20	20	100.00%
IVB	6	5	83.33%
Overall	42	41	97.62%

4 rows

Hide

```
#ctDNA post-treatment
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("III","IVA","IVB"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.postTx == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.postTx, by=list(circ_data$Stage), FUN=length)
combined_data <- data.frame(
  Stage = total_counts_by_stage$Group.1,
  Total_Count = total_counts_by_stage$x,
  Positive_Count = positive_counts_by_stage$x,
  Rate = (positive_counts_by_stage$x / total_counts_by_stage$x) * 100 # Convert to percentage
)
combined_data$Rate <- sprintf("%.2f%%", combined_data$Rate)
overall_total_count <- nrow(circ_data)
overall_positive_count <- nrow(circ_data[circ_data$ctDNA.postTx == "POSITIVE",])
overall_positivity_rate <- (overall_positive_count / overall_total_count) * 100 # Convert to percentage
overall_row <- data.frame(
  Stage = "Overall",
  Total_Count = overall_total_count,
  Positive_Count = overall_positive_count,
  Rate = sprintf("%.2f%%", overall_positivity_rate)
)
combined_data <- rbind(combined_data, overall_row)
print(combined_data)
```

Stage <fctr>	Total_Count <int>	Positive_Count <int>	Rate <chr>
III	15	2	13.33%
IVA	21	0	0.00%
IVB	6	2	33.33%
Overall	42	4	9.52%
4 rows			

#Summary Table

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]

circ_data_subset <- circ_data %>%
  select(
    Gender,
    Age,
    Smoking.Assumption,
    Alcohol.Assumption,
    Prim.Location,
    ECOG,
    Stage,
    Stage.TNM,
    p16,
    NAC,
    Treatment.Type,
    RFS.Event,
    OS.Event,
    FU.Biopsy.months) %>%
  mutate(
    Gender = factor(Gender),
    Age = as.numeric(Age),
    Smoking.Assumption = factor(Smoking.Assumption),
    Alcohol.Assumption = factor(Alcohol.Assumption),
    Prim.Location = factor(Prim.Location),
    ECOG = factor(ECOG),
    Stage = factor(Stage),
    Stage.TNM = factor(Stage.TNM),
    p16 = factor(p16),
    NAC = factor(NAC, levels = c("FALSE", "TRUE"), labels = c("No treatment", "Neoadjuvant Treatment")),
    Treatment.Type = factor(Treatment.Type),
    RFS.Event = factor(RFS.Event, levels = c("FALSE", "TRUE"), labels = c("No Recurrence", "Recurrence")),
    OS.Event = factor(OS.Event, levels = c("FALSE", "TRUE"), labels = c("Alive", "Deceased")),
    FU.Biopsy.months = as.numeric(FU.Biopsy.months))
table1 <- circ_data_subset %>%
  tbl_summary(
    statistic = list(
      all_continuous() ~ "{median} ({min} - {max})",
      all_categorical() ~ "{n} ({p}%)") %>%
    bold_labels()
  )
table1

```

Characteristic	N = 43 ¹
Gender	
Female	10 (23%)
Male	33 (77%)
Age	67 (28 - 102)
Smoking.Assumption	35 (81%)
Alcohol.Assumption	27 (63%)
Prim.Location	
Hypopharynx	5 (12%)
Larynx	6 (14%)
Oral cavity	5 (12%)
Oropharynx	27 (63%)
ECOG	
0	27 (63%)
1	14 (33%)
2	2 (4.7%)
Stage	
III	16 (37%)
IVA	21 (49%)
IVB	6 (14%)
Stage.TNM	
T1-T3N1M0	12 (28%)
T1-T4N2M0	13 (30%)
T1-T4N3M0	6 (14%)
T3N0M0	4 (9.3%)
T4aN0-N1M0	8 (19%)
¹ n (%); Median (Range)	

Characteristic	N = 43 ¹
p16	
Negative	26 (60%)
Positive	17 (40%)
NAC	
No treatment	37 (86%)
Neoadjuvant Treatment	6 (14%)
Treatment.Type	
Chemotherapy	1 (2.3%)
Radiochemotherapy	29 (67%)
Radiochemotherapy, Surgery	5 (12%)
Radioimmunotherapy	2 (4.7%)
Radiotherapy	3 (7.0%)
Radiotherapy, Surgery	2 (4.7%)
Surgery	1 (2.3%)
RFS.Event	
No Recurrence	32 (74%)
Recurrence	11 (26%)
OS.Event	
Alive	33 (77%)
Deceased	10 (23%)
FU.Biopsy.months	20 (4 - 45)
¹ n (%); Median (Range)	

Hide

```
fit1 <- as_flex_table(  
  table1,  
  include = everything(),  
  return_calls = FALSE,  
  strip_md_bold = TRUE)
```

Warning: The `strip_md_bold` argument of `as_flex_table()` is deprecated as of gtsummary 1.6.0.
This warning is displayed once every 8 hours.
Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.

Hide

fit1

Characteristic	N = 43 ¹
Gender	
Female	10 (23%)
Male	33 (77%)
Age	67 (28 - 102)
Smoking.Assumption	35 (81%)
Alcohol.Assumption	27 (63%)
Prim.Location	
Hypopharynx	5 (12%)
Larynx	6 (14%)
Oral cavity	5 (12%)
Oropharynx	27 (63%)
ECOG	
0	27 (63%)
1	14 (33%)
2	2 (4.7%)
Stage	
III	16 (37%)
IVA	21 (49%)
IVB	6 (14%)
Stage.TNM	
T1-T3N1M0	12 (28%)
T1-T4N2M0	13 (30%)
T1-T4N3M0	6 (14%)
T3N0M0	4 (9.3%)

¹n (%); Median (Range)

Characteristic	N = 43 ¹
T4aN0-N1M0	8 (19%)
p16	
Negative	26 (60%)
Positive	17 (40%)
NAC	
No treatment	37 (86%)
Neoadjuvant Treatment	6 (14%)
Treatment.Type	
Chemotherapy	1 (2.3%)
Radiochemotherapy	29 (67%)
Radiochemotherapy, Surgery	5 (12%)
Radioimmunotherapy	2 (4.7%)
Radiotherapy	3 (7.0%)
Radiotherapy, Surgery	2 (4.7%)
Surgery	1 (2.3%)
RFS.Event	
No Recurrence	32 (74%)
Recurrence	11 (26%)
OS.Event	
Alive	33 (77%)
Deceased	10 (23%)
FU.Biopsy.months	20 (4 - 45)

¹n (%); Median (Range)

Hide

```
save_as_docx(fit1, path= "~/Downloads/table1.docx")
```

#Heatmap for the clinical factors

Hide


```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data %>% arrange(Stage)
circ_datadf <- as.data.frame(circ_data)

ha <- HeatmapAnnotation(
  Stage = circ_data$Stage,
  Gender = circ_data$Gender,
  Prim.Location = circ_data$Prim.Location,
  p16 = circ_data$p16,
  NAC = circ_data$NAC,
  ctDNA.Base = circ_data$ctDNA.Base,
  ctDNA.postTx = circ_data$ctDNA.postTx,
  RFS.Event = circ_data$RFS.Event,
  OS.Event = circ_data$OS.Event,

  col = list(Stage = c("III" = "seagreen1", "IVA" = "khaki", "IVB" = "orange"),
    Gender = c("Female" = "goldenrod", "Male" = "blue4"),
    Prim.Location = c("Hypopharynx" = "purple", "Oropharynx" = "lightblue", "Larynx" = "brown2", "Oral cavity" = "lightgrey"),
    p16 = c("Negative" = "yellow", "Positive" = "brown"),
    NAC = c("TRUE" = "darkmagenta", "FALSE" = "grey"),
    ctDNA.Base = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    ctDNA.postTx = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    RFS.Event = c("TRUE" = "red3", "FALSE" = "blue"),
    OS.Event = c("TRUE" = "black", "FALSE" = "grey")
  )
)
ht <- Heatmap(matrix(nrow = 0, ncol = length(circ_data$Stage)), show_row_names = FALSE, cluster_rows = F, cluster_columns = FALSE, top_annotation = ha)
pdf("heatmap.pdf", width = 7, height = 7)
draw(ht, annotation_legend_side = "bottom")
dev.off()

```

```

null device
      1

```

#Overview plot

Hide

```
setwd("~/Downloads")
clinstage <- read.csv("OP-16933_ HNSCC OP.csv")
clinstage_df <- as.data.frame(clinstage)

# Creating the basic swimmer plot
oplot <- swimmer_plot(df=clinstage_df,
                      id='PatientName',
                      end='fu.diff.months',
                      fill='gray',
                      width=.01)

# Adding themes and scales
oplot <- oplot + theme(panel.border = element_blank())
oplot <- oplot + scale_y_continuous(breaks = seq(0, 48, by = 3))
oplot <- oplot + labs(x = "Patients", y = "Months from Diagnosis")

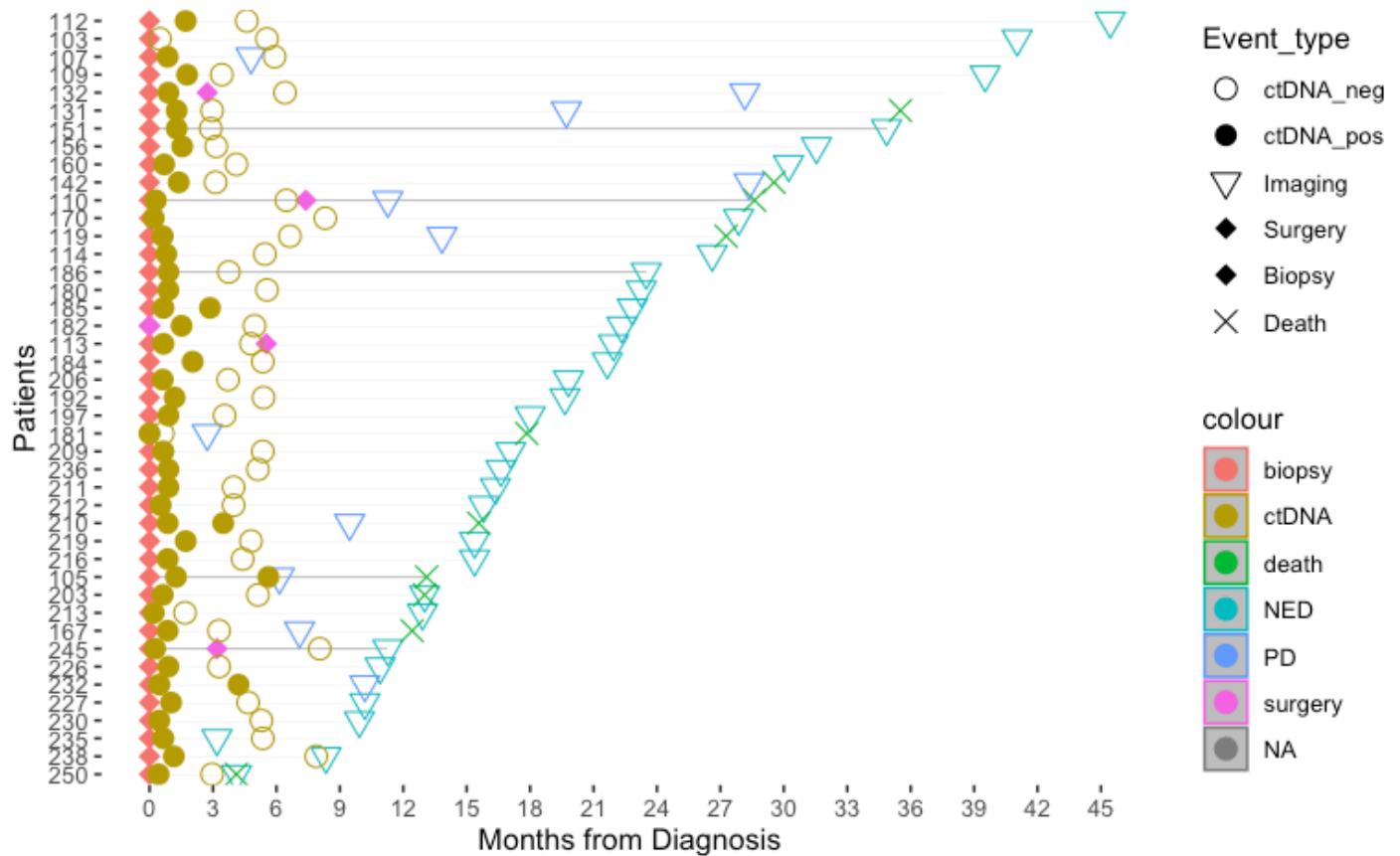
# Adding swimmer points
oplot_ev1 <- oplot + swimmer_points(df_points=clinstage_df,
                                    id='PatientName',
                                    time='date.diff.months',
                                    name_shape = 'Event_type',
                                    name_col = 'Event',
                                    size=3.5, fill='black')

# Optionally uncomment and use col='darkgreen' if needed

# Adding shape manual scale
oplot_ev1.1 <- oplot_ev1 + ggplot2::scale_shape_manual(name="Event_type",
                                                         values=c(1,16,6,18,18,4),
                                                         breaks=c('ctDNA_neg', 'ctDNA_pos',
                                                                    'Imaging', 'Surgery', 'Biopsy', 'Death'))

# Display the plot
oplot_ev1.1
```

Warning: Removed 43 rows containing missing values or values outside the scale range (`geom_point()`).


[Hide](#)

```

instage_df,
Name',
tart.months',
d.months',
x_type',
= 1.0)

