

Code ▾

AGEO PLAGAST_Clinical Final analysis 08172024

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

#ctDNA Detection rate by Stage and Window

Hide

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

Stage	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	16	6	37.50%

Stage	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
II	21	17	80.95%
III	16	14	87.50%
Overall	56	39	69.64%
4 rows			

Hide

```
#C2D1
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.C2D1 <- factor(circ_data$ctDNA.C2D1, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.C2D1 %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.C2D1 == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.C2D1, 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.C2D1 == "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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	13	3	23.08%
II	13	8	61.54%

Stage	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
III	13	9	69.23%
Overall	41	21	51.22%
4 rows			

Hide

```
#post-NAC Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.postNAC <- factor(circ_data$ctDNA.postNAC, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.postNAC %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.postNAC == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.postNAC, 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.postNAC == "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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	12	0	0.00%
II	13	3	23.08%

Stage	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
III	13	8	61.54%
Overall	41	11	26.83%
4 rows			

Hide

```
#MRD Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.MRD %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	14	0	0.00%
II	20	4	20.00%
III	14	6	42.86%

Stage	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
Overall	50	10	20.00%
4 rows			

Hide

```

#On-treatment
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.onACT <- factor(circ_data$ctDNA.onACT, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.onACT %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.onACT == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.onACT, 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.onACT == "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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	2	0	0.00%
II	2	0	0.00%
III	3	0	0.00%
Overall	7	0	0.00%

4 rows

Hide

```

#Surveillance Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.surveillance <- factor(circ_data$ctDNA.surveillance, levels=c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I", "II", "III"))
circ_data <- subset(circ_data, ctDNA.surveillance %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.surveillance == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.surveillance, 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.surveillance == "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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
I	10	0	0.00%
II	15	2	13.33%
III	10	3	30.00%
Overall	38	5	13.16%

4 rows

Hide

```

#Post-ACT Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.postACT <- factor(circ_data$ctDNA.postACT, levels=c("NEGATIVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.postACT %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.postACT == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.postACT, 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.postACT == "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	Total_Count	Positive_Count	Rate
		<int>	<chr>
I	10	0	0.00%
II	15	1	6.67%
III	7	1	14.29%
Overall	35	2	5.71%
4 rows			

Hide

```

#Post-relapse Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$ctDNA.postRelapse <- factor(circ_data$ctDNA.postRelapse, levels=c("NEGATIVE", "POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I", "II", "III"))
circ_data <- subset(circ_data, ctDNA.postRelapse %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.postRelapse == "POSITIVE", by=list(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.postRelapse, 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.postRelapse == "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	Total_Count	Positive_Count	Rate
<fctr>	<int>	<int>	<chr>
II	5	5	100.00%
III	3	3	100.00%
Overall	8	8	100.00%
3 rows			

#Demographics Table

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data_subset <- circ_data %>%
  select(
    Age,
    Gender,
    PrimSite,
    NAC,
    NAC.Regimen,
    TRG.Mandard,
    TNM,
    Stage,
    Grade,
    Lauren.Class,
    Margins,
    ACT,
    ACT.Regimen,
    DFS.Event,
    OS.Event,
    FU.months) %>%
  mutate(
    Age = as.numeric(Age),
    Gender = factor(Gender, levels = c("Male", "Female")),
    PrimSite = factor(PrimSite, levels = c("Stomach", "G/J", "Oesophagus")),
    NAC = factor(NAC, levels = c("TRUE", "FALSE"), labels = c("Neoadjuvant Therapy", "Up front Surgery")),
    NAC.Regimen = factor(NAC.Regimen),
    TRG.Mandard = factor(TRG.Mandard, levels = c("TRG1", "TRG2", "TRG3", "TRG4", "TRG5")),
    TNM = factor(TNM, levels = c("T0-TisN0M0", "T1-T2N0", "T2-T3N0-N1", "T2N1-N2", "T3N2-N3", "T4N0-N1", "T4N2-N3")),
    Stage = factor(Stage, levels = c("0", "I", "II", "III")),
    Grade = factor(Grade, levels = c("G1", "G2", "G3")),
    Lauren.Class = factor(Lauren.Class),
    Margins = factor(Margins, levels = c("R0", "R1")),
    ACT = factor(ACT, levels = c("TRUE", "FALSE"), labels = c("Adjuvant Treatment", "Observation")),
    ACT.Regimen = factor(ACT.Regimen),
    DFS.Event = factor(DFS.Event, levels = c("TRUE", "FALSE"), labels = c("Recurrence", "No Recurrence")),
    OS.Event = factor(OS.Event, levels = c("TRUE", "FALSE"), labels = c("Deceased", "Alive")),
    FU.months = as.numeric(FU.months))
table1 <- circ_data_subset %>%
 tbl_summary(
  statistic = list(
    all_continuous() ~ "{median} ({min} - {max})",
    all_categorical() ~ "{n} ({p}%)") %>%

```

```
bold_labels()
table1
```

Characteristic	N = 62¹
Age	66 (34 - 86)
Gender	
Male	39 (63%)
Female	23 (37%)
PrimSite	
Stomach	38 (61%)
G/J	24 (39%)
Oesophagus	0 (0%)
NAC	
Neoadjuvant Therapy	55 (89%)
Upfront Surgery	7 (11%)
NAC.Regimen	
	7 (11%)
Chemoimmunotherapy	7 (11%)
Chemotherapy	47 (76%)
Radiotherapy	1 (1.6%)
TRG.Mandard	
TRG1	3 (5.5%)
TRG2	8 (15%)
TRG3	21 (38%)
TRG4	16 (29%)
TRG5	7 (13%)
Unknown	7
TNM	
¹ Median (Range); n (%)	

Characteristic	N = 62¹
T0-TisN0M0	3 (4.8%)
T1-T2N0	16 (26%)
T2-T3N0-N1	17 (27%)
T2N1-N2	2 (3.2%)
T3N2-N3	8 (13%)
T4N0-N1	9 (15%)
T4N2-N3	7 (11%)
Stage	
0	3 (4.8%)
I	16 (26%)
II	26 (42%)
III	17 (27%)
Grade	
G1	11 (28%)
G2	21 (53%)
G3	8 (20%)
Unknown	22
Lauren.Class	
	5 (8.1%)
Diffuse	18 (29%)
Intestinal	34 (55%)
Mixed	5 (8.1%)
Margins	
R0	61 (98%)
R1	1 (1.6%)
ACT	
¹ Median (Range); n (%)	

Characteristic	N = 62 ¹
Adjuvant Treatment	53 (85%)
Observation	9 (15%)
ACT.Regimen	
	9 (15%)
Chemoimmunotherapy	4 (6.5%)
Chemotherapy	48 (77%)
Immunotherapy	1 (1.6%)
DFS.Event	
Recurrence	29 (47%)
No Recurrence	33 (53%)
OS.Event	
Deceased	19 (31%)
Alive	43 (69%)
FU.months	
	29 (2 - 93)

¹ Median (Range); n (%)

Hide

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

Warning: The `strip_md_bold` argument of `as_flex_table()` is deprecated as of gtsummary 1.6.0.

This warning is displayed once every 8 hours.

Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.

Hide

```
fit1
```

Characteristic	N = 62 ¹
Age	66 (34 - 86)

¹Median (Range); n (%)

Characteristic	N = 62¹
Gender	
Male	39 (63%)
Female	23 (37%)
PrimSite	
Stomach	38 (61%)
G/J	24 (39%)
Oesophagus	0 (0%)
NAC	
Neoadjuvant Therapy	55 (89%)
Upfront Surgery	7 (11%)
NAC.Regimen	
	7 (11%)
Chemoimmunotherapy	7 (11%)
Chemotherapy	47 (76%)
Radiotherapy	1 (1.6%)
TRG.Mandard	
TRG1	3 (5.5%)
TRG2	8 (15%)
TRG3	21 (38%)
TRG4	16 (29%)
TRG5	7 (13%)
Unknown	7
TNM	
T0-TisN0M0	3 (4.8%)
T1-T2N0	16 (26%)
T2-T3N0-N1	17 (27%)
T2N1-N2	2 (3.2%)
T3N2-N3	8 (13%)
T4N0-N1	9 (15%)
T4N2-N3	7 (11%)

¹Median (Range); n (%)

Characteristic	N = 62¹
Stage	
0	3 (4.8%)
I	16 (26%)
II	26 (42%)
III	17 (27%)
Grade	
G1	11 (28%)
G2	21 (53%)
G3	8 (20%)
Unknown	22
Lauren.Class	
	5 (8.1%)
Diffuse	18 (29%)
Intestinal	34 (55%)
Mixed	5 (8.1%)
Margins	
R0	61 (98%)
R1	1 (1.6%)
ACT	
Adjuvant Treatment	53 (85%)
Observation	9 (15%)
ACT.Regimen	
	9 (15%)
Chemoimmunotherapy	4 (6.5%)
Chemotherapy	48 (77%)
Immunotherapy	1 (1.6%)
DFS.Event	
Recurrence	29 (47%)
No Recurrence	33 (53%)
OS.Event	

¹Median (Range); n (%)

Characteristic	N = 62¹
Deceased	19 (31%)
Alive	43 (69%)
FU.months	29 (2 - 93)

¹Median (Range); n (%)

[Hide](#)

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

#Heatmap with Clinical & Genomics Factors

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE, ]
circ_data <- circ_data %>% arrange(Stage)
circ_datadf <- as.data.frame(circ_data)

ha <- HeatmapAnnotation(
  Stage = circ_data$Stage,
  Gender = circ_data$Gender,
  PrimSite = circ_data$PrimSite,
  NAC = circ_data$NAC,
  ACT = circ_data$ACT,
  ctDNA.Base = circ_data$ctDNA.Base,
  ctDNA.C2D1 = circ_data$ctDNA.C2D1,
  ctDNA.postNAC = circ_data$ctDNA.postNAC,
  ctDNA.MRD = circ_data$ctDNA.MRD,
  ctDNA.surveillance = circ_data$ctDNA.surveillance,
  DFS.Event = circ_data$DFS.Event,
  OS.Event = circ_data$OS.Event,

  col = list(Stage = c("0" = "seagreen1", "I" = "seagreen1", "II" = "orange", "III" = "purple"),
             Gender = c("Female" = "goldenrod" , "Male" = "blue4"),
             PrimSite = c("Stomach" = "brown", "G/J" = "darkgreen", "Oesophagus" = "orange4"),
             NAC = c("FALSE" = "cornflowerblue", "TRUE" = "darkmagenta"),
             ACT = c("TRUE" = "brown4", "FALSE" = "khaki"),
             ctDNA.Base = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
             ctDNA.C2D1 = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
             ctDNA.postNAC = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
             ctDNA.MRD = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
             ctDNA.surveillance = c("POSITIVE" = "red3", "NEGATIVE" = "blue"),
             DFS.Event = c("TRUE" = "red3", "FALSE" = "blue"),
             OS.Event = c("TRUE" = "black", "FALSE" = "gray"))
)
)

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()

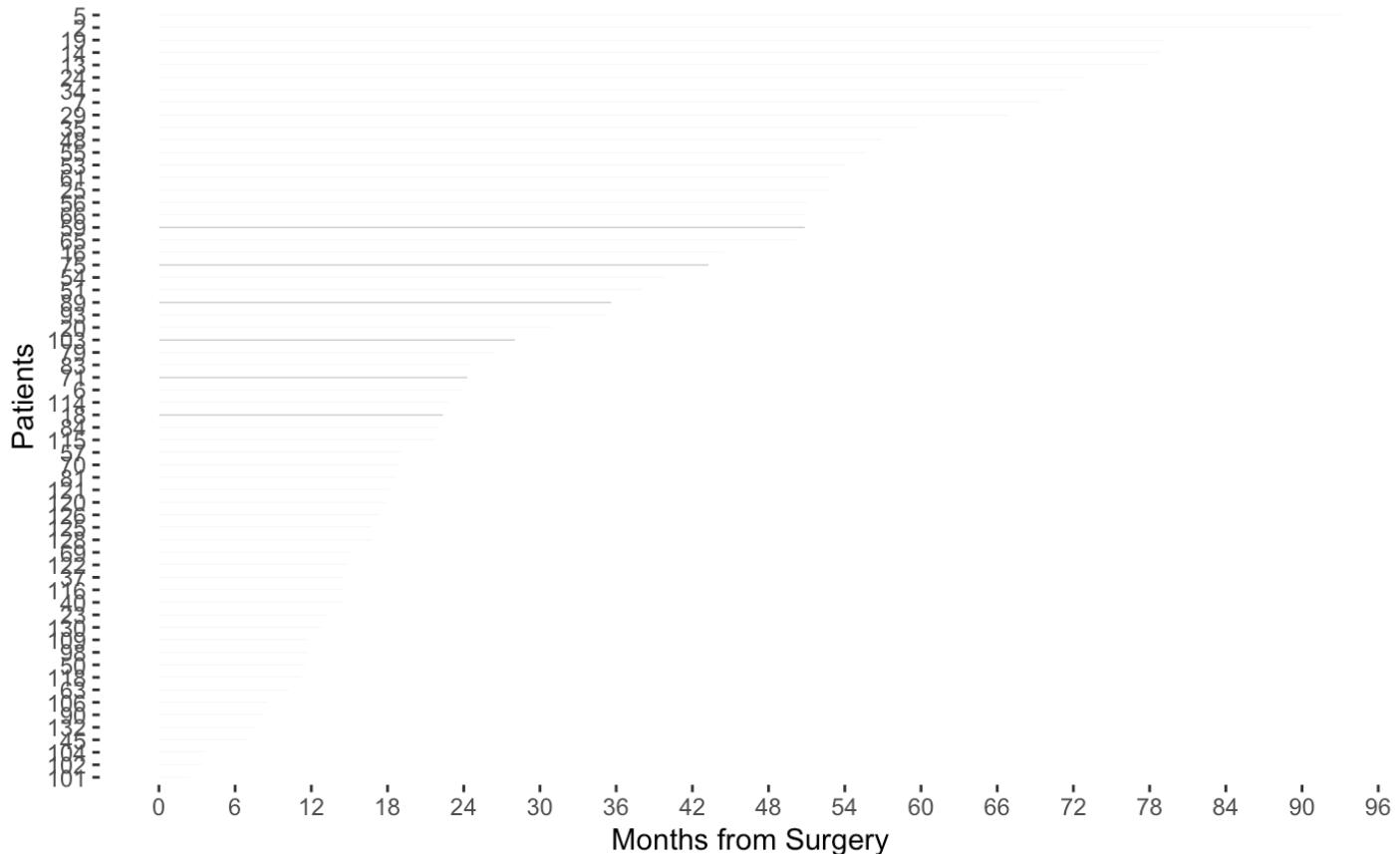
```

#Overview Plot

```
setwd("~/Downloads")
clinstage<- read.csv("PLAGAST_OP.csv")
clinstage_df<- as.data.frame(clinstage)

