

ASAN Cholangio_GL_Final analysis 092024

Code ▾

```
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(rms)

#ctDNA positivity by stage and window
```

Hide

```
#ctDNA at MRD
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.MRD %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("II","III","IV"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.MRD == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.MRD, 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.MRD == "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>
II	38	7	18.42%
III	39	10	25.64%

Stage<fctr>	Total_Count<int>	Positive_Count<int>	Rate<chr>
IV	12	5	41.67%
Overall	89	22	24.72%

4 rows

Hide

```
#ctDNA C5D1
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.C5D1!="",]
circ_data$ctDNA.C5D1 <- factor(circ_data$ctDNA.C5D1, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.C5D1 %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("II","III","IV"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.C5D1 == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.C5D1, 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.C5D1 == "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>
II	37	5	13.51%
III	39	8	20.51%
IV	12	4	33.33%

Stage <fctr>	Total_Count <int>	Positive_Count <int>	Rate <chr>
Overall	88	17	19.32%

4 rows

Hide

```
#ctDNA C8D1
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.C8D1!="",]
circ_data$ctDNA.C8D1 <- factor(circ_data$ctDNA.C8D1, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.C8D1 %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("II","III","IV"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.C8D1 == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.C8D1, 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.C8D1 == "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>
II	31	5	16.13%
III	38	8	21.05%
IV	8	2	25.00%
Overall	77	15	19.48%

4 rows

Hide

```
#ctDNA EOT
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.EOT!="",]
circ_data$ctDNA.EOT <- factor(circ_data$ctDNA.EOT, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.EOT %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("II","III","IV"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.EOT == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.EOT, 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.EOT == "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>
II	14	11	78.57%
III	20	12	60.00%
IV	7	6	85.71%
Overall	41	29	70.73%

4 rows

Hide

```
#ctDNA anytime
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.anytime!="",]
circ_data$ctDNA.anytime <- factor(circ_data$ctDNA.anytime, levels=c("NEGATIVE","POSITIVE"))
circ_data <- subset(circ_data, ctDNA.anytime %in% c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("II","III","IV"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.anytime == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.anytime, 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.anytime == "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>
II	38	17	44.74%
III	39	21	53.85%
IV	12	9	75.00%
Overall	89	47	52.81%
4 rows			

#Heatmap for clinicopathologic factors

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_datadf <- as.data.frame(circ_data)
circ_data <- circ_data %>% arrange(Stage)
circ_datadf <- as.data.frame(circ_data)

ha <- HeatmapAnnotation(
  Stage = circ_data$Stage,
  Sex = circ_data$Sex,
  PrimSite = circ_data$PrimSite,
  pT = circ_data$pT,
  Pathology = circ_data$Pathology,
  Chemo = circ_data$Chemo,
  ctDNA.MRD = circ_data$ctDNA.MRD,
  ctDNA.C5D1 = circ_data$ctDNA.C5D1,
  ctDNA.C8D1 = circ_data$ctDNA.C8D1,
  ctDNA.EOT = circ_data$ctDNA.EOT,
  ctDNA.anytime = circ_data$ctDNA.anytime,
  RecStatus = circ_data$RecStatus,
  VitalStatus = circ_data$VitalStatus,

  col = list(Stage = c("II" = "seagreen2", "III" = "orange", "IV" = "purple"),
    Sex = c("Female" = "goldenrod", "Male" = "blue4"),
    PrimSite = c("pCCA" = "darkgreen", "dCCA" = "#008BCE"),
    pT = c("T1" = "lightblue", "T2" = "orange", "T3" = "brown" ),
    Pathology = c("G1" = "yellow3", "G2" = "darkgreen", "G3" = "brown2"),
    Chemo = c("CAP" = "lightblue", "GemCis" = "orange2"),
    ctDNA.MRD = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    ctDNA.C5D1 = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    ctDNA.C8D1 = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    ctDNA.EOT = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
    ctDNA.anytime = 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

```

#Prognostic role of ctDNA at the MRD time point

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
circ_datadf <- as.data.frame(circ_data)

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

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.MRD=NEGATIVE	67	45	15.84	11.96	27.6
ctDNA.MRD=POSITIVE	22	20	8.23	7.86	15.5

Hide

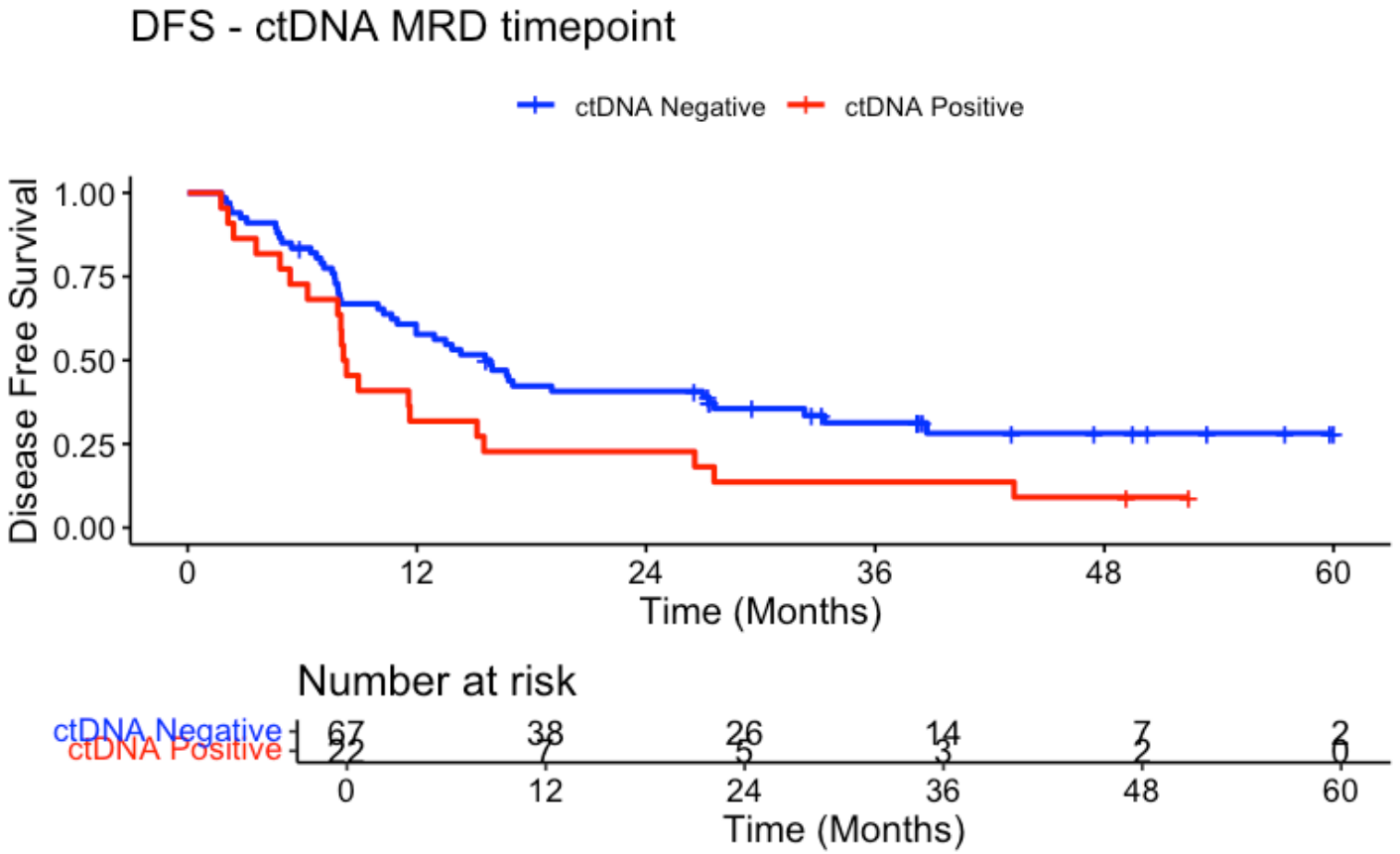
```
event_summary <- circ_data %>%
  group_by(ctDNA.MRD) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.MRD	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	67	45	0.6716418	67.16418
POSITIVE	22	20	0.9090909	90.90909

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.MRD, data = circ_data,conf.int=0.95,conf.type="l
og-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=12, palette=c("blue","red"), title="DFS - ctDNA MRD timepoint", ylab=
"Disease Free Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Po
sitive"), legend.title="")
```



Hide

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



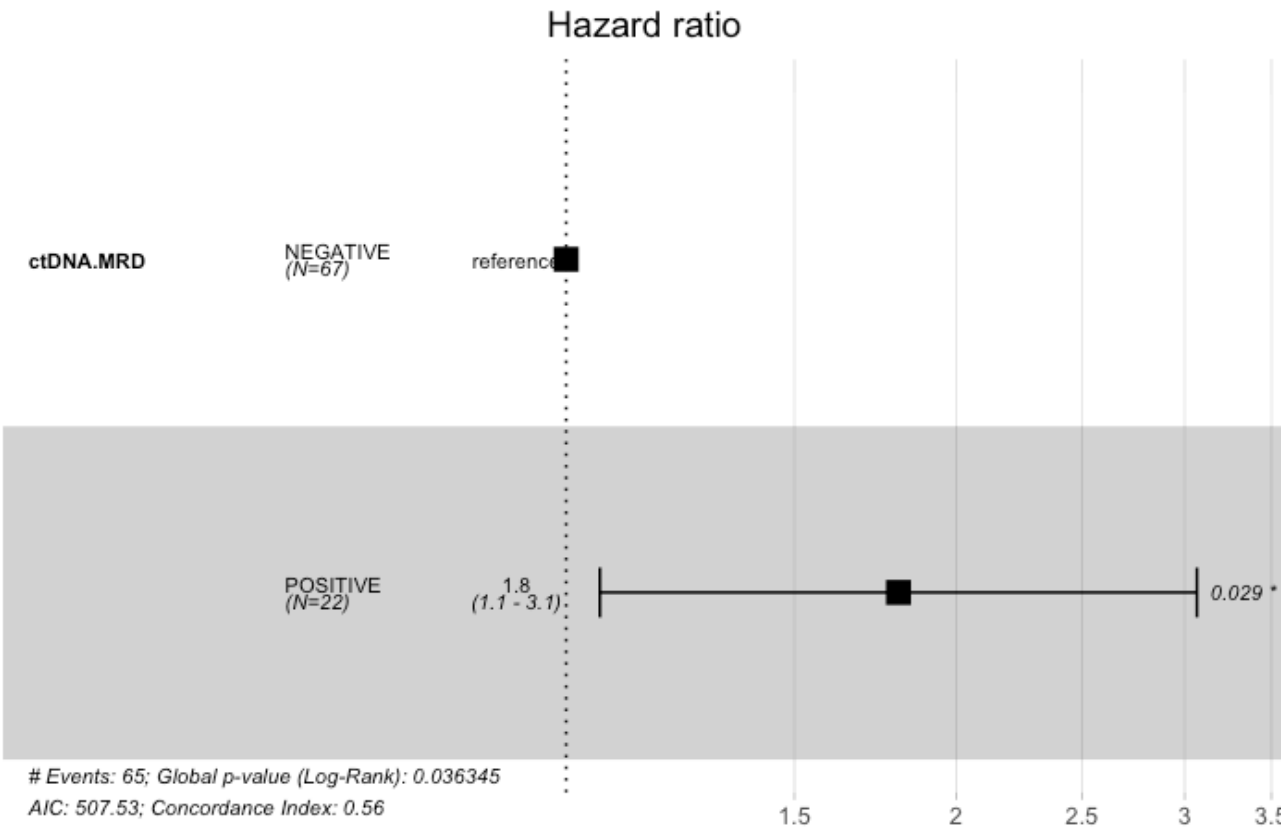
```
Call: survfit(formula = surv_object ~ ctDNA.MRD, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.MRD=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	38	28	0.577	0.0607	0.450	0.686	
24	26	11	0.407	0.0608	0.288	0.523	

ctDNA.MRD=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	7	15	0.318	0.0993	0.1418	0.511	
24	5	2	0.227	0.0893	0.0827	0.414	

Hide

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



Hide

```
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.MRDPOSITIVE	0.5902	1.8043	0.2705	2.182	0.0291 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.MRDPOSITIVE	1.804	0.5542	1.062	3.066

Concordance= 0.556 (se = 0.028)

Likelihood ratio test= 4.38 on 1 df, p=0.04

Wald test = 4.76 on 1 df, p=0.03

Score (logrank) test = 4.9 on 1 df, p=0.03

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 = 1.8 (1.06-3.07); p = 0.029"
```

#OS by ctDNA at the MRD time point

Hide

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

```
circ_data$OS.months=circ_data$OS.months-2
```

```
circ_data <- circ_data[circ_data$OS.months>=0,]
```

```
circ_datadf <- as.data.frame(circ_data)
```

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

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.MRD=NEGATIVE	67	39	38.5	28.9	NA
ctDNA.MRD=POSITIVE	22	17	30.8	22.6	NA

Hide

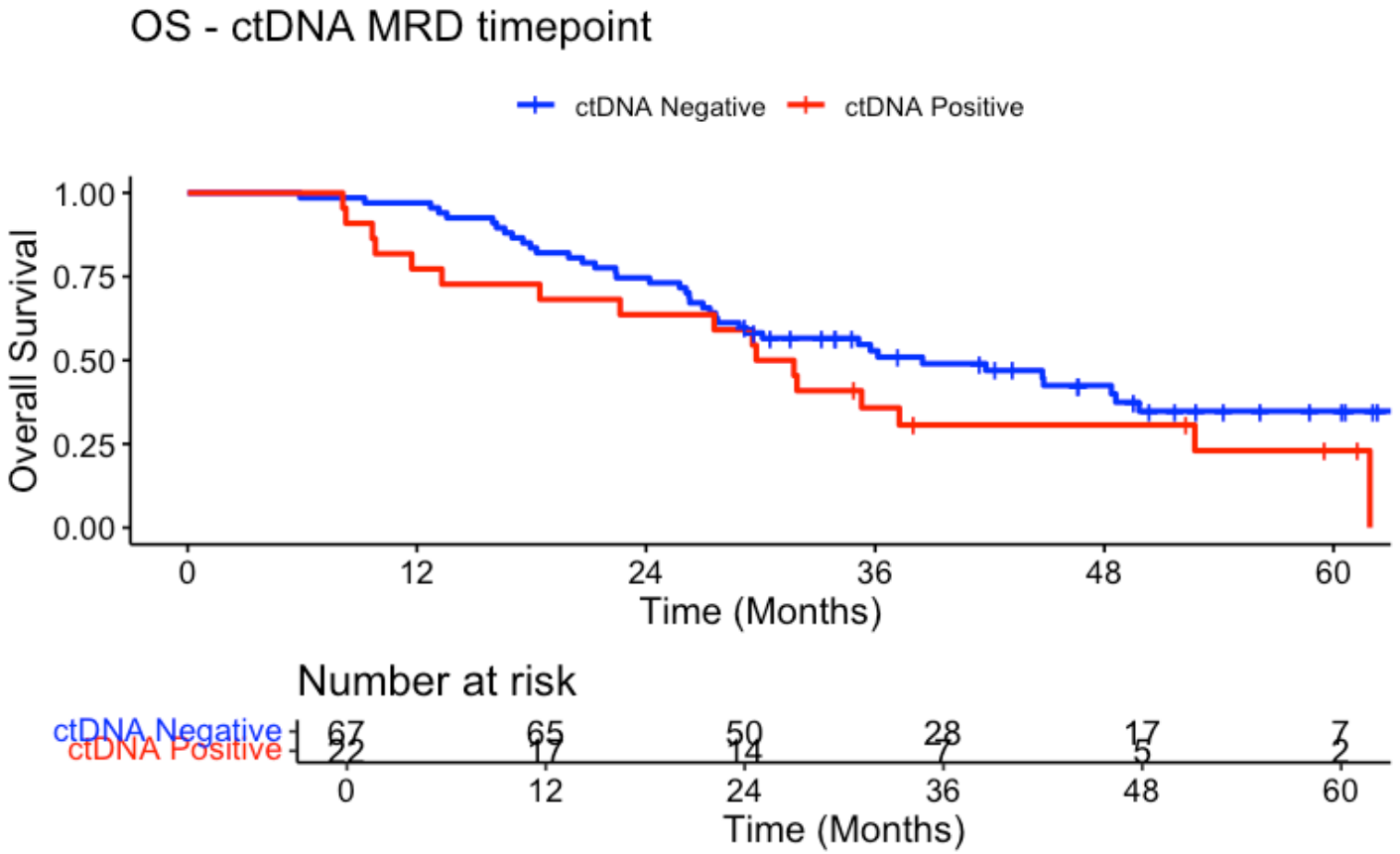
```
event_summary <- circ_data %>%
  group_by(ctDNA.MRD) %>%
  summarise(
    Total = n(),
    Events = sum(OS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.MRD	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	67	39	0.5820896	58.20896
POSITIVE	22	17	0.7727273	77.27273

2 rows

Hide

```
surv_object <-Surv(time = circ_data$OS.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.MRD, 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=12, palette=c("blue","red"), title="OS - ctDNA MRD timepoint", ylab= "Overall Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

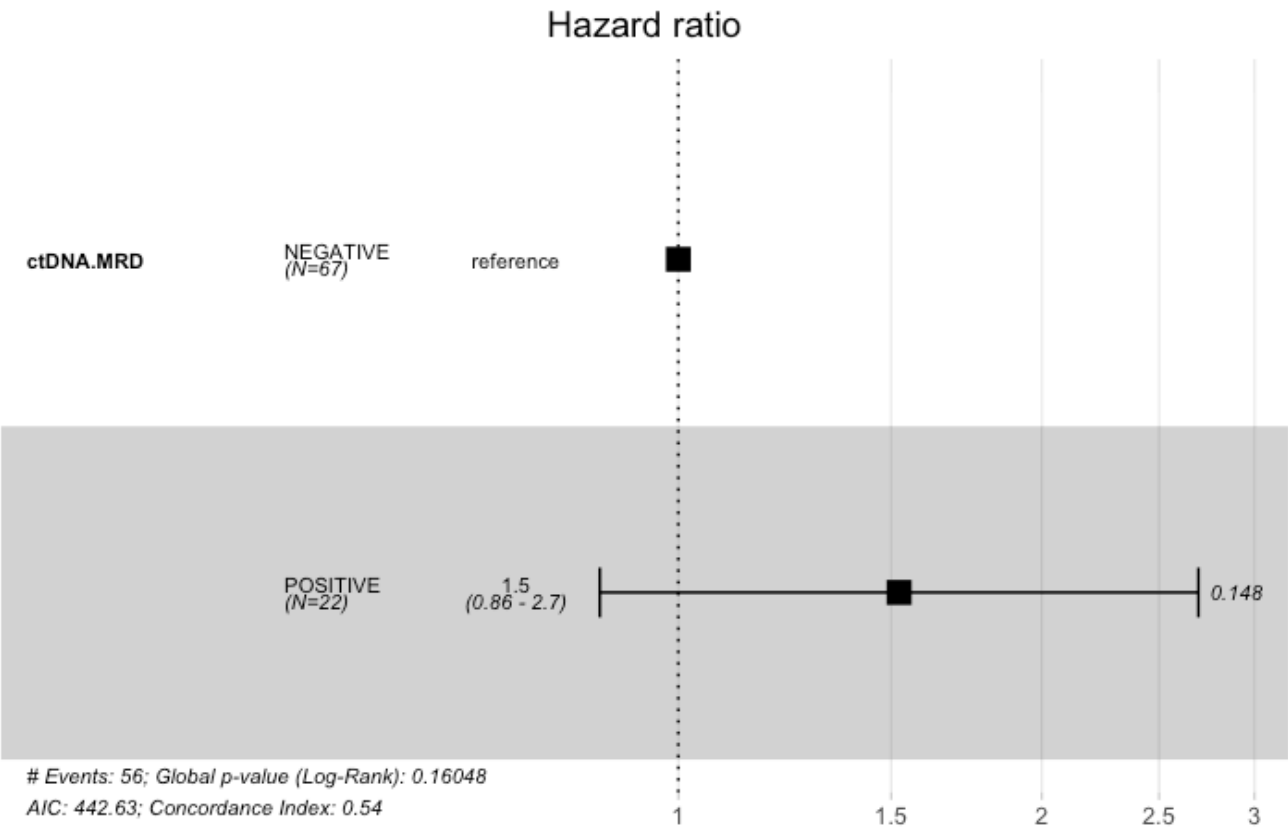
```
Call: survfit(formula = surv_object ~ ctDNA.MRD, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.MRD=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	65	2	0.970	0.0208	0.886	0.992	
24	50	15	0.746	0.0532	0.624	0.834	

ctDNA.MRD=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	17	5	0.773	0.0893	0.537	0.898	
24	14	3	0.636	0.1026	0.403	0.799	

Hide

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



Hide

```
summary(cox_fit)
```

Call:

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

n= 89, number of events= 56

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.MRDPOSITIVE	0.4209	1.5234	0.2911	1.446	0.148

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.MRDPOSITIVE	1.523	0.6564	0.861	2.695

Concordance= 0.54 (se = 0.032)

Likelihood ratio test= 1.97 on 1 df, p=0.2

Wald test = 2.09 on 1 df, p=0.1

Score (logrank) test = 2.12 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 = 1.52 (0.86-2.7); p = 0.148"
```

#Association of ctDNA MRD MTM levels with clinicopathological factors

[Hide](#)

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

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

pN

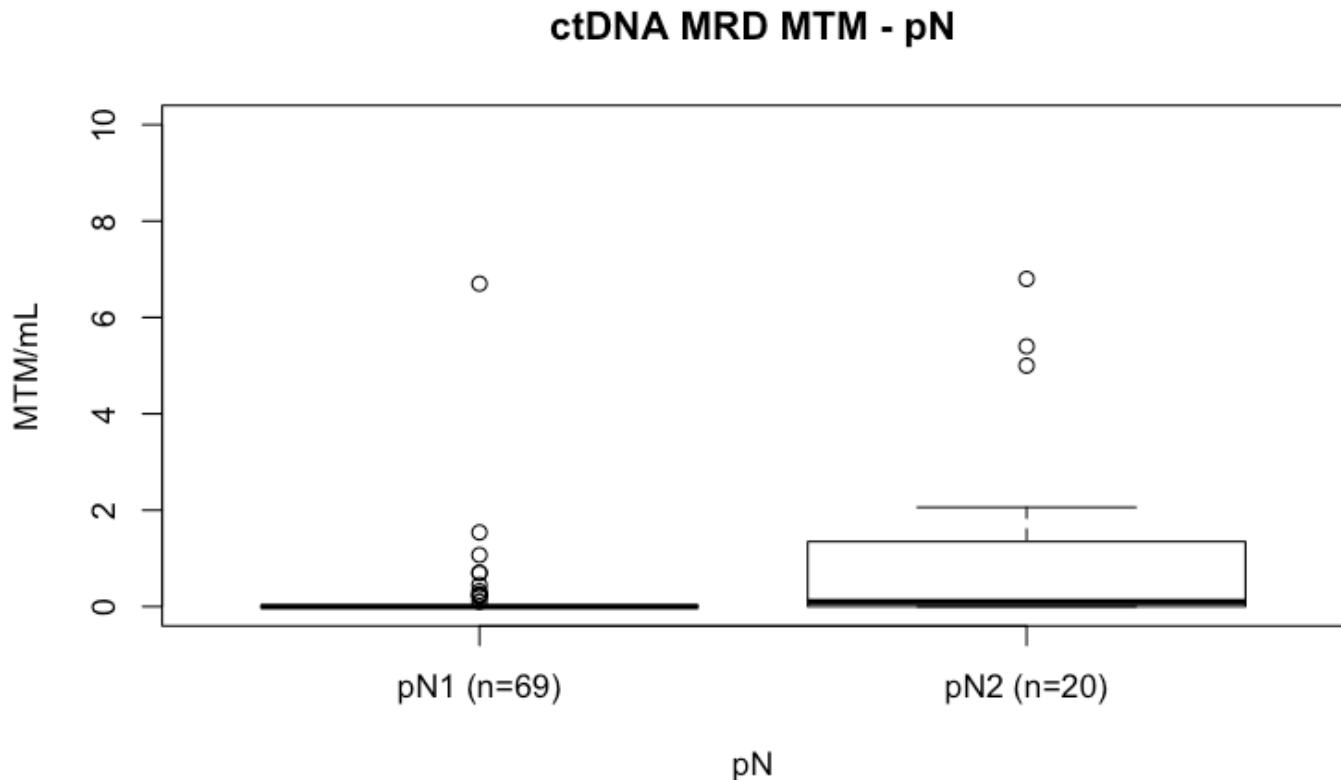
N1	N2	Total
69	20	89

[Hide](#)

```

circ_data$pN <- factor(circ_data$pN, levels = c("N1","N2"), labels = c("pN1 (n=69)","pN2 (n=20)"))
boxplot(ctDNA.MRD.MTM~pN, data=circ_data, main="ctDNA MRD MTM - pN", xlab="pN", ylab="MTM/mL", col="white",border="black", ylim=c(0, 10))

```


[Hide](#)

```

m1<-wilcox.test(ctDNA.MRD.MTM ~ pN, data=circ_data, na.rm=TRUE, exact=FALSE, conf.int=TRUE)
print(m1)

```

Wilcoxon rank sum test with continuity correction

```

data: ctDNA.MRD.MTM by pN
W = 441.5, p-value = 0.001285
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -3.700528e-01 -2.534813e-05
sample estimates:
difference in location
 -4.5909e-05

```

[Hide](#)

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

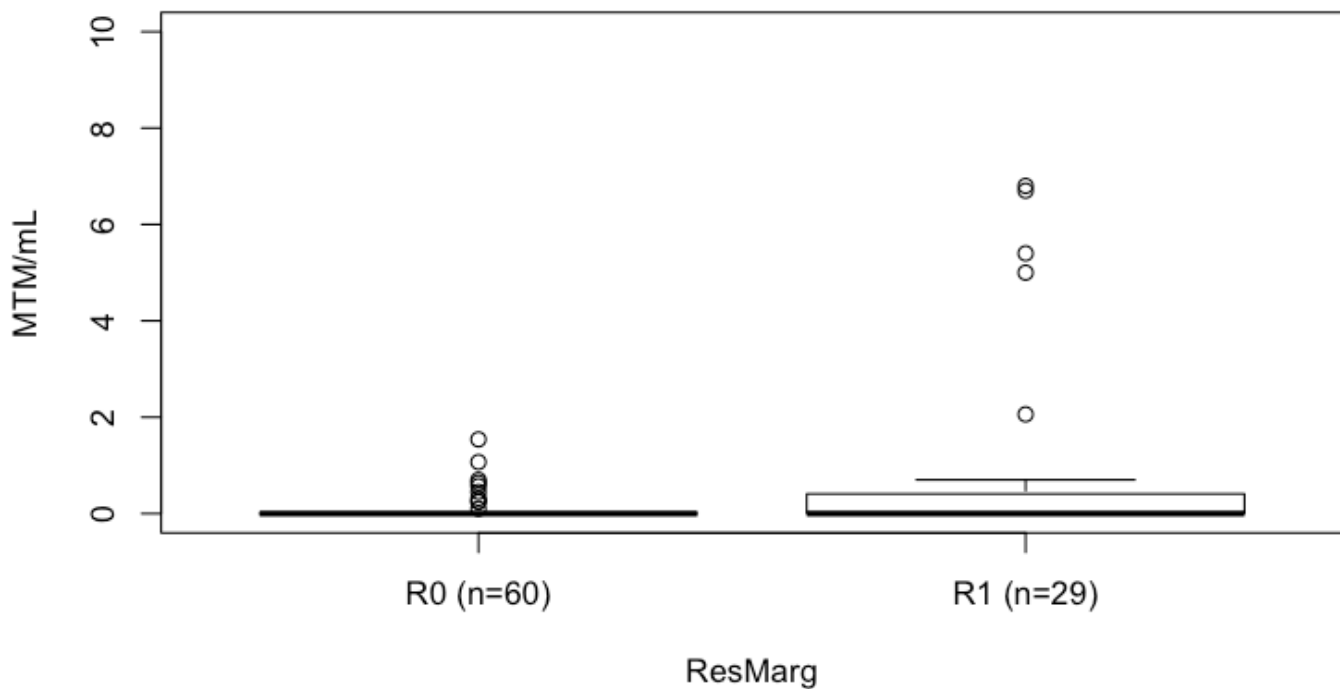
ResMarg

R0	R1	Total
60	29	89

Hide