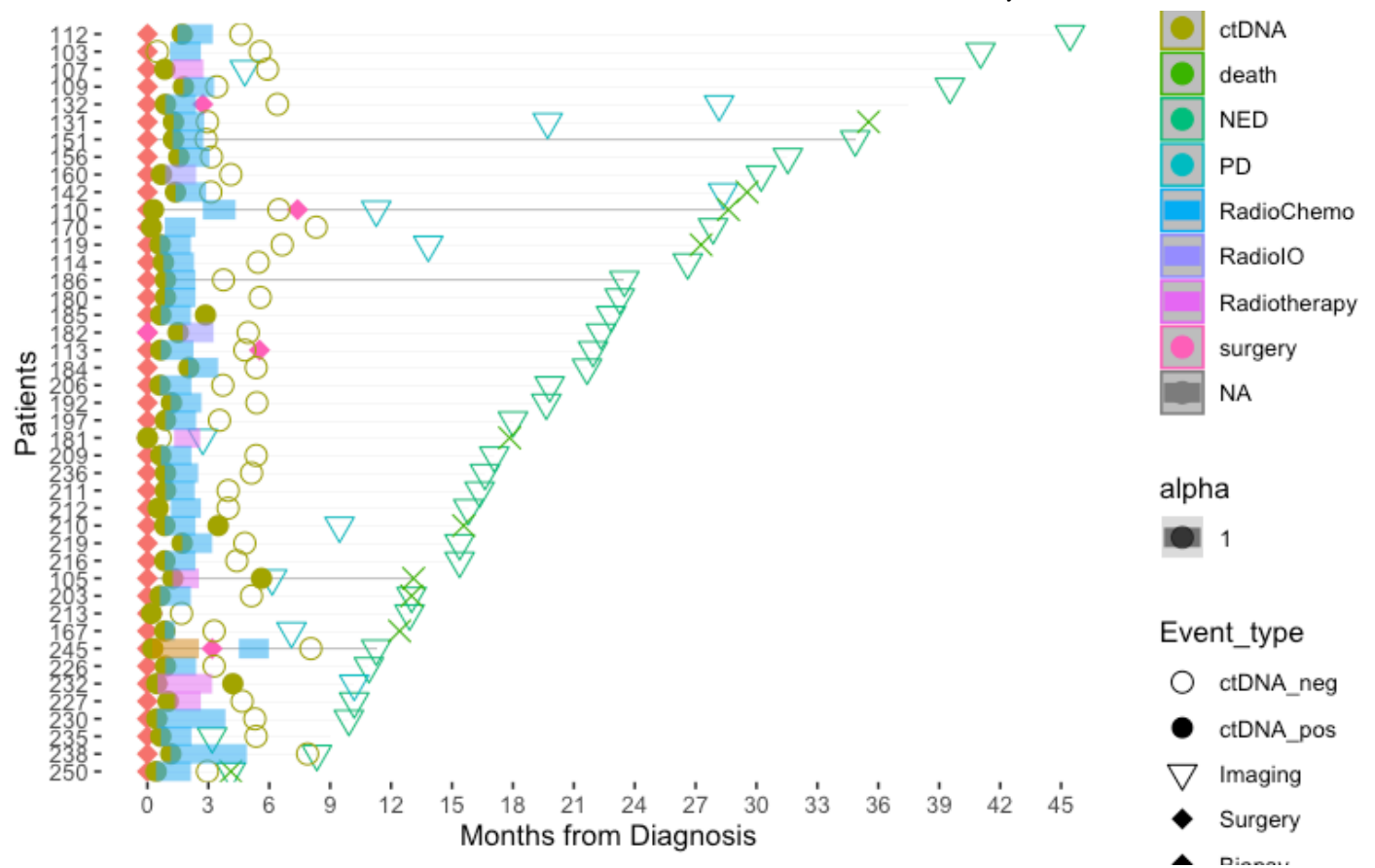
oplot_ev2 <- oplot_ev1.1 + swimmer_lines(df_lines=cl
                                         id='Patient
                                         start='Tx_s
                                         end='Tx_en
                                         name_col='T
                                         size=3.5,
                                         name_alpha

oplot_ev2 <- oplot_ev2 + guides(linetype = guide_leg
end(override.aes = list(size = 5, color = "black"))
oplot_ev2

```

Warning: Removed 43 rows containing missing values or values outside the scale range (`geom_point()`).

Warning: Removed 191 rows containing missing values or values outside the scale range (`geom_segment()`).


[Hide](#)

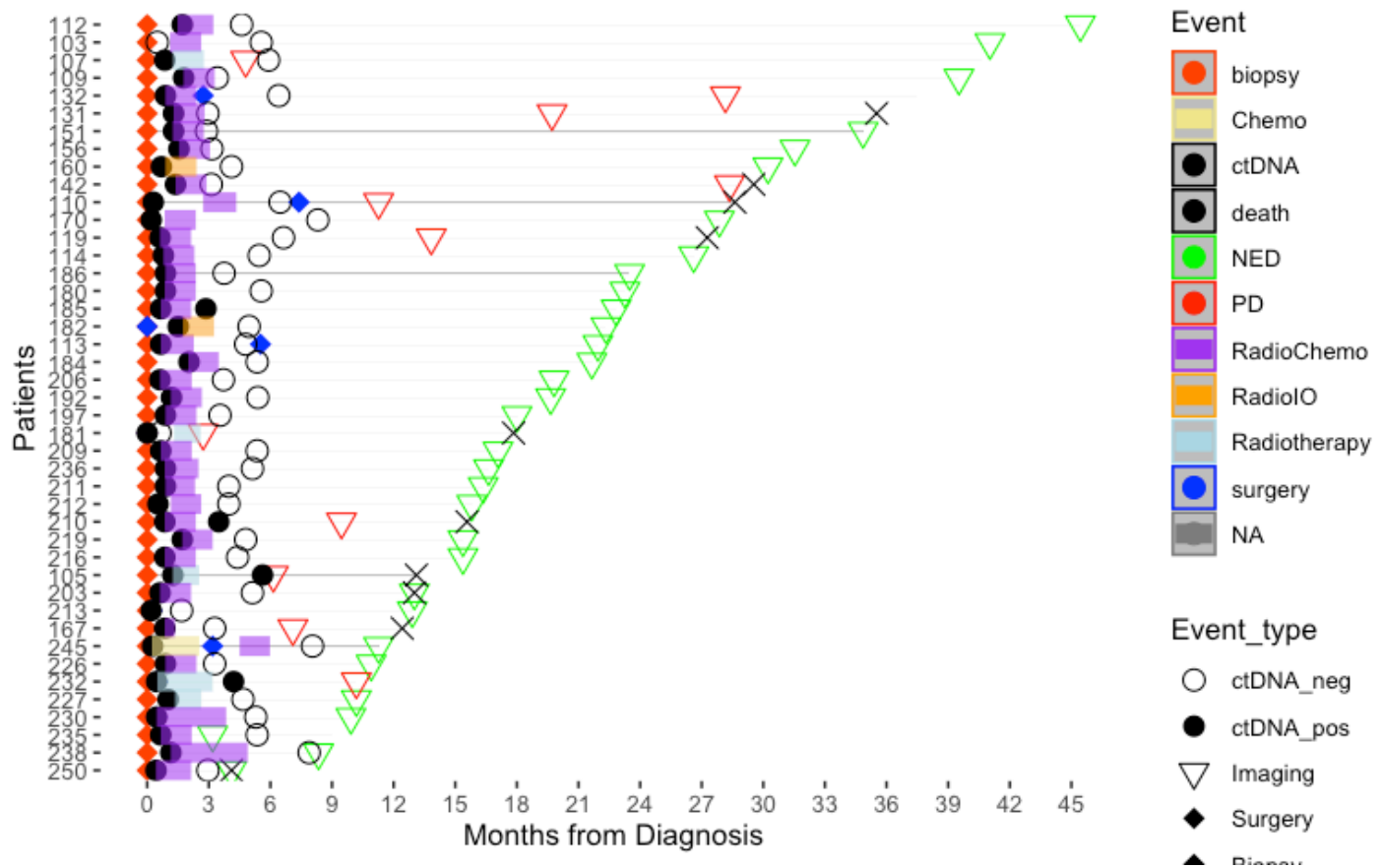
```

oplot_ev2.2 <- oplot_ev2 + ggplot2::scale_color_manual(
  name="Event",values=c( "orangered", "khaki", "black", "black", "green", "red", "purple",
    "orange", "lightblue", "blue"))
oplot_ev2.2

```

Warning: Removed 43 rows containing missing values or values outside the scale range (`geom_point()`).

Warning: Removed 191 rows containing missing values or values outside the scale range (`geom_segment()`).



#RFS in Complete Cohort (N=50)

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_datadf <- as.data.frame(circ_data)

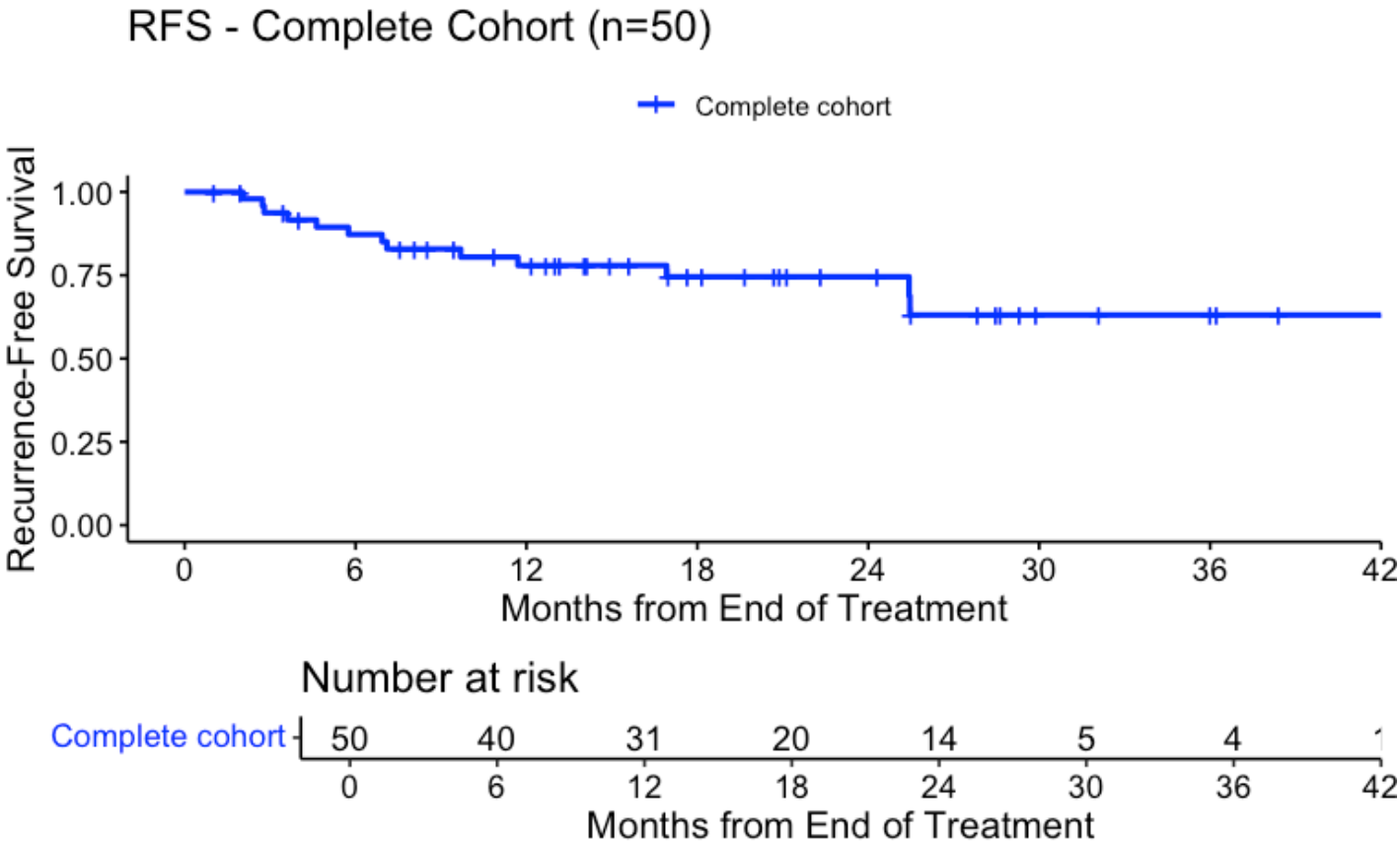
survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~HNSCC.Cohort, data = circ_data)
```

Call: `survfit(formula = Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event) ~ HNSCC.Cohort, data = circ_data)`

	n	events	median	0.95LCL	0.95UCL
[1,]	50	13	NA	25.5	NA

Hide

```
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ HNSCC.Cohort, data = circ_data, conf.int=0.95, conf.type = "log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE, break.time.by=6, palette=c("blue"), title="RFS - Complete Cohort (n=50)", ylab="Recurrence-Free Survival", xlab="Months from End of Treatment", legend.labs=c("Complete cohort"), legend.title="")
```