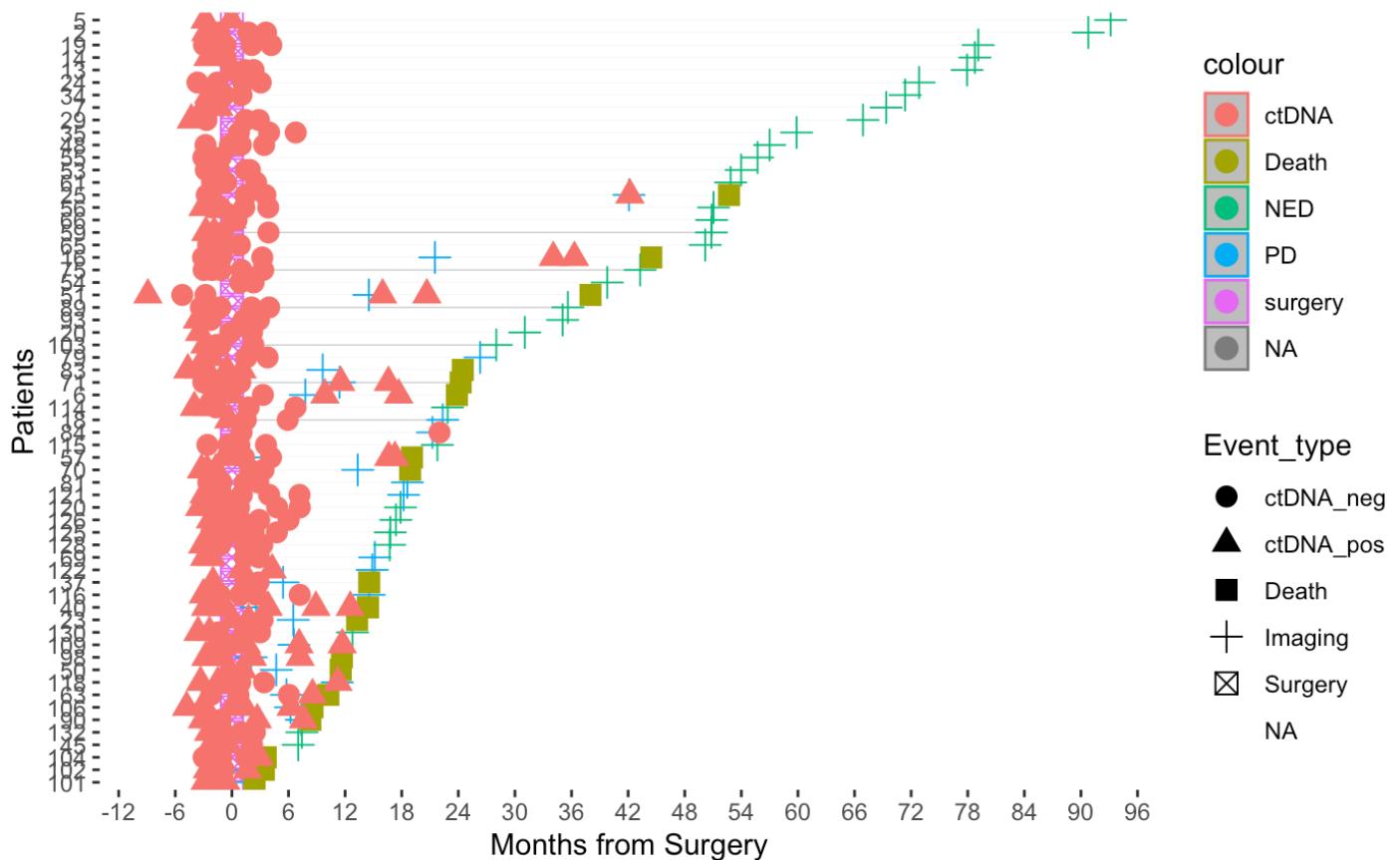
#Display the swimmer plot with the label box
oplot<-swimmer_plot(df=clinstage_df,
                      id='PatientName',
                      end='fu.diff.months',
                      fill='gray',
                      width=.01,)

oplot <- oplot + theme(panel.border = element_blank())
oplot <- oplot + scale_y_continuous(breaks = seq(-12, 96, by = 6))
oplot <- oplot + labs(x ="Patients" , y="Months from Surgery")
oplot
```

[Hide](#)

```
##plot events
oplot_ev1 <- oplot + swimmer_points(df_points=clinstage_df,
                                         id='PatientName',
                                         time='date.diff.months',
                                         name_shape ='Event_type',
                                         name_col = 'Event',
                                         size=3.5,fill='black',
                                         #col='darkgreen'
)
oplot_ev1
```

Warning: Removed 114 rows containing missing values (`geom_point()`).



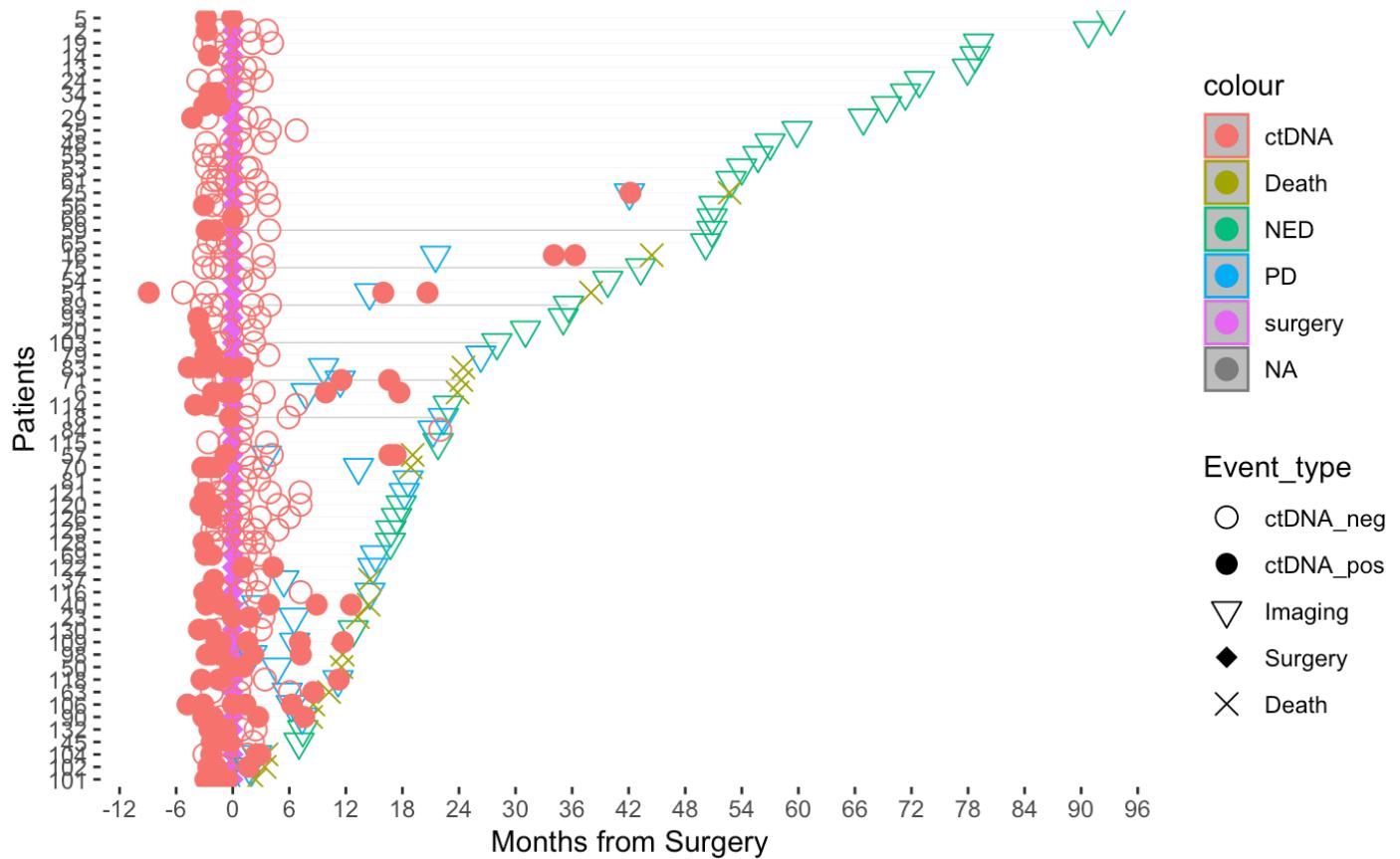
Hide

```
#Shape customization to Event_type
```

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

oplot_ev1.1
```

Warning: Removed 114 rows containing missing values (`geom_point()`).

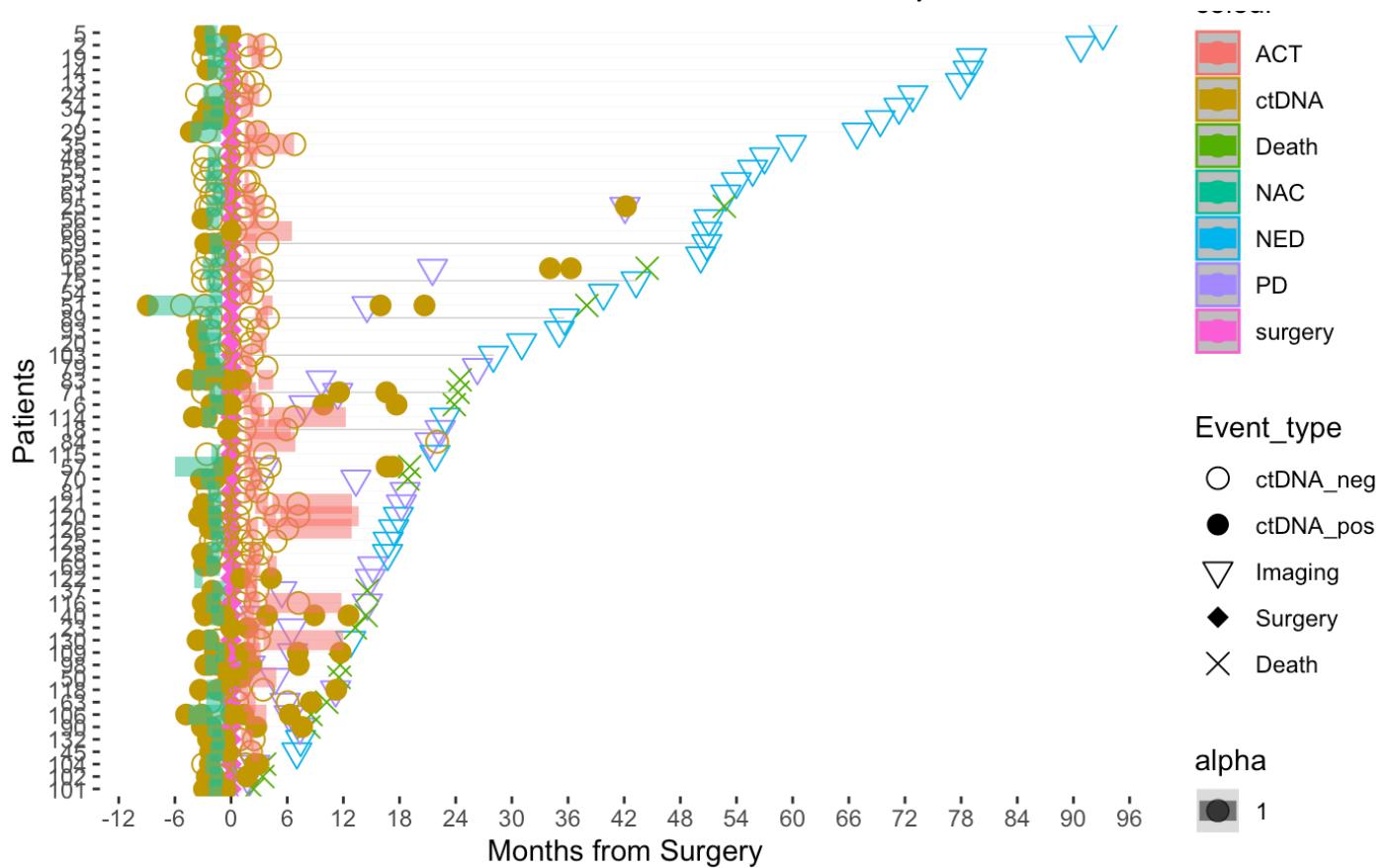


Hide

```
#plot treatment
```

```
oplot_ev2 <- oplot_ev1.1 + 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_ev2 <- oplot_ev2 + guides(linetype = guide_legend(override.aes = list(size = 5, color = "black")))
oplot_ev2
```