```
circ_data$ResMarg <- factor(circ_data$ResMarg, levels = c("R0","R1"), labels = c("R0 (n=60)","R1 (n=29)"))  
boxplot(ctDNA.MRD.MTM~ResMarg, data=circ_data, main="ctDNA MRD MTM - Resection Margin",  
        xlab="ResMarg", ylab="MTM/mL", col="white", border="black", ylim=c(0, 10))
```

ctDNA MRD MTM - Resection Margin



Hide

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


Wilcoxon rank sum test with continuity correction

data: ctDNA.MRD.MTM by ResMarg
W = 637.5, p-value = 0.007319
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
-0.1700429119 -0.0000566823
sample estimates:
difference in location
-7.789783e-05

Hide

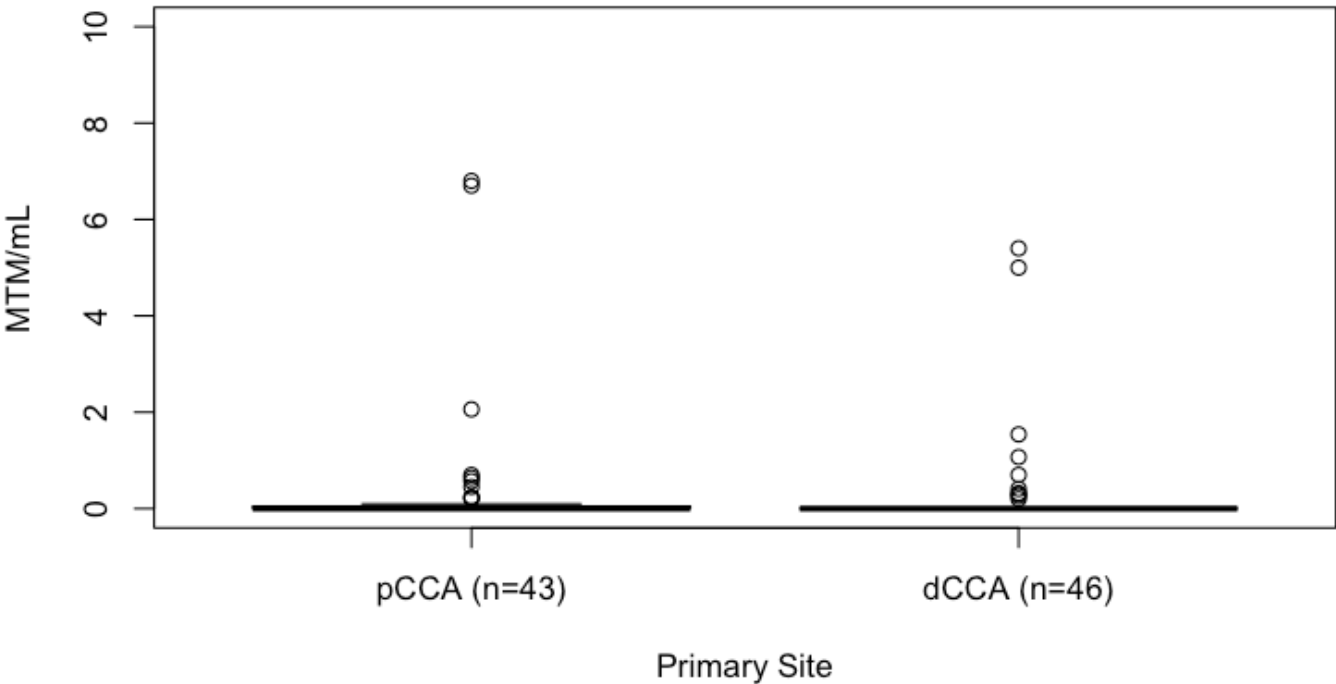
tally(~PrimSite, data=circ_data, margins = TRUE)

PrimSite		
dCCA	pCCA	Total
46	43	89

Hide

```
circ_data$PrimSite <- factor(circ_data$PrimSite, levels = c("pCCA","dCCA"), labels = c("pCCA (n=43)","dCCA (n=46)"))  
boxplot(ctDNA.MRD.MTM~PrimSite, data=circ_data, main="ctDNA MRD MTM - Primary Site",  
        xlab="Primary Site", ylab="MTM/mL", col="white", border="black", ylim=c(0, 10))
```

ctDNA MRD MTM - Primary Site



Hide

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

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.MRD.MTM by PrimSite
W = 996, p-value = 0.9438
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -2.850159e-05  9.561474e-06
sample estimates:
difference in location
      4.711544e-05
```

Hide

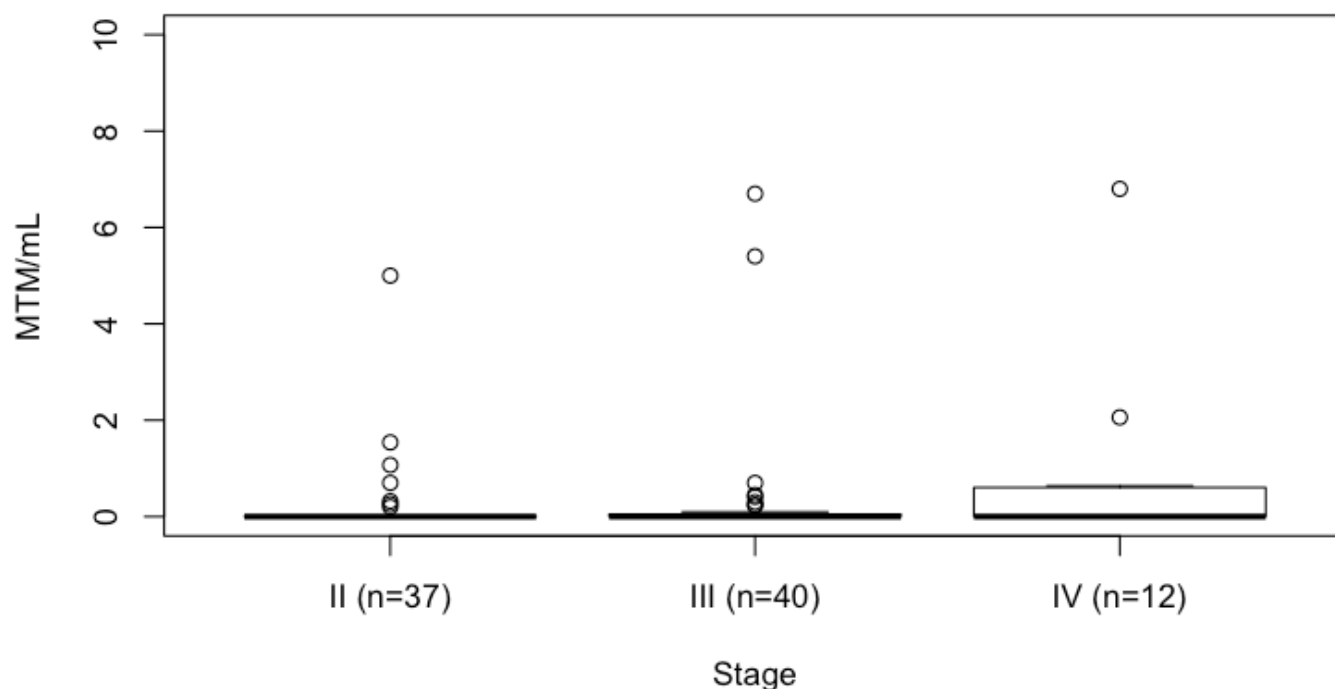
```
tally(~Stage, data=circ_data, margins = TRUE)
```

Stage				
II	III	IV	Total	
38	39	12	89	

Hide

```
circ_data$ctDNA.MRD.MTM <- as.numeric(as.character(circ_data$ctDNA.MRD.MTM))
circ_data$Stage <- factor(circ_data$Stage, levels = c("II","III","IV"), labels = c("II
(n=37)", "III (n=40)", "IV (n=12)"))
boxplot(ctDNA.MRD.MTM~Stage, data=circ_data, main="ctDNA MRD MTM - Stage",
        xlab="Stage", ylab="MTM/mL", col="white", border="black", ylim=c(0, 10))
```

ctDNA MRD MTM - Stage


[Hide](#)

```
kruskal_result <- kruskal.test(ctDNA.MRD.MTM ~ Stage, data=circ_data)
print(kruskal_result)
```

Kruskal-Wallis rank sum test

data: ctDNA.MRD.MTM by Stage

Kruskal-Wallis chi-squared = 2.8384, df = 2, p-value = 0.2419

[Hide](#)

```
pairwise_wilcox <- pairwise.wilcox.test(circ_data$ctDNA.MRD.MTM, circ_data$Stage,
                                         p.adjust.method = "BH", na.rm = TRUE)
```

Warning in wilcox.test.default(xi, xj, paired = paired, ...) :
cannot compute exact p-value with ties

Warning in wilcox.test.default(xi, xj, paired = paired, ...) :
cannot compute exact p-value with ties

Warning in wilcox.test.default(xi, xj, paired = paired, ...) :
cannot compute exact p-value with ties

[Hide](#)

```
print(pairwise_wilcox)
```

Pairwise comparisons using Wilcoxon rank sum test with continuity correction

data: circ_data\$ctDNA.MRD.MTM and circ_data\$Stage

	II (n=37)	III (n=40)
III (n=40)	0.46	-
IV (n=12)	0.30	0.39

P value adjustment method: BH

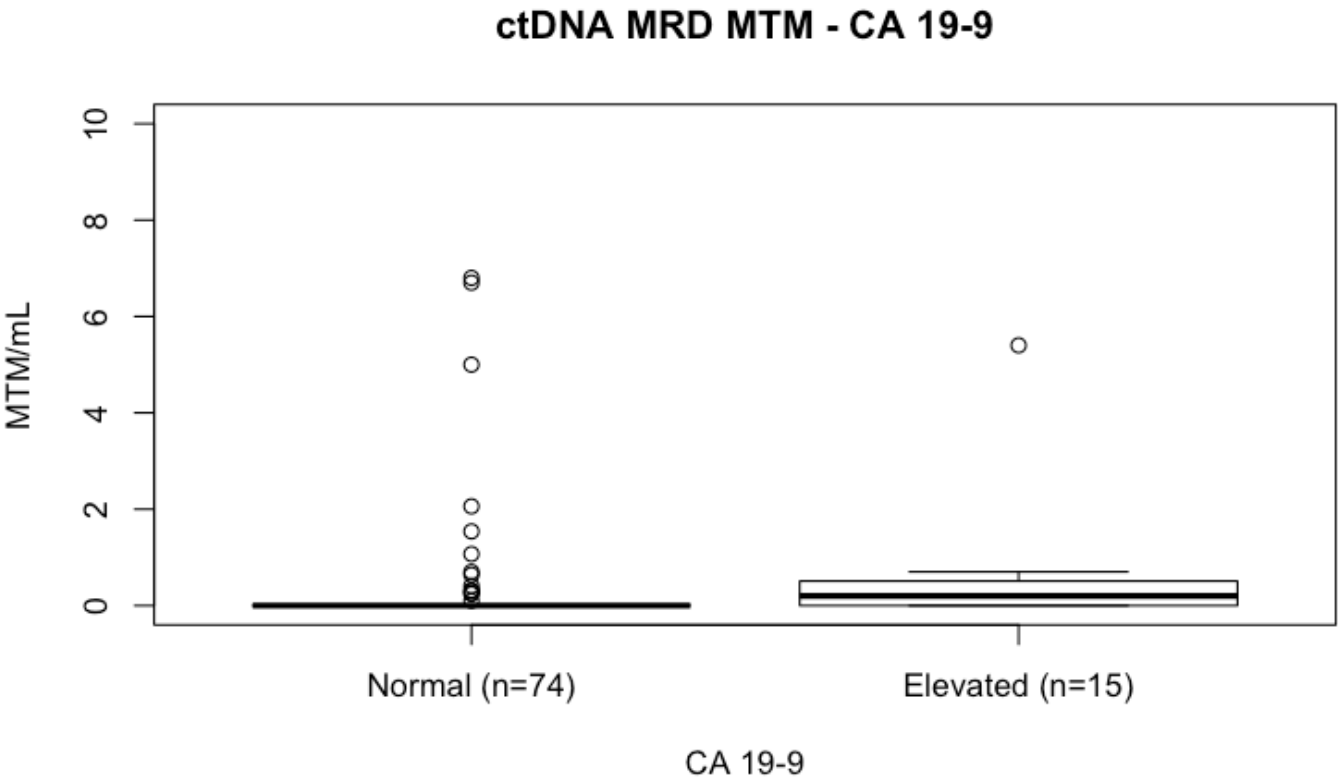
Hide

tally(~CA19.MRD, data=circ_data, margins = TRUE)

CA19.MRD		
Elevated	Normal	Total
15	74	89

Hide

```
circ_data$CA19.MRD <- factor(circ_data$CA19.MRD, levels = c("Normal","Elevated"), labels = c("Normal (n=74)","Elevated (n=15)"))
boxplot(ctDNA.MRD.MTM~CA19.MRD, data=circ_data, main="ctDNA MRD MTM - CA 19-9",
        xlab="CA 19-9", ylab="MTM/mL", col="white", border="black", ylim=c(0, 10))
```



Hide

```
m4 <- wilcox.test(ctDNA.MRD.MTM ~ CA19.MRD, data=circ_data, na.rm=TRUE, exact=FALSE, con
f.int=TRUE)
print(m4)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.MRD.MTM by CA19.MRD
W = 371.5, p-value = 0.00808
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -2.000269e-01 -4.645475e-05
sample estimates:
difference in location
 -5.786662e-06
```

#Median MTM/mL levels for ctDNA positive pts at MRD

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.MRD=="POSITIVE",]

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

Median MTM/mL post-surgery: 0.605

Hide

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

Range MTM/mL post-surgery: 0.1 - 73.04

#Association of ctDNA MRD status with clinicopathological factors

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels = c("NEGATIVE", "POSITIVE"), l
abels = c("ctDNA(-)", "ctDNA(+)))
circ_data$pN <- factor(circ_data$pN, levels = c("N1", "N2"), labels = c("pN1", "pN2"))
contingency_table <- table(circ_data$pN, circ_data$ctDNA.MRD)
chi_square_test <- chisq.test(contingency_table)
```

Warning in chisq.test(contingency_table) :
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 = 7.1944, df = 1, p-value = 0.007313
```

[Hide](#)

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

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.006385
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.403411 15.830714
sample estimates:
odds ratio
 4.647239
```

[Hide](#)

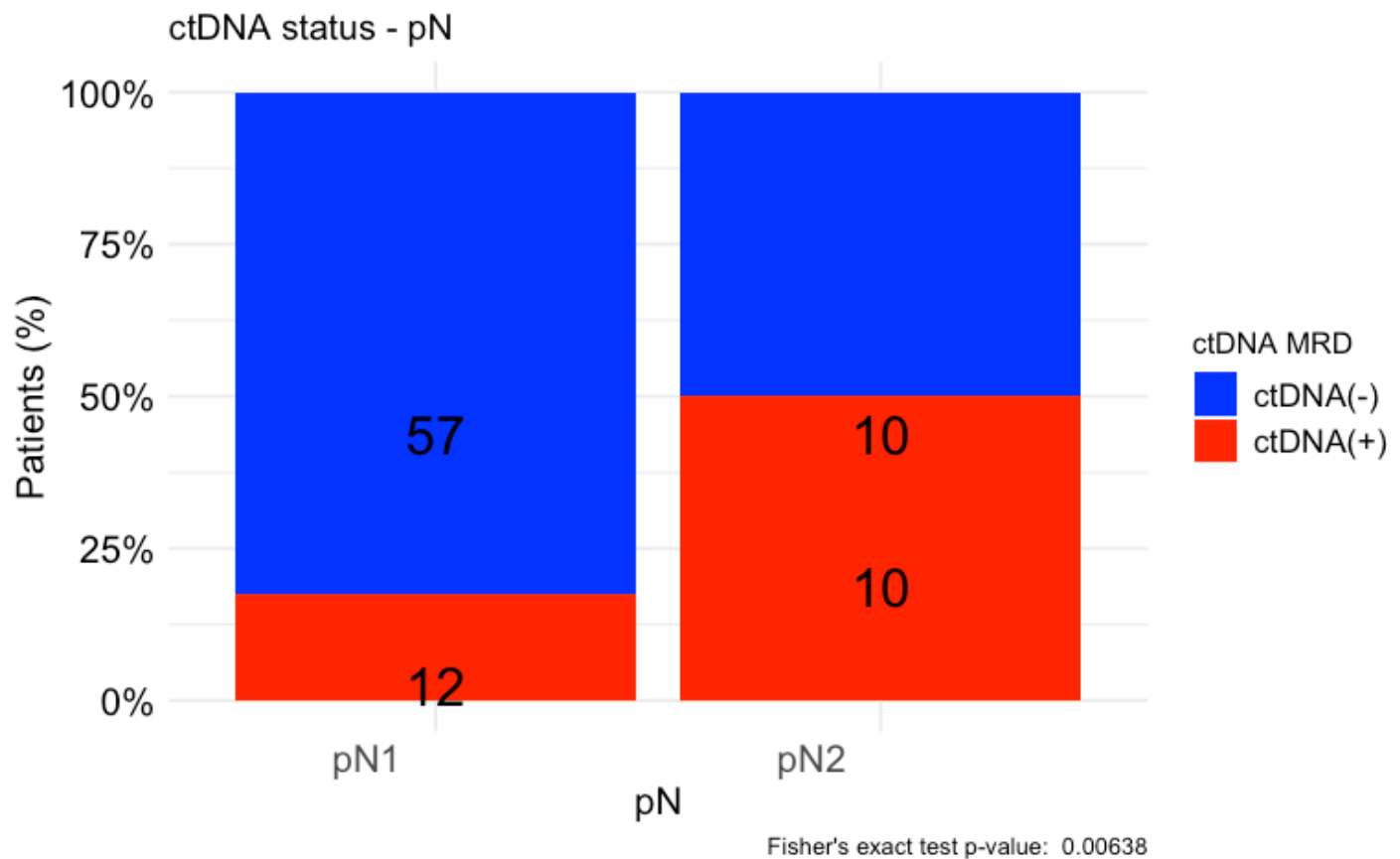
```
print(contingency_table)
```

	ctDNA(-)	ctDNA(+)
pN1	57	12
pN2	10	10

```

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 = "ctDNA status - pN",
        x = "pN",
        y = "Patients (%)",
        fill = "ctDNA MRD",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("ctDNA(-)" = "blue", "ctDNA(+)" = "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 Progression label size

```


[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels = c("NEGATIVE", "POSITIVE"), l
abels = c("ctDNA(-)", "ctDNA(+)))
circ_data$ResMarg <- factor(circ_data$ResMarg, levels = c("R0", "R1"), labels = c("R0",
"R1"))
contingency_table <- table(circ_data$ResMarg, circ_data$ctDNA.MRD)
chi_square_test <- chisq.test(contingency_table)
print(chi_square_test)
```

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

```
data: contingency_table
X-squared = 5.1569, df = 1, p-value = 0.02315
```

[Hide](#)

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


Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.01762
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.146895 10.855465
sample estimates:
odds ratio
 3.472855
```

Hide

```
print(contingency_table)
```

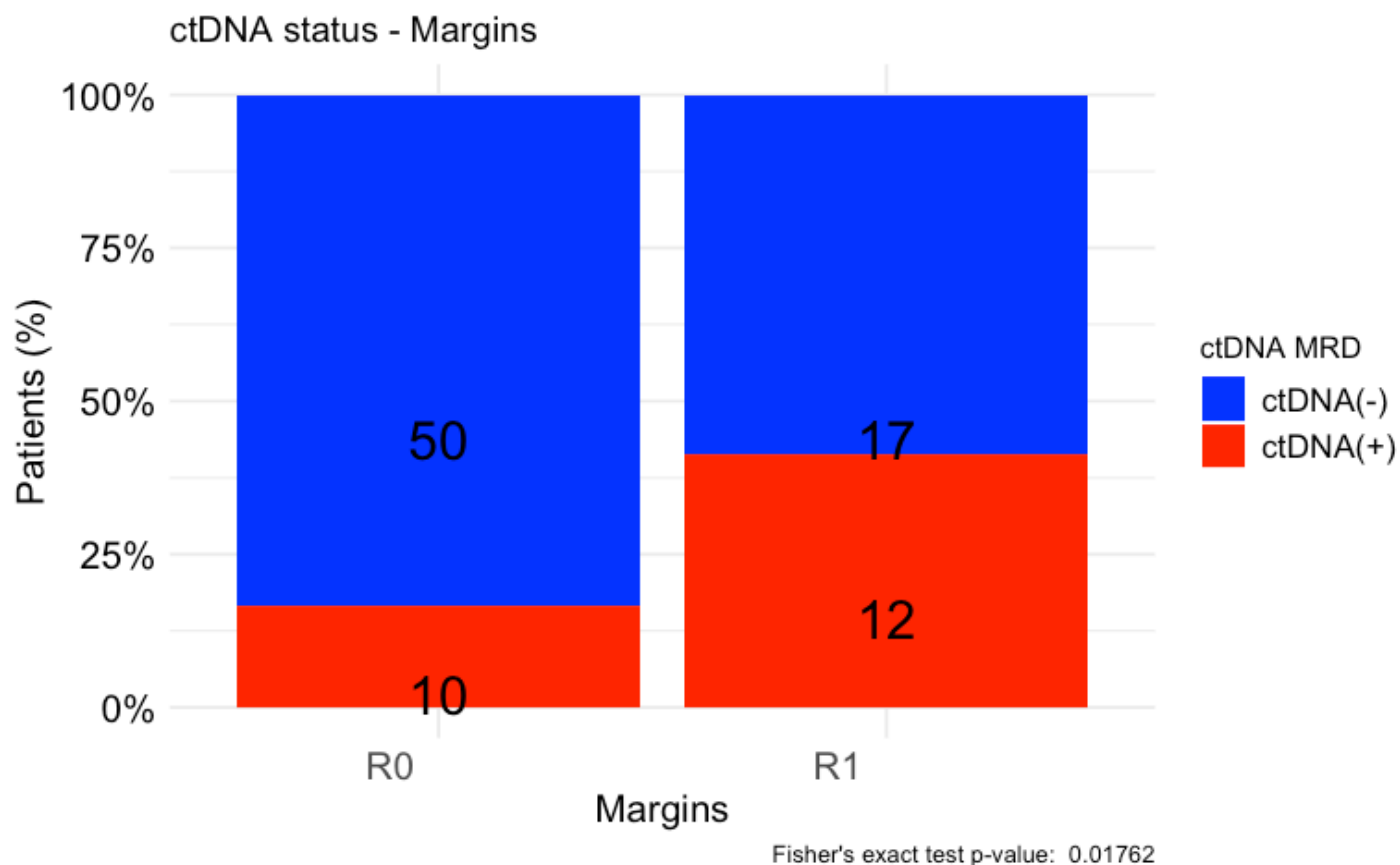
	ctDNA(-)	ctDNA(+)
R0	50	10
R1	17	12

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 = "ctDNA status - Margins",
        x = "Margins",
        y = "Patients (%)",
        fill = "ctDNA MRD",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("ctDNA(-)" = "blue", "ctDNA(+)" = "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 Progression label size

```


[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels = c("NEGATIVE", "POSITIVE"), l
abels = c("ctDNA(-)", "ctDNA(+)))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels = c("pCCA", "dCCA"), labels = c
("pCCA", "dCCA"))
contingency_table <- table(circ_data$PrimSite, circ_data$ctDNA.MRD)
chi_square_test <- chisq.test(contingency_table)
print(chi_square_test)
```

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

```
data: contingency_table
X-squared = 0, df = 1, p-value = 1
```

[Hide](#)

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

Fisher's Exact Test for Count Data

data: contingency_table
p-value = 1
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
0.3114469 2.6872455
sample estimates:
odds ratio
0.9152081

Hide

print(contingency_table)

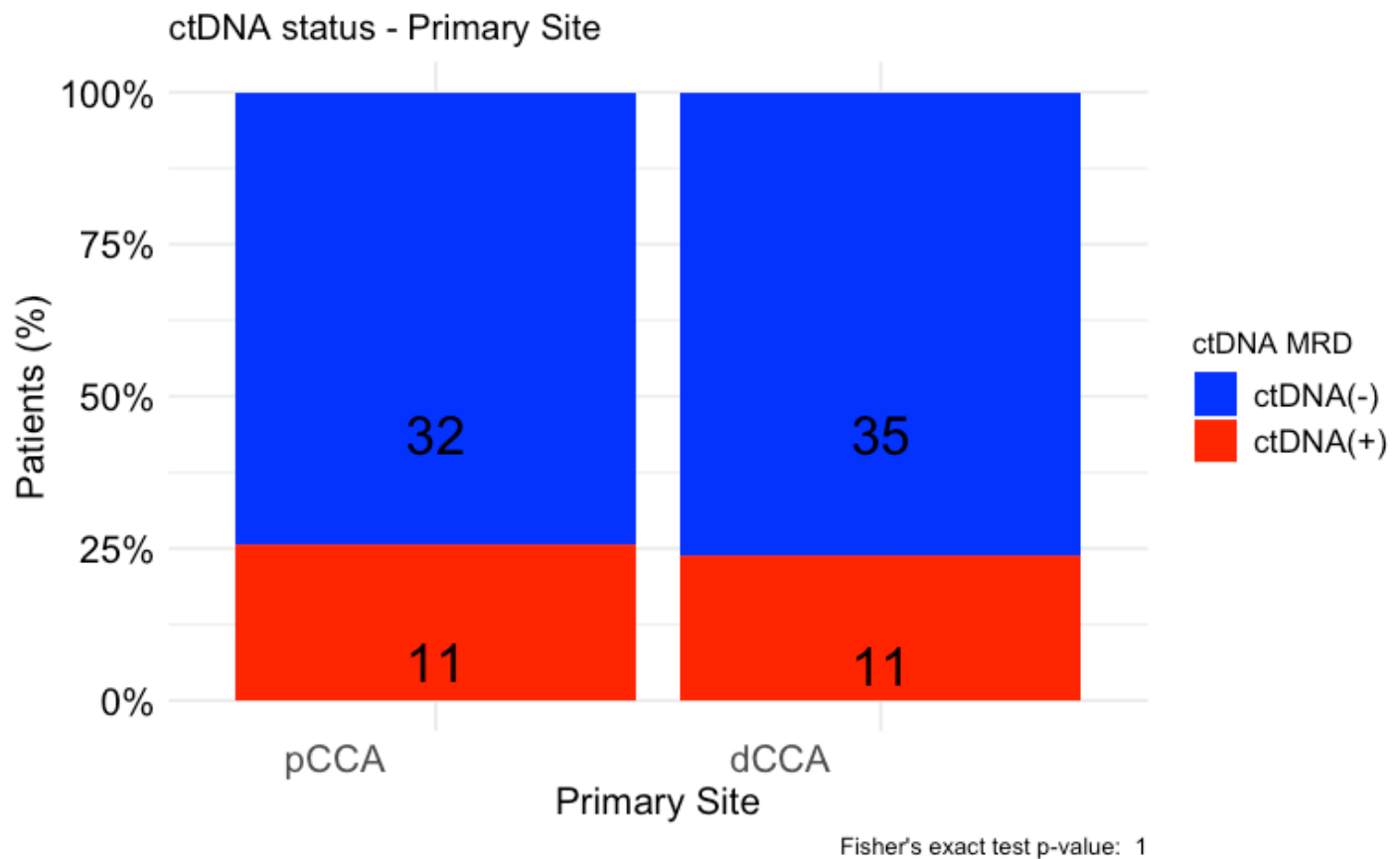
	ctDNA(-)	ctDNA(+)
pCCA	32	11
dCCA	35	11

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 = "ctDNA status - Primary Site",
        x = "Primary Site",
        y = "Patients (%)",
        fill = "ctDNA MRD",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("ctDNA(-)" = "blue", "ctDNA(+)" = "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 Progression label size

```


[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels = c("NEGATIVE", "POSITIVE"), l
abels = c("ctDNA(-)", "ctDNA(+)))
circ_data$Stage <- factor(circ_data$Stage, levels = c("II", "III", "IV"), labels = c("I
I", "III", "IV"))
contingency_table <- table(circ_data$Stage, circ_data$ctDNA.MRD)
chi_square_test <- chisq.test(contingency_table)
```

Warning in chisq.test(contingency_table) :
Chi-squared approximation may be incorrect

[Hide](#)

```
print(chi_square_test)
```

Pearson's Chi-squared test

```
data: contingency_table
X-squared = 2.68, df = 2, p-value = 0.2619
```

Hide

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

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.2546
alternative hypothesis: two.sided
```