Hide

```
summary(KM_curve, times= c(18, 24, 36))
```

```
Call: survfit(formula = surv_object ~ HNSCC.Cohort, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
18	20	11	0.746	0.0679	0.583	0.852
24	14	0	0.746	0.0679	0.583	0.852
36	4	2	0.631	0.0942	0.419	0.783

#OS in Complete Cohort (N=50)

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_datadf <- as.data.frame(circ_data)

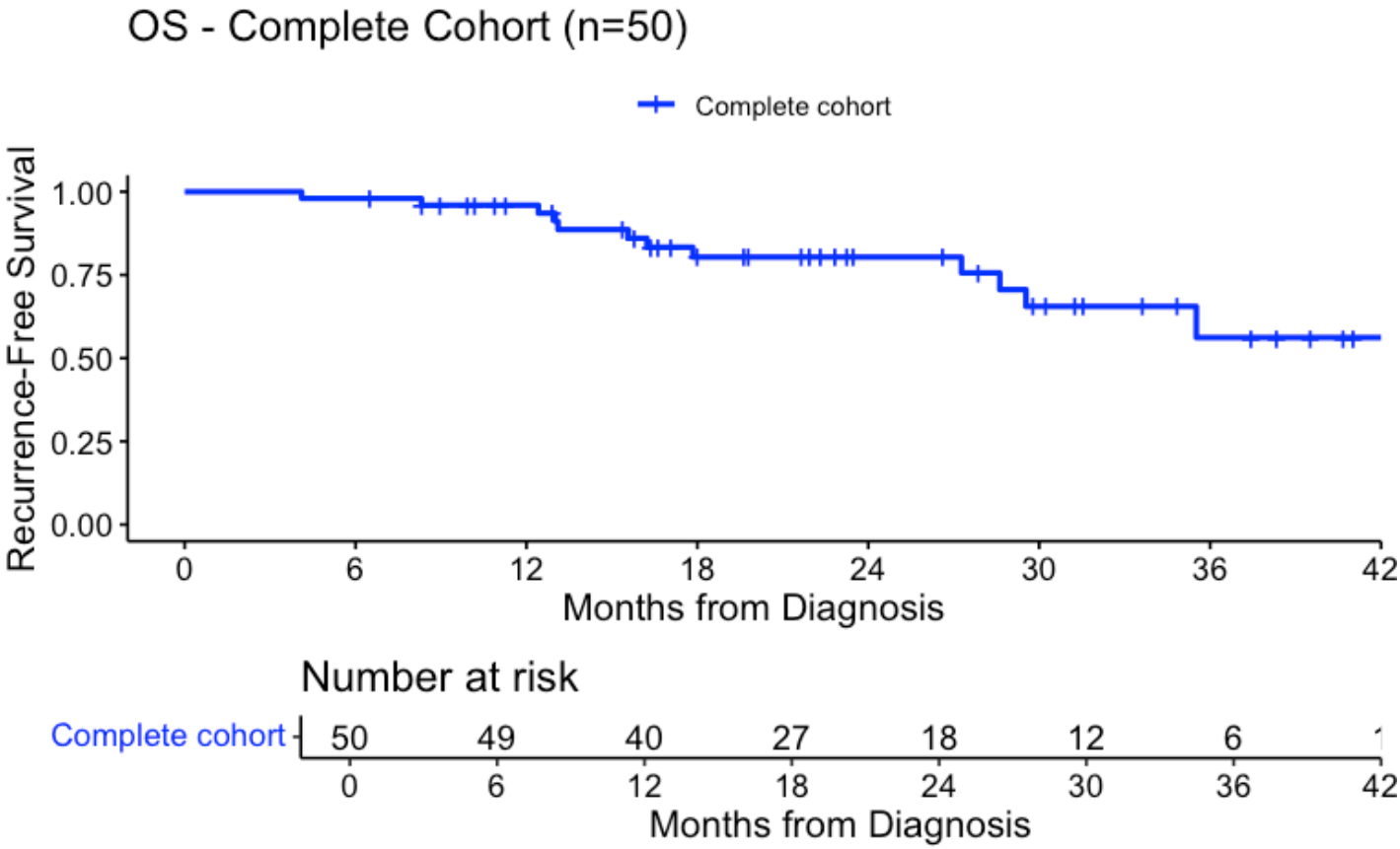
survfit(Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)~HNSCC.Cohort, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event) ~
  HNSCC.Cohort, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
[1,]	50	12	NA	29.5	NA

[Hide](#)

```
surv_object <-Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ HNSCC.Cohort, data = circ_data,conf.int=0.95,conf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE, break.time.by=6, palette=c("blue"), title="OS - Complete Cohort (n=50)", ylab= "Recurrence-Free Survival", xlab="Months from Diagnosis", legend.labs=c("Complete cohort"), legend.title="")
```



Hide

```
summary(KM_curve, times= c(18, 24, 36))
```



```
Call: survfit(formula = surv_object ~ HNSCC.Cohort, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
18	27	8	0.804	0.0632	0.643	0.898
24	18	0	0.804	0.0632	0.643	0.898
36	6	4	0.562	0.1177	0.309	0.754

#Association of Baseline ctDNA MTM levels with clinicopathological factors

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_datadf <- as.data.frame(circ_data)

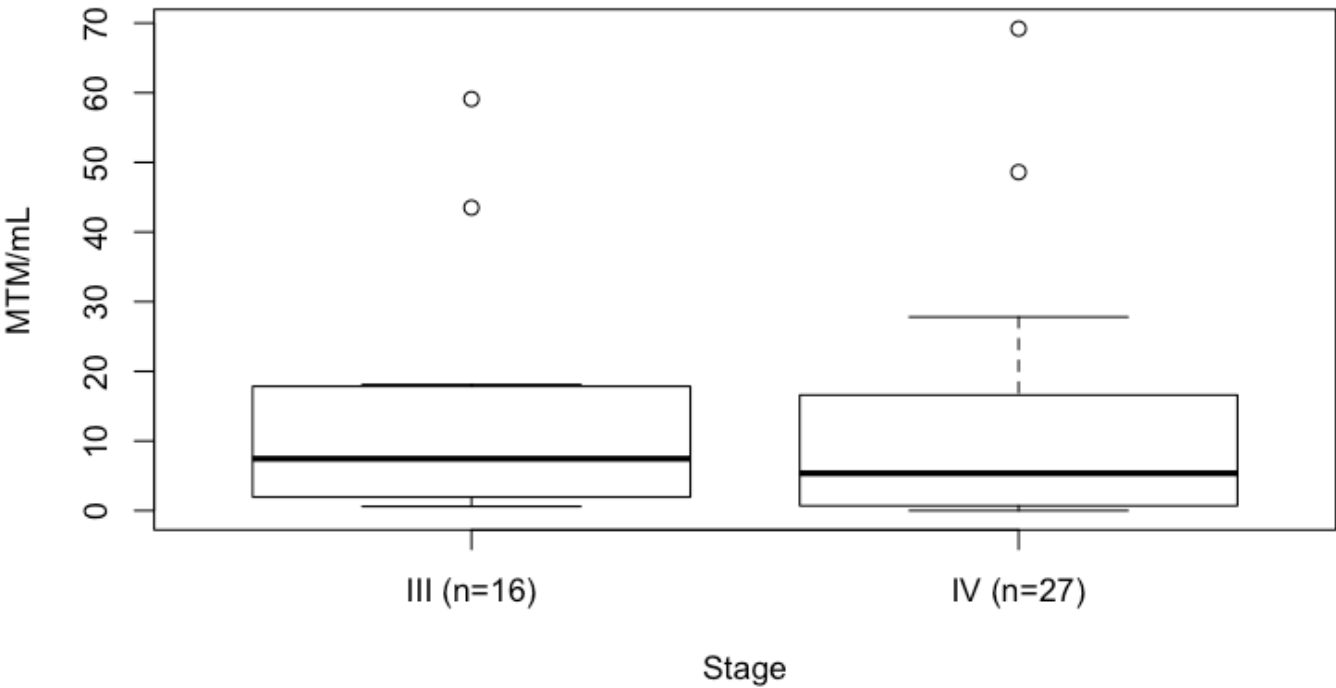
tally(~cStage, data=circ_data, margins = TRUE)
```

cStage		
III	IV	Total
16	26	42

[Hide](#)

```
circ_data$cStage <- factor(circ_data$cStage, levels = c("III","IV"), labels = c("III (n=
16)","IV (n=27)"))
boxplot(ctDNA.Base.MTM~cStage, data=circ_data, main="ctDNA pre-treatment MTM - Stage", x
lab="Stage", ylab="MTM/mL", col="white",border="black")
```

ctDNA pre-treatment MTM - Stage



Hide

```
m1<-wilcox.test(ctDNA.Base.MTM ~ cStage, data=circ_data, na.rm=TRUE, exact=FALSE, conf.i
nt=TRUE)
print(m1)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by cStage
W = 228, p-value = 0.6135
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -5.400007  6.100017
sample estimates:
difference in location
      1.0637
```

Hide

```
tally(~cT.Status, data=circ_data, margins = TRUE)
```

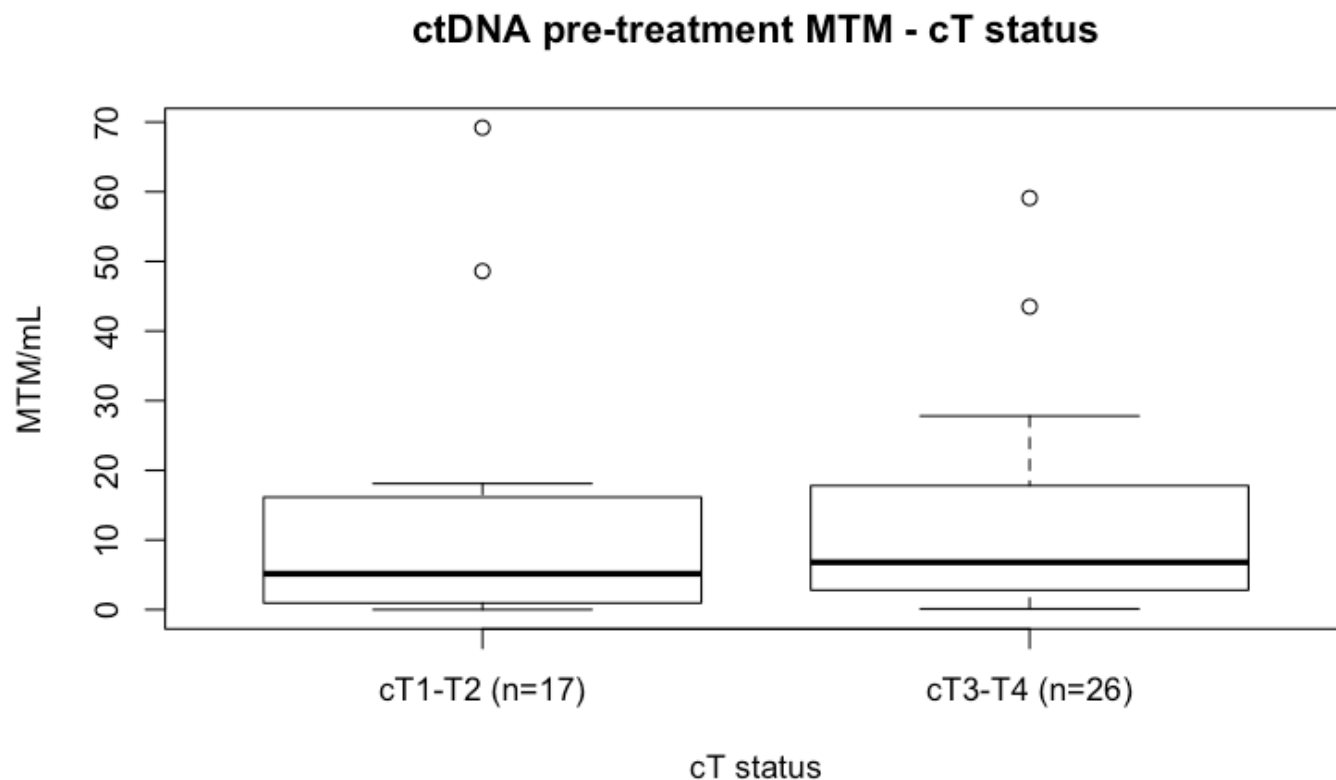
```
cT.Status
cT1-T2 cT3-T4 Total
    16    26    42
```

Hide

```

circ_data$cT.Status <- factor(circ_data$cT.Status, levels = c("cT1-T2","cT3-T4"), labels =
= c("cT1-T2 (n=17)","cT3-T4 (n=26)"))
boxplot(ctDNA.Base.MTM~cT.Status, data=circ_data, main="ctDNA pre-treatment MTM - cT sta
tus", xlab="cT status", ylab="MTM/mL", col="white",border="black")

```



Hide

```

m2<-wilcox.test(ctDNA.Base.MTM ~ cT.Status, data=circ_data, na.rm=TRUE, exact=FALSE, con
f.int=TRUE)
print(m2)

```

Wilcoxon rank sum test with continuity correction