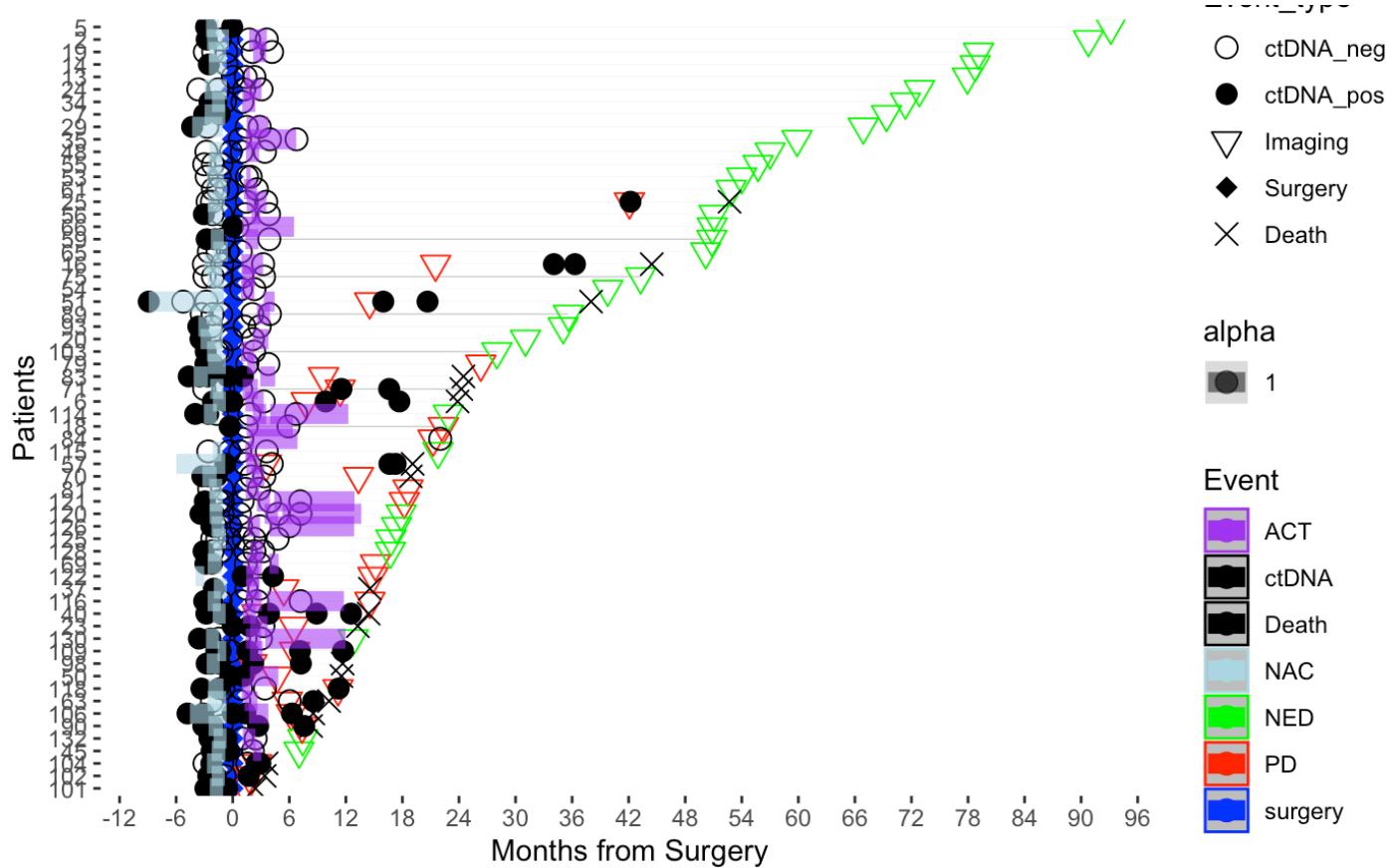
Warning: Removed 114 rows containing missing values (`geom_point()`).
 Warning: Removed 405 rows containing missing values (`geom_segment()`).



Hide

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

Warning: Removed 114 rows containing missing values (`geom_point()`).
 Warning: Removed 405 rows containing missing values (`geom_segment()`).



#RFS by ctDNA at Baseline - all stages

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_datadf <- as.data.frame(circ_data)

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

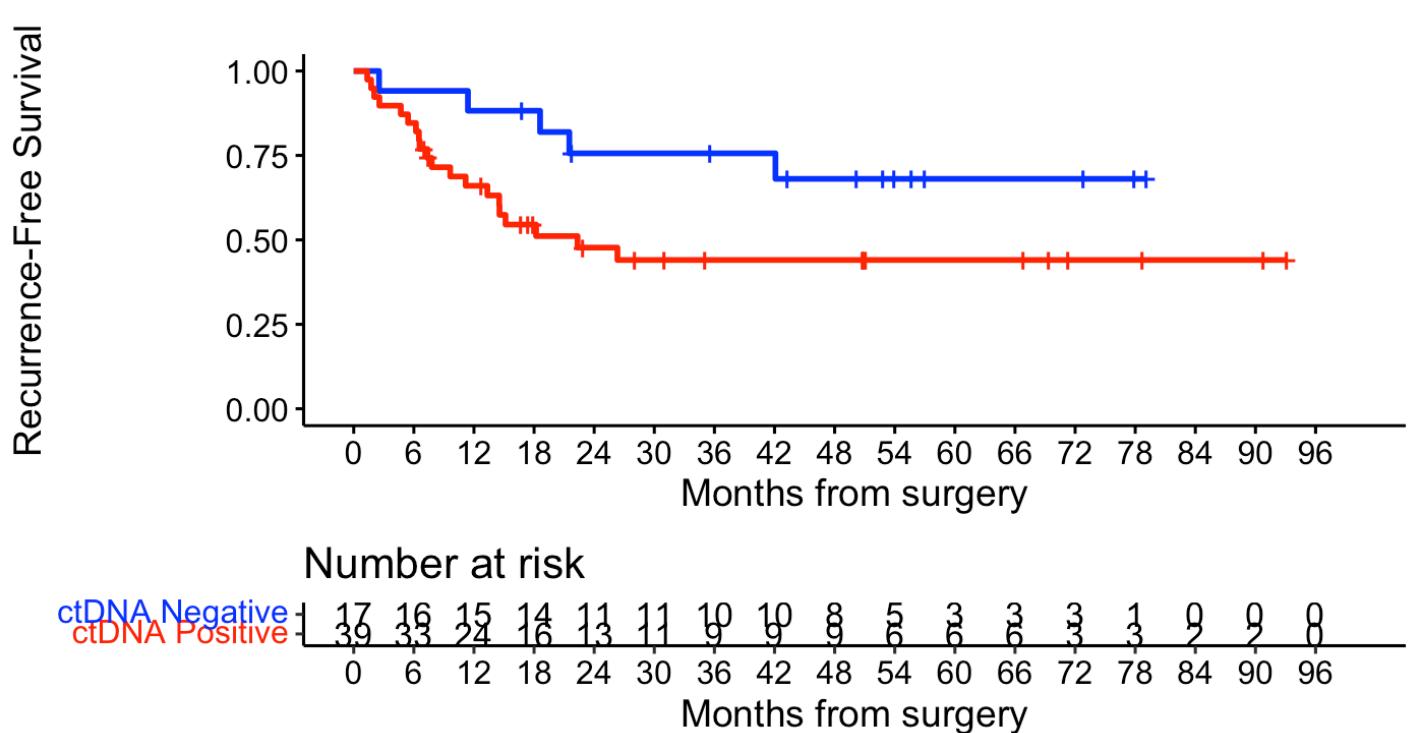
Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.Base, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.Base=NEGATIVE	17	5	NA	42.1	NA
ctDNA.Base=POSITIVE	39	20	22.3	13.3	NA

Hide

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Base, data = circ_data, conf.int=0.95, conf.type
="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE,
break.time.by=6, palette=c("blue","red"), title="RFS - ctDNA Baseline | All pts", yla
b= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negativ
e", "ctDNA Positive"), legend.title="")
```

RFS - ctDNA Baseline | All pts



Hide

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

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

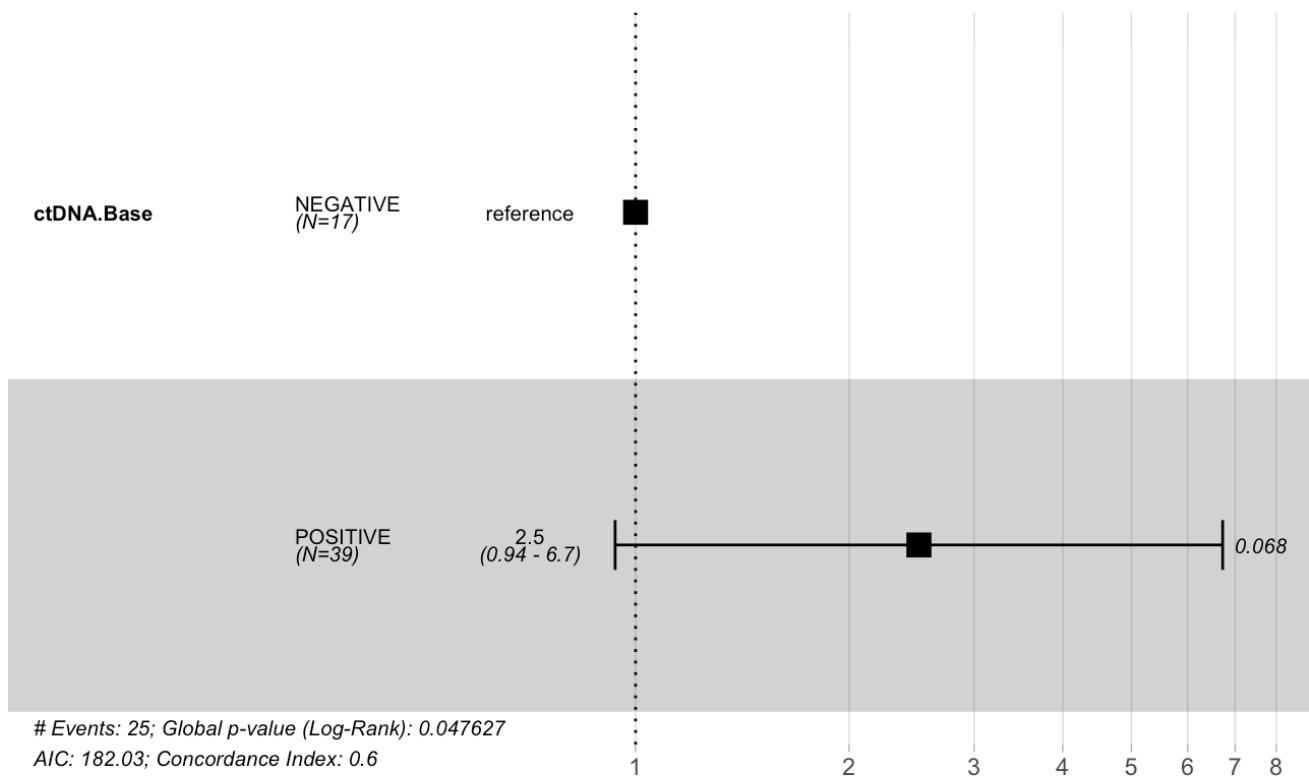
CI	ctDNA.Base=NEGATIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
01	24.000	11.000	4.000	0.756	0.106	0.473	0.9

CI	ctDNA.Base=POSITIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
01	24.0000	13.0000	19.0000	0.4772	0.0849	0.3057	0.63

Hide

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

Hazard ratio



Hide

```
summary(cox_fit)
```

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

n= 56, number of events= 25

	coef	exp(coef)	se(coef)	z	Pr(> z)		
ctDNA.Base	POSITIVE	0.9196	2.5082	0.5030	1.828	0.0675	.

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

	exp(coef)	exp(-coef)	lower .95	upper .95	
ctDNA.Base	POSITIVE	2.508	0.3987	0.9359	6.722

Concordance= 0.604 (se = 0.041)
 Likelihood ratio test= 3.92 on 1 df, p=0.05
 Wald test = 3.34 on 1 df, p=0.07
 Score (logrank) test = 3.58 on 1 df, p=0.06

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 = 2.51 (0.94-6.72); p = 0.068"
```

#OS by ctDNA at Baseline - all stages

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_datadf <- as.data.frame(circ_data)