Hide

```
print(contingency_table)
```

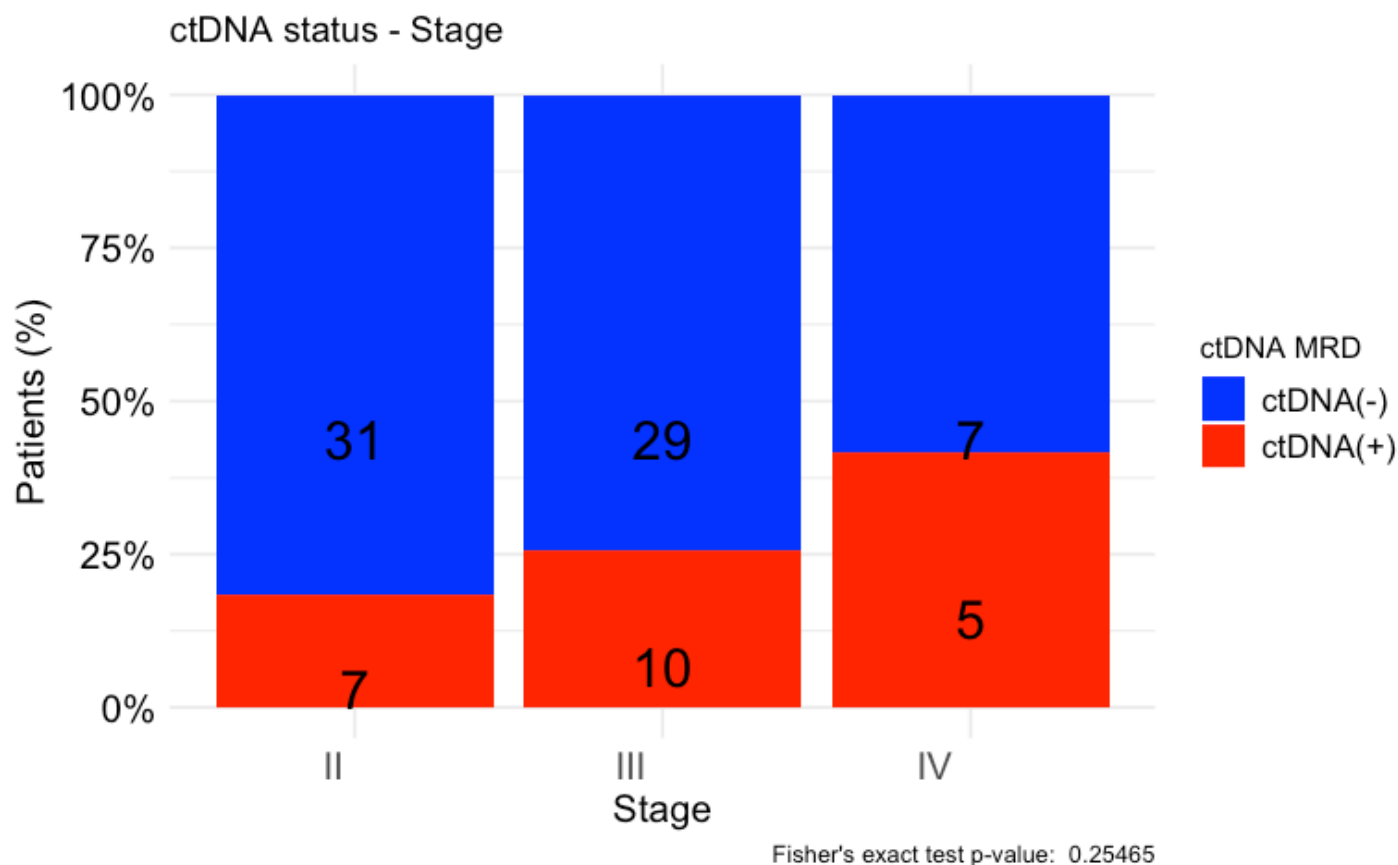
	ctDNA(-)	ctDNA(+)
II	31	7
III	29	10
IV	7	5

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 = "ctDNA status - Stage",
        x = "Stage",
        y = "Patients (%)",
        fill = "ctDNA MRD",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("ctDNA(-)" = "blue", "ctDNA(+)" = "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 Progression label size

```



[Hide](#)

```
pairwise_fisher <- function(data, factor1, factor2) {
  levels <- unique(data[[factor1]])
  results <- data.frame(Stage1 = character(), Stage2 = character(), p.value = numeric(),
stringsAsFactors = FALSE)

  for (i in 1:(length(levels) - 1)) {
    for (j in (i + 1):length(levels)) {
      subset_data <- data %>% filter(data[[factor1]] %in% c(levels[i], levels[j]))
      contingency_table_pairwise <- table(subset_data[[factor1]], subset_data[[factor
2]])
      fisher_result <- fisher.test(contingency_table_pairwise)
      results <- rbind(results, data.frame(Stage1 = levels[i], Stage2 = levels[j], p.val
ue = fisher_result$p.value))
    }
  }
  return(results)
}

# Perform pairwise comparisons
pairwise_results <- pairwise_fisher(circ_data, "Stage", "ctDNA.MRD")
print(pairwise_results)
```

Stage1<fctr>	Stage2<fctr>	p.value<dbl>
IV	III	0.3018721
IV	II	0.1292390
III	II	0.5842835
3 rows		

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels = c("NEGATIVE", "POSITIVE"), l
abels = c("ctDNA(-)", "ctDNA(+)))
circ_data$CA19.MRD <- factor(circ_data$CA19.MRD, levels = c("Normal", "Elevated"), label
s = c("Normal", "Elevated"))
contingency_table <- table(circ_data$CA19.MRD, circ_data$ctDNA.MRD)
chi_square_test <- chisq.test(contingency_table)
```

Warning in chisq.test(contingency_table) :
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 = 6.1961, df = 1, p-value = 0.0128

Hide

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

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.008766
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.282588 18.543083
sample estimates:
odds ratio
 4.787433
```

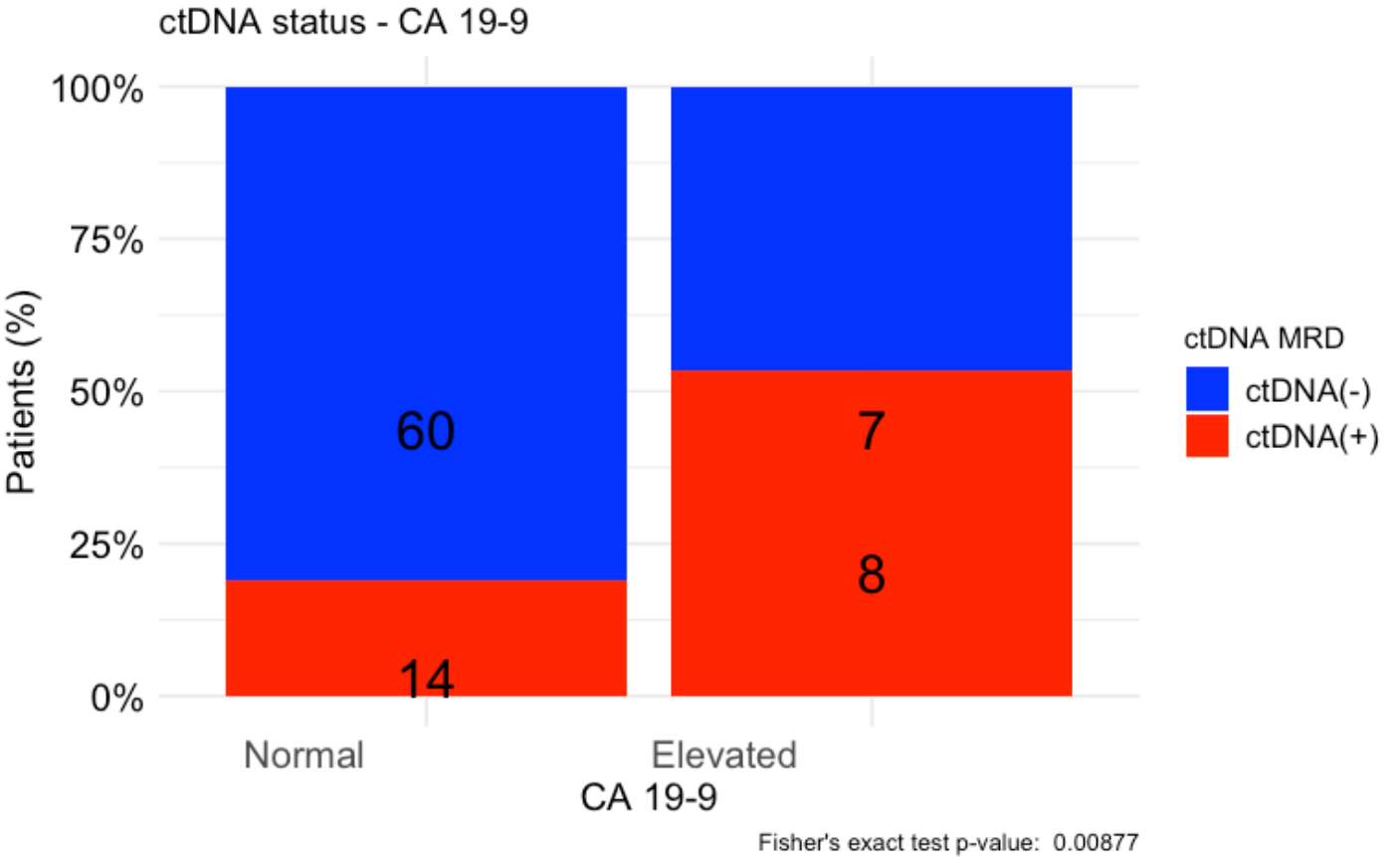
Hide

```
print(contingency_table)
```

	ctDNA(-)	ctDNA(+)
Normal	60	14
Elevated	7	8

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 = "ctDNA status - CA 19-9",
        x = "CA 19-9",
        y = "Patients (%)",
        fill = "ctDNA MRD",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("ctDNA(-)" = "blue", "ctDNA(+)" = "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 Progression label size
```



#Prognostic role of ctDNA at C5D1

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.C5D1!="",]
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C5D1, data
= circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$RFS.months, event = circ_data\$RFS.Event) ~
ctDNA.C5D1, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C5D1=NEGATIVE	71	48	15.9	12.92	27.57
ctDNA.C5D1=POSITIVE	17	17	4.8	3.09	7.86

Hide

```
event_summary <- circ_data %>%
  group_by(ctDNA.C5D1) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

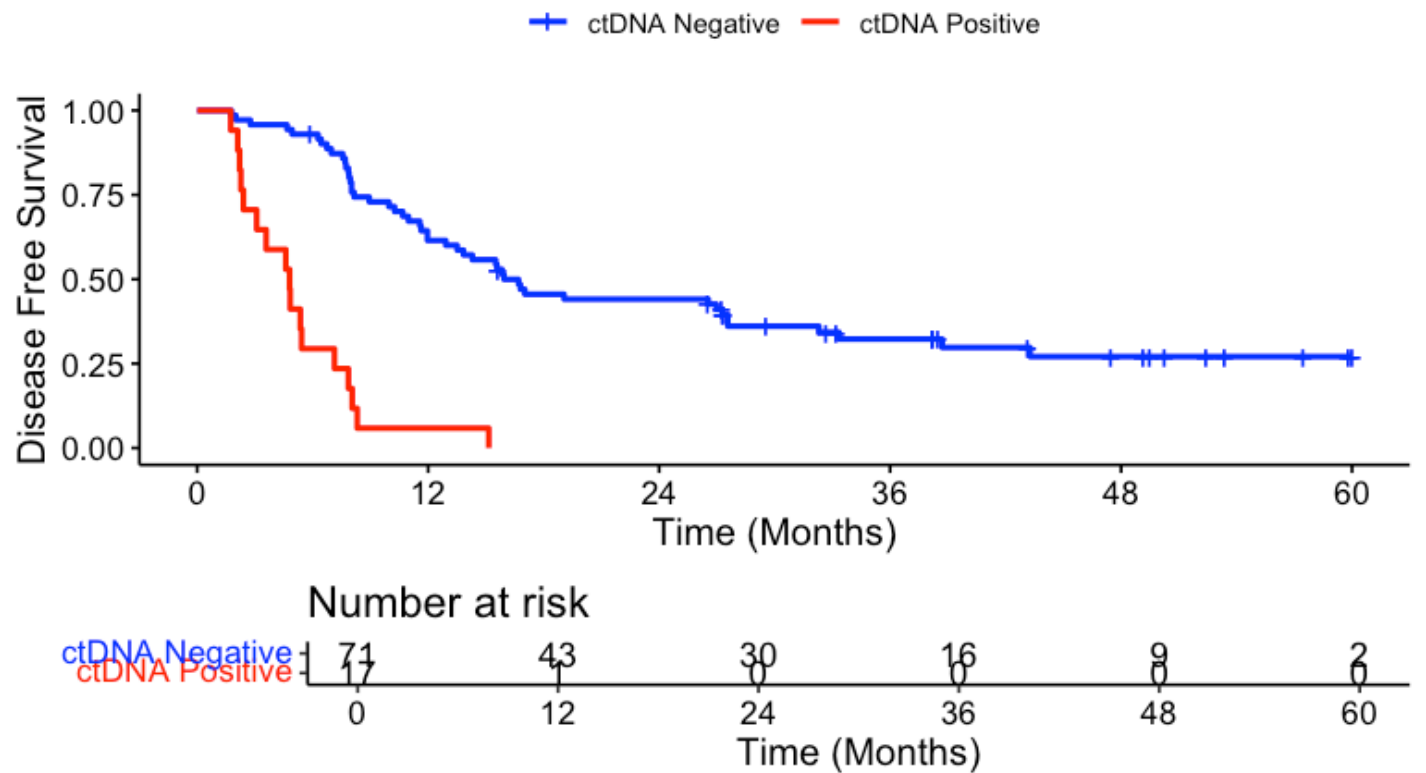
ctDNA.C5D1	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	71	48	0.6760563	67.60563
POSITIVE	17	17	1.0000000	100.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C5D1, 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=12, palette=c("blue","red"), title="DFS - ctDNA C5D1", ylab= "Disease F
ree Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"),
legend.title="")
```

DFS - ctDNA C5D1



Hide

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

```
Call: survfit(formula = surv_object ~ ctDNA.C5D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")

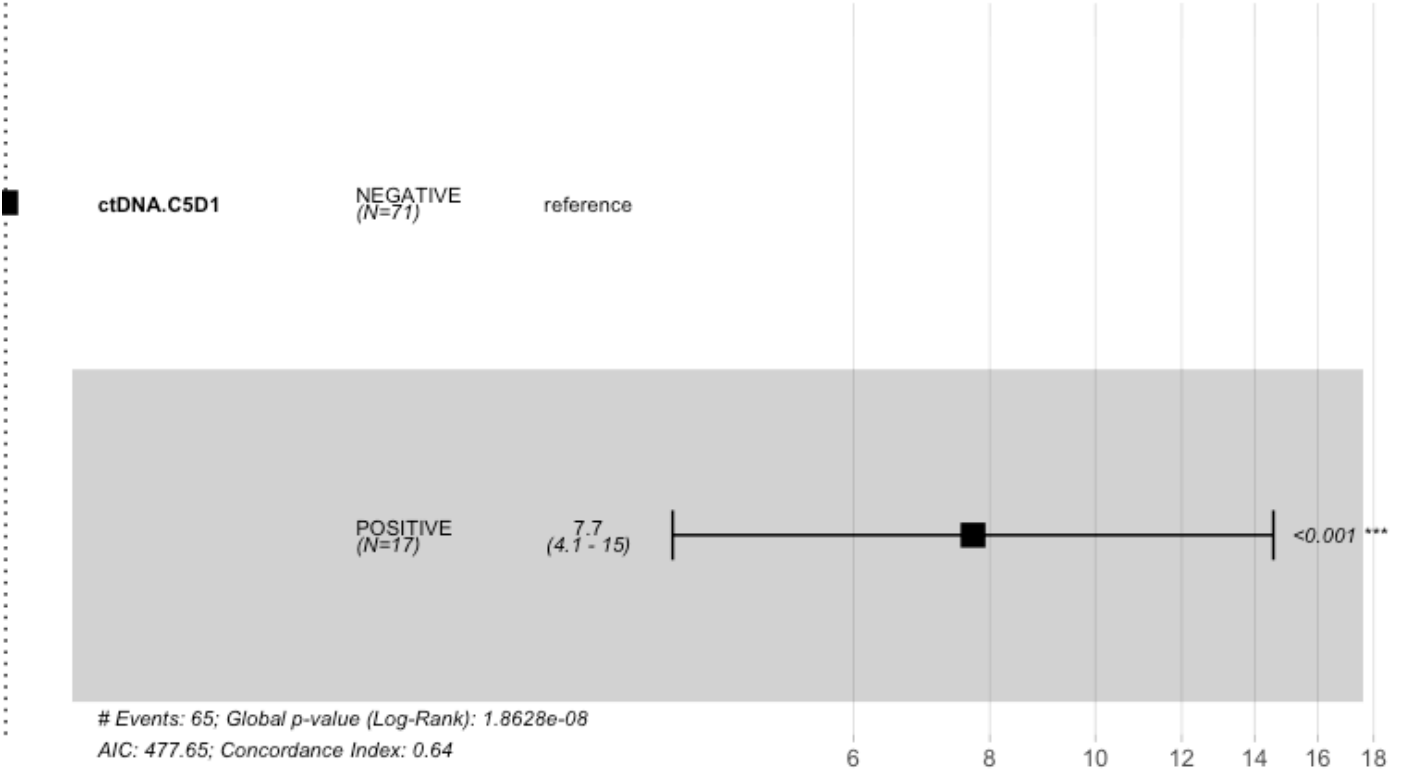
      ctDNA.C5D1=NEGATIVE
time  n.risk  n.event  survival  std.err  lower 95% CI  upper 95% CI
12      43      27    0.615    0.0581    0.491    0.717
24      30      12    0.441    0.0596    0.323    0.553

      ctDNA.C5D1=POSITIVE
time  n.risk  n.event  survival  std.err  lower 95% CI  upper 95%
CI
12.00000    1.00000    16.00000    0.05882    0.05707    0.00391    0.235
01
```

Hide

```
circ_data$ctDNA.C5D1 <- factor(circ_data$ctDNA.C5D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C5D1, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:
coxph(formula = surv_object ~ ctDNA.C5D1, data = circ_data)

n= 88, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C5D1POSITIVE	2.0439	7.7207	0.3238	6.312	2.76e-10 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C5D1POSITIVE	7.721	0.1295	4.093	14.56

Concordance= 0.638 (se = 0.026)
Likelihood ratio test= 31.63 on 1 df, p=2e-08
Wald test = 39.84 on 1 df, p=3e-10
Score (logrank) test = 53.01 on 1 df, p=3e-13

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)
```

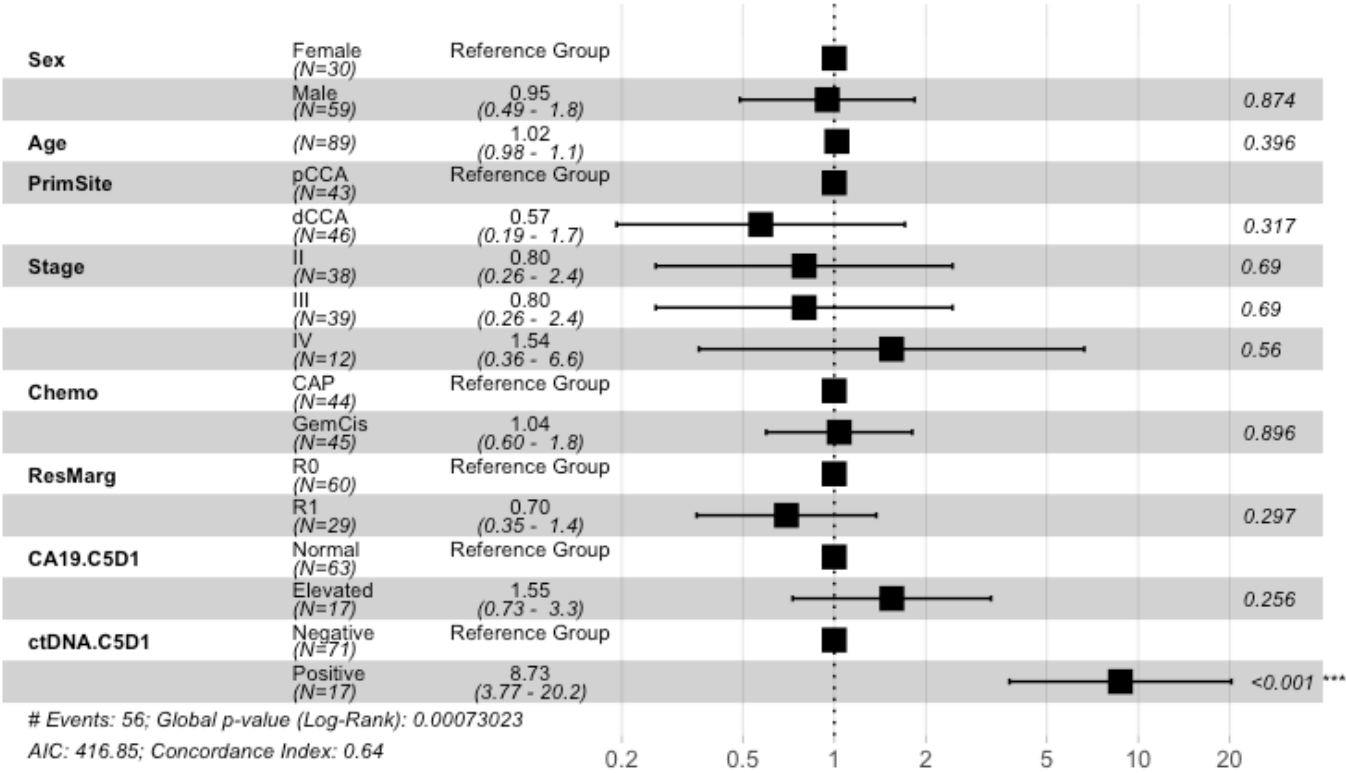
```
[1] "HR = 7.72 (4.09-14.56); p = 0"
```

#Multivariate regression model for DFS with ctDNA and CA 19-9

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
circ_datadf <- as.data.frame(circ_data)
circ_datadf$Sex <- factor(circ_datadf$Sex, levels = c("Female", "Male"), labels = c("Female", "Male"))
circ_datadf$PrimSite <- factor(circ_datadf$PrimSite, levels = c("pCCA", "dCCA"), labels = c("pCCA", "dCCA"))
circ_datadf$Chemo <- factor(circ_datadf$Chemo, levels = c("CAP", "GemCis"), labels = c("CAP", "GemCis"))
circ_datadf$ResMarg <- factor(circ_datadf$ResMarg, levels = c("R0", "R1"))
circ_datadf$Stage <- factor(circ_datadf$Stage, levels = c("II", "III", "IV"), labels = c("II", "III", "IV"))
circ_datadf$TP53 <- factor(circ_datadf$TP53, levels = c("WT", "Mut"))
circ_datadf$CA19.C5D1 <- factor(circ_datadf$CA19.C5D1, levels = c("Normal", "Elevated"))
circ_datadf$ctDNA.C5D1 <- factor(circ_datadf$ctDNA.C5D1, levels = c("NEGATIVE", "POSITIVE"), labels = c("Negative", "Positive"))
surv_object<-Surv(time = circ_datadf$RFS.months, event = circ_datadf$RFS.Event)
cox_fit <- coxph(surv_object ~ Sex + Age + PrimSite + Stage + Chemo + ResMarg + CA19.C5D1 + ctDNA.C5D1, data=circ_datadf)
ggforest(cox_fit, data = circ_datadf, main = "Multivariate Regression Model for DFS - Landmark analysis", refLabel = "Reference Group")
```

Multivariate Regression Model for DFS - Landmark analysis



Hide

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

#Univariate regression model for factors used at the C5D1 MVA

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$Sex <- factor(circ_data$Sex, levels = c("Female", "Male"), labels = c("Female", "Male")) #univariate for gender
cox_fit <- coxph(surv_object ~ Sex, data=circ_data)
summary(cox_fit)
```

Call:

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

```
n= 89, number of events= 65
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
SexMale	0.1594	1.1728	0.2689	0.593	0.553

	exp(coef)	exp(-coef)	lower .95	upper .95
SexMale	1.173	0.8527	0.6923	1.987

```
Concordance= 0.515 (se = 0.031 )
```

```
Likelihood ratio test= 0.36 on 1 df, p=0.5
```

```
Wald test = 0.35 on 1 df, p=0.6
```

```
Score (logrank) test = 0.35 on 1 df, p=0.6
```

Hide

```
cox_fit_summary <- summary(cox_fit)
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 = 1.17 (0.69-1.99); p = 0.553"
```

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
cox_fit <- coxph(surv_object ~ Age, data=circ_data) #univariate for age
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
Age	0.002393	1.002396	0.019228	0.124	0.901

	exp(coef)	exp(-coef)	lower .95	upper .95
Age	1.002	0.9976	0.9653	1.041

Concordance= 0.493 (se = 0.041)

Likelihood ratio test= 0.02 on 1 df, p=0.9

Wald test = 0.02 on 1 df, p=0.9

Score (logrank) test = 0.02 on 1 df, p=0.9

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
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 (0.97-1.04); p = 0.901"
```

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$PrimSite <- factor(circ_data$PrimSite, levels = c("pCCA", "dCCA"), labels = c("pCCA", "dCCA")) #univariate for Primary Site
cox_fit <- coxph(surv_object ~ PrimSite, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
PrimSitedCCA	-0.4558	0.6340	0.2520	-1.809	0.0705

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

	exp(coef)	exp(-coef)	lower .95	upper .95
PrimSitedCCA	0.634	1.577	0.3869	1.039

Concordance= 0.535 (se = 0.034)

Likelihood ratio test= 3.31 on 1 df, p=0.07

Wald test = 3.27 on 1 df, p=0.07

Score (logrank) test = 3.33 on 1 df, p=0.07

Hide

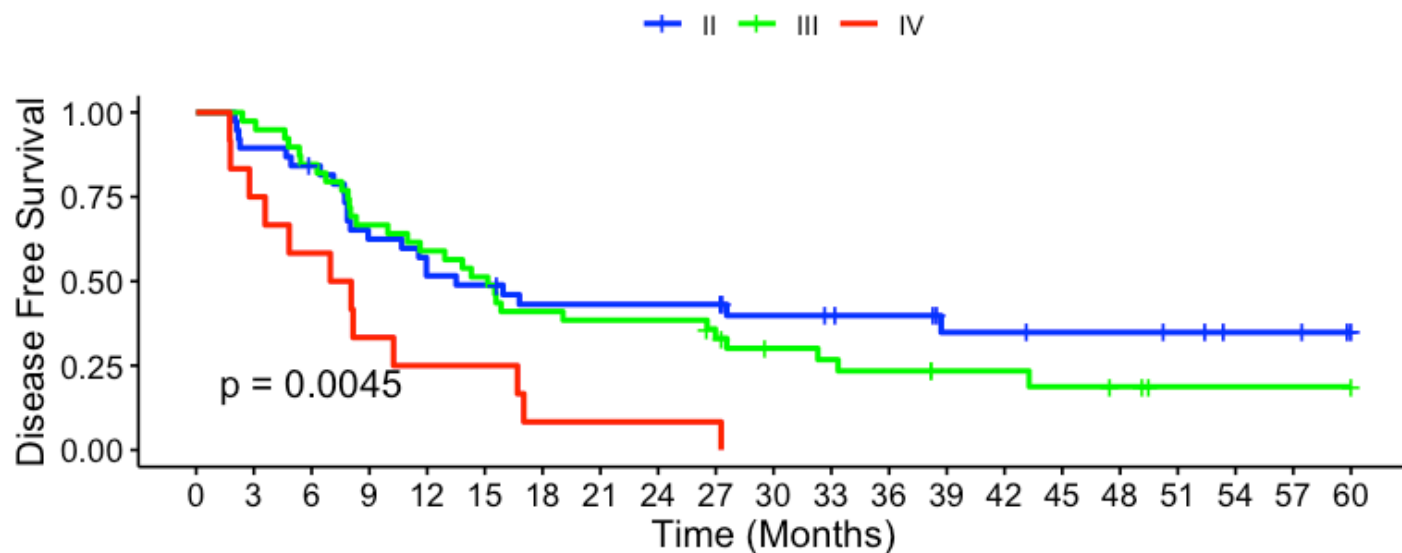
```
cox_fit_summary <- summary(cox_fit)
HR <- cox_fit_summary$coefficients[2]
lower_CI <- cox_fit_summary$conf.int[3]
upper_CI <- cox_fit_summary$conf.int[4]
p_value <- cox_fit_summary$coefficients[5]
label_text <- paste0("HR = ", round(HR, 2), " (", round(lower_CI, 2), "-", round(upper_CI, 2), "); p = ", round(p_value, 3))
print(label_text)
```

```
[1] "HR = 0.63 (0.39-1.04); p = 0.071"
```