```

data: ctDNA.Base.MTM by cT.Status
W = 179, p-value = 0.4603
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -8.499967  3.599946
sample estimates:
difference in location
 -1.59994

```

Hide

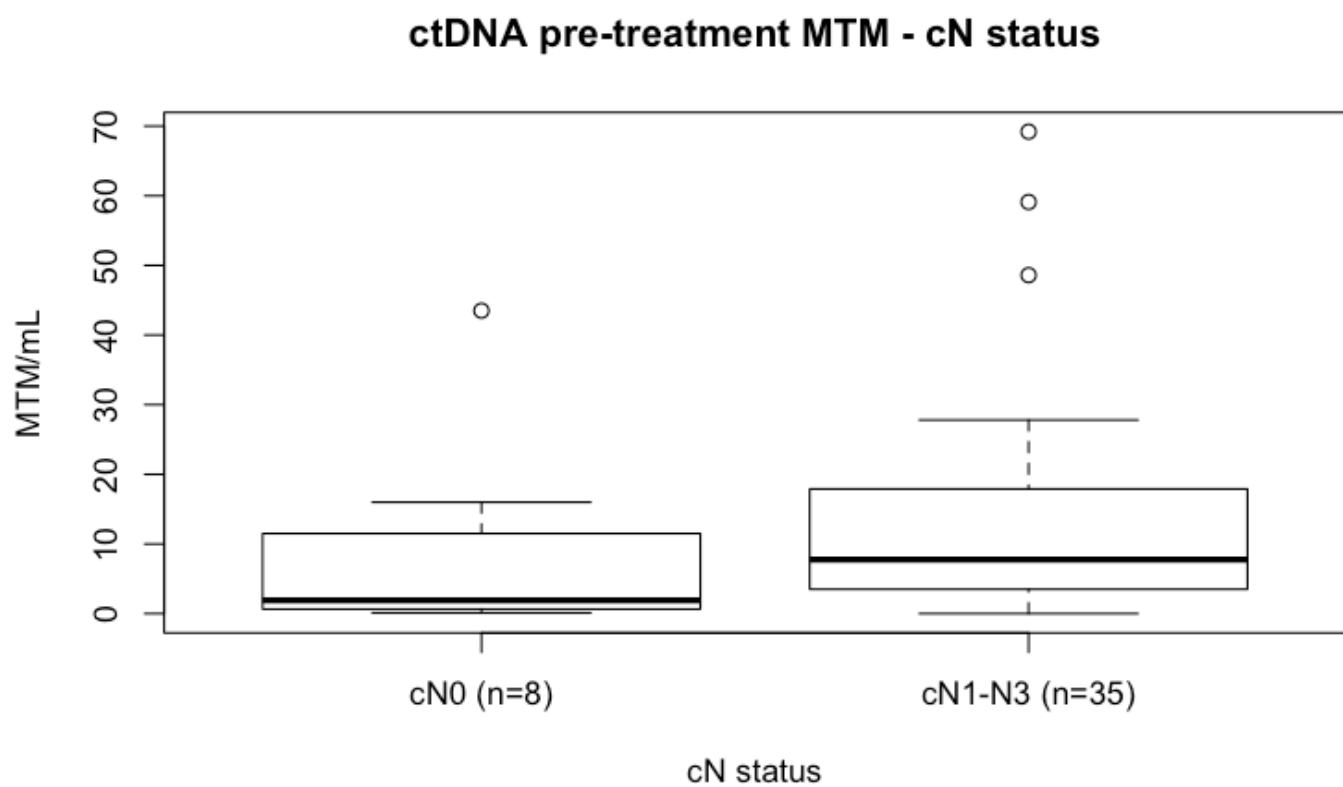
```
tally(~cN.Status, data=circ_data, margins = TRUE)
```

cN.Status

cN0	cN1-N3	Total
8	34	42

[Hide](#)

```
circ_data$cN.Status <- factor(circ_data$cN.Status, levels = c("cN0","cN1-N3"), labels =
c("cN0 (n=8)","cN1-N3 (n=35)"))
boxplot(ctDNA.Base.MTM~cN.Status, data=circ_data, main="ctDNA pre-treatment MTM - cN sta
tus", xlab="cN status", ylab="MTM/mL", col="white",border="black")
```


[Hide](#)

```
m3<-wilcox.test(ctDNA.Base.MTM ~ cN.Status, data=circ_data, na.rm=TRUE, exact=FALSE, con
f.int=TRUE)
print(m3)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by cN.Status
W = 93, p-value = 0.1733
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -12.900026  1.300003
sample estimates:
difference in location
 -3.227956
```

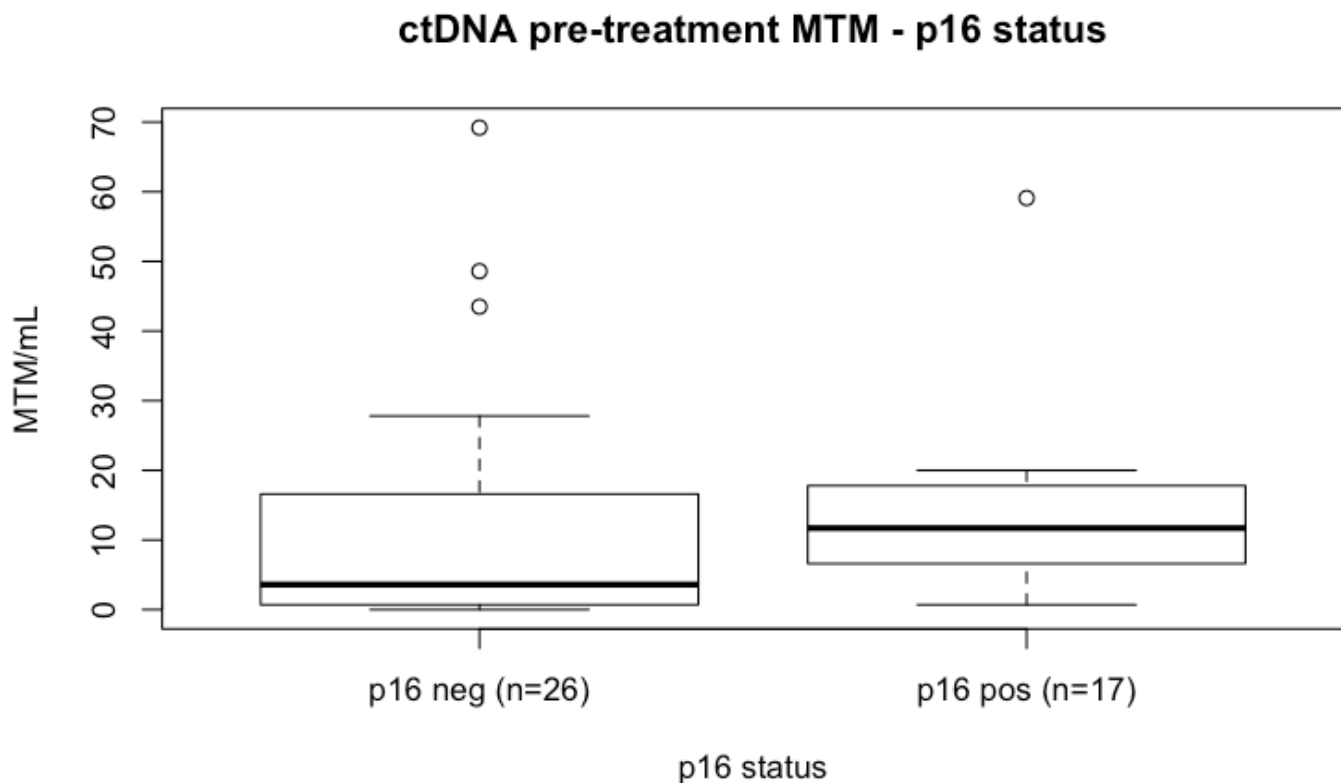
[Hide](#)

```
tally(~p16, data=circ_data, margins = TRUE)
```

p16		
Negative	Positive	Total
25	17	42

[Hide](#)

```
circ_data$p16 <- factor(circ_data$p16, levels = c("Negative","Positive"), labels = c("p16 neg (n=26)","p16 pos (n=17)"))
boxplot(ctDNA.Base.MTM~p16, data=circ_data, main="ctDNA pre-treatment MTM - p16 status",
xlab="p16 status", ylab="MTM/mL", col="white",border="black")
```



Hide

```
m4<-wilcox.test(ctDNA.Base.MTM ~ p16, data=circ_data, na.rm=TRUE, exact=FALSE, conf.int=TRUE)
print(m4)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by p16
W = 142.5, p-value = 0.07488
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -10.79993  1.59996
sample estimates:
difference in location
      -5.100049
```

Hide

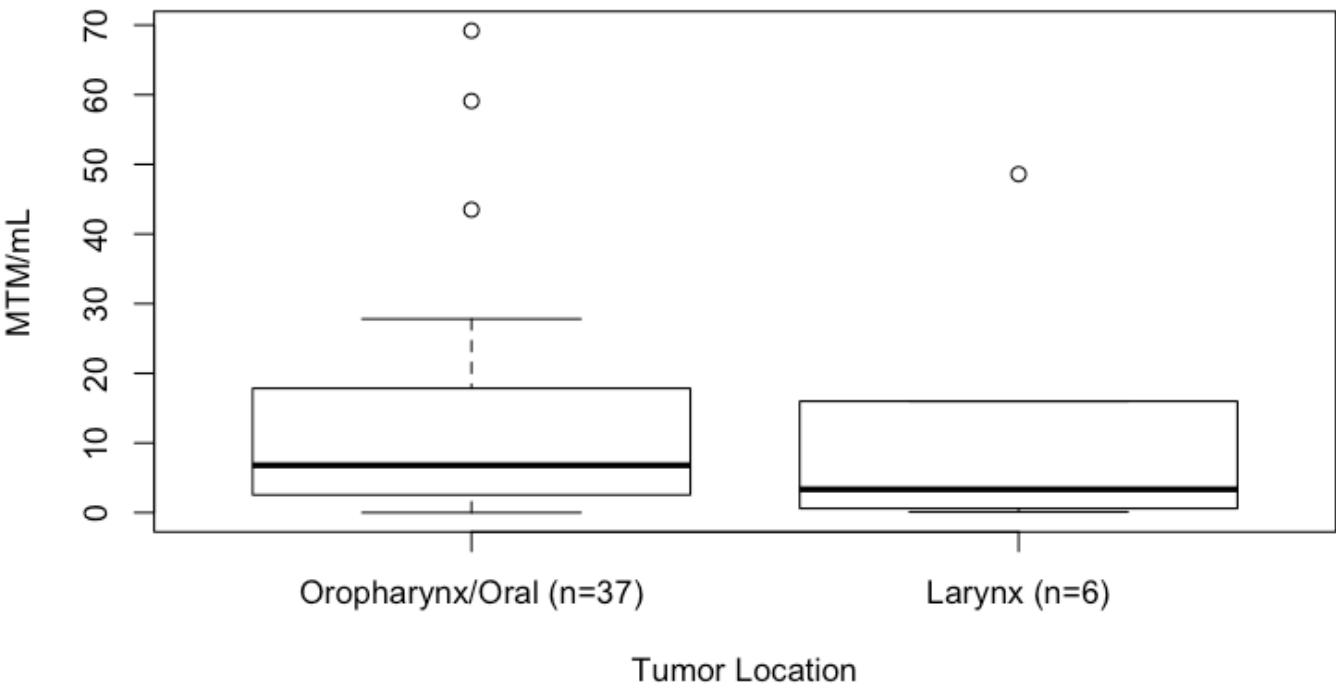
```
tally(~Location, data=circ_data, margins = TRUE)
```

Location		
	Larynx	Oropharynx/Oral
	6	36
		Total
		42

Hide

```
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral","Larynx"),
labels = c("Oropharynx/Oral (n=37)","Larynx (n=6)"))
boxplot(ctDNA.Base.MTM~Location, data=circ_data, main="ctDNA pre-treatment MTM - Tumor Location", xlab="Tumor Location", ylab="MTM/mL", col="white",border="black")
```

ctDNA pre-treatment MTM - Tumor Location



Hide

```
m4<-wilcox.test(ctDNA.Base.MTM ~ Location, data=circ_data, na.rm=TRUE, exact=FALSE, con
f.int=TRUE)
print(m4)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by Location
W = 131, p-value = 0.4186
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -7.099978 12.499975
sample estimates:
difference in location
      2.099941
```

#Median MTM/mL levels for ctDNA positive pts post-treatment

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_data <- circ_data[circ_data$ctDNA.postTx=="POSITIVE",]

median_ctDNA <- median(circ_data$ctDNA.postTx.MTM, na.rm = TRUE)
range_ctDNA <- range(circ_data$ctDNA.postTx.MTM, na.rm = TRUE)
cat("Median MTM/mL post-treatment:", median_ctDNA, "\n")
```

Median MTM/mL post-treatment: 0.95

Hide

```
cat("Range MTM/mL post-treatment:", range_ctDNA[1], "-", range_ctDNA[2], "\n")
```

Range MTM/mL post-treatment: 0.1 - 5

#Median time from end treatment to radiological recurrence for ctDNA negative pts post-treatment

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_data <- circ_data[circ_data$ctDNA.postTx=="NEGATIVE",]
circ_data <- circ_data[circ_data$RFS.Event=="TRUE",]

median_RFS <- median(circ_data$RFS.months, na.rm = TRUE)
range_RFS <- range(circ_data$RFS.months, na.rm = TRUE)
cat("Median RFS:", median_RFS, "\n")
```

Median RFS: 11.7

Hide

```
cat("Range RFS:", range_RFS[1], "-", range_RFS[2], "\n")
```

Range RFS: 2.73 - 25.46

#RFS by ctDNA post-treatment

Hide


```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.postTx, data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$RFS.months, event = circ_data\$RFS.Event) ~ ctDNA.postTx, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.postTx=NEGATIVE	38	7	NA	25.46	NA
ctDNA.postTx=POSITIVE	4	3	8.4	3.61	NA

Hide

```
event_summary <- circ_data %>%
  group_by(ctDNA.postTx) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.postTx <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	38	7	0.1842105	18.42105
POSITIVE	4	3	0.7500000	75.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.postTx, data = circ_data,conf.int=0.95,conf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = TRUE, conf.int = FALSE, risk.table = TRUE,
break.time.by=6, palette=c("blue","red"), title="RFS - ctDNA post-treatment", ylab= "Recurrence-Free Survival", xlab="Months from End of Treatment", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
summary(KM_curve, times= c(0, 12, 24, 36))
```

```
Call: survfit(formula = surv_object ~ ctDNA.postTx, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