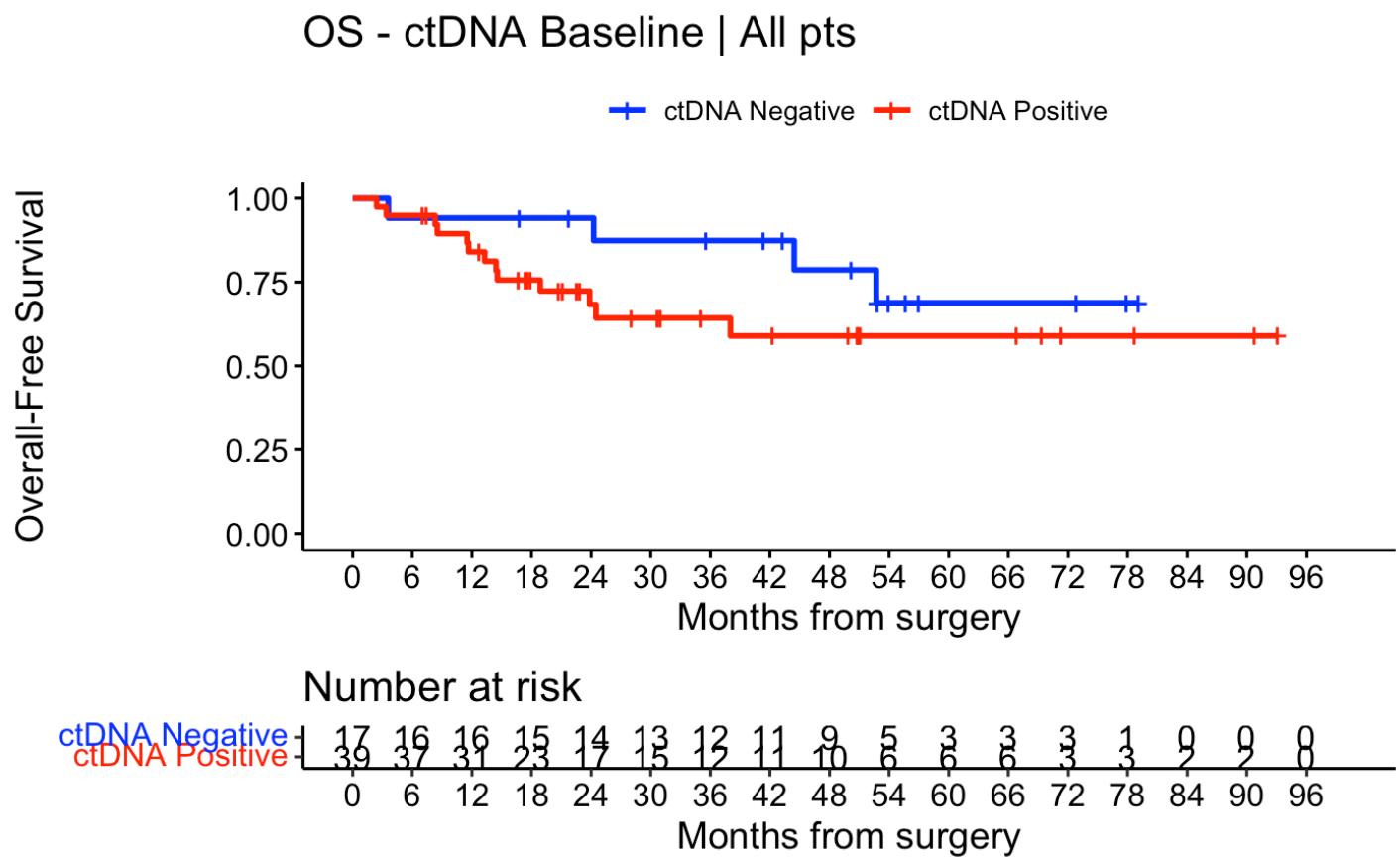
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Base, data = circ_data)
```

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.Base=NEGATIVE	17	4	NA	52.7	NA
ctDNA.Base=POSITIVE	39	13	NA	24.5	NA

[Hide](#)

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



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

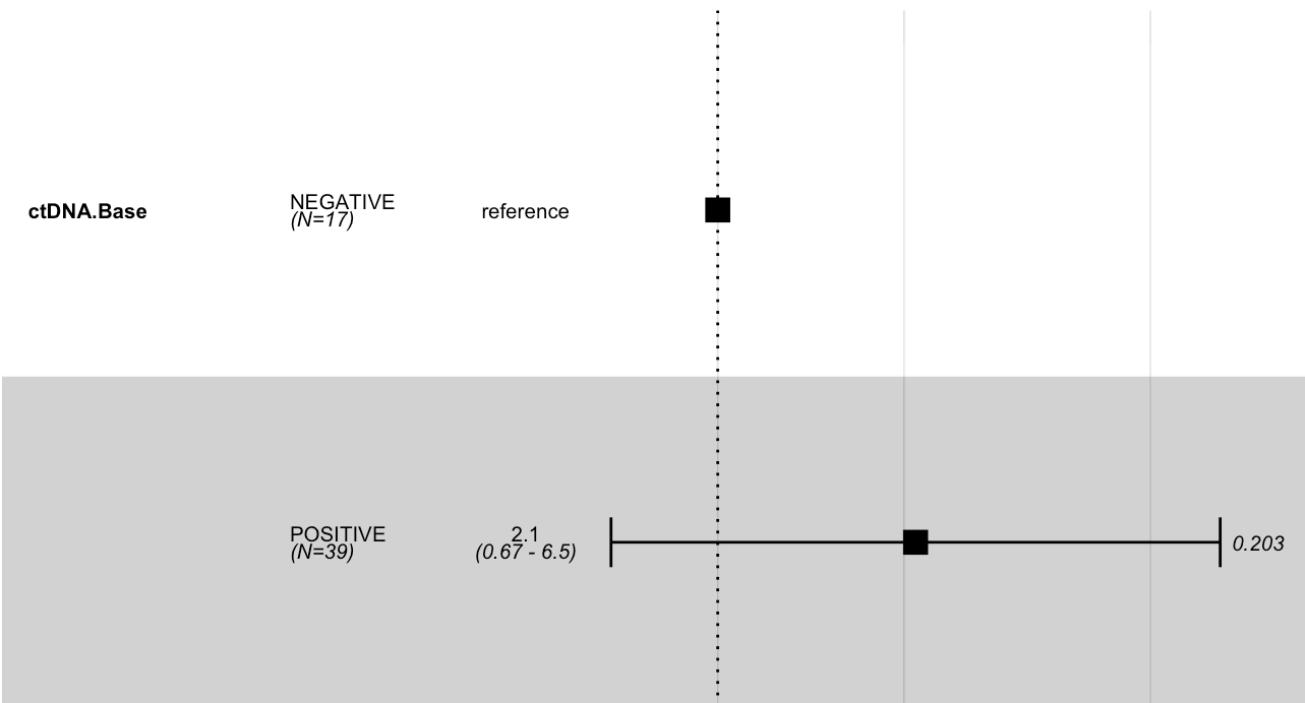
```
Call: survfit(formula = surv_object ~ ctDNA.Base, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.Base=NEGATIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
15	24.0000	14.0000	1.0000	0.9412	0.0571	0.6502	0.6502	0.99	0.99
ctDNA.Base=POSITIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
29	24.0000	17.0000	11.0000	0.6832	0.0809	0.4963	0.4963	0.81	0.81

[Hide](#)

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

Hazard ratio



Events: 17; Global p-value (Log-Rank): 0.17835

AIC: 123.79; Concordance Index: 0.6

[Hide](#)

```
summary(cox_fit)
```

Call:

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

n= 56, number of events= 17

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.BasePOSITIVE	0.7361	2.0878	0.5781	1.273	0.203

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.BasePOSITIVE	2.088	0.479	0.6724	6.483

Concordance= 0.597 (se = 0.05)

Likelihood ratio test= 1.81 on 1 df, p=0.2

Wald test = 1.62 on 1 df, p=0.2

Score (logrank) test = 1.69 on 1 df, p=0.2

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 = 2.09 (0.67-6.48); p = 0.203"

#RFS by ctDNA levels at Baseline based on AUC optimal MTM/ml level - ctDNA Baseline positive

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_data <- circ_data[circ_data$ctDNA.Base=="POSITIVE",]
circ_datadf <- as.data.frame(circ_data)

#DFS.Event
circ_data <- circ_data[complete.cases(circ_data$DFS.Event, circ_data$ctDNA.Base.MTM),]
circ_data$ctDNA.Base.MTM <- as.numeric(circ_data$ctDNA.Base.MTM)
ROC <- roc(DFS.Event ~ ctDNA.Base.MTM, data = circ_data, ci = TRUE)
```

Setting levels: control = FALSE, case = TRUE

Setting direction: controls < cases

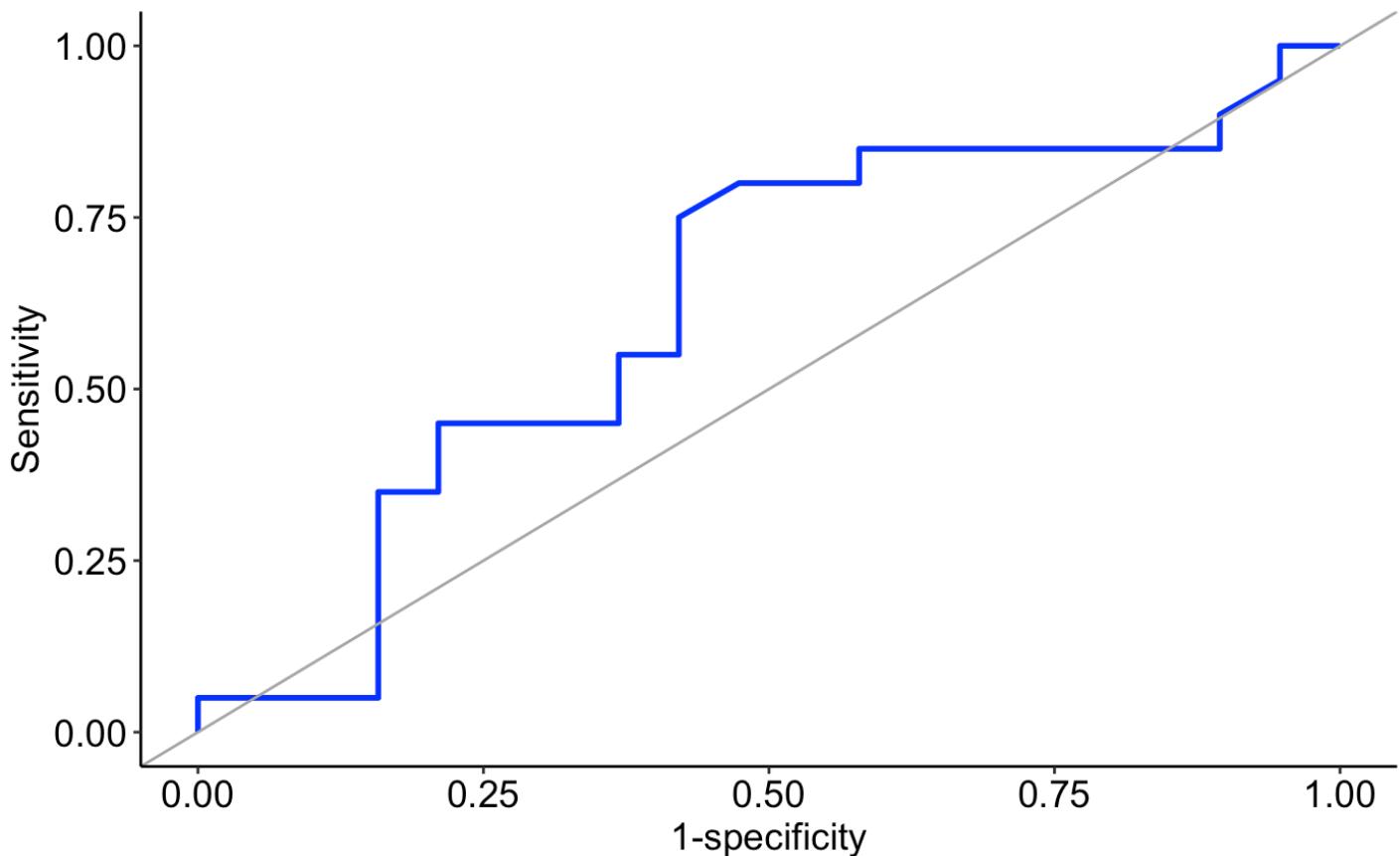
Hide

```
p<-ggroc(ROC,
  aes = c("linetype"), color = "blue", size = 1,
  legacy.axes = TRUE) +
  geom_abline(color = "dark grey", size = 0.5) +
  theme_classic()+
  ylab("Sensitivity") + theme(axis.title.x = element_text(color="black", size=14), axis.title.y = element_text(color="black", size=14), axis.text.x = element_text(colour = "black", size=14), axis.text.y = element_text(colour = "black", size=14), legend.title = element_text(), legend.text = element_text(size=14))
```

Warning in ggplot2::geom_line(aes\$aes, ...) :
Ignoring unknown parameters: `aes`

Hide

p



Hide

```
#AUC
AUC <- auc(ROC)
print(AUC)
```

Area under the curve: 0.6211

Hide

```
AUC_conf <- ci.auc(ROC)
print(AUC_conf)
```

95% CI: 0.4354–0.8067 (DeLong)

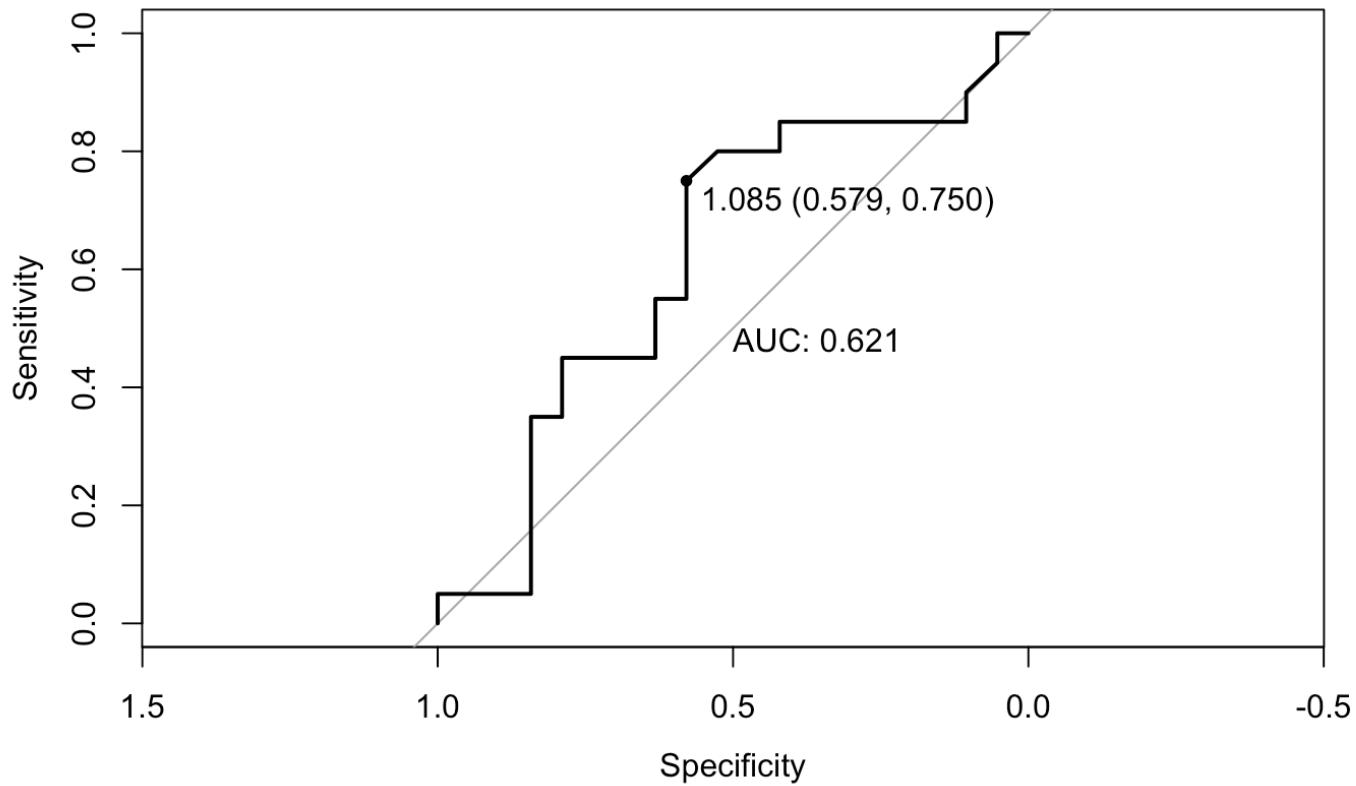
Hide