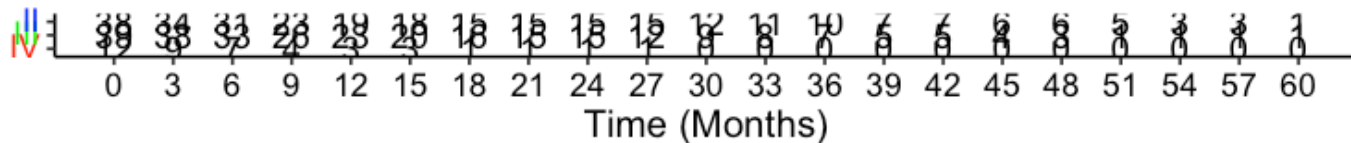
Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ Stage, 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=3, palette=c("blue","green","red"), title="DFS - ctDNA C5D1 - Stage", ylab
= "Disease Free Survival", xlab="Time (Months)", legend.labs=c("II", "III", "IV"), legend.title="")
```

DFS - ctDNA C5D1 - Stage



Number at risk

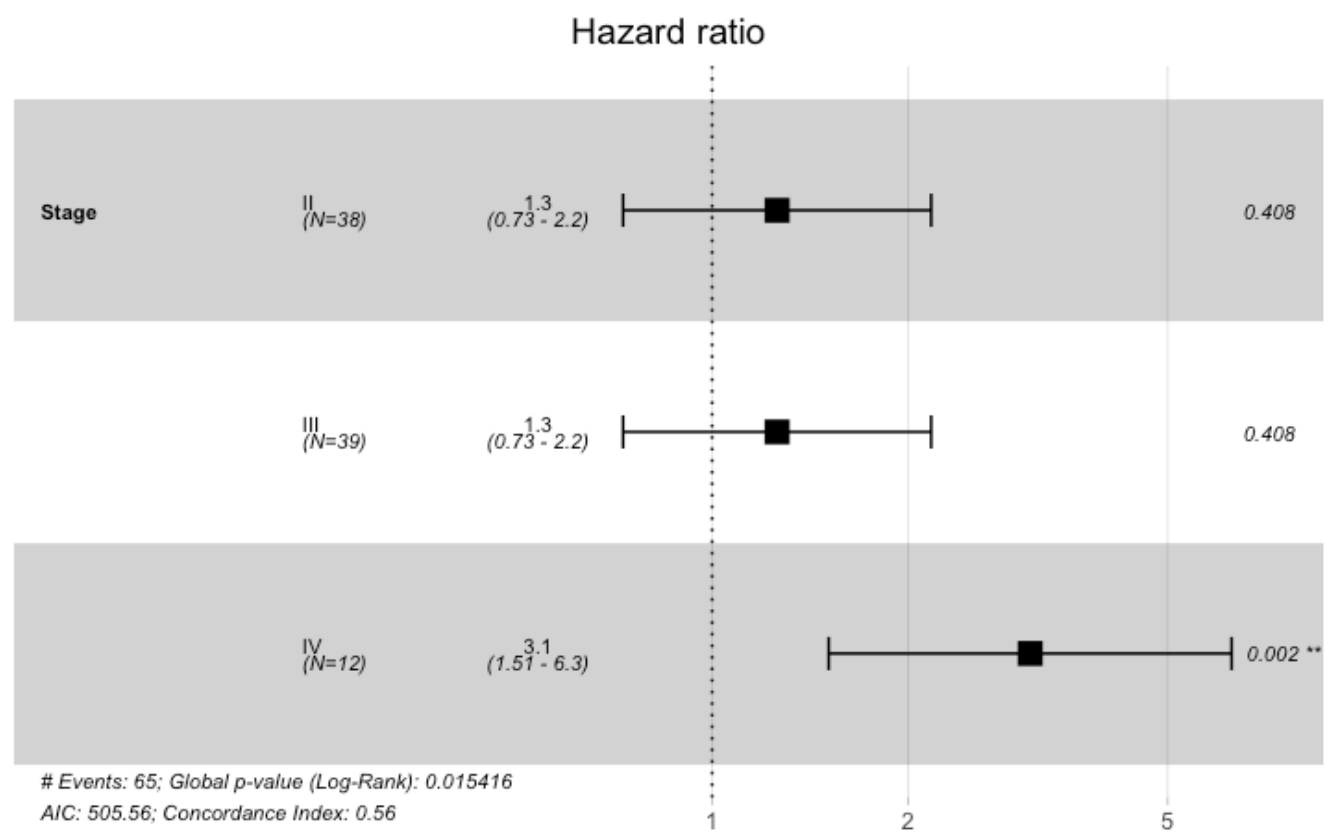


Hide

```

circ_data$Stage <- factor(circ_data$Stage, levels = c("II", "III", "IV"), labels = c("I", "III", "IV")) #univariate for Stage
cox_fit <- coxph(surv_object ~ Stage, data=circ_data)
ggforest(cox_fit,data = circ_data)

```



Hide

```
summary(cox_fit)
```

Call:
coxph(formula = surv_object ~ Stage, data = circ_data)

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
StageIII	0.2297	1.2583	0.2777	0.827	0.40805
StageIV	1.1237	3.0761	0.3631	3.095	0.00197 **

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

	exp(coef)	exp(-coef)	lower .95	upper .95
StageIII	1.258	0.7947	0.7301	2.168
StageIV	3.076	0.3251	1.5099	6.267

Concordance= 0.565 (se = 0.037)
Likelihood ratio test= 8.34 on 2 df, p=0.02
Wald test = 9.93 on 2 df, p=0.007
Score (logrank) test = 10.77 on 2 df, p=0.005

Hide


```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$Chemo <- factor(circ_data$Chemo, levels = c("CAP", "GemCis"), labels = c("CAP", "GemCis")) #univariate for Chemotherapy
cox_fit <- coxph(surv_object ~ Chemo, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
ChemoGemCis	-0.02437	0.97592	0.24846	-0.098	0.922

	exp(coef)	exp(-coef)	lower .95	upper .95
ChemoGemCis	0.9759	1.025	0.5997	1.588

Concordance= 0.507 (se = 0.034)

Likelihood ratio test= 0.01 on 1 df, p=0.9

Wald test = 0.01 on 1 df, p=0.9

Score (logrank) test = 0.01 on 1 df, p=0.9

Hide

```
cox_fit_summary <- summary(cox_fit)
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 = 0.98 (0.6-1.59); p = 0.922"
```

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$ResMarg <- factor(circ_data$ResMarg, levels = c("R0", "R1")) #univariate for R
esection margin
cox_fit <- coxph(surv_object ~ ResMarg, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
ResMargR1	0.3415	1.4071	0.2617	1.305	0.192

	exp(coef)	exp(-coef)	lower .95	upper .95
ResMargR1	1.407	0.7107	0.8425	2.35

Concordance= 0.534 (se = 0.03)

Likelihood ratio test= 1.64 on 1 df, p=0.2

Wald test = 1.7 on 1 df, p=0.2

Score (logrank) test = 1.72 on 1 df, p=0.2

[Hide](#)

```
cox_fit_summary <- summary(cox_fit)
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 = 1.41 (0.84-2.35); p = 0.192"
```

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$CA19.C5D1 <- factor(circ_data$CA19.C5D1, levels = c("Normal", "Elevated")) #univariate for CA 19-9 C5D1
cox_fit <- coxph(surv_object ~ CA19.C5D1, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ CA19.C5D1, data = circ_data)
```

n= 80, number of events= 56

(9 observations deleted due to missingness)

	coef	exp(coef)	se(coef)	z	Pr(> z)
CA19.C5D1Elevated	0.2218	1.2483	0.3264	0.679	0.497

	exp(coef)	exp(-coef)	lower .95	upper .95
CA19.C5D1Elevated	1.248	0.8011	0.6584	2.367

Concordance= 0.535 (se = 0.032)

Likelihood ratio test= 0.44 on 1 df, p=0.5

Wald test = 0.46 on 1 df, p=0.5

Score (logrank) test = 0.46 on 1 df, p=0.5

Hide

```
cox_fit_summary <- summary(cox_fit)
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 = 1.25 (0.66-2.37); p = 0.497"
```

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
surv_object <- Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
circ_data$ctDNA.C5D1 <- factor(circ_data$ctDNA.C5D1, levels = c("NEGATIVE", "POSITIVE"),
labels = c("Negative", "Positive")) #univariate for ctDNA C5D1
cox_fit <- coxph(surv_object ~ ctDNA.C5D1, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C5D1, data = circ_data)
```

```
n= 88, number of events= 65
(1 observation deleted due to missingness)
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C5D1Positive	2.0439	7.7207	0.3238	6.312	2.76e-10 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C5D1Positive	7.721	0.1295	4.093	14.56

Concordance= 0.638 (se = 0.026)

Likelihood ratio test= 31.63 on 1 df, p=2e-08

Wald test = 39.84 on 1 df, p=3e-10

Score (logrank) test = 53.01 on 1 df, p=3e-13

Hide

```
cox_fit_summary <- summary(cox_fit)
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 = 7.72 (4.09-14.56); p = 0"
```

#OS by ctDNA at C5D1

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.C5D1!="",]
circ_data$OS.months=circ_data$OS.months-2
circ_data <- circ_data[circ_data$OS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$OS.months, event = circ_data$OS.Event)~ctDNA.C5D1, data =
circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$OS.months, event = circ_data\$OS.Event) ~
ctDNA.C5D1, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C5D1=NEGATIVE	71	39	44.8	31.7	NA
ctDNA.C5D1=POSITIVE	17	17	18.4	13.1	35.3

[Hide](#)

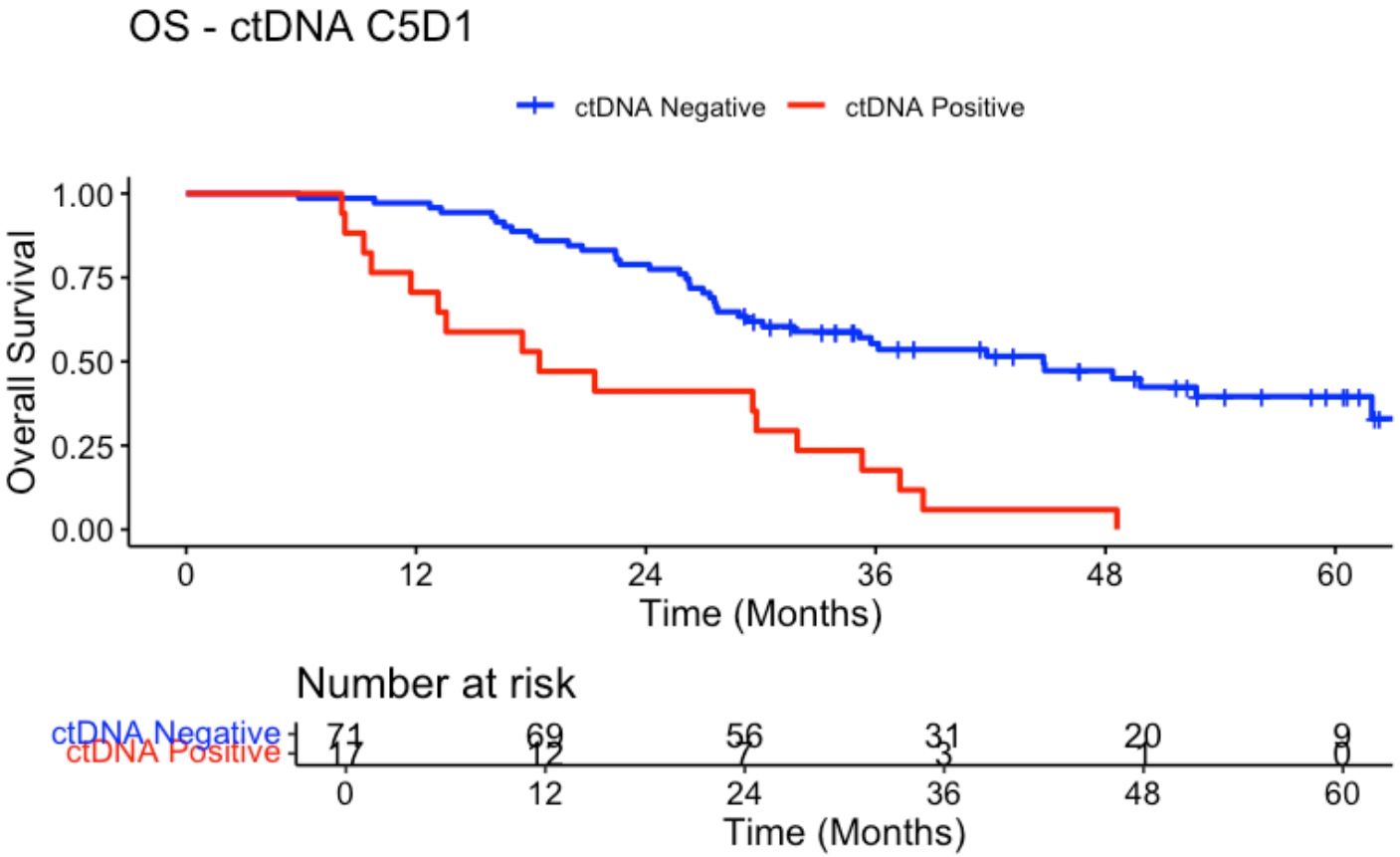
```
event_summary <- circ_data %>%
  group_by(ctDNA.C5D1) %>%
  summarise(
    Total = n(),
    Events = sum(OS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C5D1 <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	71	39	0.5492958	54.92958
POSITIVE	17	17	1.0000000	100.00000

2 rows

[Hide](#)

```
surv_object <-Surv(time = circ_data$OS.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C5D1, 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=12, palette=c("blue","red"), title="OS - ctDNA C5D1", ylab= "Overall Su
rvival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legen
d.title="")
```



Hide

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

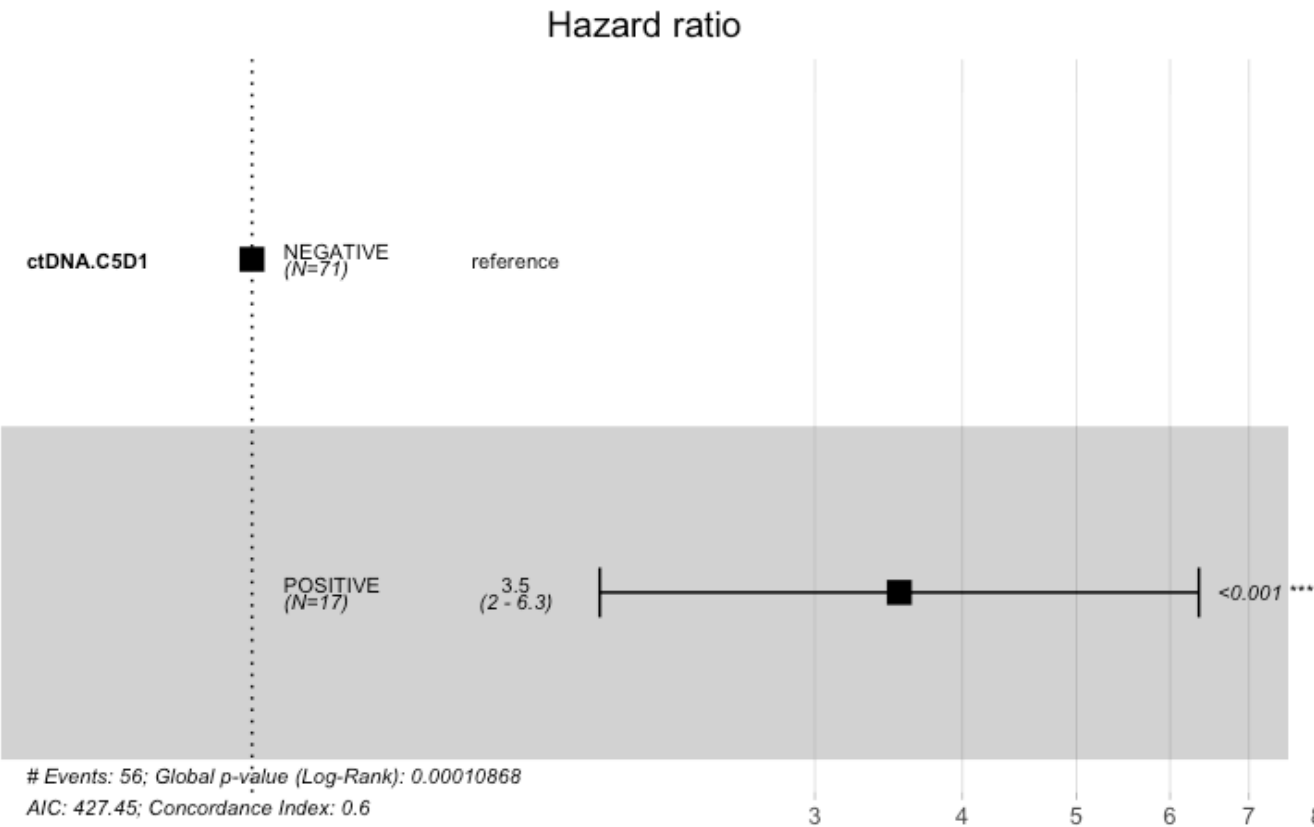
```
Call: survfit(formula = surv_object ~ ctDNA.C5D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C5D1=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	69	2	0.972	0.0196	0.892	0.993	
24	56	13	0.789	0.0484	0.674	0.867	

ctDNA.C5D1=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	12	5	0.706	0.111	0.431	0.866	
24	7	5	0.412	0.119	0.186	0.626	

Hide

```
circ_data$ctDNA.C5D1 <- factor(circ_data$ctDNA.C5D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C5D1, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C5D1, data = circ_data)
```

n= 88, number of events= 56

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C5D1POSITIVE	1.2634	3.5375	0.2983	4.235	2.29e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C5D1POSITIVE	3.537	0.2827	1.971	6.348

Concordance= 0.603 (se = 0.029)

Likelihood ratio test= 14.98 on 1 df, p=1e-04

Wald test = 17.93 on 1 df, p=2e-05

Score (logrank) test = 20.38 on 1 df, p=6e-06

[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 = 3.54 (1.97-6.35); p = 0"
```

#Prognostic role of ctDNA at C8D1

[Hide](#)

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.C8D1!="",]
```

```
circ_data$RFS.months=circ_data$RFS.months-2
```

```
circ_data <- circ_data[circ_data$RFS.months>=0,]
```

```
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
```

```
circ_datadf <- as.data.frame(circ_data)
```

```
survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C8D1, data = circ_data)
```


Call: `survfit(formula = Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event) ~ ctDNA.C8D1, data = circ_data)`

	n	events	median	0.95LCL	0.95UCL
ctDNA.C8D1=NEGATIVE	62	40	17.02	13.84	33.4
ctDNA.C8D1=POSITIVE	15	15	7.13	5.36	15.2

Hide

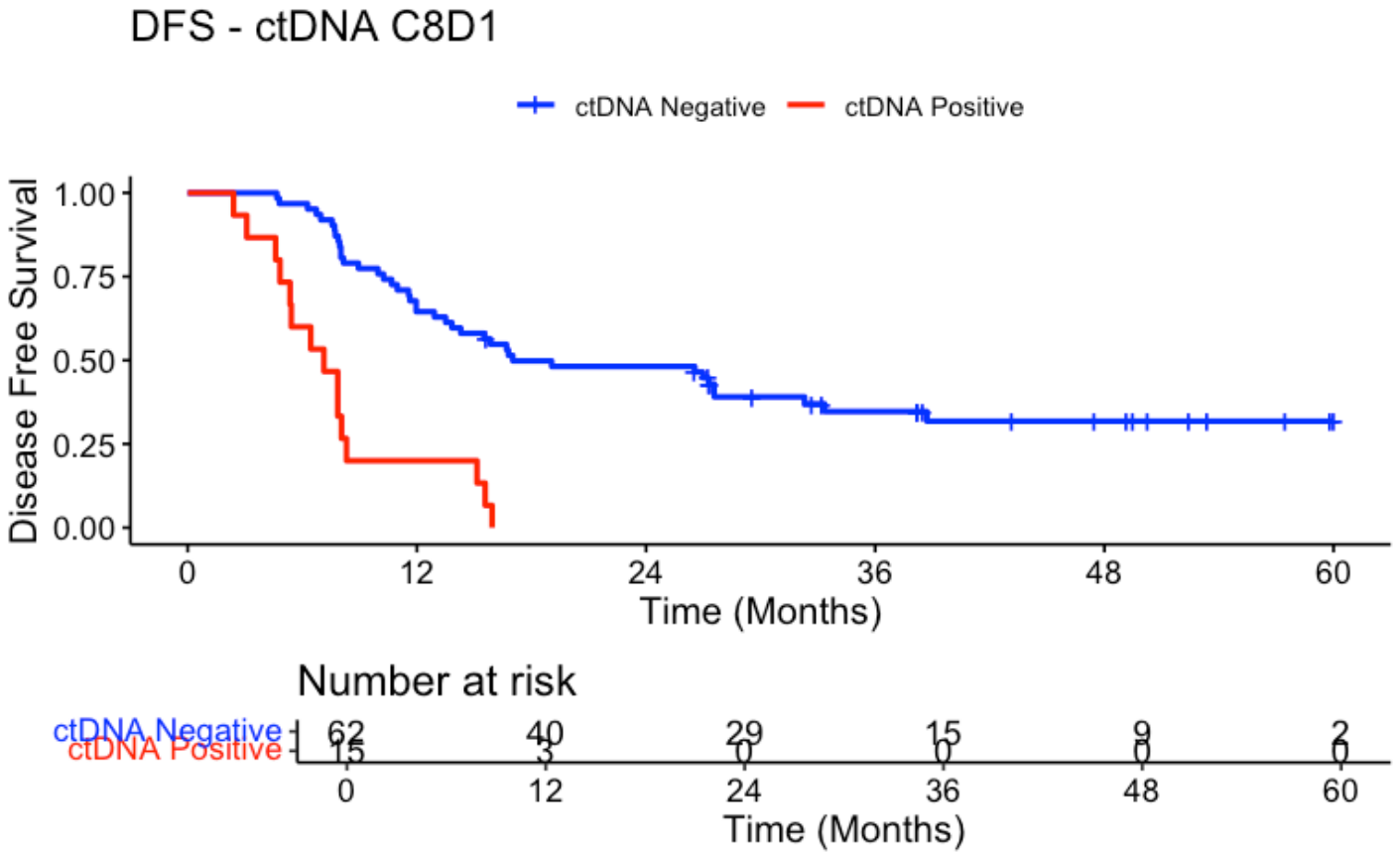
```
event_summary <- circ_data %>%
  group_by(ctDNA.C8D1) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C8D1	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	62	40	0.6451613	64.51613
POSITIVE	15	15	1.0000000	100.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C8D1, 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=12, palette=c("blue","red"), title="DFS - ctDNA C8D1", ylab= "Disease Free Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

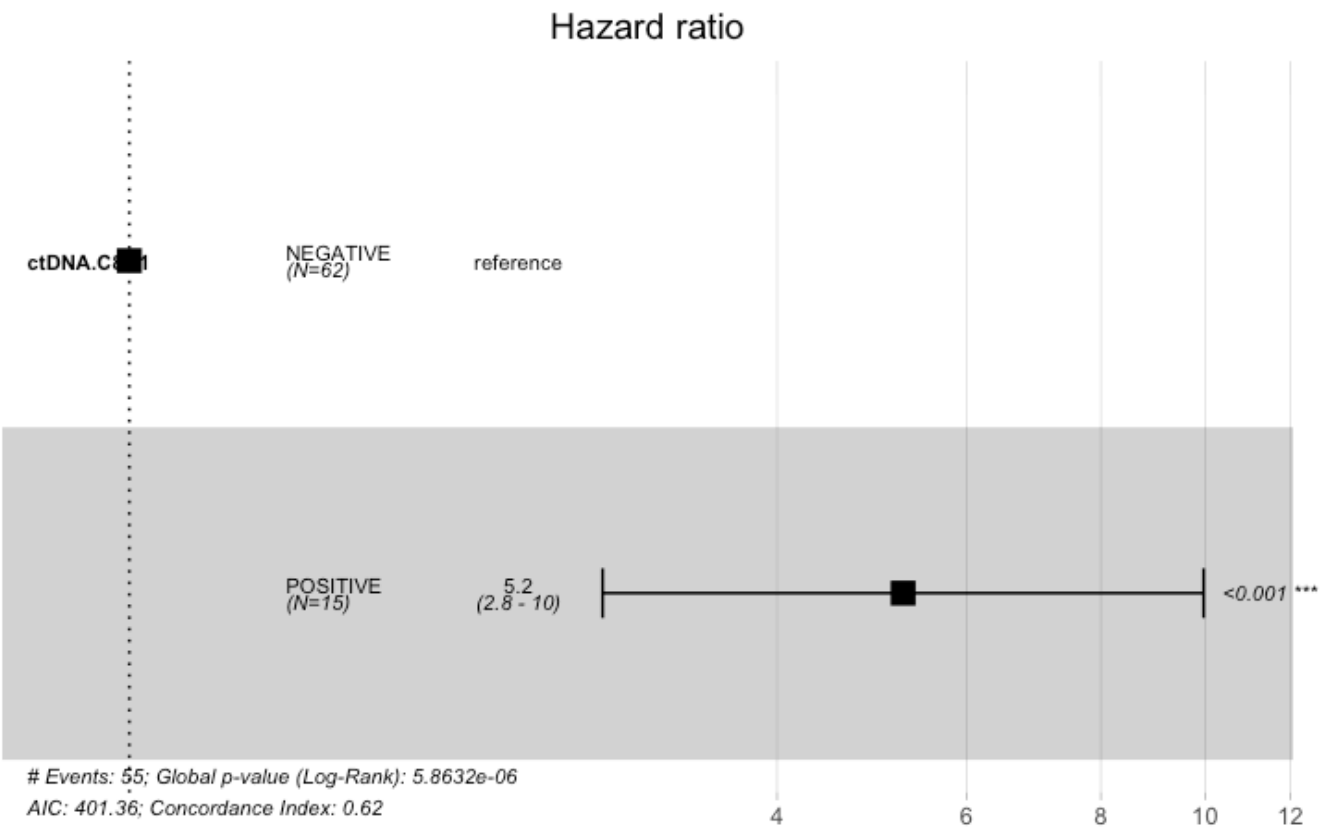
```
Call: survfit(formula = surv_object ~ ctDNA.C8D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C8D1=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	40	22	0.645	0.0608	0.513	0.750	
24	29	10	0.481	0.0637	0.353	0.599	

ctDNA.C8D1=POSITIVE							
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
39	12.0000	3.0000	12.0000	0.2000	0.1033	0.0489	0.42

Hide

```
circ_data$ctDNA.C8D1 <- factor(circ_data$ctDNA.C8D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C8D1, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C8D1, data = circ_data)
```

```
n= 77, number of events= 55
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C8D1POSITIVE	1.6561	5.2389	0.3281	5.047	4.49e-07 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C8D1POSITIVE	5.239	0.1909	2.754	9.967

Concordance= 0.624 (se = 0.028)

Likelihood ratio test= 20.53 on 1 df, p=6e-06

Wald test = 25.47 on 1 df, p=4e-07

Score (logrank) test = 31.38 on 1 df, p=2e-08

[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 = 5.24 (2.75-9.97); p = 0"
```

#OS by ctDNA at C8D1

[Hide](#)

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

```
circ_data <- circ_data[circ_data$ctDNA.C8D1!="",]
```

```
circ_data$OS.months=circ_data$OS.months-2
```

```
circ_data <- circ_data[circ_data$OS.months>=0,]
```

```
circ_datadf <- as.data.frame(circ_data)
```

```
survfit(Surv(time = circ_data$OS.months, event = circ_data$OS.Event)~ctDNA.C8D1, data =  
circ_data)
```

Call: `survfit(formula = Surv(time = circ_data$OS.months, event = circ_data$OS.Event) ~ ctDNA.C8D1, data = circ_data)`

	n	events	median	0.95LCL	0.95UCL
ctDNA.C8D1=NEGATIVE	62	32	48.4	30.1	NA
ctDNA.C8D1=POSITIVE	15	15	29.8	17.5	41.8

Hide

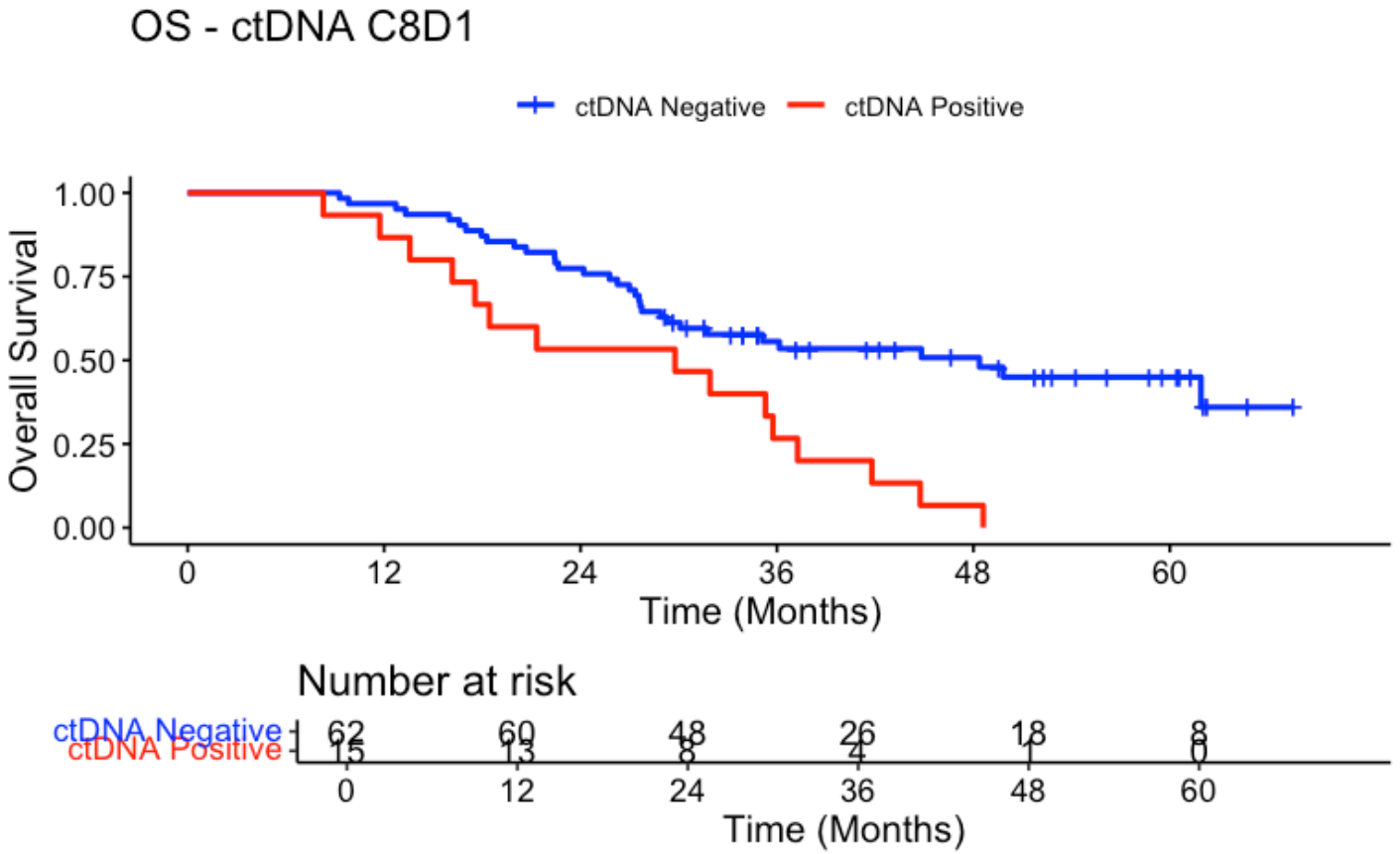
```
event_summary <- circ_data %>%
  group_by(ctDNA.C8D1) %>%
  summarise(
    Total = n(),
    Events = sum(OS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C8D1	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	62	32	0.516129	51.6129
POSITIVE	15	15	1.000000	100.0000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$OS.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C8D1, 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=12, palette=c("blue","red"), title="OS - ctDNA C8D1", ylab= "Overall Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