```

ctDNA.postTx=NEGATIVE
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0      38      0    1.000  0.0000      1.000      1.000
 12      26      4    0.881  0.0561      0.712      0.954
 24      10      1    0.832  0.0712      0.633      0.929
 36       3      2    0.647  0.1279      0.346      0.836

ctDNA.postTx=POSITIVE
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0       4      0     1.00   0.000      1.00000      1.000
 12       1      3    0.25   0.217      0.00895      0.665
```

Hide

```
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hide

```
summary(cox_fit)
```

```
Call:
```

```
coxph(formula = surv_object ~ ctDNA.postTx, data = circ_data)
```

```
n= 42, number of events= 10
```

```

              coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.postTxPOSITIVE 2.0148    7.4995  0.7402 2.722  0.00649 **
```

```
----
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

              exp(coef) exp(-coef) lower .95 upper .95
ctDNA.postTxPOSITIVE      7.5      0.1333      1.758      32
```

```
Concordance= 0.652 (se = 0.078 )
```

```
Likelihood ratio test= 5.72 on 1 df, p=0.02
```

```
Wald test = 7.41 on 1 df, p=0.006
```

```
Score (logrank) test = 10.2 on 1 df, p=0.001
```

Hide

```

cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
print(label_text)

```

```
[1] "HR = 7.5 (1.76-32); p = 0.006"
```

[Hide](#)

```

#Fisher test for RFS percentages at 12, 24, and 36 months
dfs_times <- c(12, 24, 36)
p_values <- sapply(dfs_times, function(time) {
  neg_count <- sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$RFS.months >= time &
    circ_data$RFS.Event == 0)
  pos_count <- sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$RFS.months >= time &
    circ_data$RFS.Event == 0)
  neg_total <- sum(circ_data$ctDNA.postTx == "NEGATIVE")
  pos_total <- sum(circ_data$ctDNA.postTx == "POSITIVE")

  neg_surv <- neg_total - sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$RFS.Event
    == 1 & circ_data$RFS.months < time)
  pos_surv <- pos_total - sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$RFS.Event
    == 1 & circ_data$RFS.months < time)

  surv_matrix <- matrix(c(neg_surv, pos_surv, neg_total - neg_surv, pos_total - pos_surv), nrow = 2)
  test_result <- fisher.test(surv_matrix)
  return(test_result$p.value)
})
names(p_values) <- paste0("p-value at ", dfs_times, " months")
print(p_values)

```

p-value at 12 months	p-value at 24 months	p-value at 36 months
0.01125704	0.01763602	0.03618333

#RFS by ctDNA post-treatment - Recurrence at year 1

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)~ctDNA.postTx,
data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$RFS.months, event = circ_data\$RFS.Event.1y) ~
ctDNA.postTx, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.postTx=NEGATIVE	38	4	NA	NA	NA
ctDNA.postTx=POSITIVE	4	3	8.4	3.61	NA

Hide

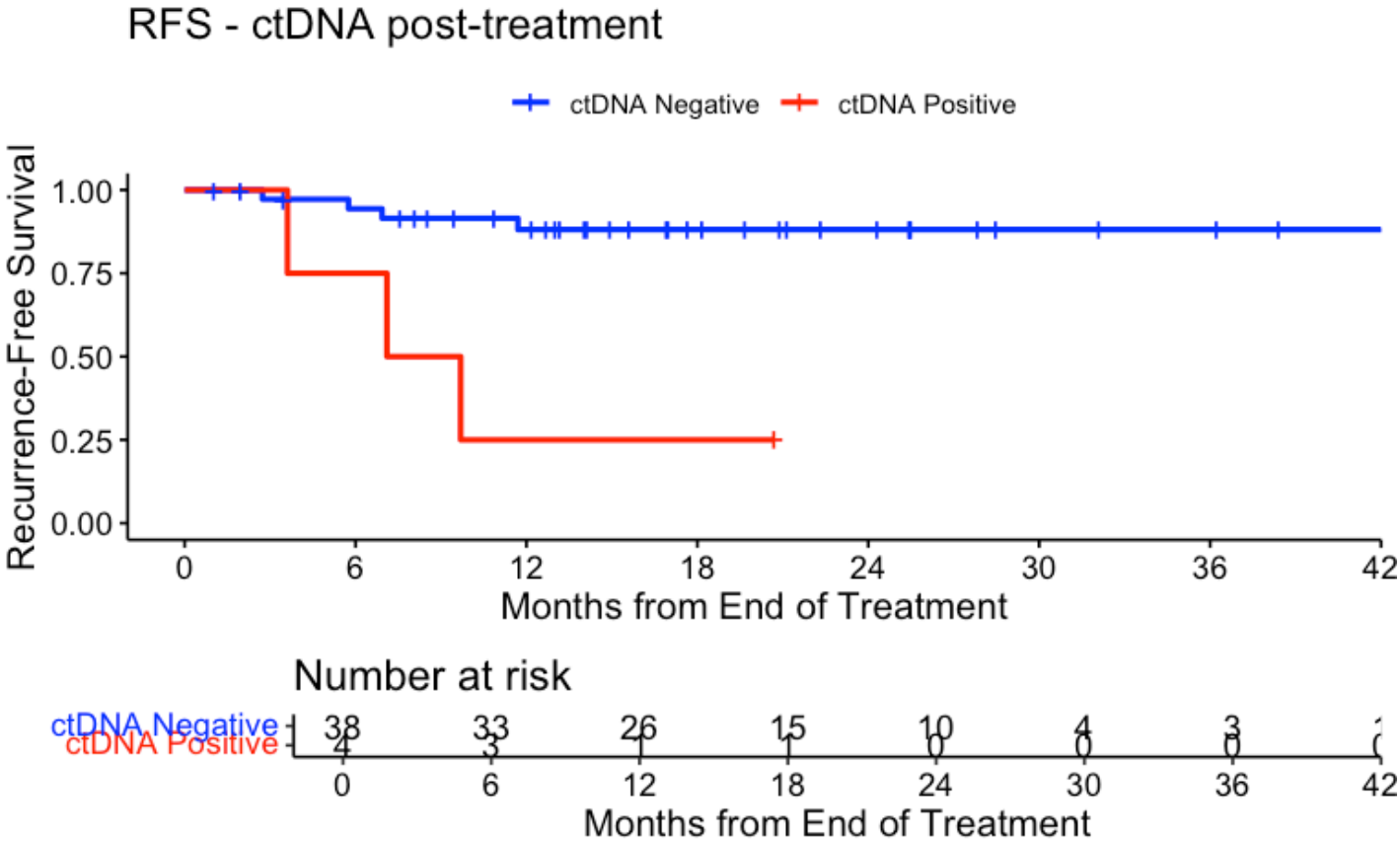
```
event_summary <- circ_data %>%
  group_by(ctDNA.postTx) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event.1y),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.postTx <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	38	4	0.1052632	10.52632
POSITIVE	4	3	0.7500000	75.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
KM_curve <- survfit(surv_object ~ ctDNA.postTx, data = circ_data,conf.int=0.95,conf.type
="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=6, palette=c("blue","red"), title="RFS - ctDNA post-treatment", ylab=
"Recurrence-Free Survival", xlab="Months from End of Treatment", legend.labs=c("ctDNA Ne
gative", "ctDNA Positive"), legend.title="")
```



Hide

```
summary(KM_curve, times= c(0, 12, 24, 36))
```

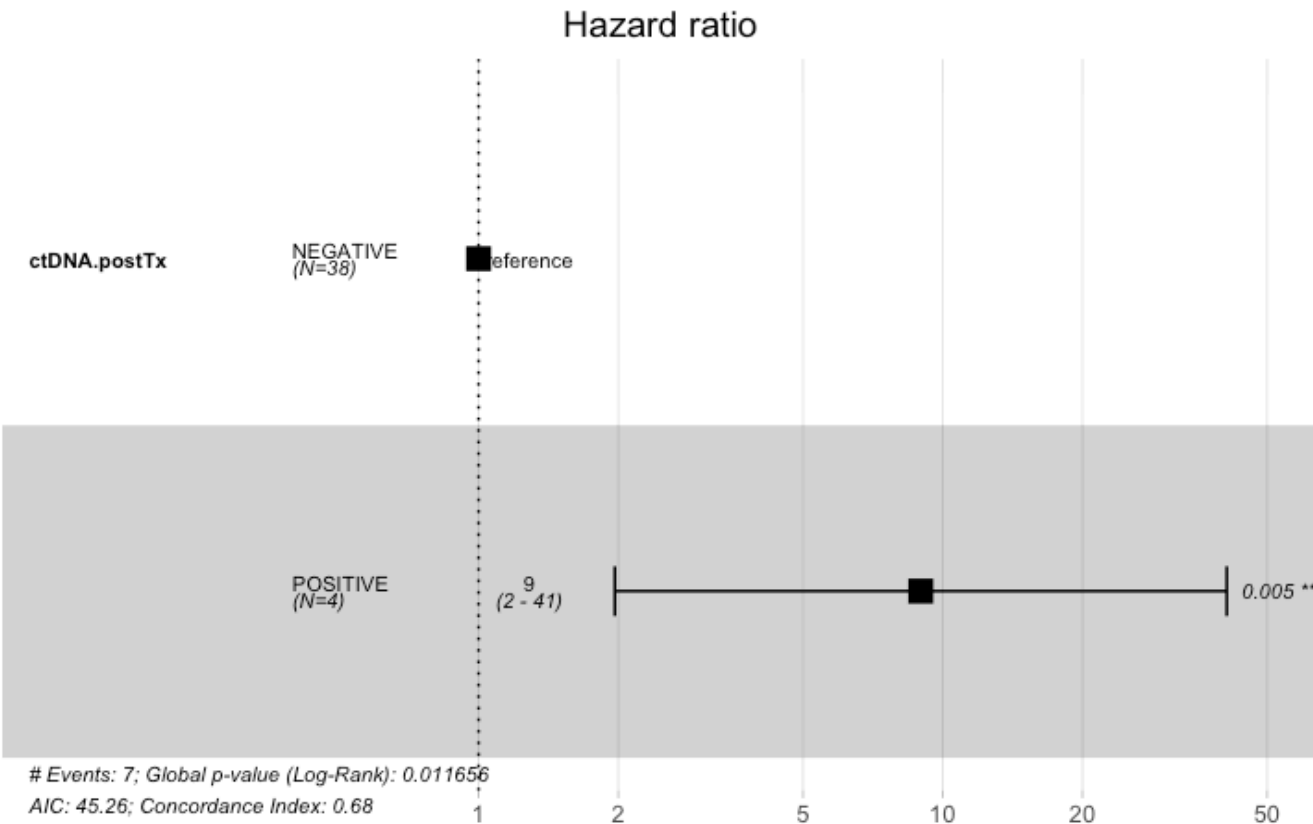
```
Call: survfit(formula = surv_object ~ ctDNA.postTx, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.postTx=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	38	0	1.000	0.0000	1.000	1.000	
12	26	4	0.881	0.0561	0.712	0.954	
24	10	0	0.881	0.0561	0.712	0.954	
36	3	0	0.881	0.0561	0.712	0.954	

ctDNA.postTx=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	4	0	1.00	0.000	1.00000	1.000	
12	1	3	0.25	0.217	0.00895	0.665	

Hide

```
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.postTx, data = circ_data)
```

```
n= 42, number of events= 7
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.postTxPOSITIVE	2.1937	8.9685	0.7745	2.832	0.00462 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.postTxPOSITIVE	8.969	0.1115	1.966	40.92

Concordance= 0.675 (se = 0.088)

Likelihood ratio test= 6.36 on 1 df, p=0.01

Wald test = 8.02 on 1 df, p=0.005

Score (logrank) test = 11.69 on 1 df, p=6e-04

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 8.97 (1.97-40.92); p = 0.005"
```

[Hide](#)

```
#Fisher test for RFS percentages at 12, 24, and 36 months
dfs_times <- c(12, 24, 36)
p_values <- sapply(dfs_times, function(time) {
  neg_count <- sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$RFS.months >= time &
    circ_data$RFS.Event.1y == 0)
  pos_count <- sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$RFS.months >= time &
    circ_data$RFS.Event.1y == 0)
  neg_total <- sum(circ_data$ctDNA.postTx == "NEGATIVE")
  pos_total <- sum(circ_data$ctDNA.postTx == "POSITIVE")

  neg_surv <- neg_total - sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$RFS.Event.1y == 1 & circ_data$RFS.months < time)
  pos_surv <- pos_total - sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$RFS.Event.1y == 1 & circ_data$RFS.months < time)

  surv_matrix <- matrix(c(neg_surv, pos_surv, neg_total - neg_surv, pos_total - pos_surv), nrow = 2)
  test_result <- fisher.test(surv_matrix)
  return(test_result$p.value)
})
names(p_values) <- paste0("p-value at ", dfs_times, " months")
print(p_values)
```

p-value at 12 months	p-value at 24 months	p-value at 36 months
0.01125704	0.01125704	0.01125704

#Barplot for Recurrence % at year 1 by ctDNA post-treatment

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]

circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels = c("NEGATIVE", "POSITIVE"), labels = c("Negative", "Positive"))
circ_data$RFS.Event.1y <- factor(circ_data$RFS.Event.1y, levels = c("FALSE", "TRUE"), labels = c("No Recurrence", "Recurrence"))
contingency_table <- table(circ_data$ctDNA.postTx, circ_data$RFS.Event.1y)
chi_square_test <- chisq.test(contingency_table)
```

```
Warning in stats::chisq.test(x, y, ...) :
  Chi-squared approximation may be incorrect
```