```
res.roc <- roc(circ_data$DFS.Event, circ_data$ctDNA.Base.MTM)
```

```
Setting levels: control = FALSE, case = TRUE
Setting direction: controls < cases
```

Hide

```
plot.roc(res.roc, print.auc = TRUE, print.thres = "best")
```



Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_data <- circ_data[circ_data$ctDNA.Base=="POSITIVE",]
circ_datadf <- as.data.frame(circ_data)

# Create a new variable based on these quartiles

circ_data$ctDNA.6mMTM.Q <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.6mMTM.Q = case_when(
    ctDNA.Base.MTM < 1.085 ~ 1,
    ctDNA.Base.MTM >= 1.085 ~ 2
  ))

survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.6mMTM.Q, data = circ_data)

```

Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.6mMTM.Q, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.6mMTM.Q=1	16	5	NA	26.32	NA
ctDNA.6mMTM.Q=2	23	15	14.5	7.79	NA

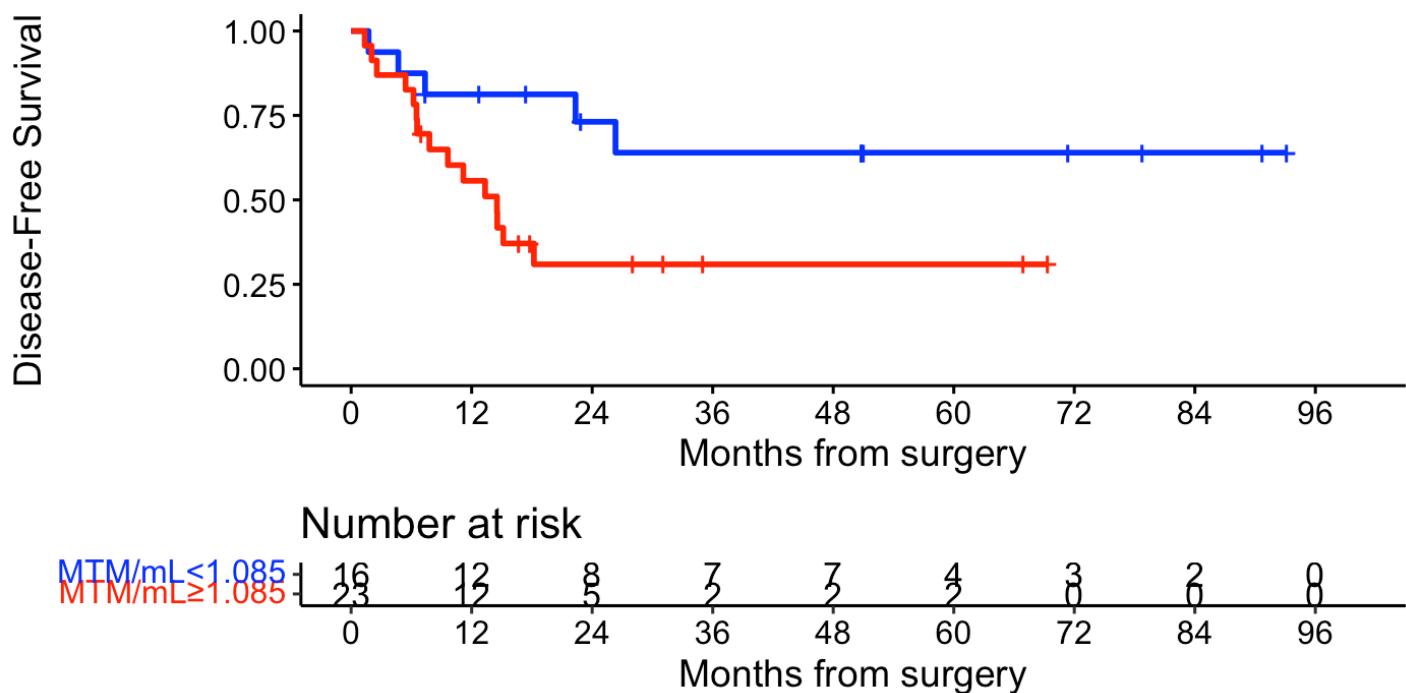
[Hide](#)

```

surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.6mMTM.Q, 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 MTM/mL groups at Baseline",
ylab= "Disease-Free Survival", xlab="Months from surgery", legend.labs=c("MTM/mL<1.085",
"MTM/mL≥1.085"), legend.title="")

```

DFS - ctDNA MTM/mL groups at Baseline



Hide

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

```
Call: survfit(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

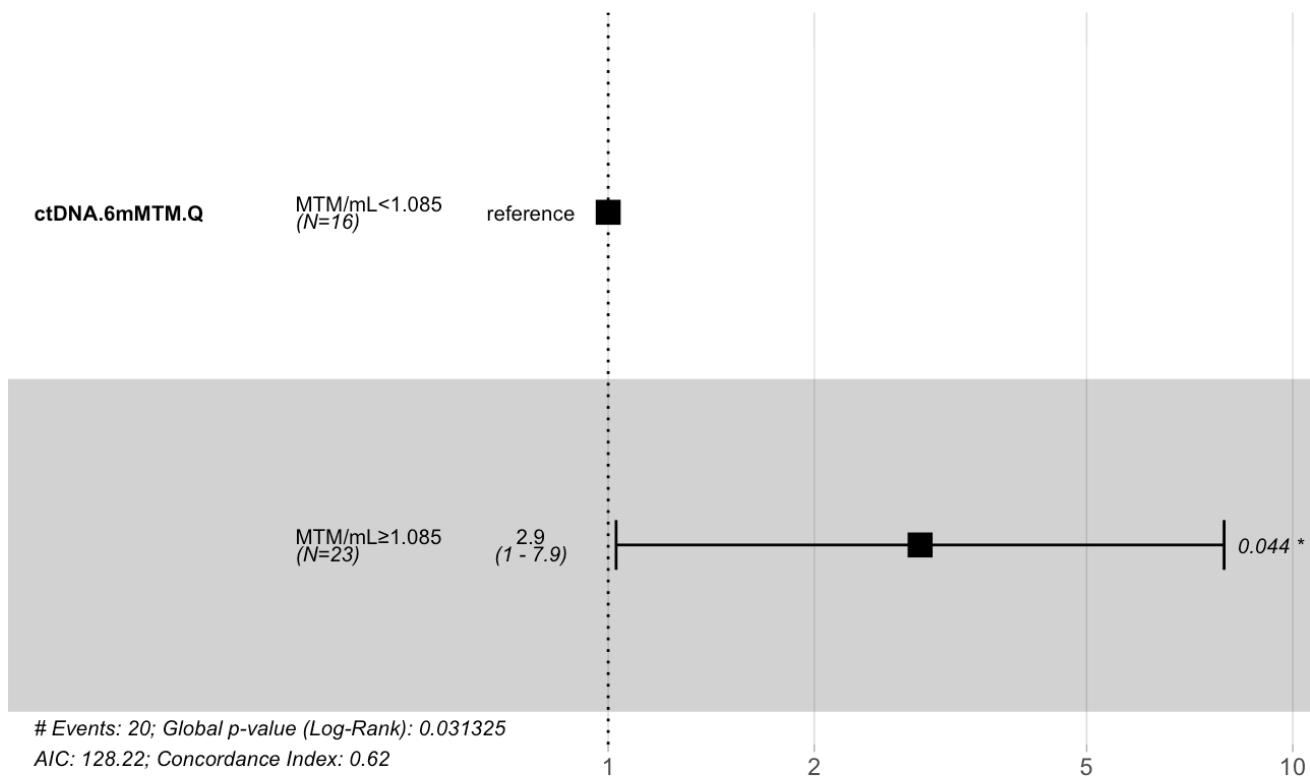
CI	ctDNA.6mMTM.Q=1						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.000	8.000	4.000	0.731	0.117	0.427	0.8	

CI	ctDNA.6mMTM.Q=2						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.000	5.000	15.000	0.309	0.103	0.129	0.5	

Hide

```
circ_data$ctDNA.6mMTM.Q <- factor(circ_data$ctDNA.6mMTM.Q, levels=c("1","2"), labels = c
("MTM/mL<1.085", "MTM/mL≥1.085"))
cox_fit <- coxph(surv_object ~ ctDNA.6mMTM.Q, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:
`coxph(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data)`

`n= 39, number of events= 20`

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.6mMTM.QMTM/mL≥1.085	1.0492	2.8553	0.5217	2.011	0.0443 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.6mMTM.QMTM/mL≥1.085	2.855	0.3502	1.027	7.939

Concordance= 0.617 (se = 0.056)

Likelihood ratio test= 4.64 on 1 df, p=0.03

Wald test = 4.04 on 1 df, p=0.04

Score (logrank) test = 4.4 on 1 df, p=0.04

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 = 2.86 (1.03-7.94); p = 0.044"
```

#OS by ctDNA levels at Baseline based on AUC optimal MTM/mL level from RFS model - ctDNA Baseline positive

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!="",]
circ_data <- circ_data[circ_data$ctDNA.Base=="POSITIVE",]
circ_datadf <- as.data.frame(circ_data)

# Create a new variable based on these quartiles

circ_data$ctDNA.6mMTM.Q <- NA
circ_data <- circ_data %>%
  mutate(ctDNA.6mMTM.Q = case_when(
    ctDNA.Base.MTM < 1.085 ~ 1,
    ctDNA.Base.MTM >= 1.085 ~ 2
  ))

survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.6mMTM.Q, data = circ_data)
```

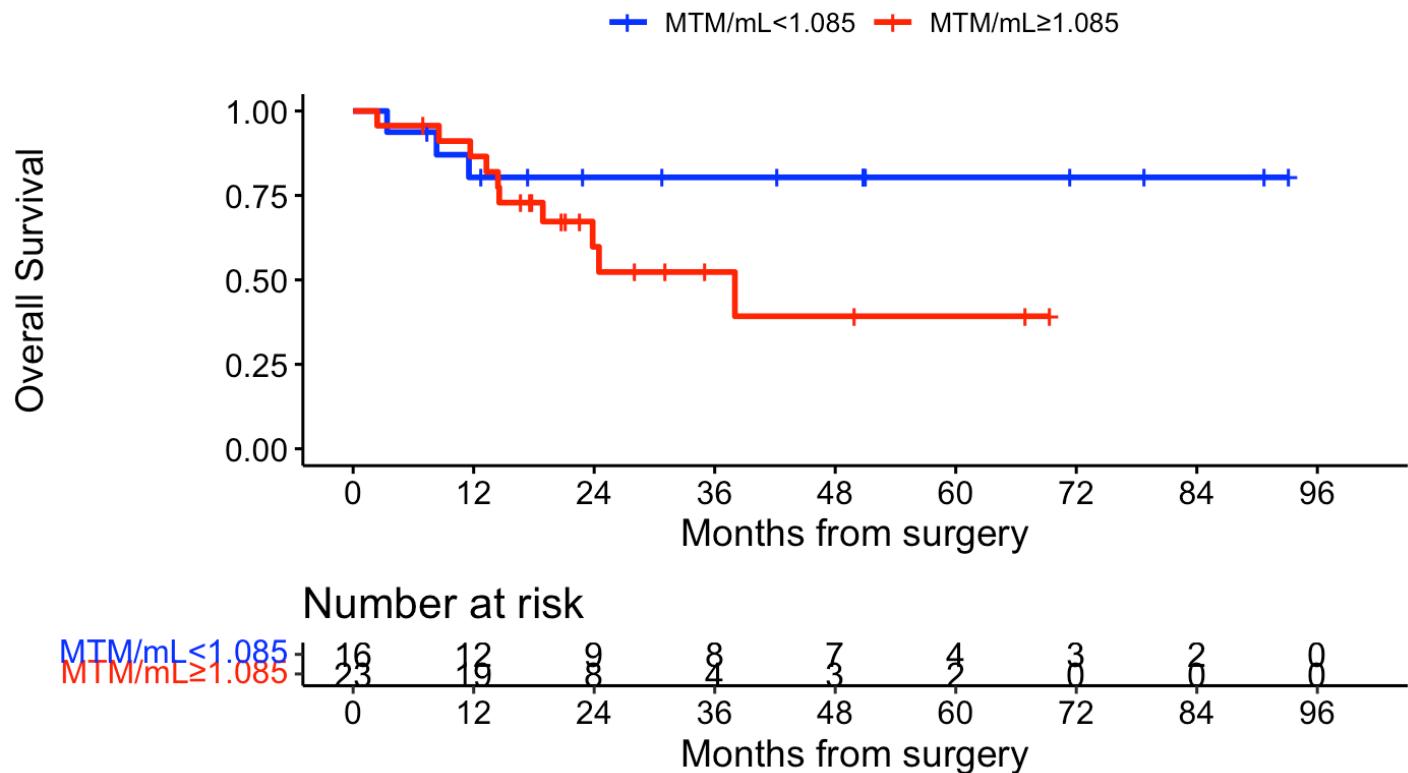
Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~ ctDNA.6mMTM.Q, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.6mMTM.Q=1	16	3	NA	NA	NA
ctDNA.6mMTM.Q=2	23	10	38	18.9	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.6mMTM.Q, 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 MTM/mL groups at Baseline",
ylab= "Overall Survival", xlab="Months from surgery", legend.labs=c("MTM/mL<1.085",
"MTM/mL≥1.085"), legend.title="")
```

OS - ctDNA MTM/mL groups at Baseline



Hide

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