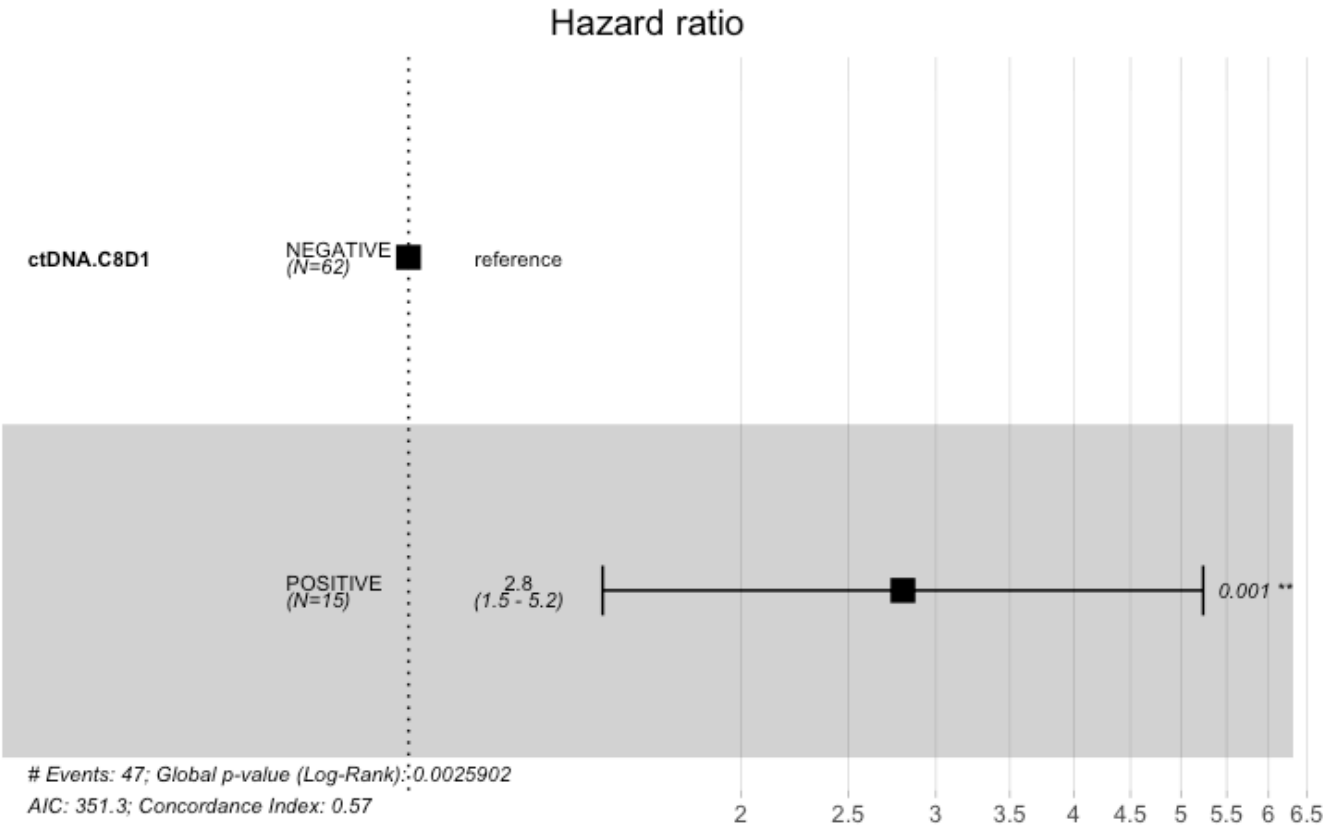
```
Call: survfit(formula = surv_object ~ ctDNA.C8D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C8D1=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	60	2	0.968	0.0224	0.877	0.992	
24	48	12	0.774	0.0531	0.649	0.860	

ctDNA.C8D1=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	13	2	0.867	0.0878	0.564	0.965	
24	8	5	0.533	0.1288	0.263	0.744	

Hide

```
circ_data$ctDNA.C8D1 <- factor(circ_data$ctDNA.C8D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C8D1, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C8D1, data = circ_data)
```

```
n= 77, number of events= 47
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C8D1POSITIVE	1.0300	2.8011	0.3192	3.227	0.00125 **

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C8D1POSITIVE	2.801	0.357	1.499	5.236

Concordance= 0.574 (se = 0.031)

Likelihood ratio test= 9.08 on 1 df, p=0.003

Wald test = 10.42 on 1 df, p=0.001

Score (logrank) test = 11.35 on 1 df, p=8e-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 = 2.8 (1.5-5.24); p = 0.001"
```

#Prognostic role of ctDNA anytime post-surgery

[Hide](#)

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

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

```
circ_data$RFS.months=circ_data$RFS.months-2
```

```
circ_data <- circ_data[circ_data$RFS.months>=0,]
```

```
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60
```

```
circ_datadf <- as.data.frame(circ_data)
```

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



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

	n	events	median	0.95LCL	0.95UCL
ctDNA.anytime=NEGATIVE	42	20	38.71	19.06	NA
ctDNA.anytime=POSITIVE	47	45	8.05	7.56	12

Hide

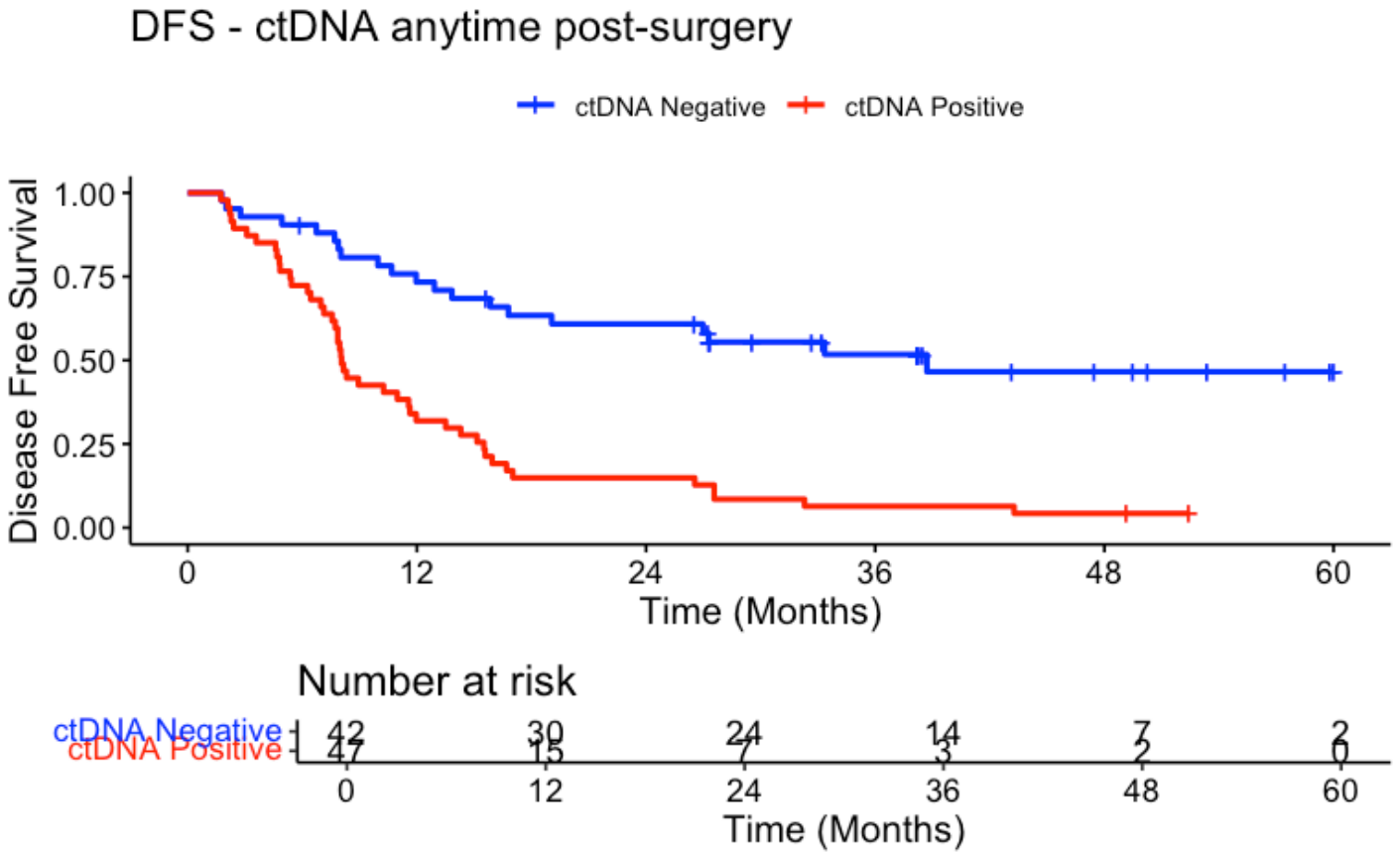
```
event_summary <- circ_data %>%
  group_by(ctDNA.anytime) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.anytime	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	42	20	0.4761905	47.61905
POSITIVE	47	45	0.9574468	95.74468

2 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.anytime, 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=12, palette=c("blue","red"), title="DFS - ctDNA anytime post-surgery", ylab= "Disease Free Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

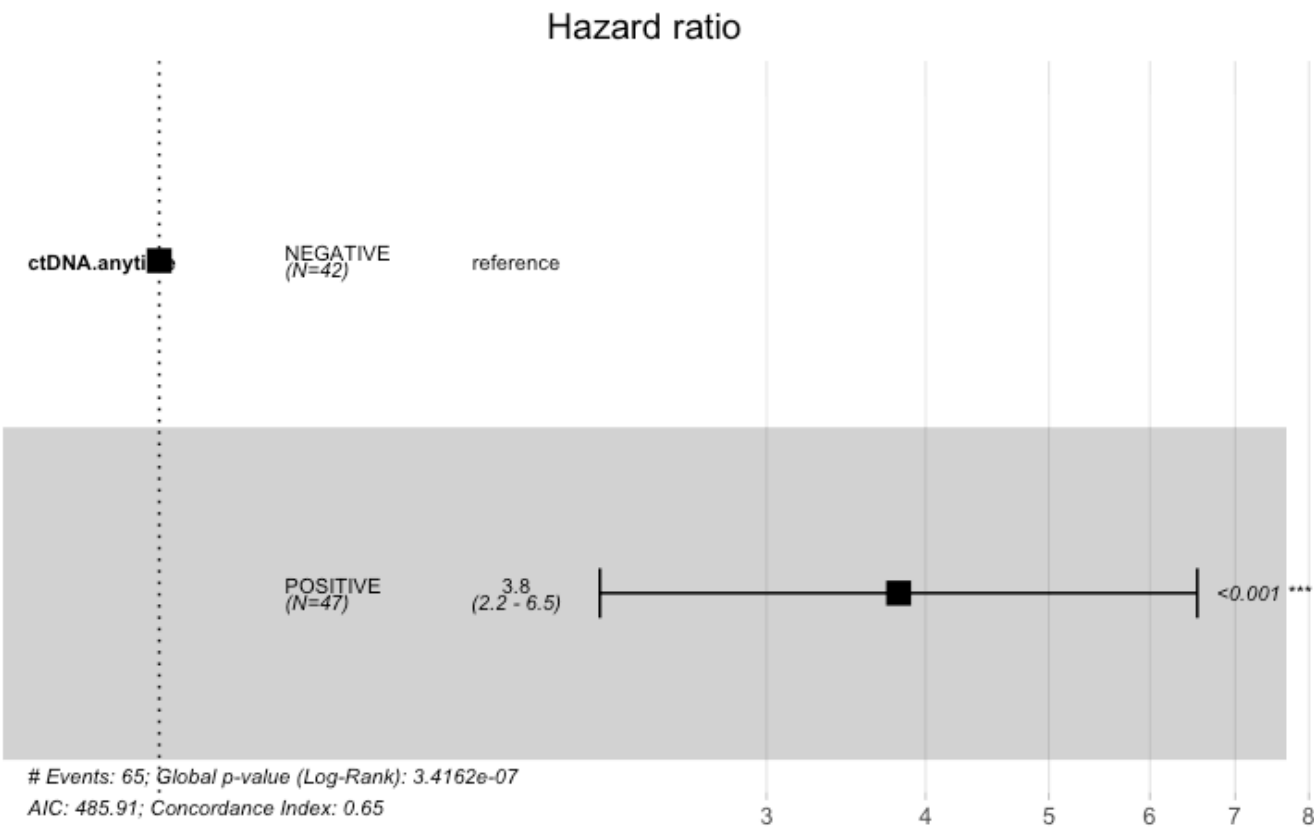
```
Call: survfit(formula = surv_object ~ ctDNA.anytime, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.anytime=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	30	11	0.734	0.0689	0.571	0.843	
24	24	5	0.609	0.0766	0.442	0.739	

ctDNA.anytime=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	15	32	0.319	0.0680	0.1928	0.453	
24	7	8	0.149	0.0519	0.0655	0.264	

Hide

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



Hide

```
summary(cox_fit)
```

Call:

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

n= 89, number of events= 65

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.anytimePOSITIVE	1.3374	3.8090	0.2757	4.851	1.23e-06 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.anytimePOSITIVE	3.809	0.2625	2.219	6.538

Concordance= 0.649 (se = 0.03)

Likelihood ratio test= 26 on 1 df, p=3e-07

Wald test = 23.54 on 1 df, p=1e-06

Score (logrank) test = 26.68 on 1 df, p=2e-07

[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 = 3.81 (2.22-6.54); p = 0"
```

#OS by ctDNA anytime post-surgery

[Hide](#)

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

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

```
circ_data$OS.months=circ_data$OS.months-2
```

```
circ_data <- circ_data[circ_data$OS.months>=0,]
```

```
circ_datadf <- as.data.frame(circ_data)
```

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

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.anytime=NEGATIVE	42	15	NA	48.4	NA
ctDNA.anytime=POSITIVE	47	41	27.6	21.3	35.3

Hide

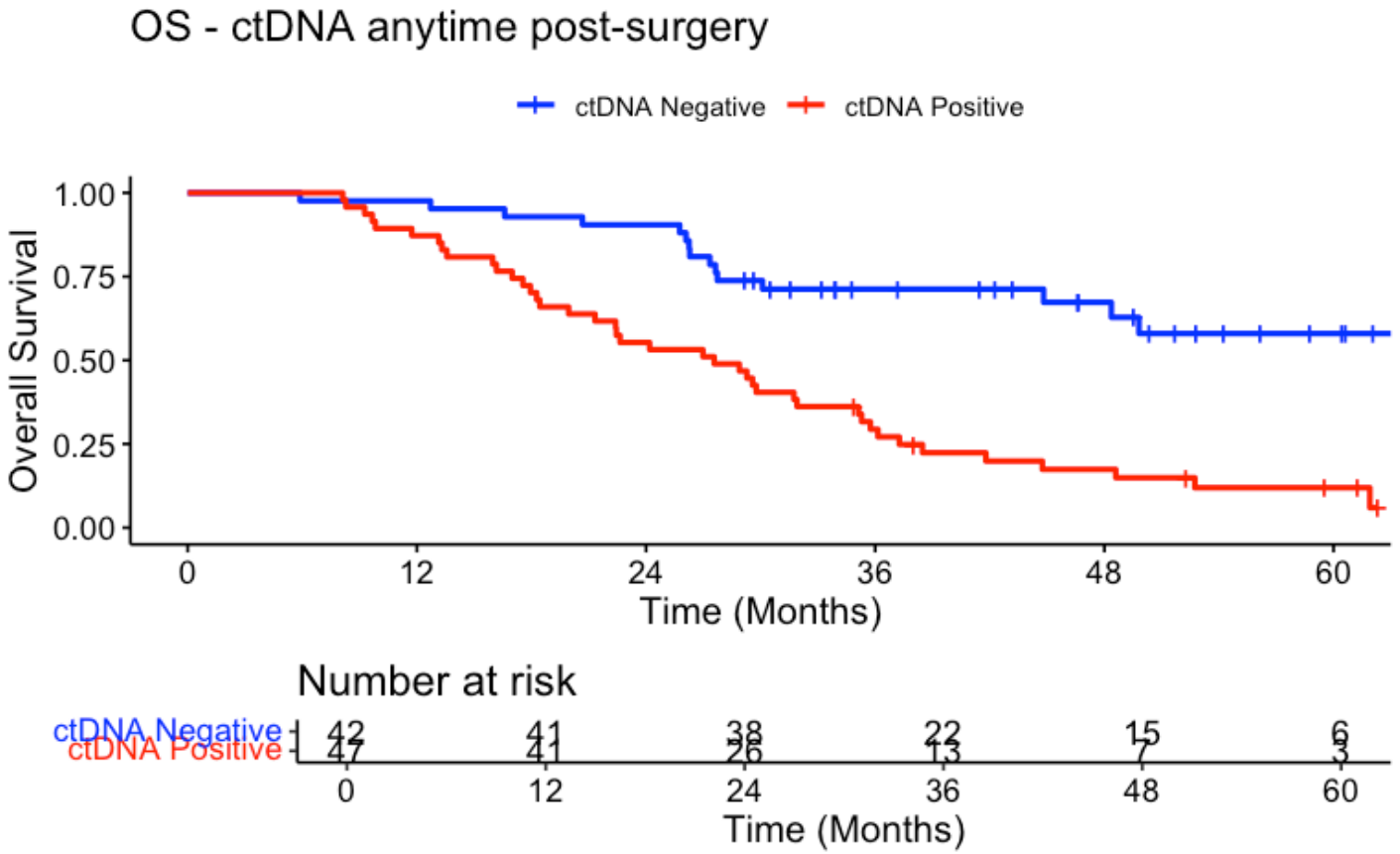
```
event_summary <- circ_data %>%
  group_by(ctDNA.anytime) %>%
  summarise(
    Total = n(),
    Events = sum(OS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.anytime	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	42	15	0.3571429	35.71429
POSITIVE	47	41	0.8723404	87.23404

2 rows

Hide

```
surv_object <-Surv(time = circ_data$OS.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.anytime, 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=12, palette=c("blue","red"), title="OS - ctDNA anytime post-surgery", ylab= "Overall Survival", xlab="Time (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

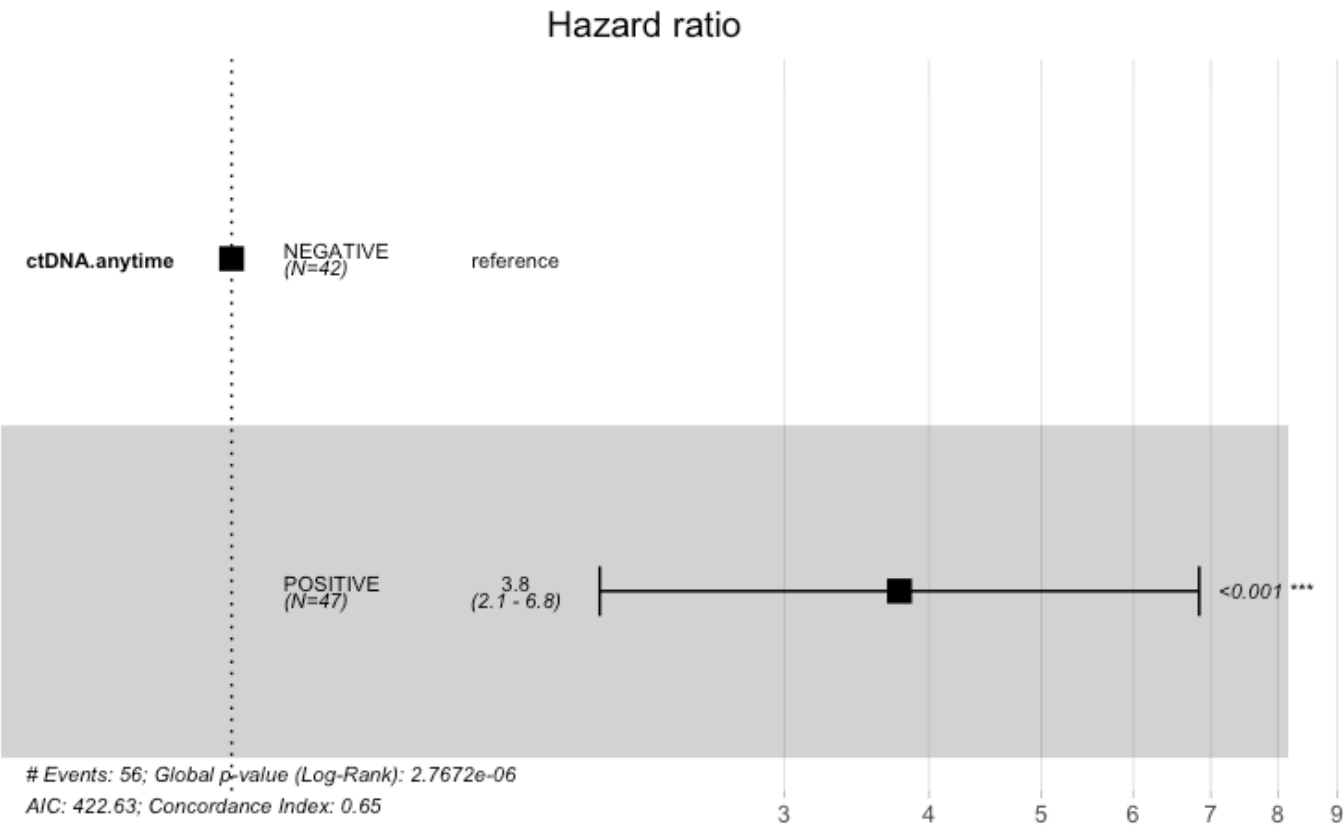
```
Call: survfit(formula = surv_object ~ ctDNA.anytime, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.anytime=NEGATIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	41	1	0.976	0.0235	0.843	0.997	
24	38	3	0.905	0.0453	0.766	0.963	

ctDNA.anytime=POSITIVE							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	41	6	0.872	0.0487	0.738	0.941	
24	26	15	0.553	0.0725	0.401	0.681	

Hide

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



Hide

```
summary(cox_fit)
```

Call:

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

n= 89, number of events= 56

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.anytimePOSITIVE	1.3272	3.7706	0.3039	4.367	1.26e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.anytimePOSITIVE	3.771	0.2652	2.078	6.841

Concordance= 0.648 (se = 0.033)

Likelihood ratio test= 21.97 on 1 df, p=3e-06

Wald test = 19.07 on 1 df, p=1e-05

Score (logrank) test = 21.89 on 1 df, p=3e-06

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 = 3.77 (2.08-6.84); p = 0"
```

#Prognostic role of ctDNA Dynamics

Hide

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

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

```
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
```

```
circ_data$RFS.months=circ_data$RFS.months-2
```

```
circ_data <- circ_data[circ_data$RFS.months>=0,]
```

```
circ_data$RFS.months[circ_data$RFS.months > 48] <- 48
```

```
circ_datadf <- as.data.frame(circ_data)
```

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



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

	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=Converted Negative	12	10	13.57	8.94	NA
ctDNA.Dynamics=Converted Positive	11	11	5.43	4.60	NA
ctDNA.Dynamics=Persistently Negative	56	34	19.06	13.84	NA
ctDNA.Dynamics=Persistently Positive	10	10	5.10	2.40	NA

Hide

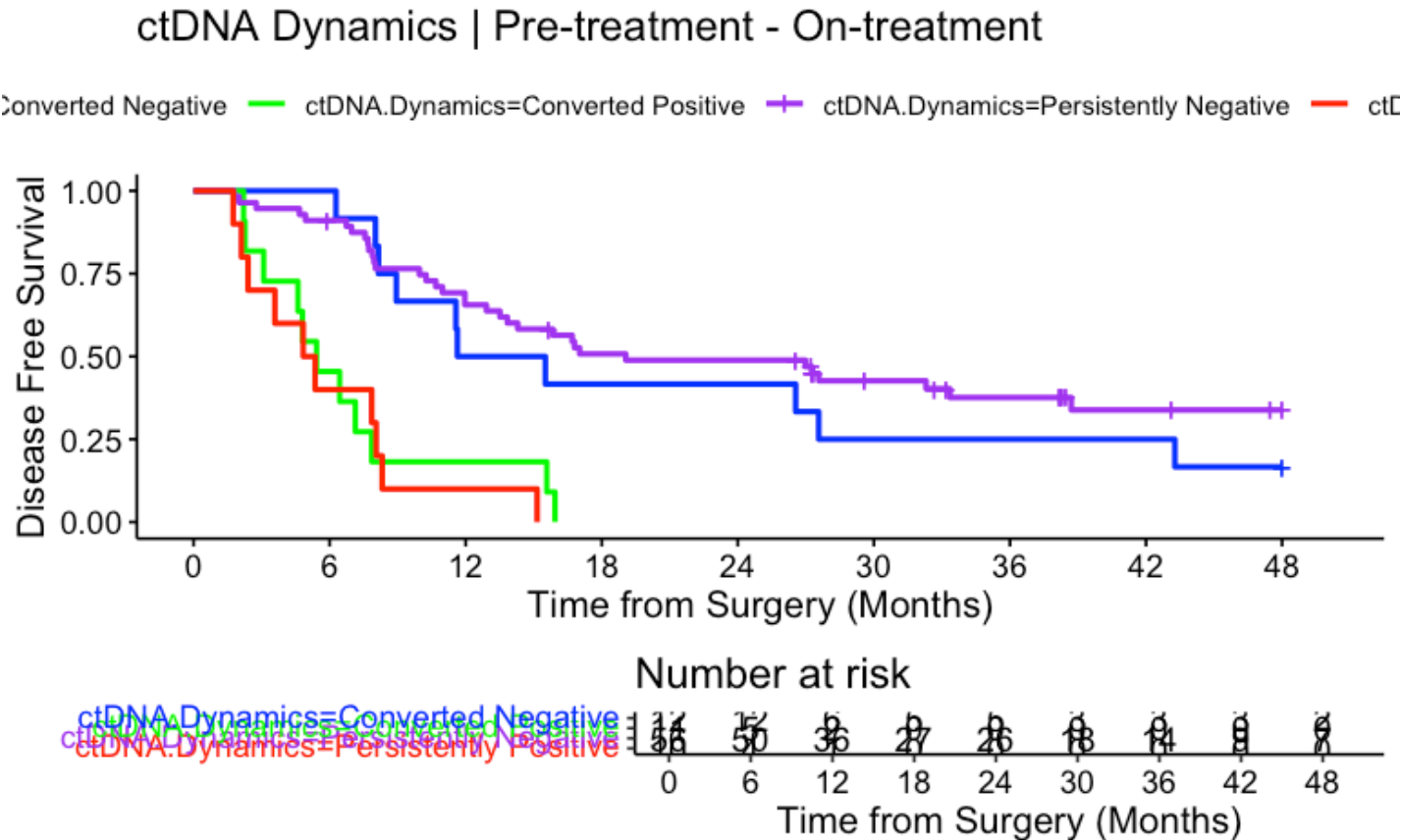
```
event_summary <- circ_data %>%
  group_by(ctDNA.Dynamics) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.Dynamics	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
Converted Negative	12	10	0.8333333	83.33333
Converted Positive	11	11	1.0000000	100.00000
Persistently Negative	56	34	0.6071429	60.71429
Persistently Positive	10	10	1.0000000	100.00000

4 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Dynamics, 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","green","purple","red"), title="ctDNA Dynamics | Pre-treatment - On-treatment", ylab= "Disease Free Survival", xlab="Time from Surgery (Months)", legend.title="")
```