[Hide](#)

```
print(chi_square_test)
```


Pearson's Chi-squared test with Yates' continuity correction

```
data: contingency_table
X-squared = 6.6868, df = 1, p-value = 0.009713
```

[Hide](#)

```
fisher_exact_test <- fisher.test(contingency_table)
print(fisher_exact_test)
```

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.01126
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.421207 1359.954433
sample estimates:
odds ratio
 22.02536
```

[Hide](#)

```
print(contingency_table)
```

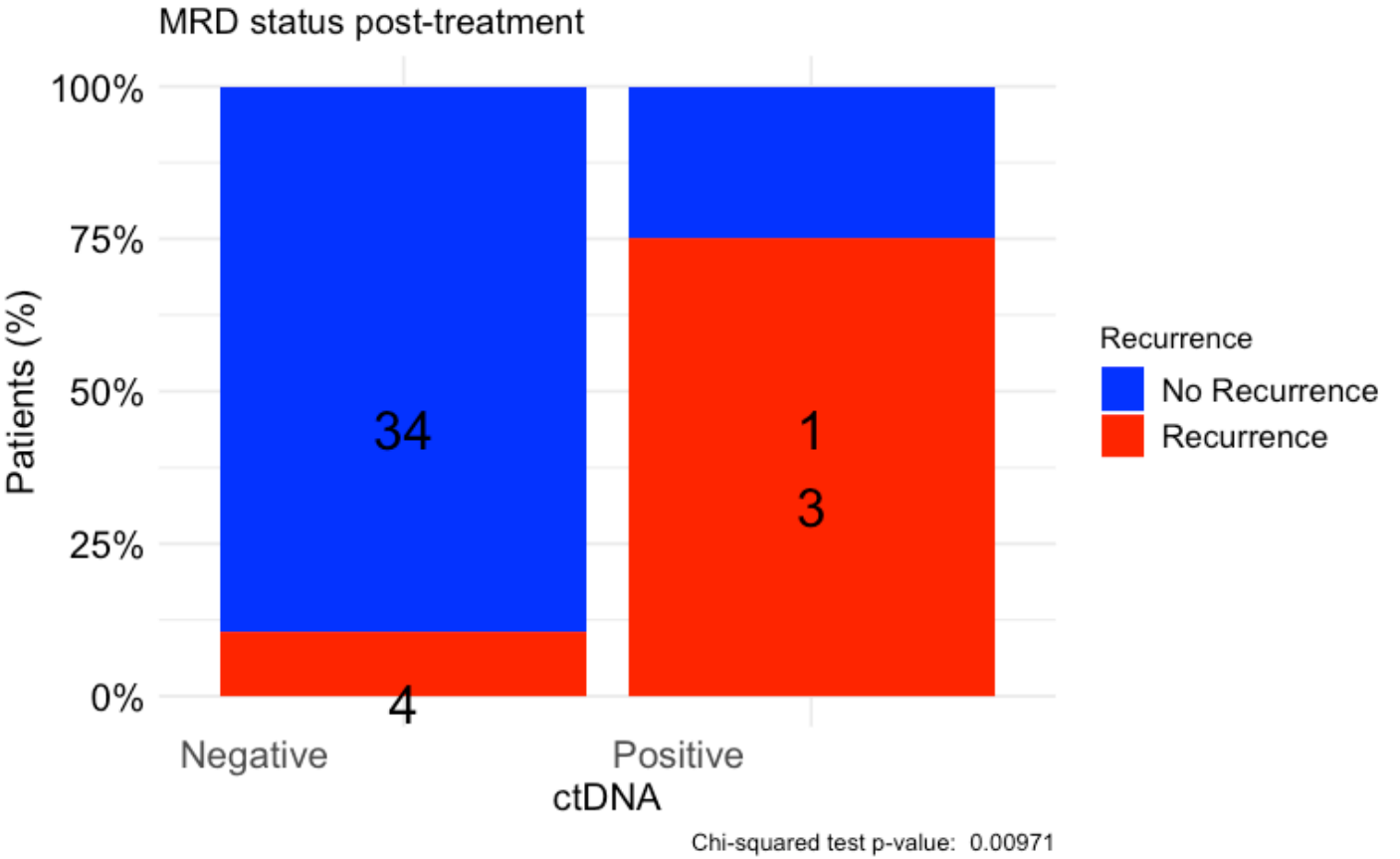
	No Recurrence	Recurrence
Negative	34	4
Positive	1	3

[Hide](#)

```

table_df <- as.data.frame(contingency_table)
table_df$Total <- ave(table_df$Freq, table_df$Var1, FUN = sum)
table_df$Percentage <- table_df$Freq / table_df$Total
table_df$MiddlePercentage <- table_df$Percentage / 2
ggplot(table_df, aes(x = Var1, y = Percentage, fill = Var2)) +
  geom_bar(stat = "identity") +
  geom_text(aes(y = MiddlePercentage, label = Freq), position = "stack", color = "black", vjust = 1.5, size = 7) +
  theme_minimal() +
  labs(title = "MRD status post-treatment",
        x = "ctDNA",
        y = "Patients (%)",
        fill = "Recurrence",
        caption = paste("Chi-squared test p-value: ", format.pval(chi_square_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("No Recurrence" = "blue", "Recurrence" = "red")) + # define custom colors
  theme(axis.text.x = element_text(angle = 0, hjust = 1.5, size = 14), # increase x-axis text size
        axis.text.y = element_text(size = 14, color = "black"), # increase y-axis text size
        axis.title.x = element_text(size = 14, color = "black"), # increase x-axis label size
        axis.title.y = element_text(size = 14, color = "black"), # increase y-axis label size
        legend.text = element_text(size = 12, color = "black")) # increase Recurrence label size

```



#OS by ctDNA post-treatment

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)~ctDNA.postTx, data = circ_data)
```

Call: `survfit(formula = Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event) ~ ctDNA.postTx, data = circ_data)`

	n	events	median	0.95LCL	0.95UCL
ctDNA.postTx=NEGATIVE	38	8	NA	29.5	NA
ctDNA.postTx=POSITIVE	4	2	15.6	13.1	NA

Hide

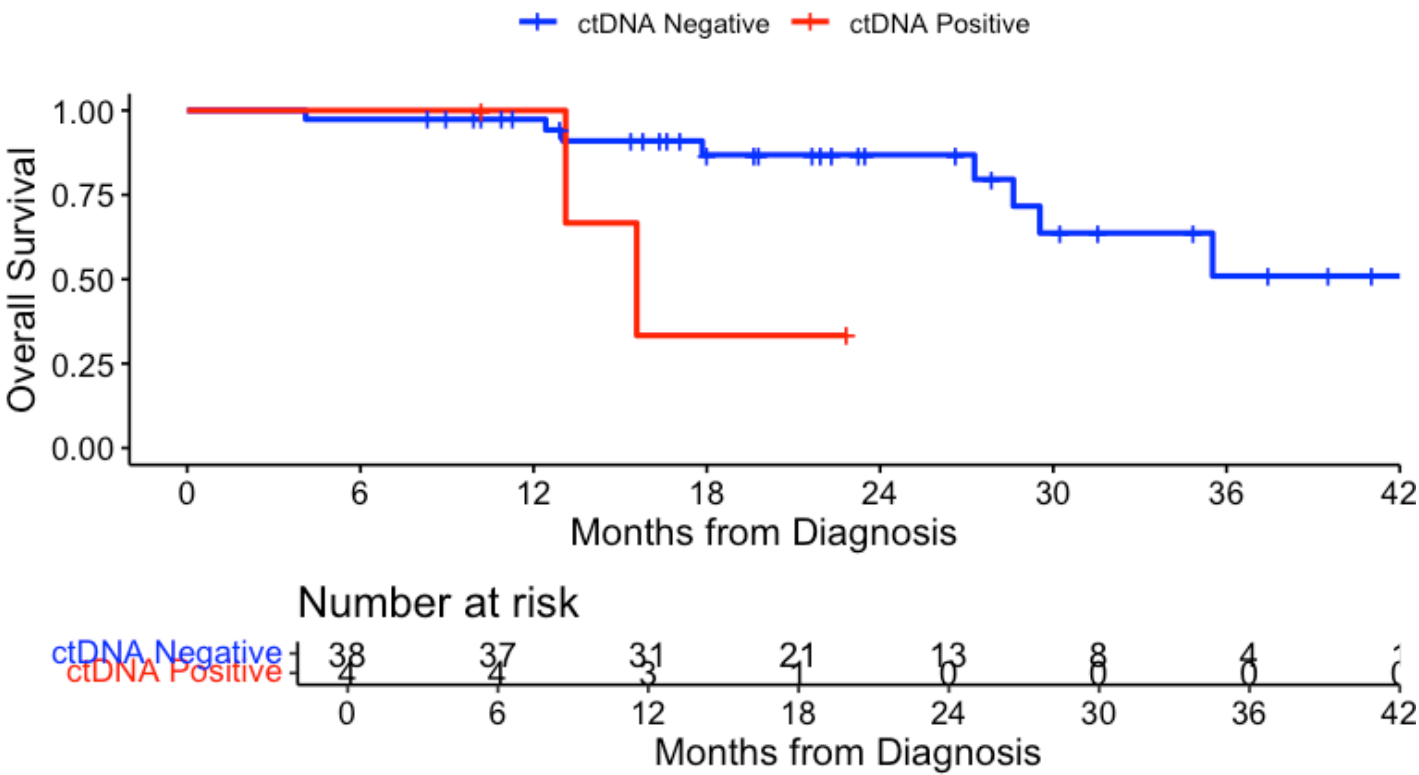
```
event_summary <- circ_data %>%
  group_by(ctDNA.postTx) %>%
  summarise(
    Total = n(),
    Events = sum(OS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.postTx <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	38	8	0.2105263	21.05263
POSITIVE	4	2	0.5000000	50.00000
2 rows				

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.postTx, data = circ_data, conf.int=0.95, conf.type
="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE,
break.time.by=6, palette=c("blue","red"), title="OS - ctDNA post-treatment", ylab= "Overall Survival",
xlab="Months from Diagnosis", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

OS - ctDNA post-treatment



Hide

```
summary(KM_curve, times= c(0, 24, 36))
```

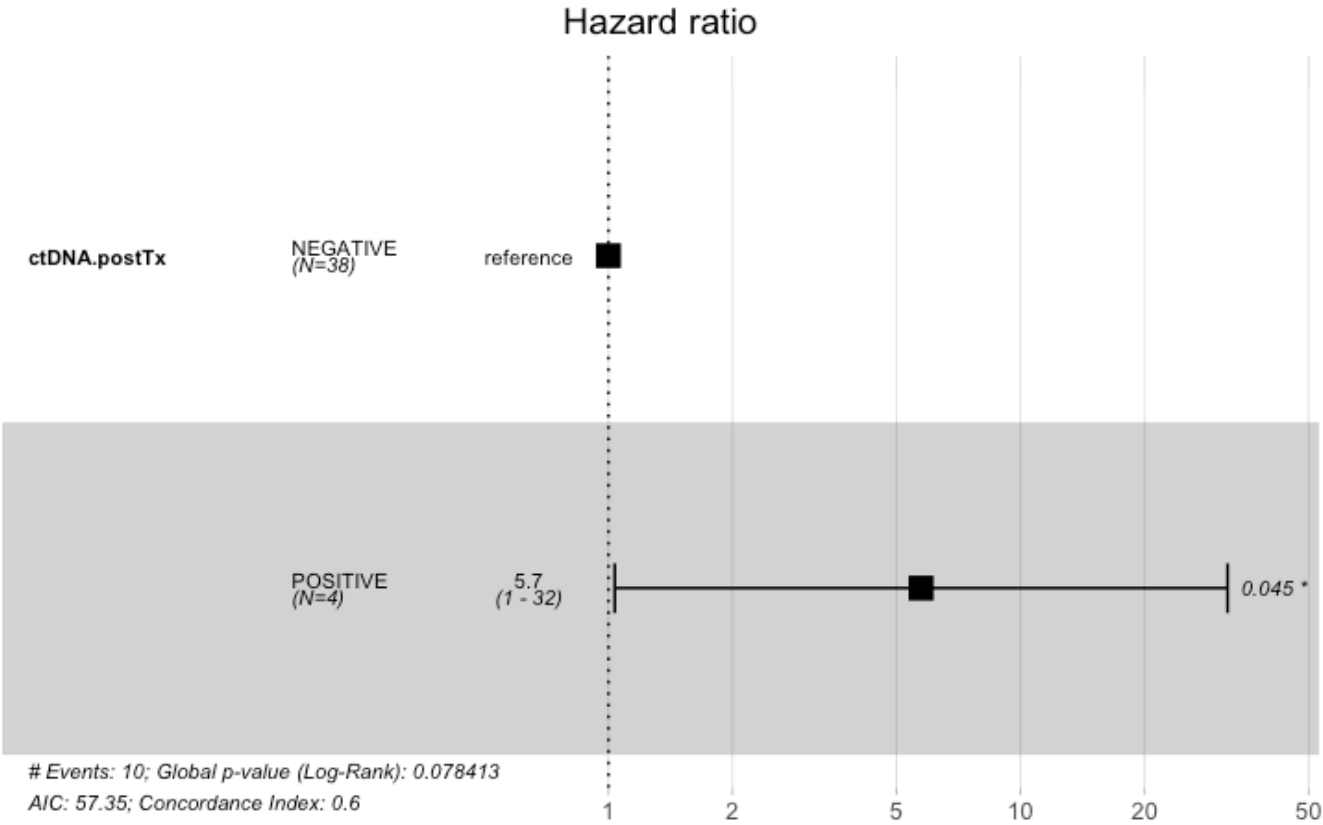
Call: `survfit(formula = surv_object ~ ctDNA.postTx, data = circ_data, conf.int = 0.95, conf.type = "log-log")`

ctDNA.postTx=NEGATIVE								
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI		
0	38	0	1.000	0.0000	1.000	1.000		
24	13	4	0.868	0.0625	0.681	0.949		
36	4	4	0.509	0.1509	0.203	0.752		

ctDNA.postTx=POSITIVE								
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
1	0	4	0	1	0	1		

Hide