```
Call: survfit(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

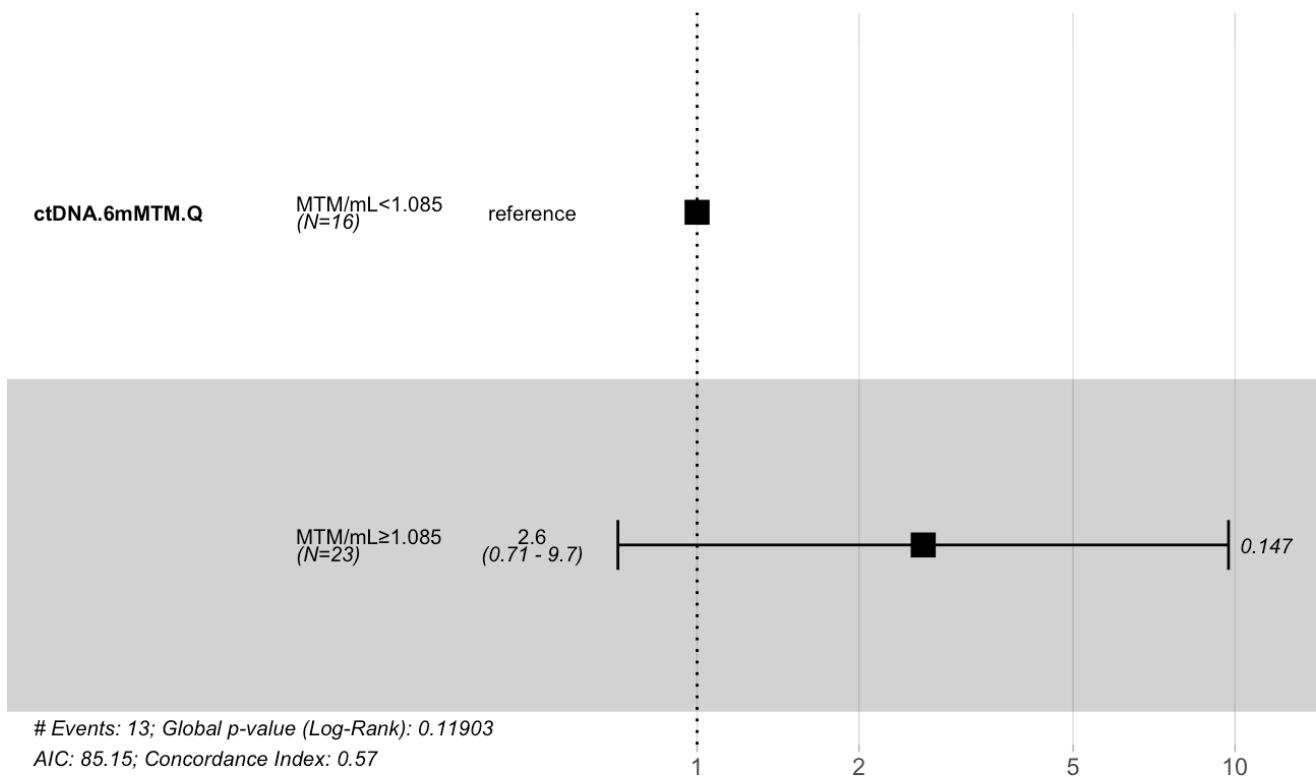
CI	ctDNA.6mMTM.Q=1						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
32	24.000	9.000	3.000	0.804	0.102	0.506	0.9

CI	ctDNA.6mMTM.Q=2						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
81	24.000	8.000	8.000	0.598	0.115	0.342	0.7

Hide

```
circ_data$ctDNA.6mMTM.Q <- factor(circ_data$ctDNA.6mMTM.Q, levels=c("1","2"), labels = c
("MTM/mL<1.085", "MTM/mL≥1.085"))
cox_fit <- coxph(surv_object ~ ctDNA.6mMTM.Q, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data)
```

```
n= 39, number of events= 13
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.6mMTM.QMTM/mL ≥ 1.085	0.9680	2.6326	0.6668	1.452	0.147

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.6mMTM.QMTM/mL ≥ 1.085	2.633	0.3798	0.7125	9.727

```
Concordance= 0.566 (se = 0.076 )
```

```
Likelihood ratio test= 2.43 on 1 df, p=0.1
```

```
Wald test = 2.11 on 1 df, p=0.1
```

```
Score (logrank) test = 2.26 on 1 df, p=0.1
```

Hide

```
cox_fit_summary <- summary(cox_fit)

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

[1] "HR = 2.63 (0.71-9.73); p = 0.147"

#RFS by ctDNA on-NAT - all stages

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.C2D1!="",]
circ_datadf <- as.data.frame(circ_data)

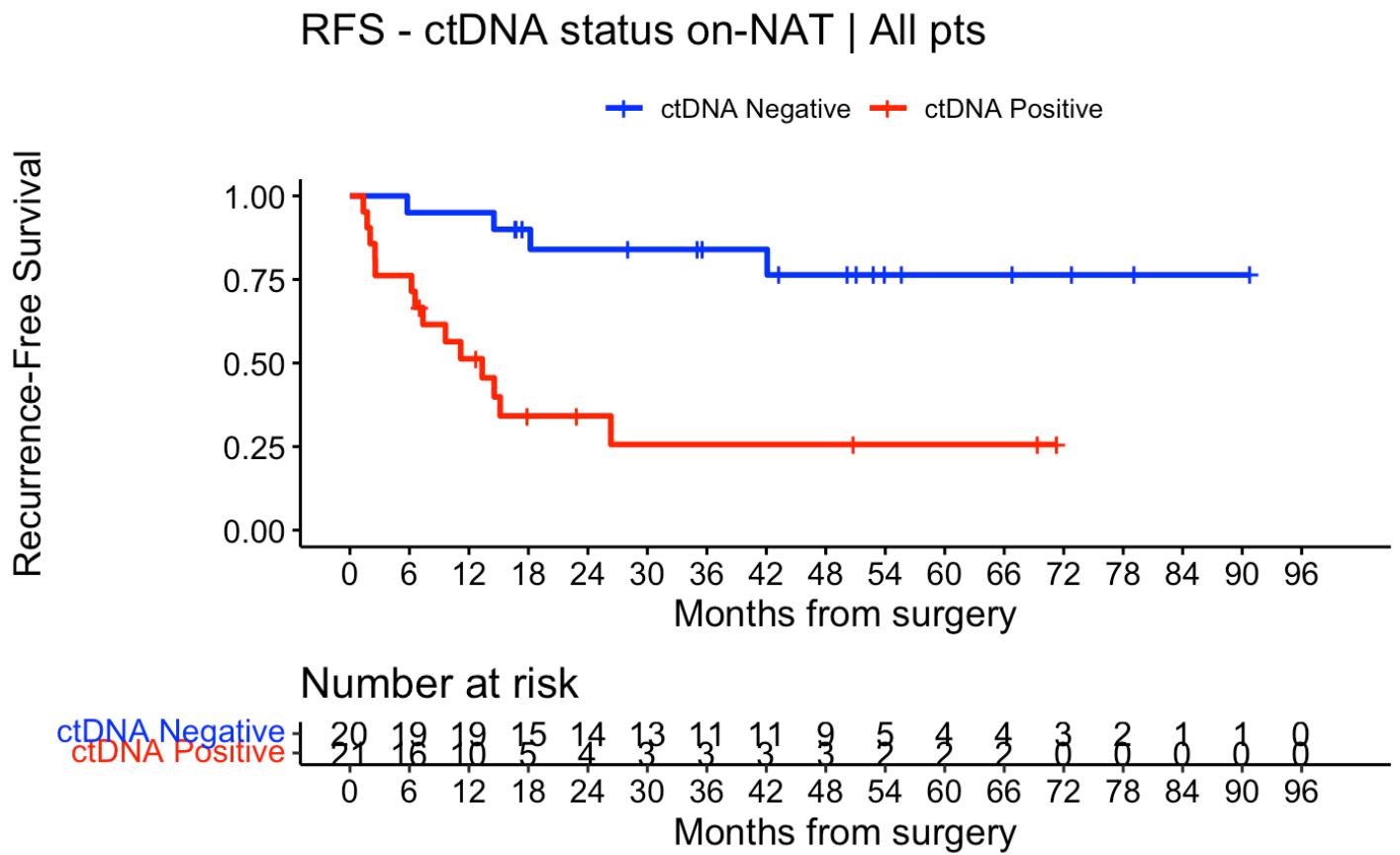
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.C2D1, data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.C2D1, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1=NEGATIVE	20	4	NA	NA	NA
ctDNA.C2D1=POSITIVE	21	14	13.3	6.57	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1, data = circ_data, conf.int=0.95, conf.type ="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE, break.time.by=6, palette=c("blue","red"), title="RFS - ctDNA status on-NAT | All pts", ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



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

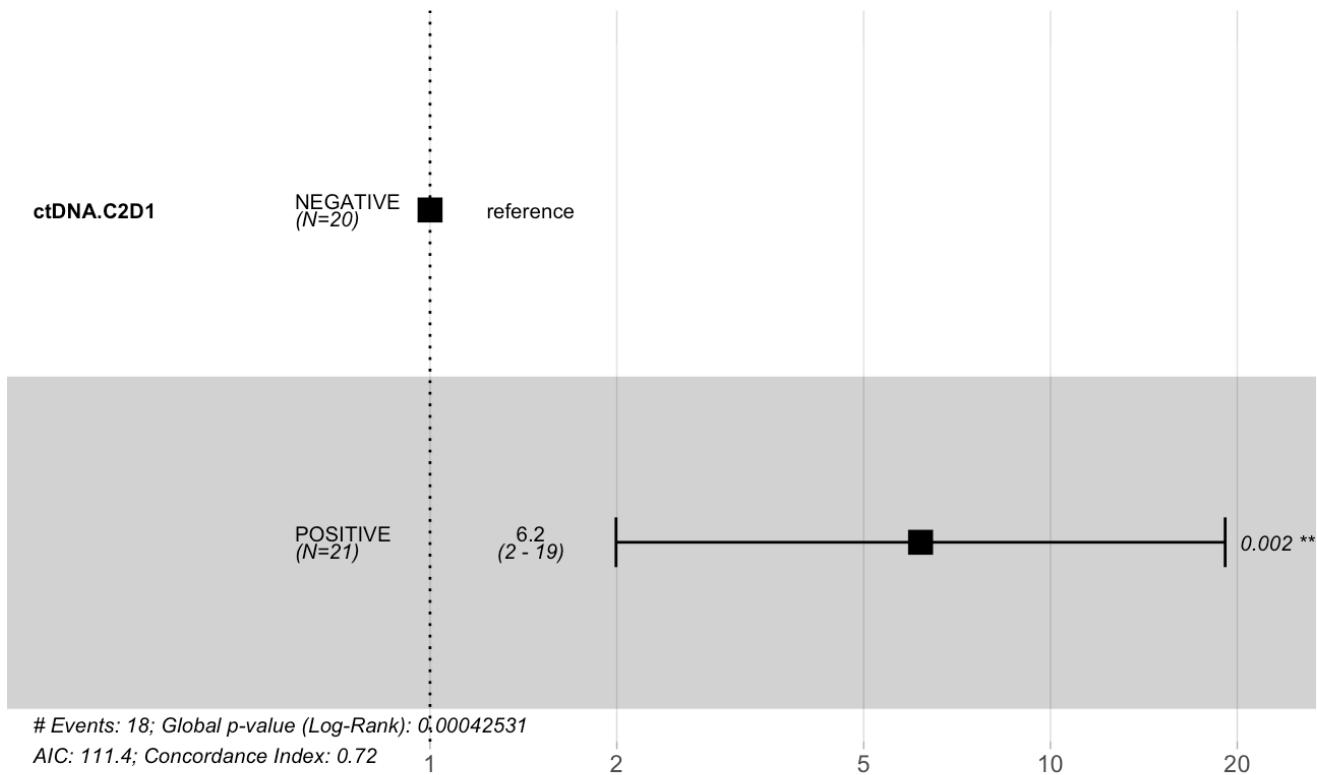
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C2D1=NEGATIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
59	24.0000	14.0000	3.0000	0.8400	0.0853	0.5792	0.5792	0.94	
ctDNA.C2D1=POSITIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
50	24.00	4.000	13.000	0.342	0.109	0.146	0.146	0.5	