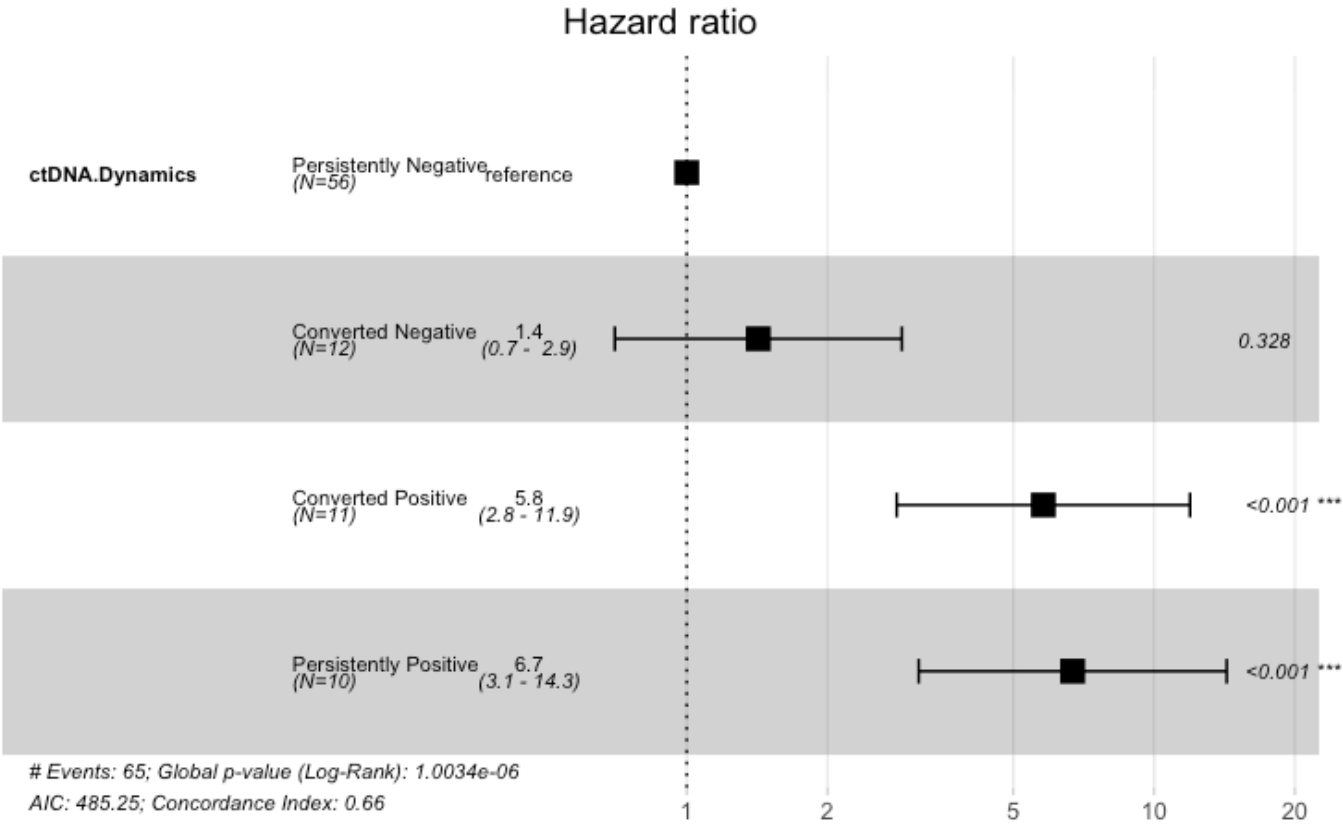
```
summary(KM_curve, times= c(12, 24))
```

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

ctDNA.Dynamics=Converted Negative							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	6	6	0.500	0.144	0.208	0.736	
24	5	1	0.417	0.142	0.152	0.665	
ctDNA.Dynamics=Converted Positive							
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95%
17	12.0000	2.0000	9.0000	0.1818	0.1163	0.0285	0.44
ctDNA.Dynamics=Persistently Negative							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	36	19	0.656	0.0640	0.515	0.765	
24	26	9	0.489	0.0677	0.351	0.613	
ctDNA.Dynamics=Persistently Positive							
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95%
13	12.00000	1.00000	9.00000	0.10000	0.09487	0.00572	0.358

Hide

```
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("Persistently Negative","Converted Negative","Converted Positive", "Persistently Positive"))
cox_fit <- coxph(surv_object ~ ctDNA.Dynamics, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

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

n= 89, number of events= 65

              coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.DynamicsConverted Negative    0.3528    1.4230   0.3607  0.978    0.328
ctDNA.DynamicsConverted Positive    1.7567    5.7935   0.3684  4.768 1.86e-06 ***
ctDNA.DynamicsPersistently Positive  1.9013    6.6944   0.3869  4.914 8.94e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
ctDNA.DynamicsConverted Negative    1.423    0.7027    0.7017    2.886
ctDNA.DynamicsConverted Positive    5.794    0.1726    2.8140   11.928
ctDNA.DynamicsPersistently Positive  6.694    0.1494    3.1358   14.291

Concordance= 0.659 (se = 0.031 )
Likelihood ratio test= 30.66 on 3 df,  p=1e-06
Wald test               = 35.98 on 3 df,  p=8e-08
Score (logrank) test = 44.86 on 3 df,  p=1e-09
```

#OS ctDNA Dynamics

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data$OS.months=circ_data$OS.months-2
circ_data <- circ_data[circ_data$OS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

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

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=Converted Negative	12	7	52.7	27.57	NA
ctDNA.Dynamics=Converted Positive	11	11	21.3	16.17	NA
ctDNA.Dynamics=Persistently Negative	56	28	48.4	29.28	NA
ctDNA.Dynamics=Persistently Positive	10	10	24.0	9.66	NA

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

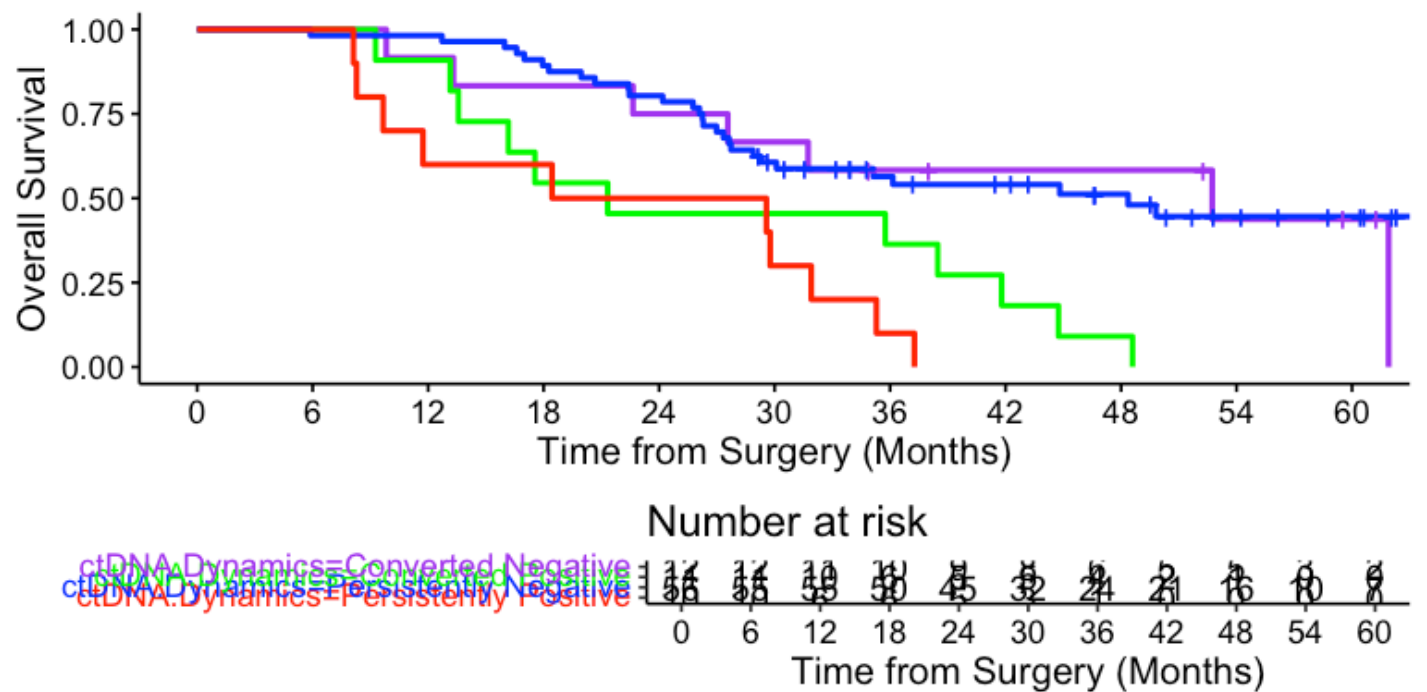
ctDNA.Dynamics <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
Converted Negative	12	7	0.5833333	58.33333
Converted Positive	11	11	1.0000000	100.00000
Persistently Negative	56	28	0.5000000	50.00000
Persistently Positive	10	10	1.0000000	100.00000

4 rows

```
surv_object <- Surv(time = circ_data$OS.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Dynamics, 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("purple","green","blue","red"), title="MRD Dynamics | Pre-treatment - On-treatment", ylab= "Overall Survival", xlab="Time from Surgery (Months)", legend.title="")
```

MRD Dynamics | Pre-treatment - On-treatment

Converted Negative — ctDNA.Dynamics=Converted Positive + ctDNA.Dynamics=Persistently Negative — ctDNA.Dynamics=Persistently Positive



Hide

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

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

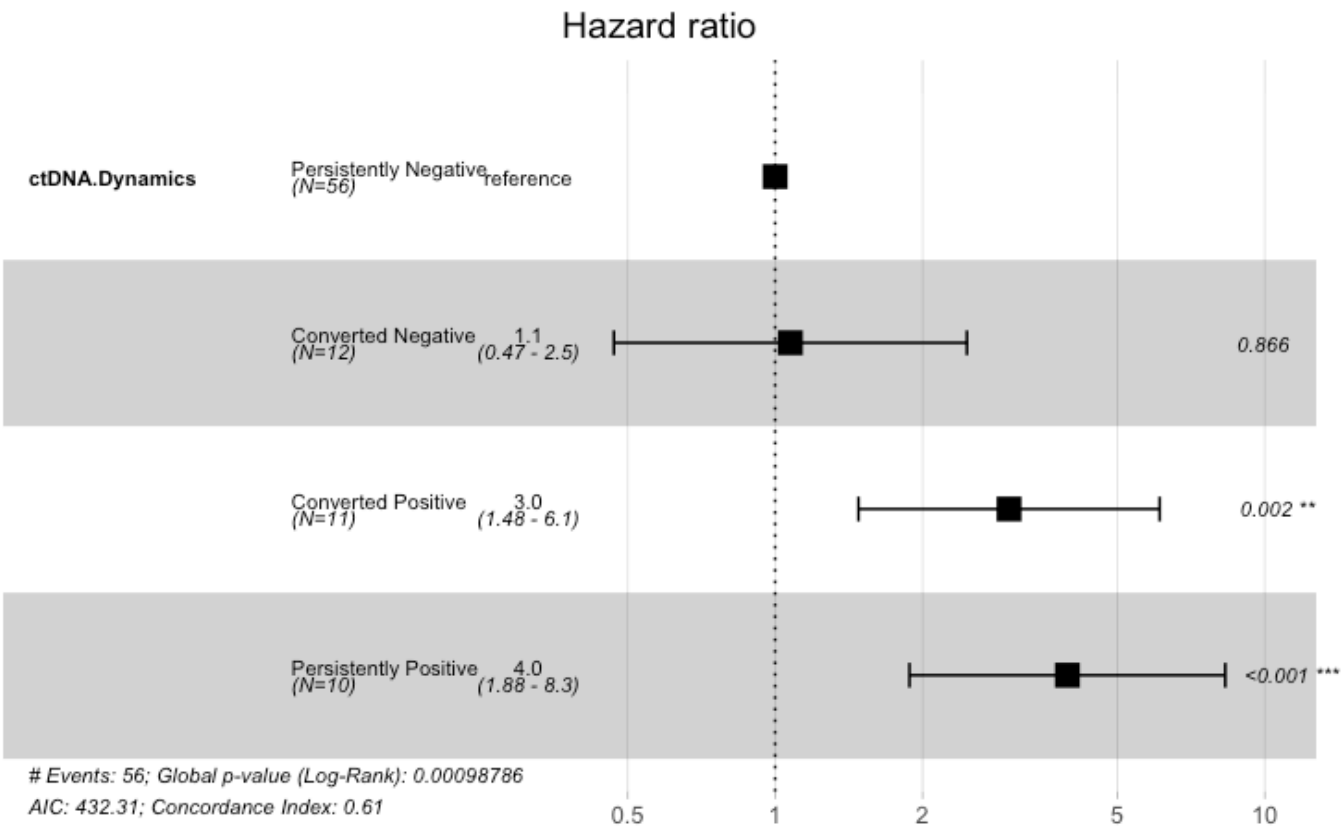
ctDNA.Dynamics=Converted Negative							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	11	1	0.917	0.0798	0.539	0.988	
24	9	2	0.750	0.1250	0.408	0.912	

ctDNA.Dynamics=Converted Positive							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	10	1	0.909	0.0867	0.508	0.987	
24	5	5	0.455	0.1501	0.167	0.707	

ctDNA.Dynamics=Persistently Negative							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	55	1	0.982	0.0177	0.880	0.997	
24	45	10	0.804	0.0531	0.673	0.886	

ctDNA.Dynamics=Persistently Positive							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
12	6	4	0.6	0.155	0.253	0.827	
24	5	1	0.5	0.158	0.184	0.753	

```
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("Persistently Negative","Converted Negative","Converted Positive", "Persistently Positive"))
cox_fit <- coxph(surv_object ~ ctDNA.Dynamics, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



```
summary(cox_fit)
```


Call:

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

n= 89, number of events= 56

		coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.DynamicsConverted	Negative	0.07171	1.07435	0.42352	0.169	0.865537
ctDNA.DynamicsConverted	Positive	1.10013	3.00455	0.36133	3.045	0.002329 **
ctDNA.DynamicsPersistently	Positive	1.37436	3.95256	0.37891	3.627	0.000287 ***

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

		exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.DynamicsConverted	Negative	1.074	0.9308	0.4684	2.464
ctDNA.DynamicsConverted	Positive	3.005	0.3328	1.4799	6.100
ctDNA.DynamicsPersistently	Positive	3.953	0.2530	1.8808	8.306

Concordance= 0.61 (se = 0.035)

Likelihood ratio test= 16.29 on 3 df, p=0.001

Wald test = 18.53 on 3 df, p=3e-04

Score (logrank) test = 21 on 3 df, p=1e-04

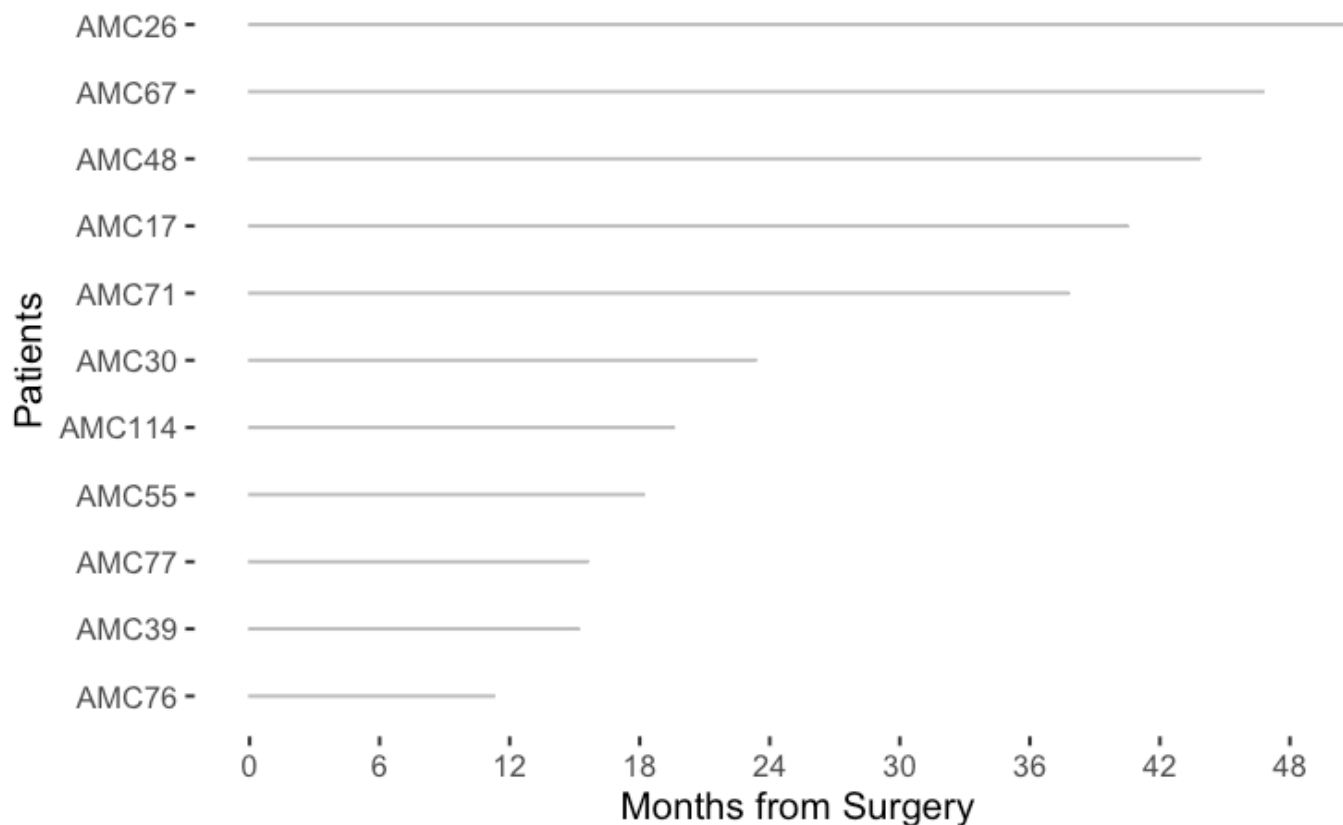
#OP for pts converted positive during ACT

Hide

```
setwd("~/Downloads")
clinstage<- read.csv("ASAN_Cholangio_OP.csv")
clinstage_df<- as.data.frame(clinstage)
clinstage_df <- clinstage_df[clinstage_df$ctDNA.Dynamics=="Converted Positive",]

##Overview plot - stratified by Stage
oplot_stratify <-swimmer_plot(df=clinstage_df,
                             id='PatientName',
                             end='fu.diff.months',
                             #name_fill='Arm',
                             col="gray",
                             alpha=0.75,
                             width=.01,
                             base_size = 14)

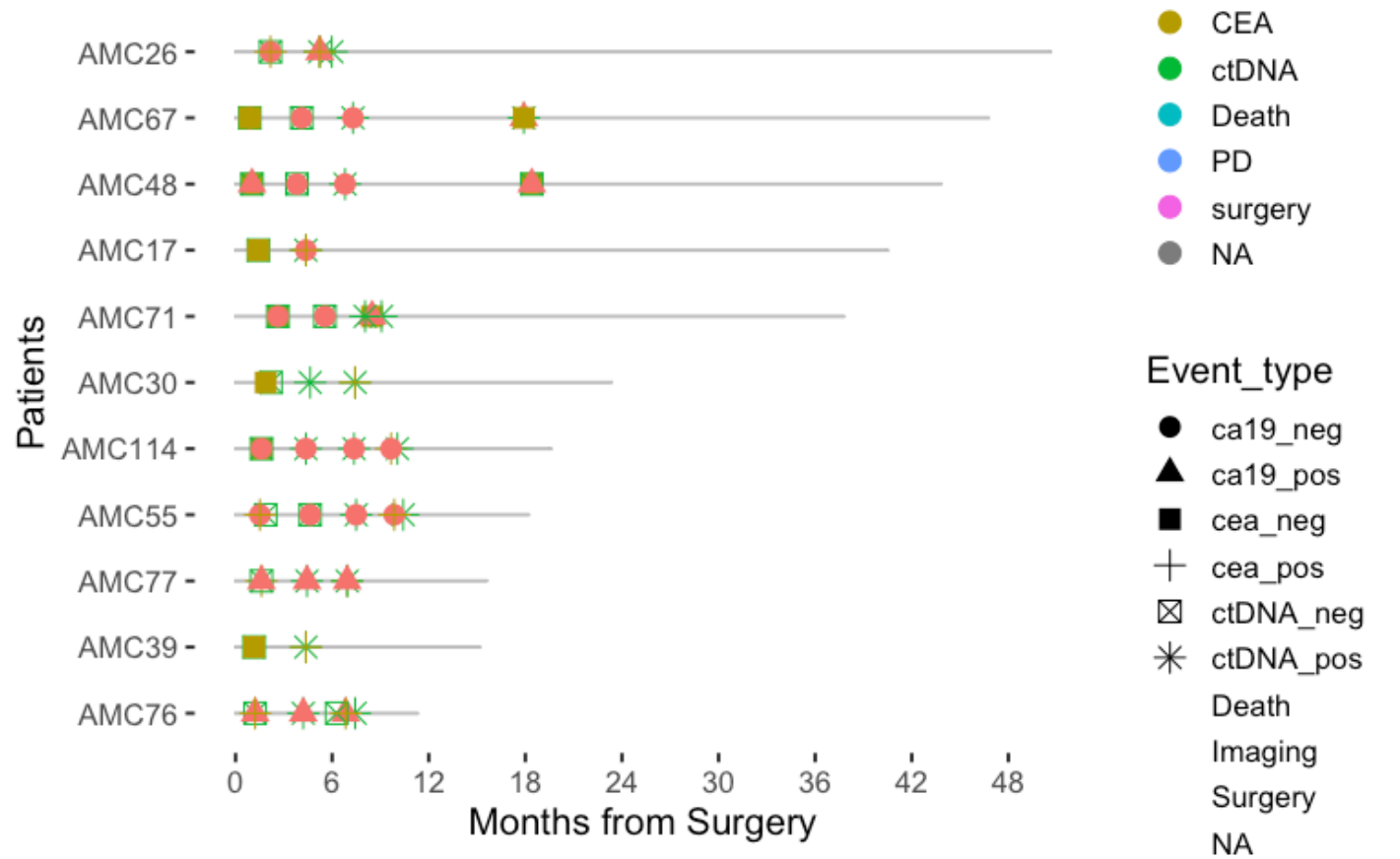
oplot_stratify <- opplot_stratify + theme(panel.border = element_blank())
oplot_stratify <- opplot_stratify + scale_y_continuous(breaks = seq(0, 108, by = 6))
oplot_stratify <- opplot_stratify + labs(x ="Patients" , y="Months from Surgery")
oplot_stratify
```


[Hide](#)

```
##plot events
oplot_ev3 <- oplot_stratify + swimmer_points(df_points=clinstage_df,
                                             id='PatientName',
                                             time='date.diff.months',
                                             name_shape = 'Event_type',
                                             name_col = 'Event',
                                             size=3.5,fill='black',
                                             #col='darkgreen'
)
oplot_ev3
```

Warning: The shape palette can deal with a maximum of 6 discrete values because more than 6 becomes difficult to discriminate
 you have requested 9 values. Consider specifying shapes manually if you need that many have them.

Warning: Removed 55 rows containing missing values or values outside the scale range (`geom_point()`).



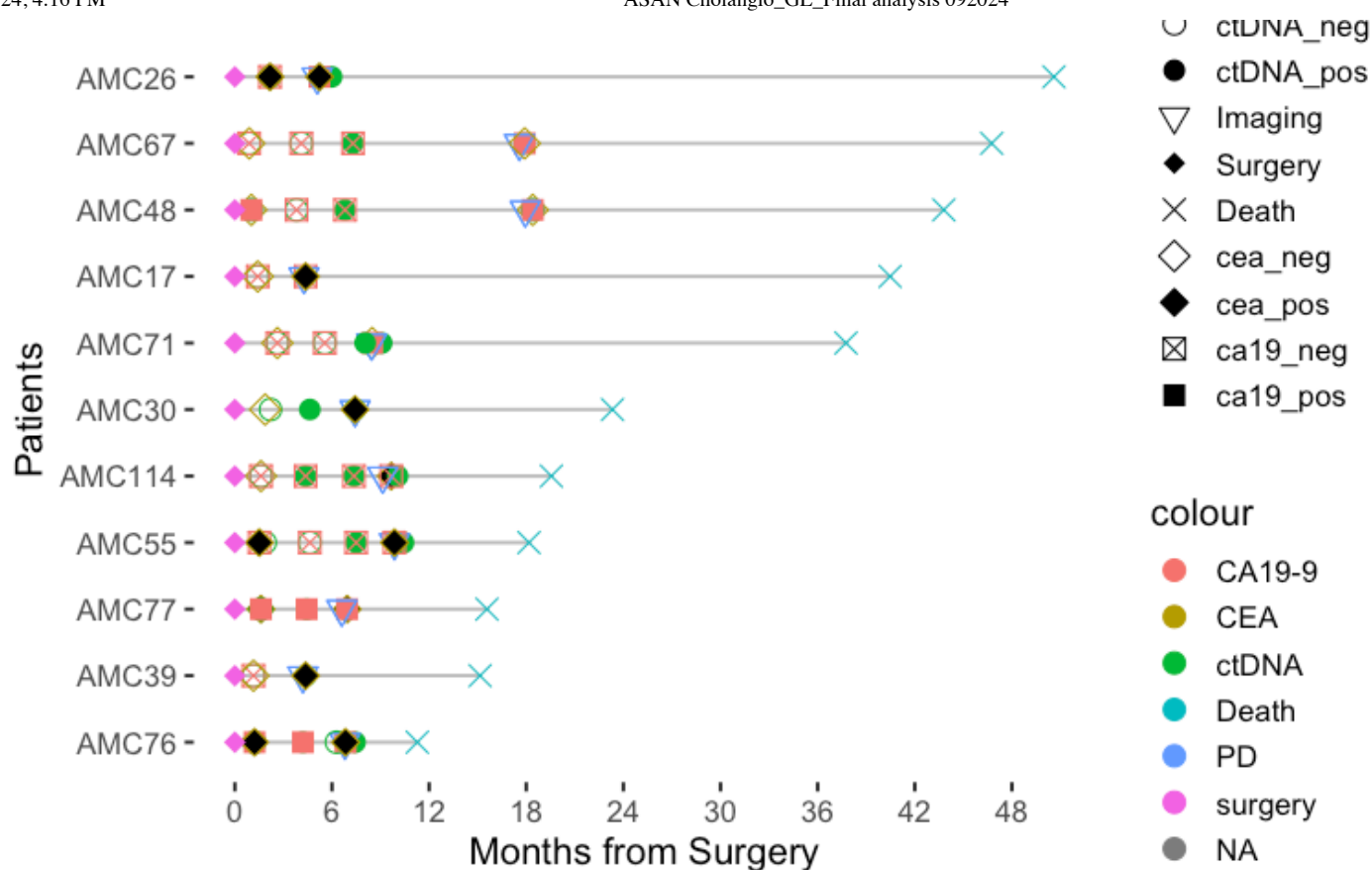
Hide