```
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE", "POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
ggforest(cox_fit, data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:
coxph(formula = surv_object ~ ctDNA.postTx, data = circ_data)

n= 42, number of events= 10

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.postTxPOSITIVE	1.7472	5.7385	0.8736	2	0.0455 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.postTxPOSITIVE	5.738	0.1743	1.036	31.8

Concordance= 0.6 (se = 0.074)
Likelihood ratio test= 3.1 on 1 df, p=0.08
Wald test = 4 on 1 df, p=0.05
Score (logrank) test = 5.11 on 1 df, p=0.02

Hide

```
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
print(label_text)
```

```
[1] "HR = 5.74 (1.04-31.8); p = 0.045"
```

[Hide](#)

```
#Fisher test for RFS percentages at 12, 24, and 36 months
dfs_times <- c(12, 24, 36)
p_values <- sapply(dfs_times, function(time) {
  neg_count <- sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$FU.Biopsy.months >=
time & circ_data$OS.Event == 0)
  pos_count <- sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$FU.Biopsy.months >=
time & circ_data$OS.Event == 0)
  neg_total <- sum(circ_data$ctDNA.postTx == "NEGATIVE")
  pos_total <- sum(circ_data$ctDNA.postTx == "POSITIVE")

  neg_surv <- neg_total - sum(circ_data$ctDNA.postTx == "NEGATIVE" & circ_data$OS.Event
== 1 & circ_data$FU.Biopsy.months < time)
  pos_surv <- pos_total - sum(circ_data$ctDNA.postTx == "POSITIVE" & circ_data$OS.Event
== 1 & circ_data$FU.Biopsy.months < time)

  surv_matrix <- matrix(c(neg_surv, pos_surv, neg_total - neg_surv, pos_total - pos_sur
v), nrow = 2)
  test_result <- fisher.test(surv_matrix)
  return(test_result$p.value)
})
names(p_values) <- paste0("p-value at ", dfs_times, " months")
print(p_values)
```

```
p-value at 12 months p-value at 24 months p-value at 36 months
1.00000000 0.09099437 0.23559367
```

```
#Barplot for Death % by ctDNA post-treatment
```

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]

circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels = c("NEGATIVE", "POSITIVE"), labels = c("Negative", "Positive"))
circ_data$OS.Event <- factor(circ_data$OS.Event, levels = c("FALSE", "TRUE"), labels = c("Alive", "Deceased"))
contingency_table <- table(circ_data$ctDNA.postTx, circ_data$OS.Event)
chi_square_test <- chisq.test(contingency_table)
```

Warning in stats::chisq.test(x, y, ...) :
Chi-squared approximation may be incorrect

[Hide](#)

```
print(chi_square_test)
```

Pearson's Chi-squared test with Yates' continuity correction

```
data: contingency_table
X-squared = 0.45678, df = 1, p-value = 0.4991
```

[Hide](#)

```
fisher_exact_test <- fisher.test(contingency_table)
print(fisher_exact_test)
```

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.2356
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.2285925 57.0740405
sample estimates:
odds ratio
 3.604577
```

[Hide](#)

```
print(contingency_table)
```

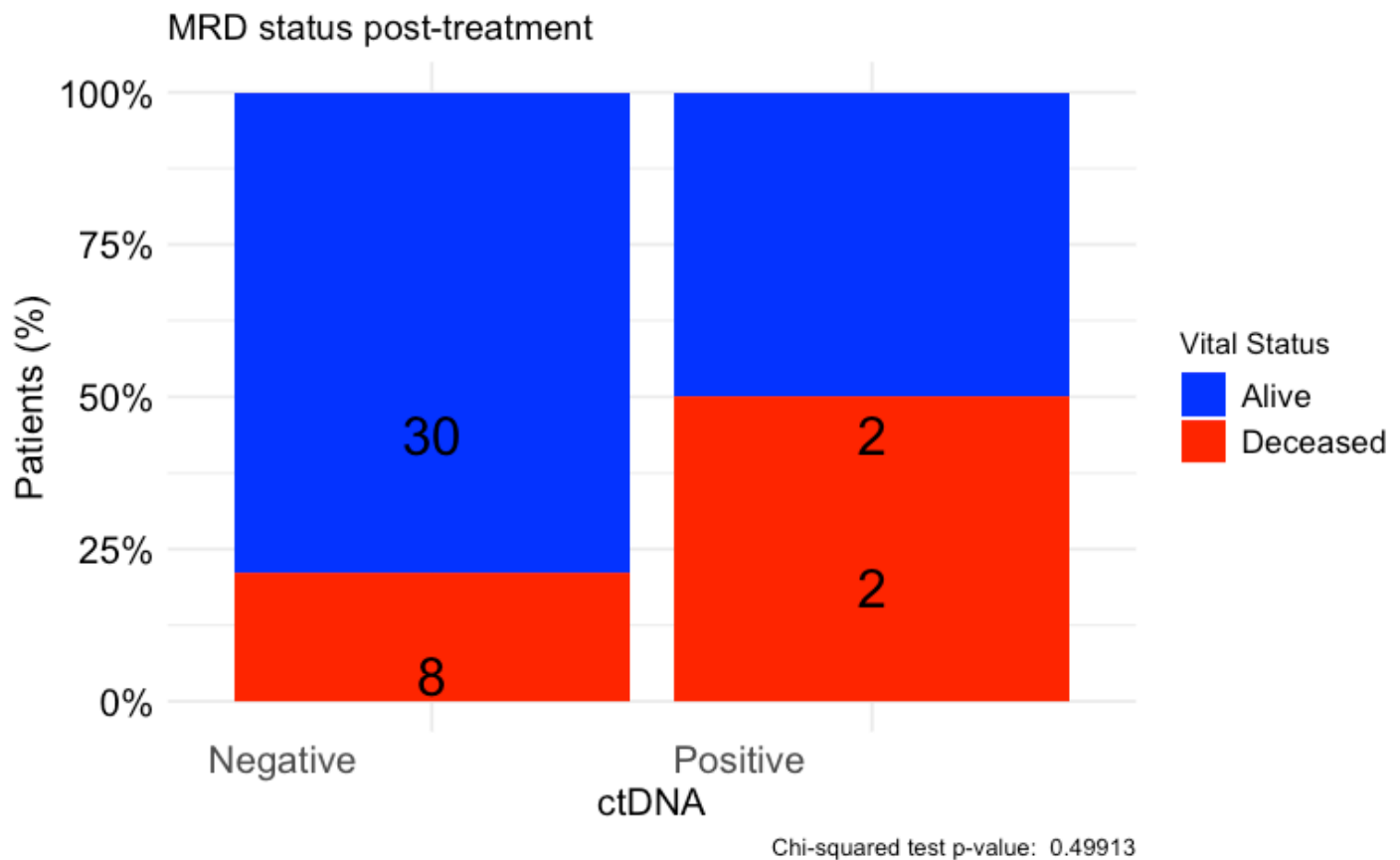

	Alive	Deceased
Negative	30	8
Positive	2	2

Hide

```

table_df <- as.data.frame(contingency_table)
table_df$Total <- ave(table_df$Freq, table_df$Var1, FUN = sum)
table_df$Percentage <- table_df$Freq / table_df$Total
table_df$MiddlePercentage <- table_df$Percentage / 2
ggplot(table_df, aes(x = Var1, y = Percentage, fill = Var2)) +
  geom_bar(stat = "identity") +
  geom_text(aes(y = MiddlePercentage, label = Freq), position = "stack", color = "black",
    vjust = 1.5, size = 7) +
  theme_minimal() +
  labs(title = "MRD status post-treatment",
    x = "ctDNA",
    y = "Patients (%)",
    fill = "Vital Status",
    caption = paste("Chi-squared test p-value: ", format.pval(chi_square_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("Alive" = "blue", "Deceased" = "red")) + # define custom colors
  theme(axis.text.x = element_text(angle = 0, hjust = 1.5, size = 14), # increase x-axis text size
    axis.text.y = element_text(size = 14, color = "black"), # increase y-axis text size
    axis.title.x = element_text(size = 14, color = "black"), # increase x-axis label size
    axis.title.y = element_text(size = 14, color = "black"), # increase y-axis label size
    legend.text = element_text(size = 12, color = "black")) # increase Recurrence label size

```



#Multivariate cox regression for RFS

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$Gender <- factor(circ_data$Gender, levels = c("Female", "Male"), labels = c("Female", "Male"))
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Larynx"))
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c("0", "1/2"))
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV"))
circ_data$NAC <- factor(circ_data$NAC, levels = c("TRUE", "FALSE"))
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative"))
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.postTx + Location + cStage + ECOG.score + p16, data = circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for RFS", refLabel = "Reference Group")
test.ph <- cox.zph(cox_fit)
```

#Univariate RFS cox regression for variables included in MVA

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive")) #univariate for ctDNA post-treatment
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
summary(cox_fit)
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
print(label_text)

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Larynx")) #univariate for Primary location
cox_fit <- coxph(surv_object ~ Location, data=circ_data)
summary(cox_fit)
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
print(label_text)

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV")) #univariate for Stage
cox_fit <- coxph(surv_object ~ cStage, data=circ_data)
summary(cox_fit)

```

```

cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_C
I, 2), "); p = ", round(p_value, 3))
print(label_text)

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c
("0", "1/2")) #univariate for ECOG
cox_fit <- coxph(surv_object ~ ECOG.score, data=circ_data)
summary(cox_fit)
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_C
I, 2), "); p = ", round(p_value, 3))
print(label_text)

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative")) #univariate f
or p16
cox_fit <- coxph(surv_object ~ p16, data=circ_data)
summary(cox_fit)
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_C
I, 2), "); p = ", round(p_value, 3))
print(label_text)

```

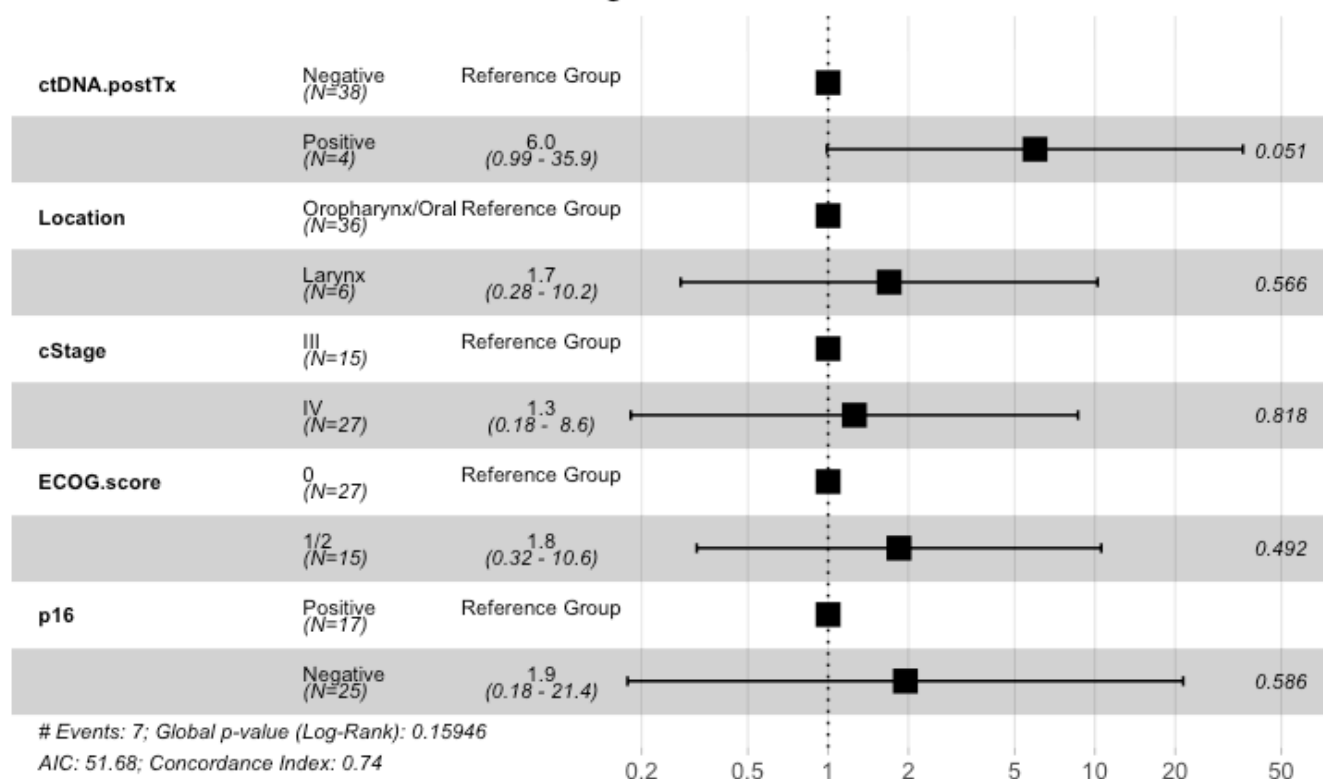
#Multivariate cox regression for RFS - Recurrence at year 1