[Hide](#)

```
circ_data$ctDNA.C2D1 <- factor(circ_data$ctDNA.C2D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hazard ratio



[Hide](#)

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C2D1, data = circ_data)

n= 41, number of events= 18

            coef exp(coef)  se(coef)      z Pr(>|z|)    
ctDNA.C2D1POSITIVE 1.8205     6.1750    0.5767  3.157   0.0016 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

            exp(coef) exp(-coef) lower .95 upper .95    
ctDNA.C2D1POSITIVE     6.175      0.1619     1.994     19.12   

Concordance= 0.718 (se = 0.046 )
Likelihood ratio test= 12.42 on 1 df,  p=4e-04
Wald test             = 9.97 on 1 df,  p=0.002
Score (logrank) test = 12.64 on 1 df,  p=4e-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 = 6.17 (1.99-19.12); p = 0.002"
```

#OS by ctDNA on-NAT - all stages

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.C2D1!=""]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.C2D1, data = circ_data)
```

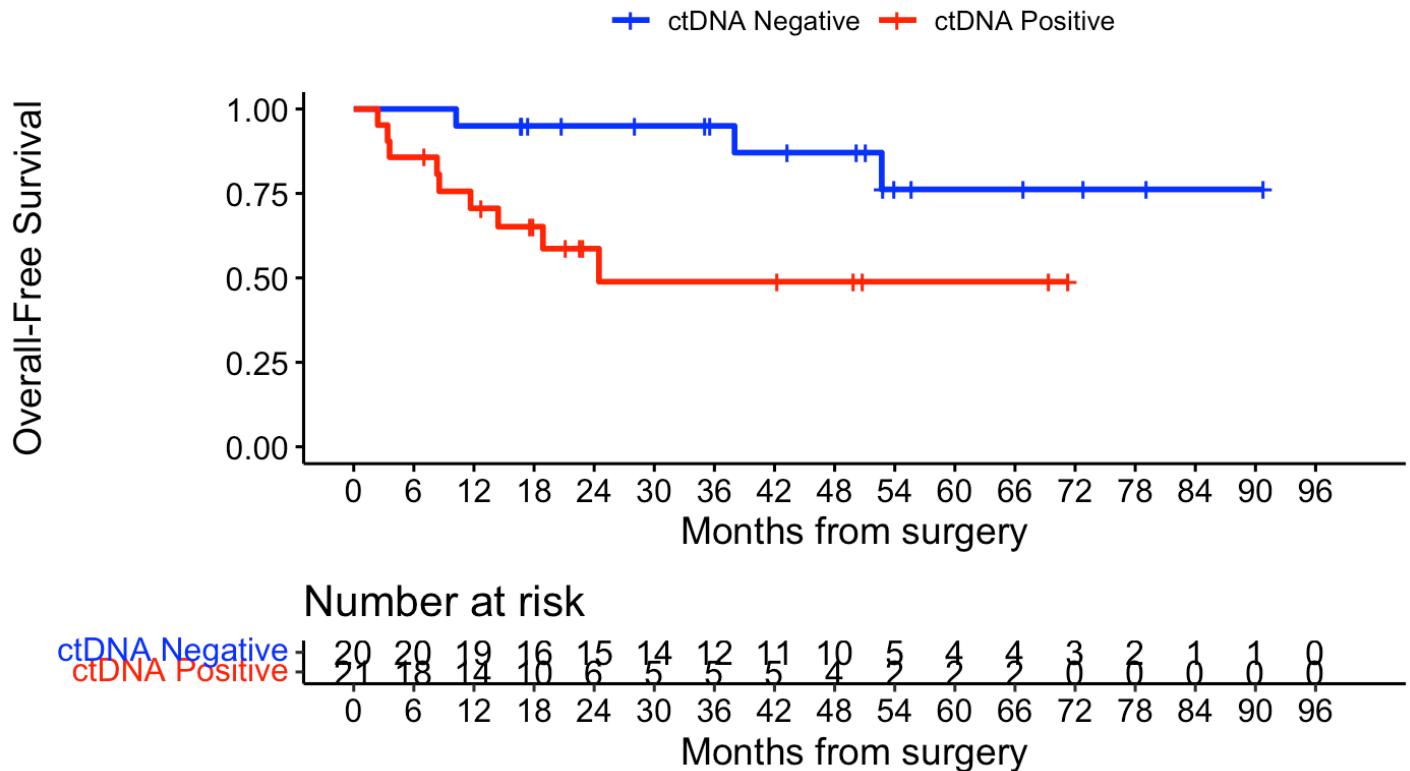
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.C2D1, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1=NEGATIVE	20	3	NA	NA	NA
ctDNA.C2D1=POSITIVE	21	9	24.5	14.4	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1, data = circ_data, conf.int=0.95, conf.type
="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE,
break.time.by=6, palette=c("blue","red"), title="OS - ctDNA status on-NAT | All pts",
ylab= "Overall-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negative",
"ctDNA Positive"), legend.title="")
```

OS - ctDNA status on-NAT | All pts

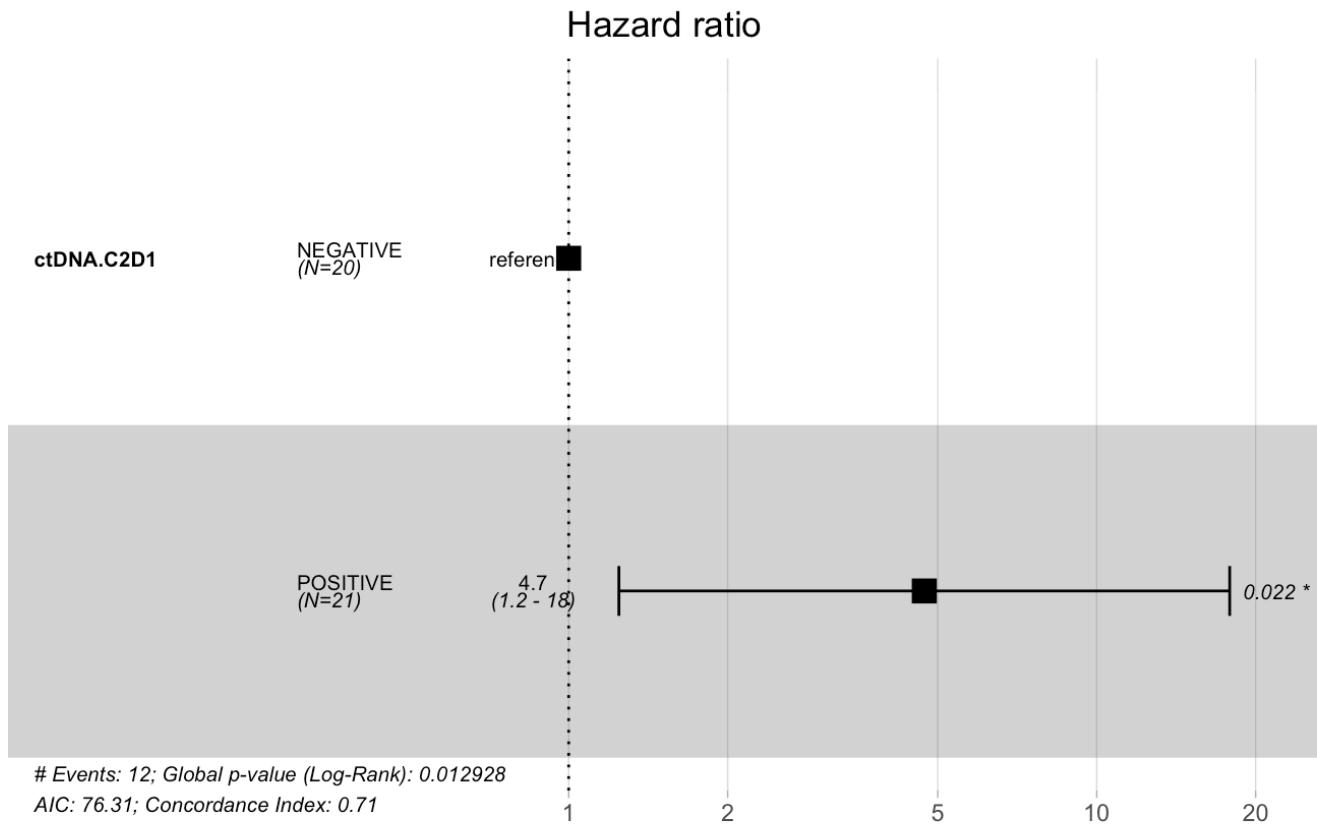


```
summary(KM_curve, times= c(24))
```

```
Call: survfit(formula = surv_object ~ ctDNA.C2D1, data = circ_data,  
conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C2D1=NEGATIVE							
	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95%
CI	24.0000	15.0000	1.0000	0.9500	0.0487	0.6947	0.99
28							
ctDNA.C2D1=POSITIVE							
	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95%
CI	24.0000	6.0000	8.0000	0.586	0.114	0.335	0.7
71							

```
circ_data$ctDNA.C2D1 <- factor(circ_data$ctDNA.C2D1, levels=c("NEGATIVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1, data=circ_data)
qqforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

Call:
`coxph(formula = surv_object ~ ctDNA.C2D1, data = circ_data)`

`n= 41, number of events= 12`

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C2D1POSITIVE	1.5506	4.7145	0.6795	2.282	0.0225 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C2D1POSITIVE	4.714	0.2121	1.245	17.86

Concordance= 0.706 (se = 0.054)
 Likelihood ratio test= 6.18 on 1 df, p=0.01
 Wald test = 5.21 on 1 df, p=0.02
 Score (logrank) test = 6.2 on 1 df, p=0.01

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 4.71 (1.24-17.86); p = 0.022"
```

#DFS by ctDNA Clearance during NAT

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

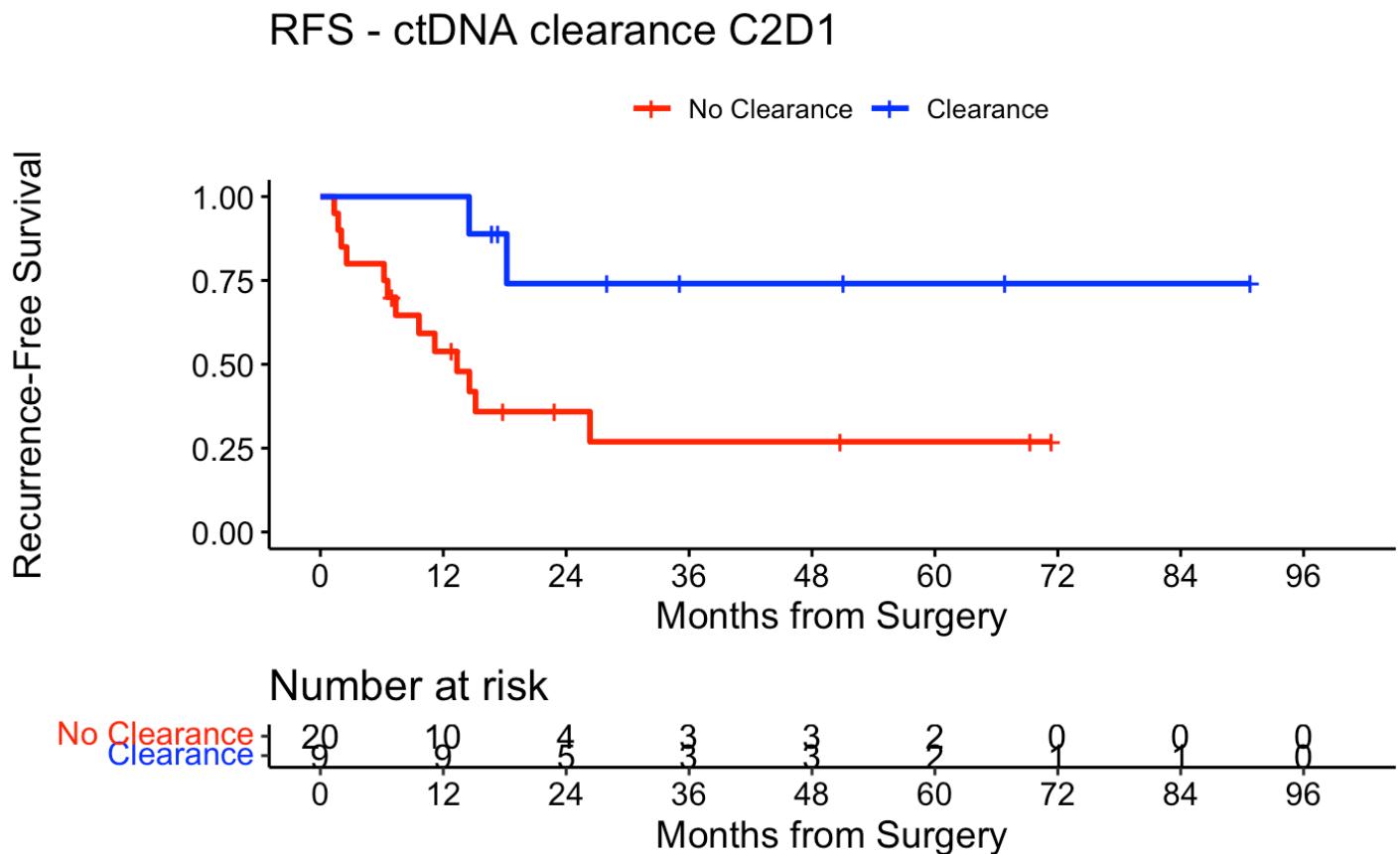
circ_data$ctDNA.C2D1.Clearance <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C2D1.Clearance = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" ~ "TRUE",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" ~ "FALSE",
  ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.C2D1.Clearance),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.C2D1.Clearance, data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.C2D1.Clearance, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1.Clearance=FALSE	20	13	13.3	7.36	NA
ctDNA.C2D1.Clearance=TRUE	9	2	NA	18.20	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1.Clearance, 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("red","blue"), title="RFS - ctDNA clearance C2D1", ylab="Recurrence-Free Survival", xlab="Months from Surgery", legend.labs=c("No Clearance", "Clearance"), legend.title="")
```