```
#Shape customization to Event_type

oplot_ev3.1 <- oplot_ev3 + ggplot2::scale_shape_manual(name="Event_type",values=c(1,16,
6,18,4, 5, 23, 7, 15),breaks=c('ctDNA_neg', 'ctDNA_pos', 'Imaging', 'Surgery', 'Death',
"cea_neg", "cea_pos", "ca19_neg", "ca19_pos"))

oplot_ev3.1
```

Warning: Removed 11 rows containing missing values or values outside the scale range (`geom_point()`).

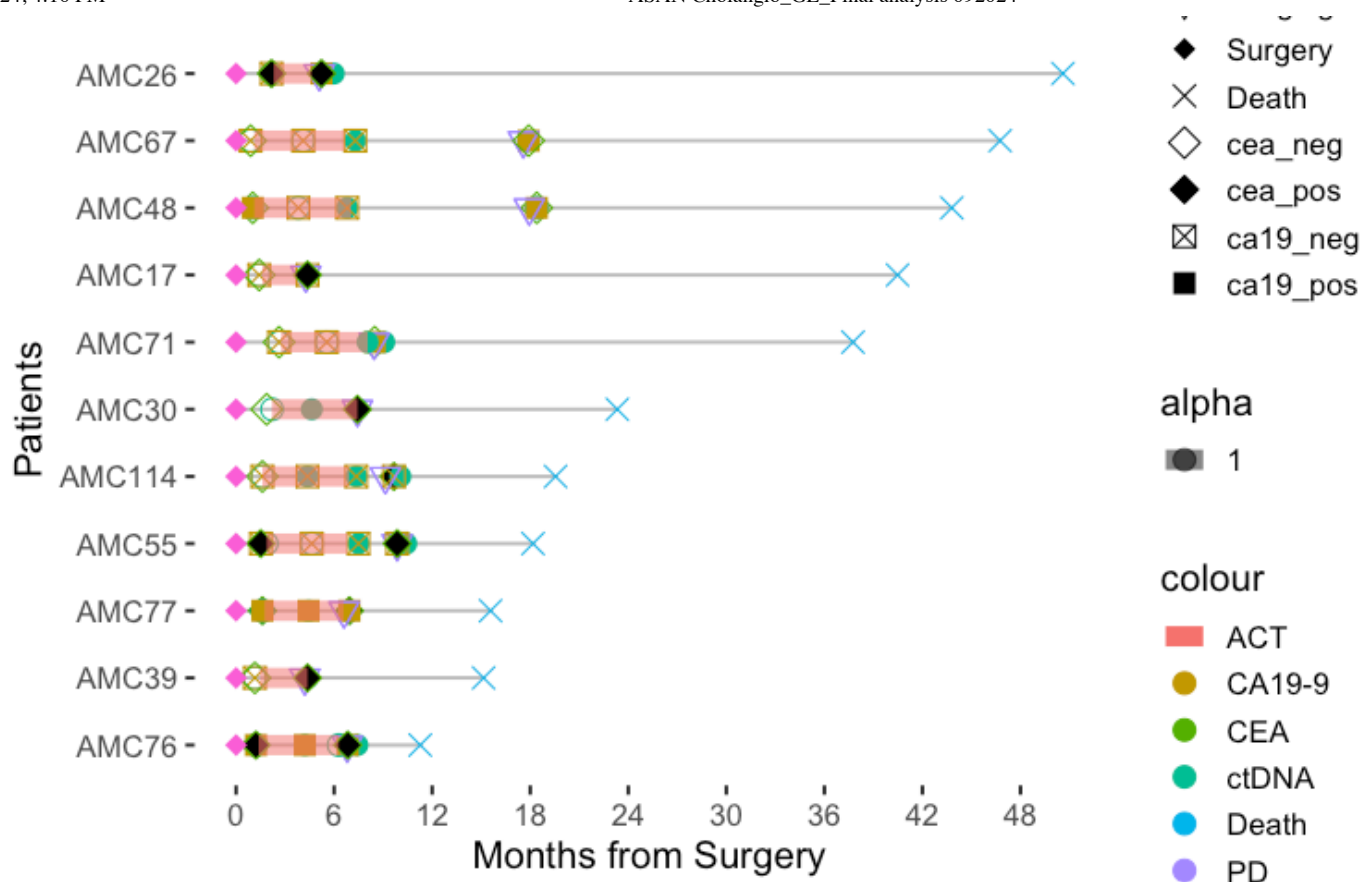

[Hide](#)

```
#plot treatment
```

```
oplot_ev4 <- oplot_ev3.1 + swimmer_lines(df_lines=clinstage_df,
                                         id='PatientName',
                                         start='Tx_start.months',
                                         end='Tx_end.months',
                                         name_col='Tx_type',
                                         size=3.5,
                                         name_alpha = 1.0)
oplot_ev4 <- oplot_ev4 + guides(linetype = guide_legend(override.aes = list(size = 5, color = "black")))
oplot_ev4
```

Warning: Removed 11 rows containing missing values or values outside the scale range (`geom_point()`).

Warning: Removed 133 rows containing missing values or values outside the scale range (`geom_segment()`).

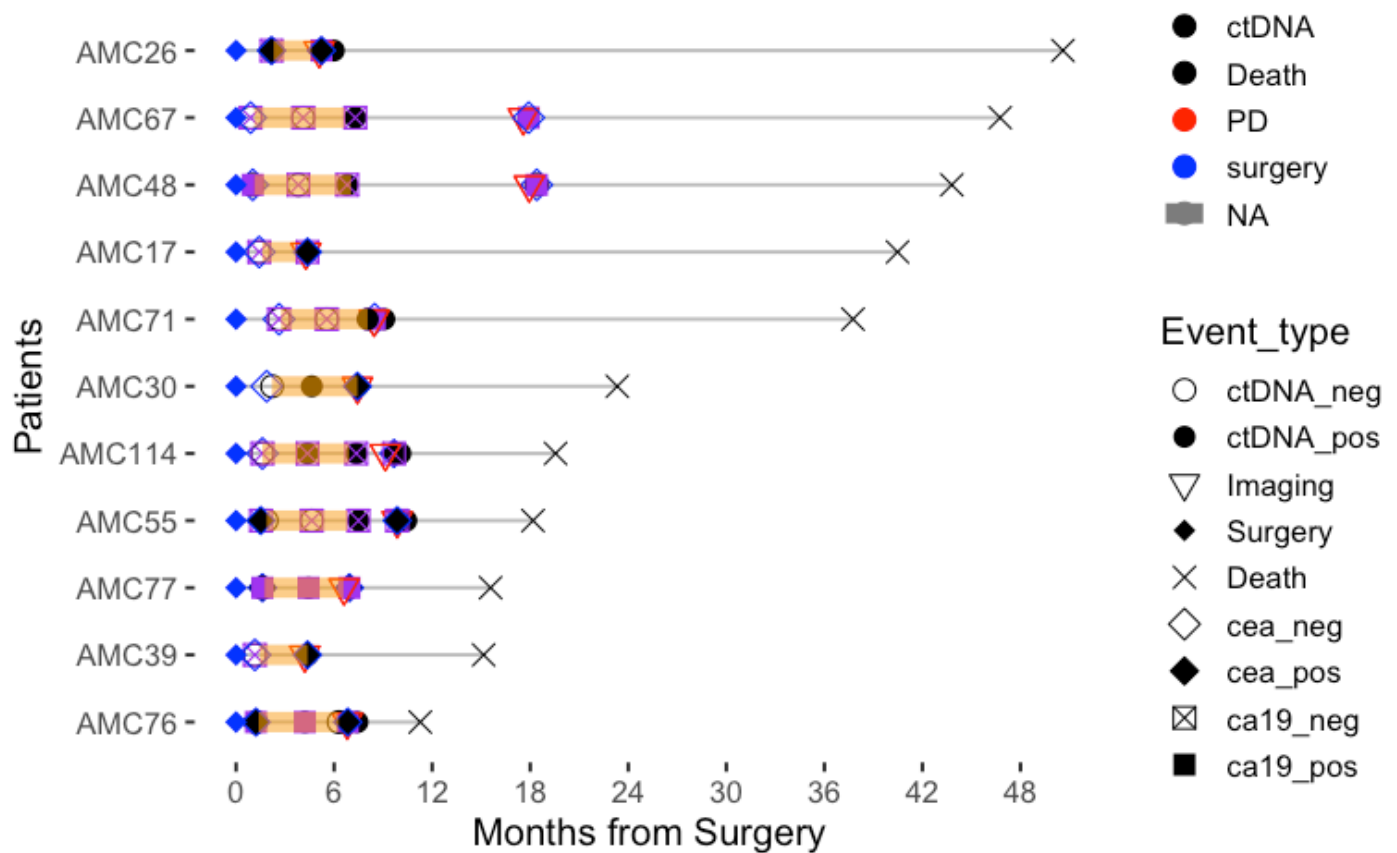


Hide

```
#colour customization
# orange=ACT, Black=Death, Red=PD, ctDNA negative=white, ctDNA positive=black, Surgery=blue, TURBT=gray
oplot_ev4.2 <- oplot_ev4 + ggplot2::scale_color_manual(name="Event", values=c( "orange",
"purple", "blue", "black", "black", "red", "blue", "blue"))
oplot_ev4.2
```

Warning: Removed 11 rows containing missing values or values outside the scale range (`geom_point()`).

Warning: Removed 133 rows containing missing values or values outside the scale range (`geom_segment()`).



#ctDNA clearance proportions by chemotherapy regimen

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- circ_data[circ_data$ctDNA.MRD=="POSITIVE",]

circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels = c("Converted Negative", "Persistently Positive"), labels = c("Clearance", "No Clearance"))
circ_data$Chemo <- factor(circ_data$Chemo, levels = c("CAP", "GemCis"))
contingency_table <- table(circ_data$Chemo, circ_data$ctDNA.Dynamics)
chi_square_test <- chisq.test(contingency_table)
```

Warning in chisq.test(contingency_table) :
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 = 0.26481, df = 1, p-value = 0.6068
```

Hide

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

Fisher's Exact Test for Count Data

```
data: contingency_table
p-value = 0.4149
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.3023403 20.3968347
sample estimates:
odds ratio
 2.24368
```

Hide

```
print(contingency_table)
```

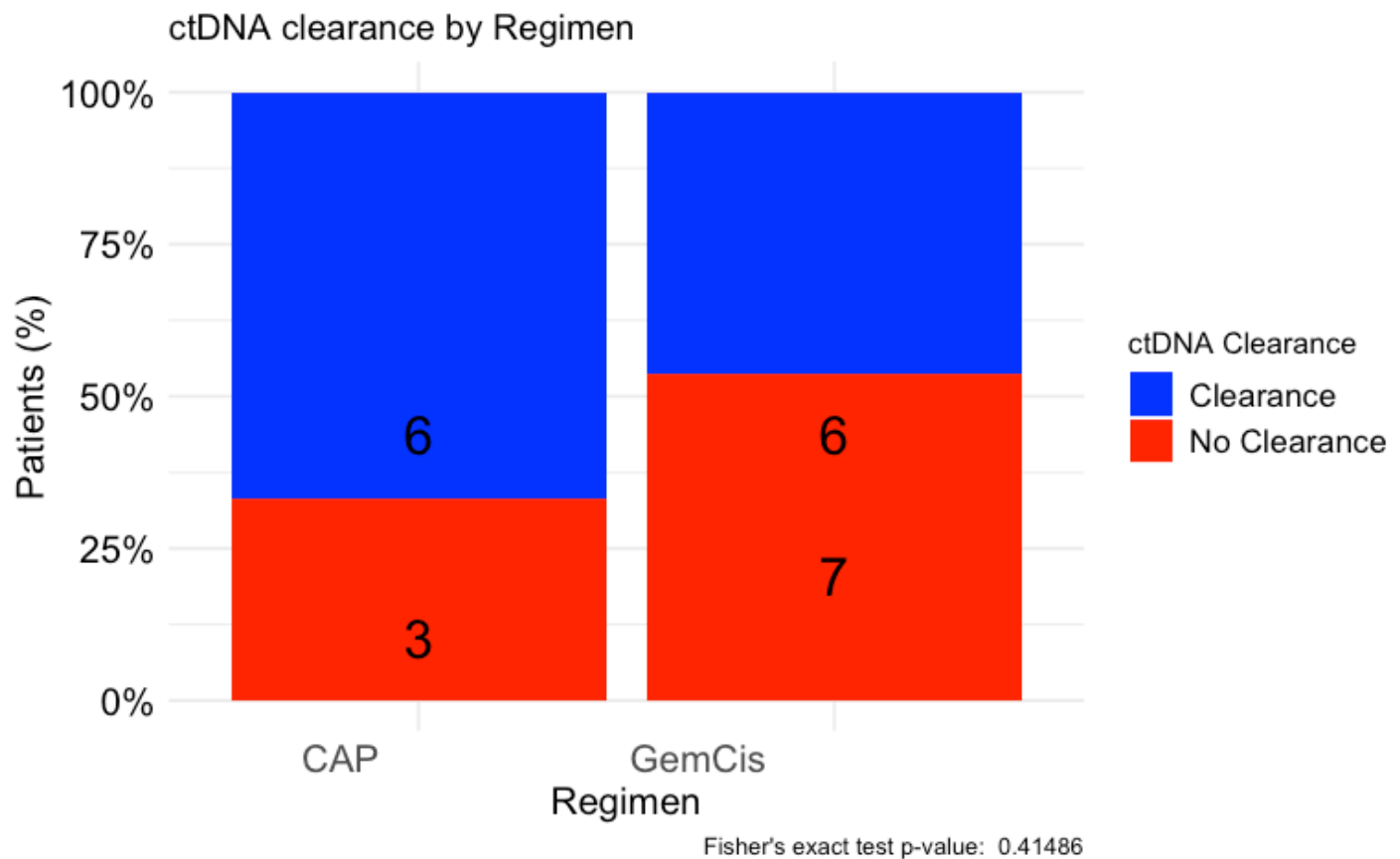
	Clearance	No Clearance
CAP	6	3
GemCis	6	7

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 = "ctDNA clearance by Regimen",
        x = "Regimen",
        y = "Patients (%)",
        fill = "ctDNA Clearance",
        caption = paste("Fisher's exact test p-value: ", format.pval(fisher_exact_test$p.value))) +
  scale_y_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("Clearance" = "blue", "No Clearance" = "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 Progression label size

```

#Prognostic role of ctDNA C5D1 on Chemotherapy - 4 groups

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60

circ_data$ctDNA.C5D1.Chemo <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C5D1.Chemo = case_when(
    Chemo == "CAP" & ctDNA.C5D1 == "NEGATIVE" ~ 1,
    Chemo == "CAP" & ctDNA.C5D1 == "POSITIVE" ~ 2,
    Chemo == "GemCis" & ctDNA.C5D1 == "NEGATIVE" ~ 3,
    Chemo == "GemCis" & ctDNA.C5D1 == "POSITIVE" ~ 4
  ))
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C5D1.Chemo,
data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event) ~
      ctDNA.C5D1.Chemo, data = circ_data)
```

1 observation deleted due to missingness

	n	events	median	0.95LCL	0.95UCL
ctDNA.C5D1.Chemo=1	36	24	15.58	11.57	NA
ctDNA.C5D1.Chemo=2	8	8	3.95	2.27	NA
ctDNA.C5D1.Chemo=3	35	24	16.79	11.96	38.7
ctDNA.C5D1.Chemo=4	9	9	4.83	3.59	NA

Hide

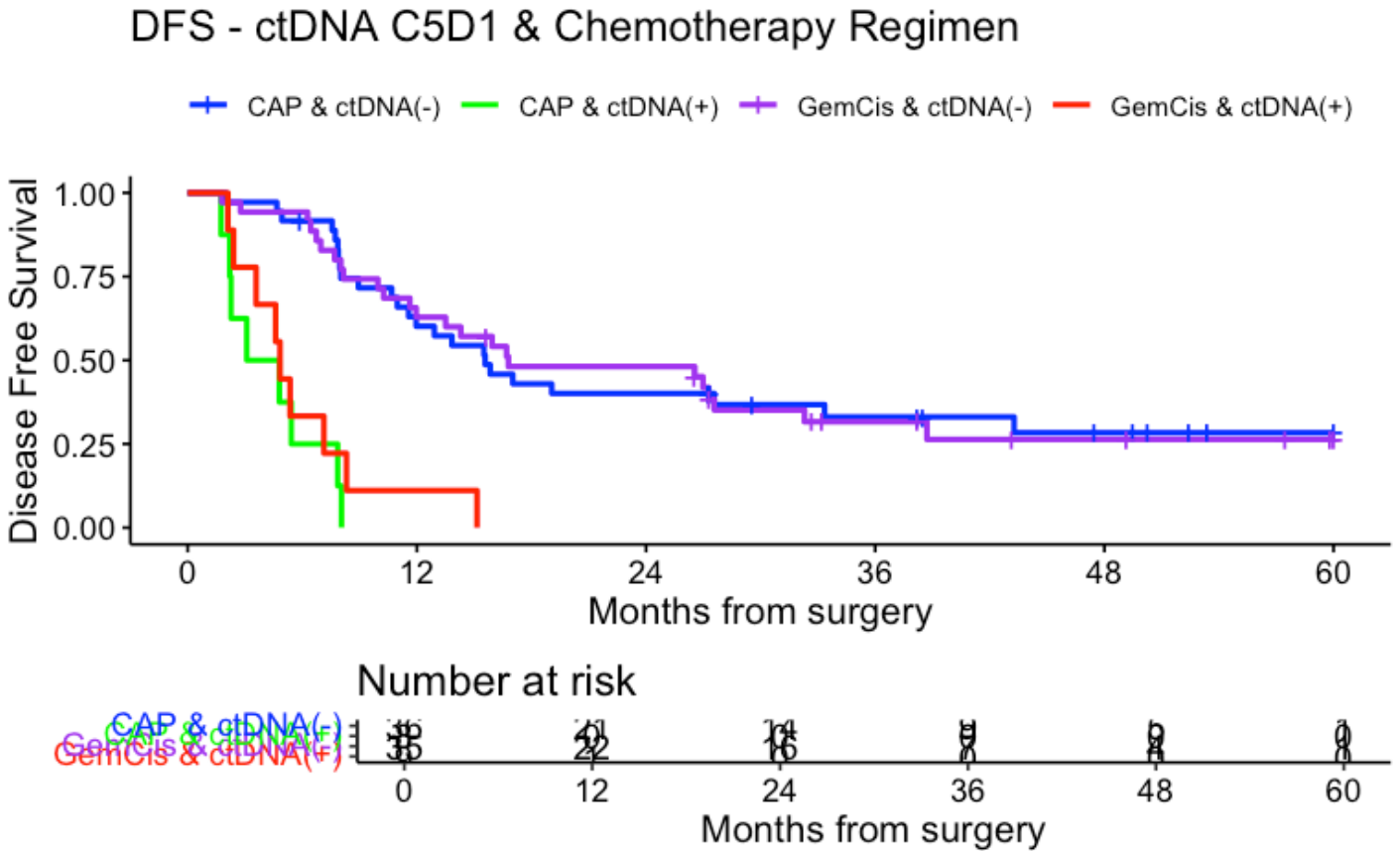
```
event_summary <- circ_data %>%
  group_by(ctDNA.C5D1.Chemo) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C5D1.Chemo	Total	Events	Fraction	Percentage
<dbl>	<int>	<int>	<dbl>	<dbl>
1	36	24	0.6666667	66.66667
2	8	8	1.0000000	100.00000
3	35	24	0.6857143	68.57143
4	9	9	1.0000000	100.00000
NA	1	0	0.0000000	0.00000

5 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C5D1.Chemo, 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=12, palette=c("blue","green","purple", "red"), title="DFS - ctDNA C5D1
& Chemotherapy Regimen", ylab= "Disease Free Survival", xlab="Months from surgery", lege
nd.labs=c("CAP & ctDNA(-)", "CAP & ctDNA(+)","GemCis & ctDNA(-)", "GemCis & ctDNA(+)"),
legend.title="")
```



Hide

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

Call: `survfit(formula = surv_object ~ ctDNA.C5D1.Chemo, data = circ_data, conf.int = 0.95, conf.type = "log-log")`

1 observation deleted due to missingness

ctDNA.C5D1.Chemo=1							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	36	0	1.000	0.0000	1.000	1.000	
12	21	14	0.602	0.0827	0.422	0.741	
24	14	7	0.401	0.0829	0.241	0.556	

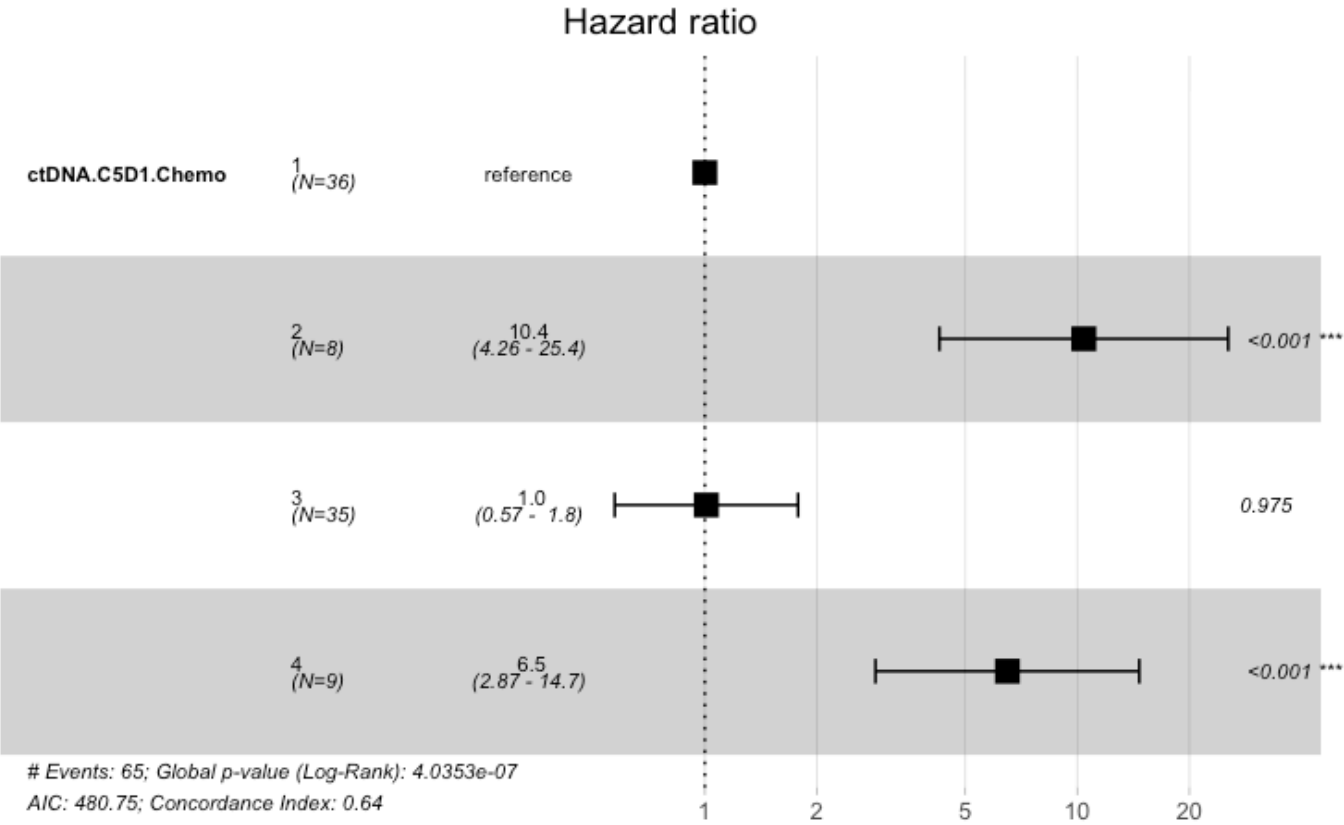
ctDNA.C5D1.Chemo=2							
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	0	8	0	1	0	1	

ctDNA.C5D1.Chemo=3							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	35	0	1.000	0.0000	1.000	1.000	
12	22	13	0.629	0.0817	0.448	0.765	
24	16	5	0.481	0.0851	0.309	0.634	

ctDNA.C5D1.Chemo=4							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	9	0	1.000	0.000	1.00000	1.000	
12	1	8	0.111	0.105	0.00613	0.388	

Hide

```
circ_data$ctDNA.C5D1.Chemo <- factor(circ_data$ctDNA.C5D1.Chemo, levels=c("1","2","3","4"))
cox_fit <- coxph(surv_object ~ ctDNA.C5D1.Chemo, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ ctDNA.C5D1.Chemo, data = circ_data)

n= 88, number of events= 65
(1 observation deleted due to missingness)

              coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.C5D1.Chemo2  2.34287  10.41111  0.45548  5.144 2.69e-07 ***
ctDNA.C5D1.Chemo3  0.00918   1.00922  0.28914  0.032   0.975
ctDNA.C5D1.Chemo4  1.87019   6.48950  0.41611  4.494 6.97e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
ctDNA.C5D1.Chemo2    10.411    0.09605    4.2637    25.422
ctDNA.C5D1.Chemo3     1.009    0.99086    0.5726     1.779
ctDNA.C5D1.Chemo4     6.490    0.15409    2.8709    14.669

Concordance= 0.639 (se = 0.035 )
Likelihood ratio test= 32.54 on 3 df,  p=4e-07
Wald test               = 40.73 on 3 df,  p=7e-09
Score (logrank) test = 55.87 on 3 df,  p=4e-12
```

Hide

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

#Prognostic role of ctDNA C8D1 on Chemotherapy - 4 groups

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C8D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60

circ_data$ctDNA.C8D1.Chemo <- NA #first we create the variable for the ctDNA & NAC combi
nation, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C8D1.Chemo = case_when(
    Chemo == "CAP" & ctDNA.C8D1 == "NEGATIVE" ~ 1,
    Chemo == "CAP" & ctDNA.C8D1 == "POSITIVE" ~ 2,
    Chemo == "GemCis" & ctDNA.C8D1 == "NEGATIVE" ~ 3,
    Chemo == "GemCis" & ctDNA.C8D1 == "POSITIVE" ~ 4
  ))
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C8D1.Chemo,
data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$RFS.months, event = circ_data\$RFS.Event) ~
ctDNA.C8D1.Chemo, data = circ_data)

12 observations deleted due to missingness

	n	events	median	0.95LCL	0.95UCL
ctDNA.C8D1.Chemo=1	31	20	15.84	11.57	NA
ctDNA.C8D1.Chemo=2	6	6	7.86	5.43	NA
ctDNA.C8D1.Chemo=3	31	20	26.98	13.51	NA
ctDNA.C8D1.Chemo=4	9	9	6.44	4.83	NA

Hide

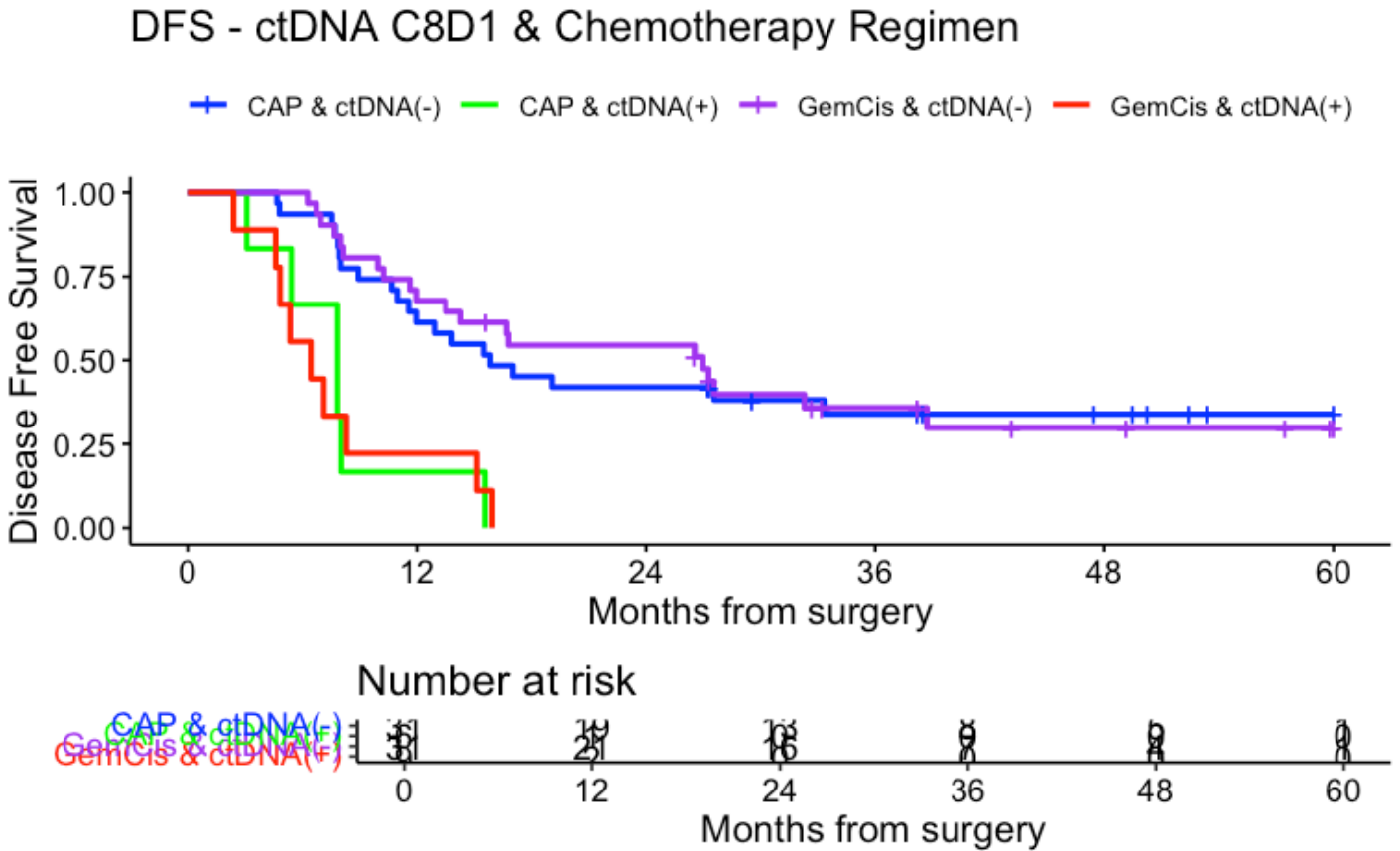
```
event_summary <- circ_data %>%
  group_by(ctDNA.C8D1.Chemo) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C8D1.Chemo	Total	Events	Fraction	Percentage
<dbl>	<int>	<int>	<dbl>	<dbl>
1	31	20	0.6451613	64.51613
2	6	6	1.0000000	100.00000
3	31	20	0.6451613	64.51613
4	9	9	1.0000000	100.00000
NA	12	10	0.8333333	83.33333