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$Gender <- factor(circ_data$Gender, levels = c("Female", "Male"), labels = c("Female", "Male"))
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Larynx"))
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c("0", "1/2"))
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV"))
circ_data$NAC <- factor(circ_data$NAC, levels = c("TRUE", "FALSE"))
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative"))
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
cox_fit <- coxph(surv_object ~ ctDNA.postTx + Location + cStage + ECOG.score + p16, data=circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for RFS", refLabel = "Reference Group")
```

Multivariate Regression Model for RFS



Hide

```
test.ph <- cox.zph(cox_fit)
```

#Univariate RFS cox regression for variables included in MVA - Recurrence at year 1

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive")) #univariate for ctDNA post-treatment
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.postTx, data = circ_data)
```

```
n= 42, number of events= 7
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.postTxPositive	2.1937	8.9685	0.7745	2.832	0.00462 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.postTxPositive	8.969	0.1115	1.966	40.92

Concordance= 0.675 (se = 0.088)

Likelihood ratio test= 6.36 on 1 df, p=0.01

Wald test = 8.02 on 1 df, p=0.005

Score (logrank) test = 11.69 on 1 df, p=6e-04

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 8.97 (1.97-40.92); p = 0.005"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
```

```
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Larynx")) #univariate for Primary location
```

```
cox_fit <- coxph(surv_object ~ Location, data=circ_data)
```

```
summary(cox_fit)
```


Call:

```
coxph(formula = surv_object ~ Location, data = circ_data)
```

```
n= 42, number of events= 7
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
LocationLarynx	0.9057	2.4737	0.8369	1.082	0.279

	exp(coef)	exp(-coef)	lower .95	upper .95
LocationLarynx	2.474	0.4042	0.4797	12.76

Concordance= 0.586 (se = 0.089)

Likelihood ratio test= 1.01 on 1 df, p=0.3

Wald test = 1.17 on 1 df, p=0.3

Score (logrank) test = 1.25 on 1 df, p=0.3

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 2.47 (0.48-12.76); p = 0.279"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
```

```
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV")) #univariate for Stage
```

```
cox_fit <- coxph(surv_object ~ cStage, data=circ_data)
```

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ cStage, data = circ_data)
```

```
n= 42, number of events= 7
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
cStageIV	-0.2073	0.8128	0.7650	-0.271	0.786

	exp(coef)	exp(-coef)	lower .95	upper .95
cStageIV	0.8128	1.23	0.1815	3.64

Concordance= 0.519 (se = 0.094)

Likelihood ratio test= 0.07 on 1 df, p=0.8

Wald test = 0.07 on 1 df, p=0.8

Score (logrank) test = 0.07 on 1 df, p=0.8

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 0.81 (0.18-3.64); p = 0.786"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
```

```
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c("0", "1/2")) #univariate for ECOG
```

```
cox_fit <- coxph(surv_object ~ ECOG.score, data=circ_data)
```

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ECOG.score, data = circ_data)
```

n= 42, number of events= 7

	coef	exp(coef)	se(coef)	z	Pr(> z)
ECOG.score1/2	1.1173	3.0567	0.7654	1.46	0.144

	exp(coef)	exp(-coef)	lower .95	upper .95
ECOG.score1/2	3.057	0.3272	0.6819	13.7

Concordance= 0.643 (se = 0.092)

Likelihood ratio test= 2.14 on 1 df, p=0.1

Wald test = 2.13 on 1 df, p=0.1

Score (logrank) test = 2.36 on 1 df, p=0.1

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

#Extract values for HR, 95% CI, and p-value

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 3.06 (0.68-13.7); p = 0.144"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event.1y)
```

```
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative")) #univariate for p16
```

```
cox_fit <- coxph(surv_object ~ p16, data=circ_data)
```

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ p16, data = circ_data)
```

```
n= 42, number of events= 7
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
p16Negative	1.459	4.304	1.081	1.35	0.177

	exp(coef)	exp(-coef)	lower .95	upper .95
p16Negative	4.304	0.2324	0.5173	35.8

```
Concordance= 0.654 (se = 0.059 )
```

```
Likelihood ratio test= 2.49 on 1 df, p=0.1
```

```
Wald test = 1.82 on 1 df, p=0.2
```

```
Score (logrank) test = 2.17 on 1 df, p=0.1
```

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 4.3 (0.52-35.8); p = 0.177"
```

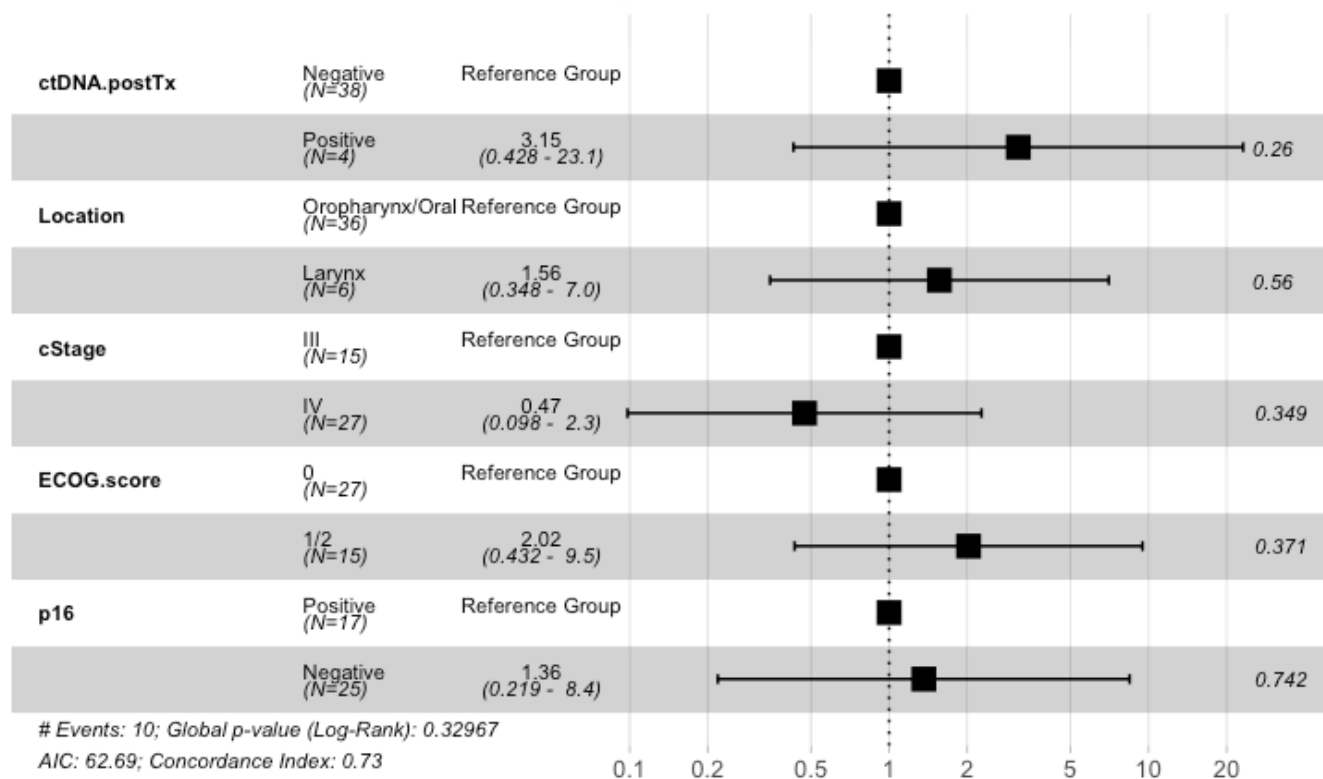
```
#Multivariate cox regression for OS
```

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$Gender <- factor(circ_data$Gender, levels = c("Female", "Male"), labels = c("Female", "Male"))
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Larynx"))
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c("0", "1/2"))
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV"))
circ_data$NAC <- factor(circ_data$NAC, levels = c("TRUE", "FALSE"))
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative"))
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.postTx + Location + cStage + ECOG.score + p16, data=circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for OS", refLabel = "Reference Group")
```

Multivariate Regression Model for OS


[Hide](#)

```
test.ph <- cox.zph(cox_fit)
```

#Univariate OS cox regression for variables included in MVA

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
circ_data$ctDNA.postTx <- factor(circ_data$ctDNA.postTx, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive")) #univariate for ctDNA post-treatment
cox_fit <- coxph(surv_object ~ ctDNA.postTx, data=circ_data)
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ ctDNA.postTx, data = circ_data)

n= 42, number of events= 10

              coef exp(coef) se(coef) z Pr(>|z|)
ctDNA.postTxPositive 1.7472    5.7385  0.8736 2  0.0455 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
ctDNA.postTxPositive    5.738    0.1743    1.036    31.8

Concordance= 0.6 (se = 0.074 )
Likelihood ratio test= 3.1 on 1 df,  p=0.08
Wald test               = 4 on 1 df,  p=0.05
Score (logrank) test = 5.11 on 1 df,  p=0.02
```

Hide

```
cox_fit_summary <- summary(cox_fit)

#Extract values for HR, 95% CI, and p-value
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_C
I, 2), "); p = ", round(p_value, 3))
print(label_text)
```

```
[1] "HR = 5.74 (1.04-31.8); p = 0.045"
```

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("OP-16933_HNSCC Clinical data.csv")
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
circ_data$Location <- factor(circ_data$Location, levels = c("Oropharynx/Oral", "Laryn
x")) #univariate for Primary location
cox_fit <- coxph(surv_object ~ Location, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ Location, data = circ_data)
```

```
n= 42, number of events= 10
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
LocationLarynx	0.7183	2.0509	0.6926	1.037	0.3

	exp(coef)	exp(-coef)	lower .95	upper .95
LocationLarynx	2.051	0.4876	0.5277	7.97

Concordance= 0.553 (se = 0.07)

Likelihood ratio test= 0.97 on 1 df, p=0.3

Wald test = 1.08 on 1 df, p=0.3

Score (logrank) test = 1.12 on 1 df, p=0.3

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 2.05 (0.53-7.97); p = 0.3"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
```

```
circ_data$cStage <- factor(circ_data$cStage, levels = c("III", "IV")) #univariate for Stage
```

```
cox_fit <- coxph(surv_object ~ cStage, data=circ_data)
```

```
summary(cox_fit)
```


Call:

```
coxph(formula = surv_object ~ cStage, data = circ_data)
```

```
n= 42, number of events= 10
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
cStageIV	-0.5638	0.5690	0.6561	-0.859	0.39

	exp(coef)	exp(-coef)	lower .95	upper .95
cStageIV	0.569	1.757	0.1573	2.059

Concordance= 0.565 (se = 0.09)

Likelihood ratio test= 0.71 on 1 df, p=0.4

Wald test = 0.74 on 1 df, p=0.4

Score (logrank) test = 0.76 on 1 df, p=0.4

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 0.57 (0.16-2.06); p = 0.39"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
```

```
circ_data$ECOG.score <- factor(circ_data$ECOG.score, levels = c("0", "1"), labels = c("0", "1/2")) #univariate for ECOG
```

```
cox_fit <- coxph(surv_object ~ ECOG.score, data=circ_data)
```

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ECOG.score, data = circ_data)
```

n= 42, number of events= 10

	coef	exp(coef)	se(coef)	z	Pr(> z)
ECOG.score1/2	0.9052	2.4724	0.6403	1.414	0.157

	exp(coef)	exp(-coef)	lower .95	upper .95
ECOG.score1/2	2.472	0.4045	0.7048	8.673

Concordance= 0.701 (se = 0.08)

Likelihood ratio test= 1.94 on 1 df, p=0.2

Wald test = 2 on 1 df, p=0.2

Score (logrank) test = 2.13 on 1 df, p=0.1

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

#Extract values for HR, 95% CI, and p-value

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 2.47 (0.7-8.67); p = 0.157"
```

[Hide](#)

```
rm(list=ls())
```

```
setwd("~/Downloads")
```

```
circ_data <- read.csv("0P-16933_HNSCC Clinical data.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.available=="TRUE",]
```

```
circ_data <- circ_data[circ_data$ctDNA.postTx!="",]
```

```
surv_object <- Surv(time = circ_data$FU.Biopsy.months, event = circ_data$OS.Event)
```

```
circ_data$p16 <- factor(circ_data$p16, levels = c("Positive", "Negative")) #univariate for p16
```

```
cox_fit <- coxph(surv_object ~ p16, data=circ_data)
```

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ p16, data = circ_data)
```

```
n= 42, number of events= 10
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
p16Negative	0.5330	1.7041	0.6911	0.771	0.441

	exp(coef)	exp(-coef)	lower .95	upper .95
p16Negative	1.704	0.5868	0.4397	6.604

Concordance= 0.583 (se = 0.092)

Likelihood ratio test= 0.63 on 1 df, p=0.4

Wald test = 0.59 on 1 df, p=0.4

Score (logrank) test = 0.61 on 1 df, p=0.4

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
```

```
#Extract values for HR, 95% CI, and p-value
```

```
HR <- cox_fit_summary$coefficients[2]
```

```
lower_CI <- cox_fit_summary$conf.int[3]
```

```
upper_CI <- cox_fit_summary$conf.int[4]
```

```
p_value <- cox_fit_summary$coefficients[5]
```

```
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
```

```
print(label_text)
```

```
[1] "HR = 1.7 (0.44-6.6); p = 0.441"
```