5 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C8D1.Chemo, 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=12, palette=c("blue","green","purple", "red"), title="DFS - ctDNA C8D1
& Chemotherapy Regimen", ylab= "Disease Free Survival", xlab="Months from surgery", lege
nd.labs=c("CAP & ctDNA(-)", "CAP & ctDNA(+)","GemCis & ctDNA(-)", "GemCis & ctDNA(+)"),
legend.title="")
```



Hide

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



```
Call: survfit(formula = surv_object ~ ctDNA.C8D1.Chemo, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

12 observations deleted due to missingness

ctDNA.C8D1.Chemo=1

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	31	0	1.000	0.0000	1.000	1.000
12	19	12	0.613	0.0875	0.420	0.758
24	13	6	0.419	0.0886	0.247	0.583

ctDNA.C8D1.Chemo=2

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	6	0	1.000	0.000	1.00000	1.000
12	1	5	0.167	0.152	0.00772	0.517

ctDNA.C8D1.Chemo=3

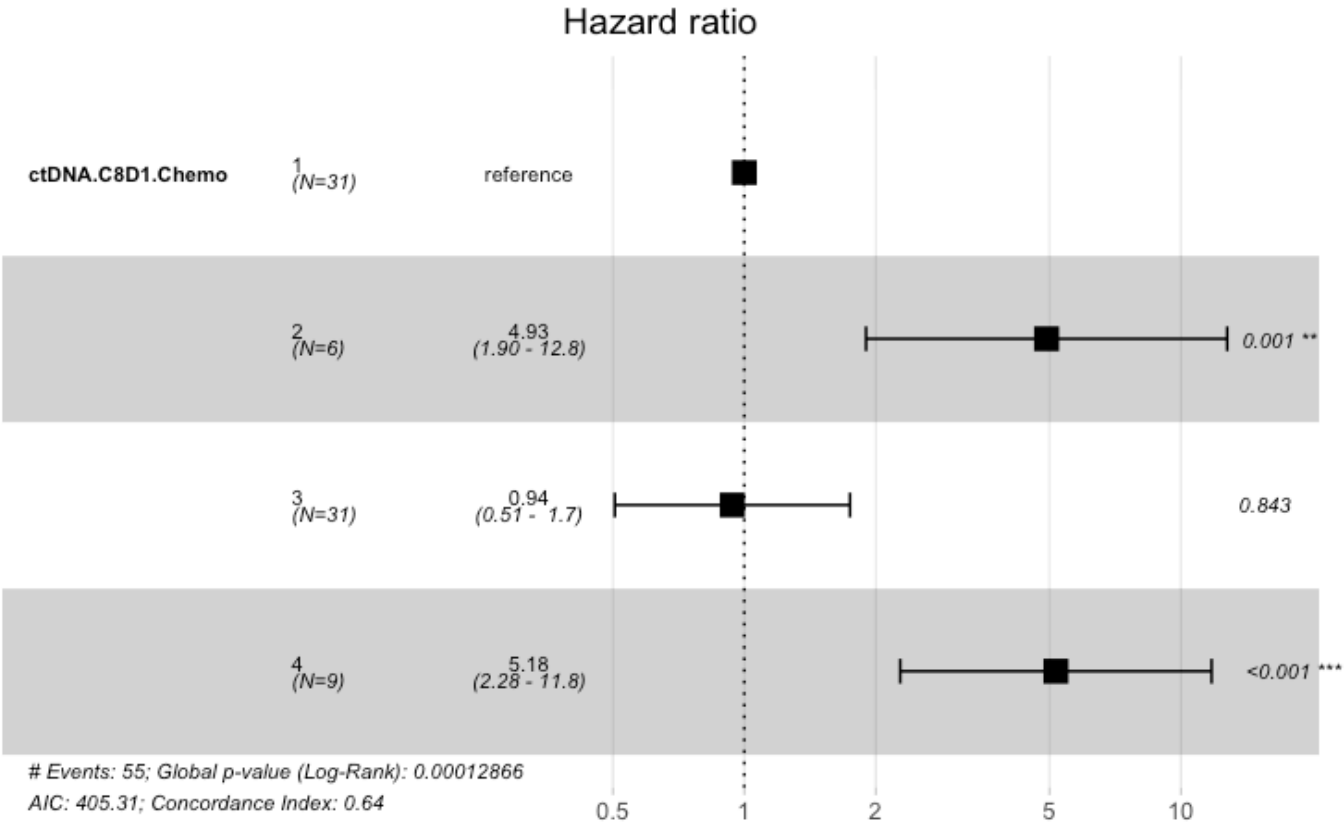
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	31	0	1.000	0.000	1.000	1.000
12	21	10	0.677	0.084	0.484	0.812
24	16	4	0.545	0.090	0.355	0.700

ctDNA.C8D1.Chemo=4

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	9	0	1.000	0.000	1.0000	1.000
12	2	7	0.222	0.139	0.0337	0.513

Hide

```
circ_data$ctDNA.C8D1.Chemo <- factor(circ_data$ctDNA.C8D1.Chemo, levels=c
("1","2","3","4"))
cox_fit <- coxph(surv_object ~ ctDNA.C8D1.Chemo, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:
coxph(formula = surv_object ~ ctDNA.C8D1.Chemo, data = circ_data)

n= 77, number of events= 55
(12 observations deleted due to missingness)

	coef	exp(coef)	se(coef)	z	Pr(> z)	
ctDNA.C8D1.Chemo2	1.59481	4.92738	0.48599	3.282	0.00103	**
ctDNA.C8D1.Chemo3	-0.06279	0.93914	0.31645	-0.198	0.84271	
ctDNA.C8D1.Chemo4	1.64389	5.17525	0.41910	3.922	8.76e-05	***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C8D1.Chemo2	4.9274	0.2029	1.9008	12.773
ctDNA.C8D1.Chemo3	0.9391	1.0648	0.5051	1.746
ctDNA.C8D1.Chemo4	5.1753	0.1932	2.2761	11.767

Concordance= 0.638 (se = 0.037)
Likelihood ratio test= 20.58 on 3 df, p=1e-04
Wald test = 25.53 on 3 df, p=1e-05
Score (logrank) test = 31.47 on 3 df, p=7e-07

Hide

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

#Prognostic role of ctDNA C5D1 on Primary Site - 4 groups

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C5D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60

circ_data$ctDNA.C5D1.PrimSite <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C5D1.PrimSite = case_when(
    PrimSite == "dCCA" & ctDNA.C5D1 == "NEGATIVE" ~ 1,
    PrimSite == "dCCA" & ctDNA.C5D1 == "POSITIVE" ~ 2,
    PrimSite == "pCCA" & ctDNA.C5D1 == "NEGATIVE" ~ 3,
    PrimSite == "pCCA" & ctDNA.C5D1 == "POSITIVE" ~ 4
  ))
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C5D1.PrimSi
te, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event) ~
ctDNA.C5D1.PrimSite, data = circ_data)
```

1 observation deleted due to missingness

	n	events	median	0.95LCL	0.95UCL
ctDNA.C5D1.PrimSite=1	37	20	16.79	11.96	NA
ctDNA.C5D1.PrimSite=2	8	8	3.88	2.27	NA
ctDNA.C5D1.PrimSite=3	34	28	15.71	11.63	27.3
ctDNA.C5D1.PrimSite=4	9	9	4.80	3.59	NA

Hide

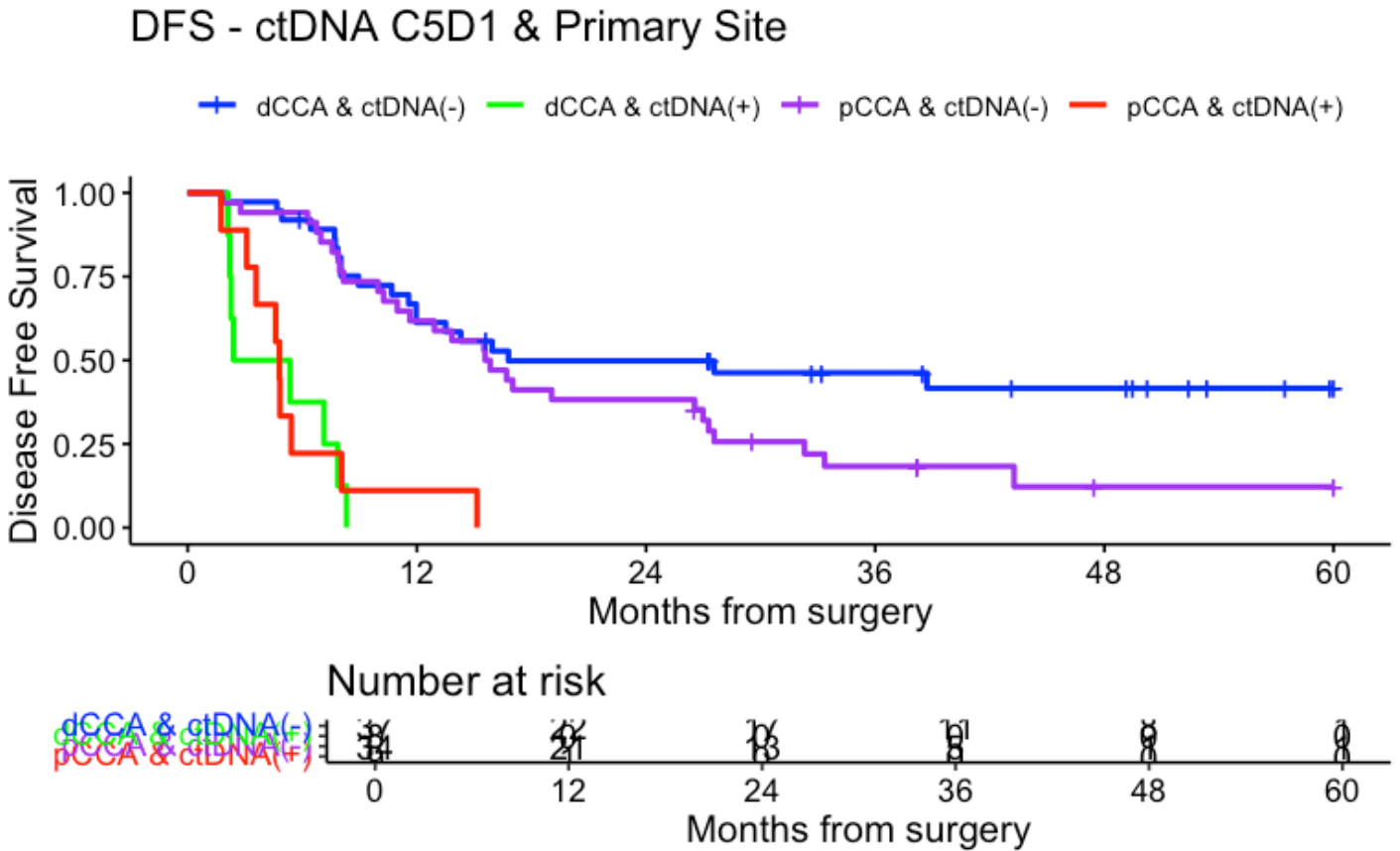
```
event_summary <- circ_data %>%
  group_by(ctDNA.C5D1.PrimSite) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C5D1.PrimSite	Total	Events	Fraction	Percentage
<dbl>	<int>	<int>	<dbl>	<dbl>
1	37	20	0.5405405	54.05405
2	8	8	1.0000000	100.00000
3	34	28	0.8235294	82.35294
4	9	9	1.0000000	100.00000
NA	1	0	0.0000000	0.00000

5 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C5D1.PrimSite, 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=12, palette=c("blue","green","purple", "red"), title="DFS - ctDNA C5D1 & Primary Site", ylab= "Disease Free Survival", xlab="Months from surgery", legend.labs=c("dCCA & ctDNA(-)", "dCCA & ctDNA(+)","pCCA & ctDNA(-)", "pCCA & ctDNA(+)"), legend.title="")
```



Hide

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

Call: `survfit(formula = surv_object ~ ctDNA.C5D1.PrimSite, data = circ_data, conf.int = 0.95, conf.type = "log-log")`

1 observation deleted due to missingness

ctDNA.C5D1.PrimSite=1							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	37	0	1.000	0.0000	1.000	1.000	
12	22	14	0.613	0.0811	0.435	0.749	
24	17	4	0.498	0.0838	0.327	0.648	

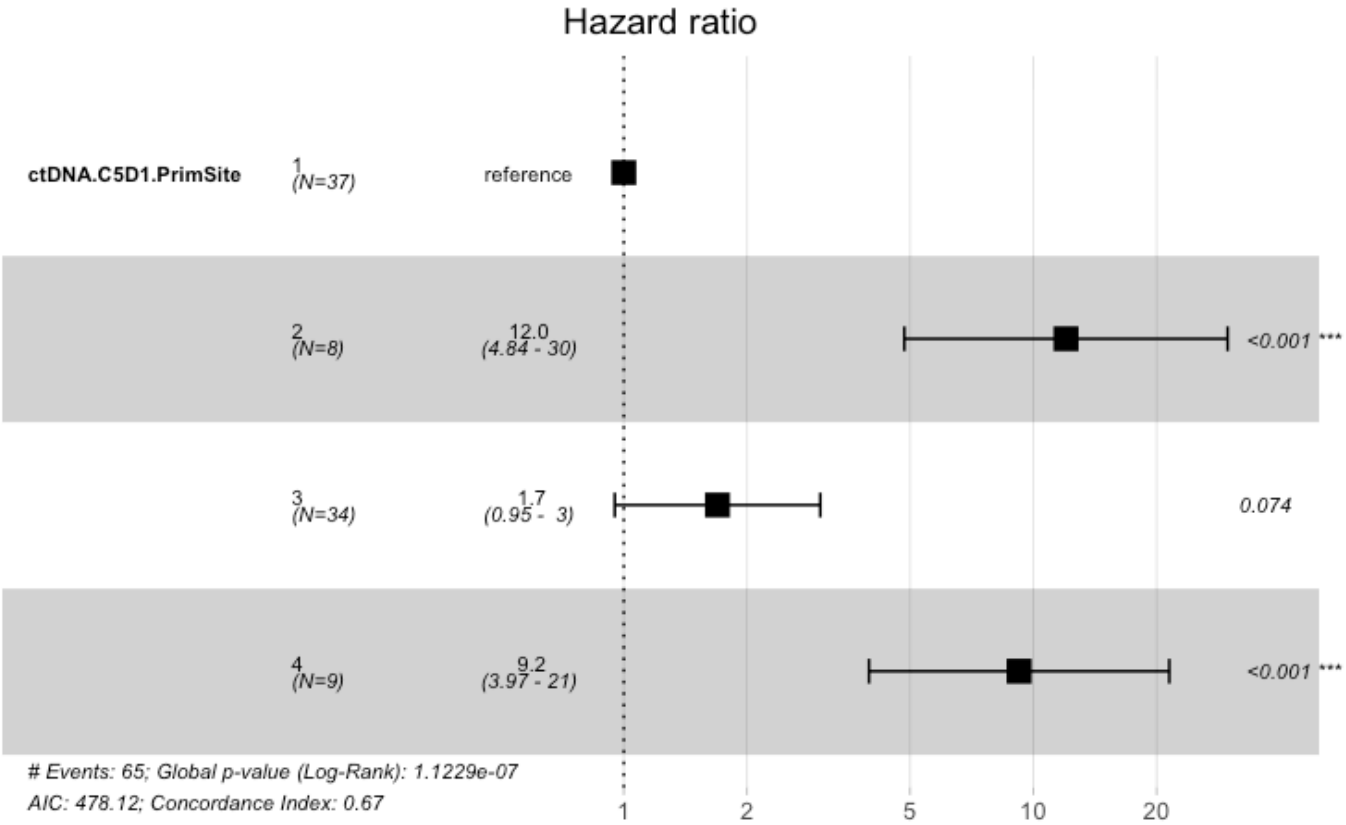
ctDNA.C5D1.PrimSite=2							
CI	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
1	0	8	0	1	0	1	

ctDNA.C5D1.PrimSite=3							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	34	0	1.000	0.0000	1.000	1.000	
12	21	13	0.618	0.0833	0.434	0.757	
24	13	8	0.382	0.0833	0.223	0.540	

ctDNA.C5D1.PrimSite=4							
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI	
0	9	0	1.000	0.000	1.00000	1.000	
12	1	8	0.111	0.105	0.00613	0.388	

Hide

```
circ_data$ctDNA.C5D1.PrimSite <- factor(circ_data$ctDNA.C5D1.PrimSite, levels=c("1","2","3","4"))
cox_fit <- coxph(surv_object ~ ctDNA.C5D1.PrimSite, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:
coxph(formula = surv_object ~ ctDNA.C5D1.PrimSite, data = circ_data)

n= 88, number of events= 65
(1 observation deleted due to missingness)

	coef	exp(coef)	se(coef)	z	Pr(> z)	
ctDNA.C5D1.PrimSite2	2.4864	12.0174	0.4637	5.362	8.24e-08	***
ctDNA.C5D1.PrimSite3	0.5269	1.6937	0.2948	1.787	0.0739	.
ctDNA.C5D1.PrimSite4	2.2230	9.2352	0.4307	5.162	2.45e-07	***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C5D1.PrimSite2	12.017	0.08321	4.8428	29.821
ctDNA.C5D1.PrimSite3	1.694	0.59042	0.9503	3.019
ctDNA.C5D1.PrimSite4	9.235	0.10828	3.9706	21.480

Concordance= 0.665 (se = 0.035)
Likelihood ratio test= 35.17 on 3 df, p=1e-07
Wald test = 41.85 on 3 df, p=4e-09
Score (logrank) test = 56.32 on 3 df, p=4e-12

Hide

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

#Prognostic role of ctDNA C8D1 on Primary Site - 4 groups

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("ASAN_ClinicalData_GL_082023.csv")
circ_data <- subset(circ_data, !is.na(ctDNA.C8D1))
circ_data$RFS.months=circ_data$RFS.months-2
circ_data <- circ_data[circ_data$RFS.months>=0,]
circ_data$RFS.months[circ_data$RFS.months > 60] <- 60

circ_data$ctDNA.C8D1.PrimSite <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C8D1.PrimSite = case_when(
    PrimSite == "dCCA" & ctDNA.C8D1 == "NEGATIVE" ~ 1,
    PrimSite == "dCCA" & ctDNA.C8D1 == "POSITIVE" ~ 2,
    PrimSite == "pCCA" & ctDNA.C8D1 == "NEGATIVE" ~ 3,
    PrimSite == "pCCA" & ctDNA.C8D1 == "POSITIVE" ~ 4
  ))
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)~ctDNA.C8D1.PrimSi
te, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event) ~
ctDNA.C8D1.PrimSite, data = circ_data)
```

12 observations deleted due to missingness

	n	events	median	0.95LCL	0.95UCL
ctDNA.C8D1.PrimSite=1	31	15	38.71	13.51	NA
ctDNA.C8D1.PrimSite=2	8	8	7.50	6.44	NA
ctDNA.C8D1.PrimSite=3	31	25	15.84	10.98	27.6
ctDNA.C8D1.PrimSite=4	7	7	5.43	4.60	NA

Hide

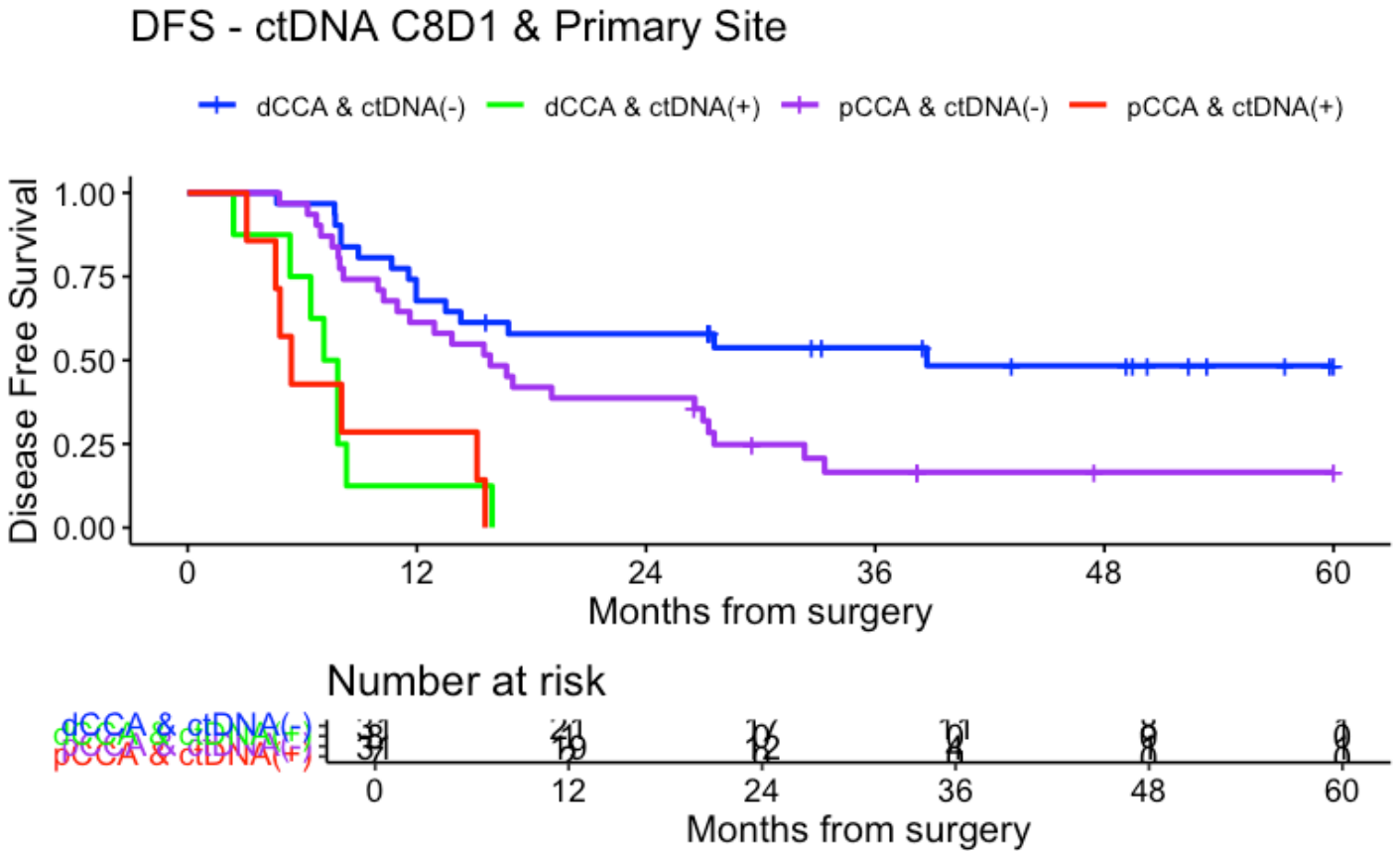

```
event_summary <- circ_data %>%
  group_by(ctDNA.C8D1.PrimSite) %>%
  summarise(
    Total = n(),
    Events = sum(RFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.C8D1.PrimSite	Total	Events	Fraction	Percentage
<dbl>	<int>	<int>	<dbl>	<dbl>
1	31	15	0.4838710	48.38710
2	8	8	1.0000000	100.00000
3	31	25	0.8064516	80.64516
4	7	7	1.0000000	100.00000
NA	12	10	0.8333333	83.33333

5 rows

Hide

```
surv_object <-Surv(time = circ_data$RFS.months, event = circ_data$RFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C8D1.PrimSite, 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=12, palette=c("blue","green","purple", "red"), title="DFS - ctDNA C8D1 & Primary Site", ylab= "Disease Free Survival", xlab="Months from surgery", legend.labs=c("dCCA & ctDNA(-)", "dCCA & ctDNA(+)","pCCA & ctDNA(-)", "pCCA & ctDNA(+)"), legend.title="")
```



Hide

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

```
Call: survfit(formula = surv_object ~ ctDNA.C8D1.PrimSite, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

12 observations deleted due to missingness

ctDNA.C8D1.PrimSite=1

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	31	0	1.000	0.000	1.000	1.000
12	21	10	0.677	0.084	0.484	0.812
24	17	3	0.579	0.089	0.387	0.730

ctDNA.C8D1.PrimSite=2

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	8	0	1.000	0.000	1.00000	1.000
12	1	7	0.125	0.117	0.00659	0.423

ctDNA.C8D1.PrimSite=3

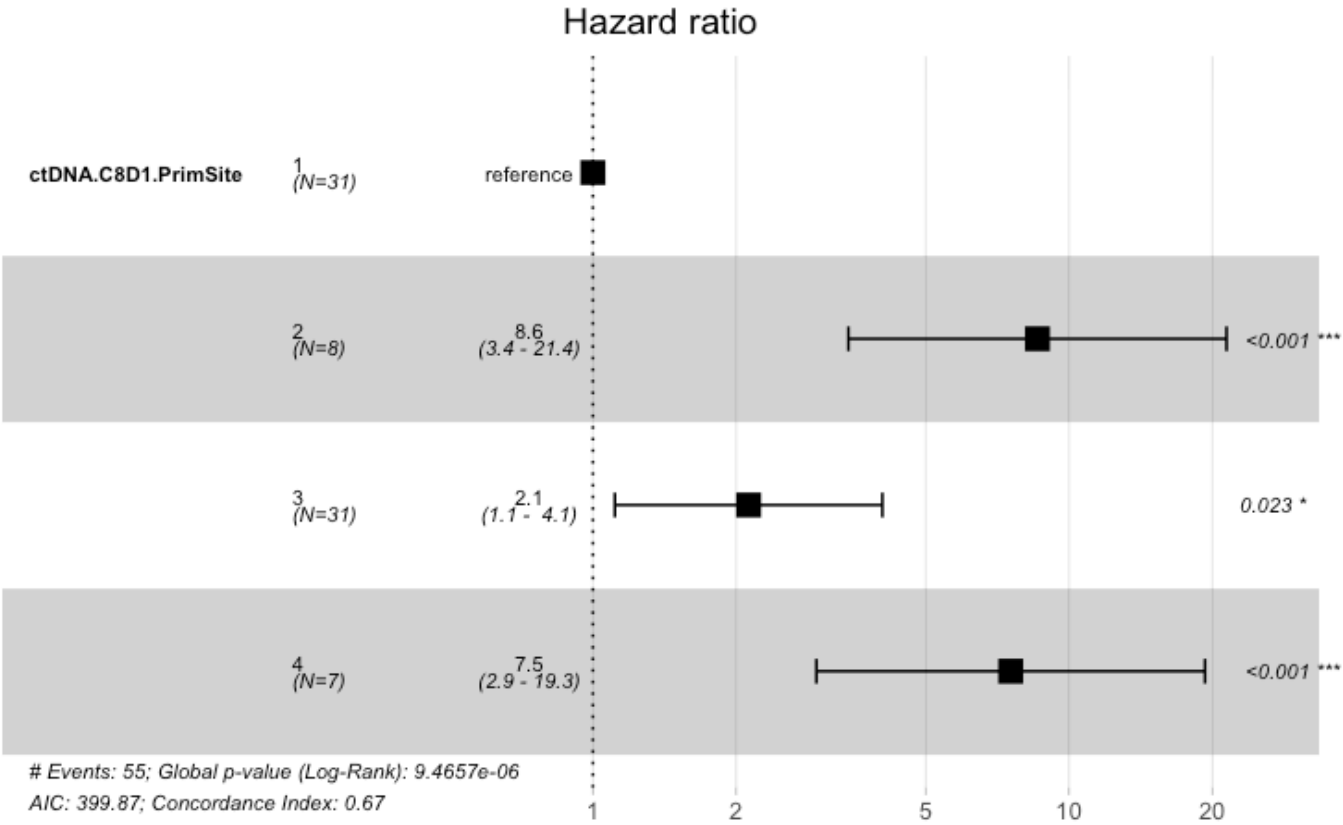
time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	31	0	1.000	0.0000	1.00	1.000
12	19	12	0.613	0.0875	0.42	0.758
24	12	7	0.387	0.0875	0.22	0.551

ctDNA.C8D1.PrimSite=4

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
0	7	0	1.000	0.000	1.0000	1.000
12	2	5	0.286	0.171	0.0411	0.612

[Hide](#)

```
circ_data$ctDNA.C8D1.PrimSite <- factor(circ_data$ctDNA.C8D1.PrimSite, levels=c
("1","2","3","4"))
cox_fit <- coxph(surv_object ~ ctDNA.C8D1.PrimSite, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ ctDNA.C8D1.PrimSite, data = circ_data)

n= 77, number of events= 55
(12 observations deleted due to missingness)

              coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.C8D1.PrimSite2 2.1501    8.5860  0.4665 4.609 4.05e-06 ***
ctDNA.C8D1.PrimSite3 0.7532    2.1239  0.3301 2.282  0.0225  *
ctDNA.C8D1.PrimSite4 2.0206    7.5431  0.4794 4.215 2.50e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
ctDNA.C8D1.PrimSite2    8.586    0.1165    3.441    21.424
ctDNA.C8D1.PrimSite3    2.124    0.4708    1.112    4.056
ctDNA.C8D1.PrimSite4    7.543    0.1326    2.948    19.303

Concordance= 0.671 (se = 0.036 )
Likelihood ratio test= 26.02  on 3 df,   p=9e-06
Wald test              = 28.78  on 3 df,   p=2e-06
Score (logrank) test = 35.81  on 3 df,   p=8e-08
```

[Hide](#)

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