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

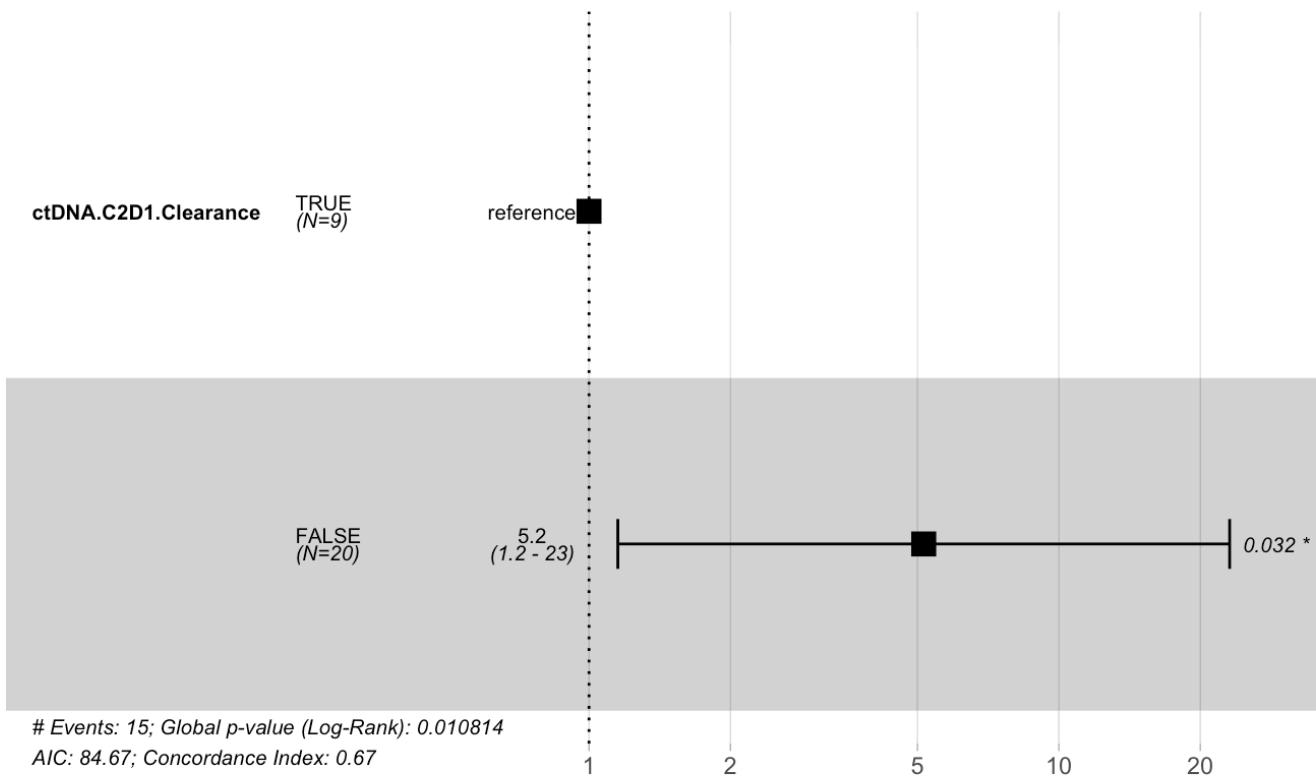
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C2D1.Clearance=FALSE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
72	24.000	4.000	12.000	0.359	0.114	0.153	0.5		
ctDNA.C2D1.Clearance=TRUE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
30	24.000	5.000	2.000	0.741	0.161	0.289	0.9		

Hide

```
circ_data$ctDNA.C2D1.Clearance <- factor(circ_data$ctDNA.C2D1.Clearance, levels=c("TRUE", "FALSE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1.Clearance, data=circ_data)
ggforest(cox_fit, data = circ_data)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data)

n= 29, number of events= 15

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C2D1.Clearance	1.6403	5.1567	0.7649	2.144	0.032 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C2D1.Clearance	5.157	0.1939	1.152	23.09

Concordance= 0.667 (se = 0.046)

Likelihood ratio test= 6.5 on 1 df, p=0.01

Wald test = 4.6 on 1 df, p=0.03

Score (logrank) test = 5.66 on 1 df, p=0.02

Hide

cox_fit_summary <- summary(cox_fit)

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

HR <- cox_fit_summary\$coefficients[2]

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

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

p_value <- cox_fit_summary\$coefficients[5]

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

print(label_text)

[1] "HR = 5.16 (1.15-23.09); p = 0.032"

#OS by ctDNA Clearance during NAT

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.C2D1.Clearance <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C2D1.Clearance = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" ~ "TRUE",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" ~ "FALSE",
  ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.C2D1.Clearance),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.C2D1.Clearance, data = circ_data)

```

Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~ ctDNA.C2D1.Clearance, data = circ_data)

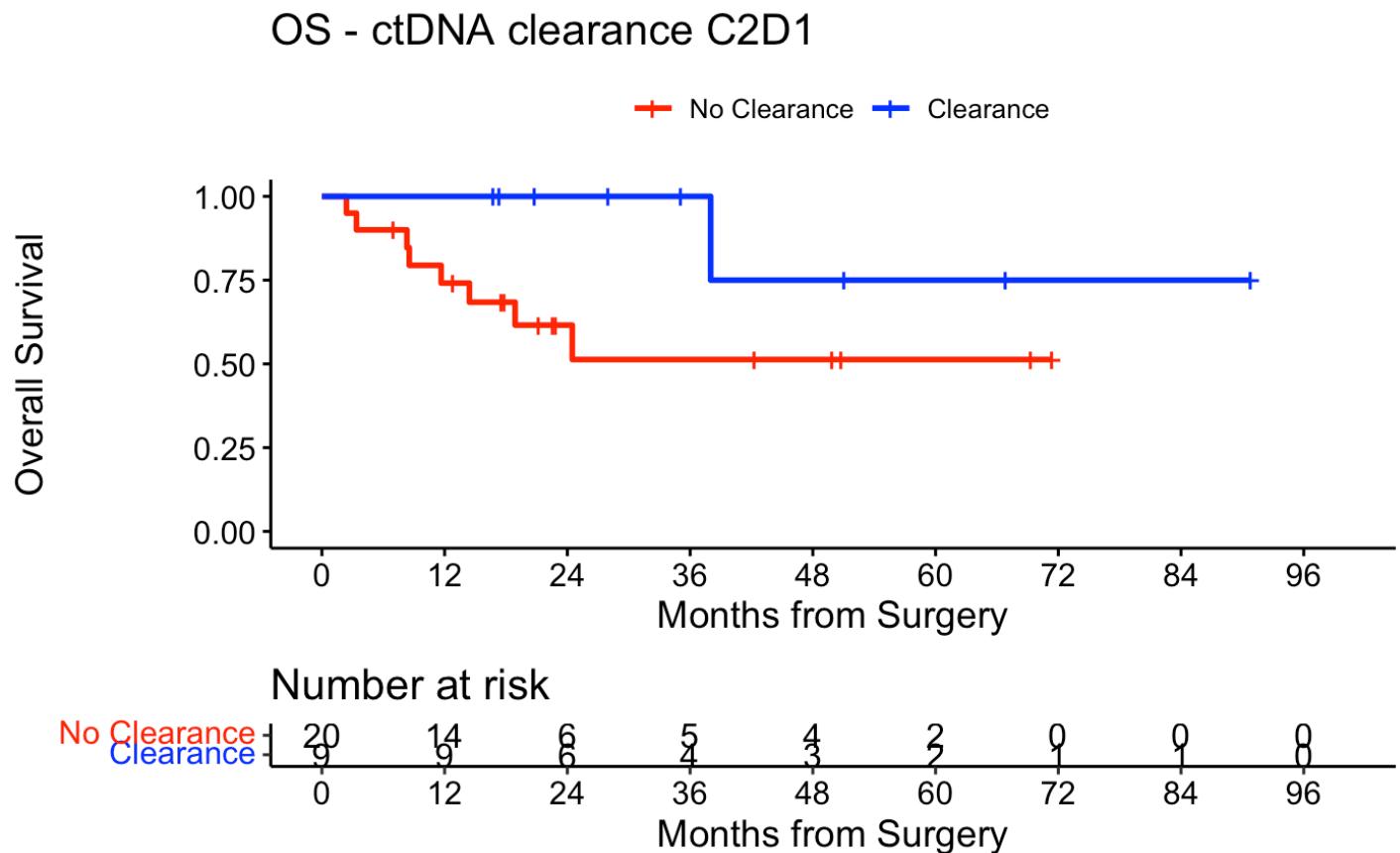
	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1.Clearance=FALSE	20	8	NA	18.9	NA
ctDNA.C2D1.Clearance=TRUE	9	1	NA	38.0	NA

[Hide](#)

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1.Clearance, 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("red","blue"), title="OS - ctDNA clearance C2D1", ylab="Overall Survival", xlab="Months from Surgery", legend.labs=c("No Clearance", "Clearance"), legend.title="")

```



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

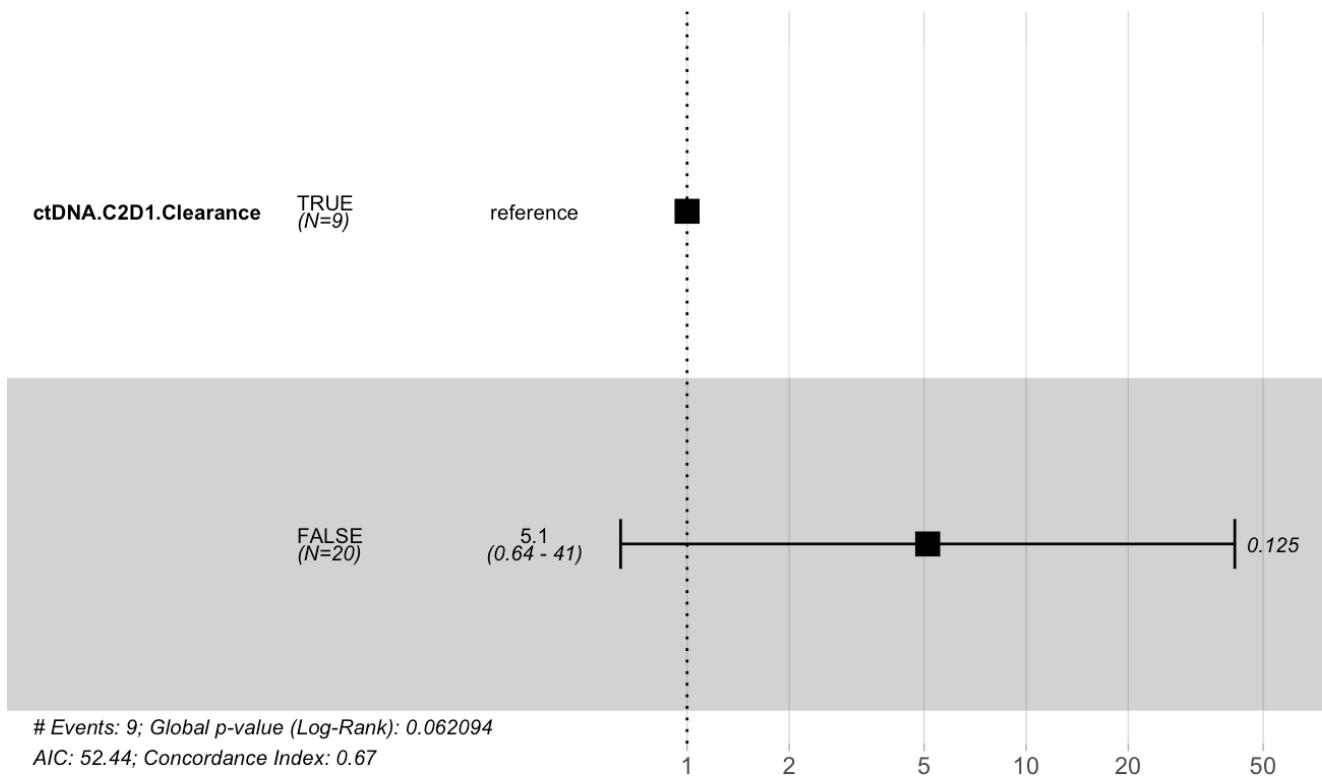
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C2D1.Clearance=FALSE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
98	24.000	6.000	7.000	0.616	0.116	0.353	0.7		
ctDNA.C2D1.Clearance=TRUE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
NA	24	6	0	1	0	NA			

[Hide](#)

```
circ_data$ctDNA.C2D1.Clearance <- factor(circ_data$ctDNA.C2D1.Clearance, levels=c("TRUE", "FALSE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1.Clearance, data=circ_data)
ggforest(cox_fit, data = circ_data)
```

Hazard ratio

[Hide](#)

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data)

n= 29, number of events= 9

            coef exp(coef)  se(coef)      z Pr(>|z|)    
ctDNA.C2D1.ClearanceFALSE 1.634      5.126     1.064  1.536     0.125    
                                         exp(coef) exp(-coef) lower .95 upper .95    
ctDNA.C2D1.ClearanceFALSE      5.126     0.1951    0.6369    41.25    

Concordance= 0.669 (se = 0.047 )
Likelihood ratio test= 3.48 on 1 df,  p=0.06
Wald test             = 2.36 on 1 df,  p=0.1
Score (logrank) test = 2.92 on 1 df,  p=0.09
```

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 5.13 (0.64-41.25); p = 0.125"
```

#RFS by ctDNA post-NAT - all stages

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

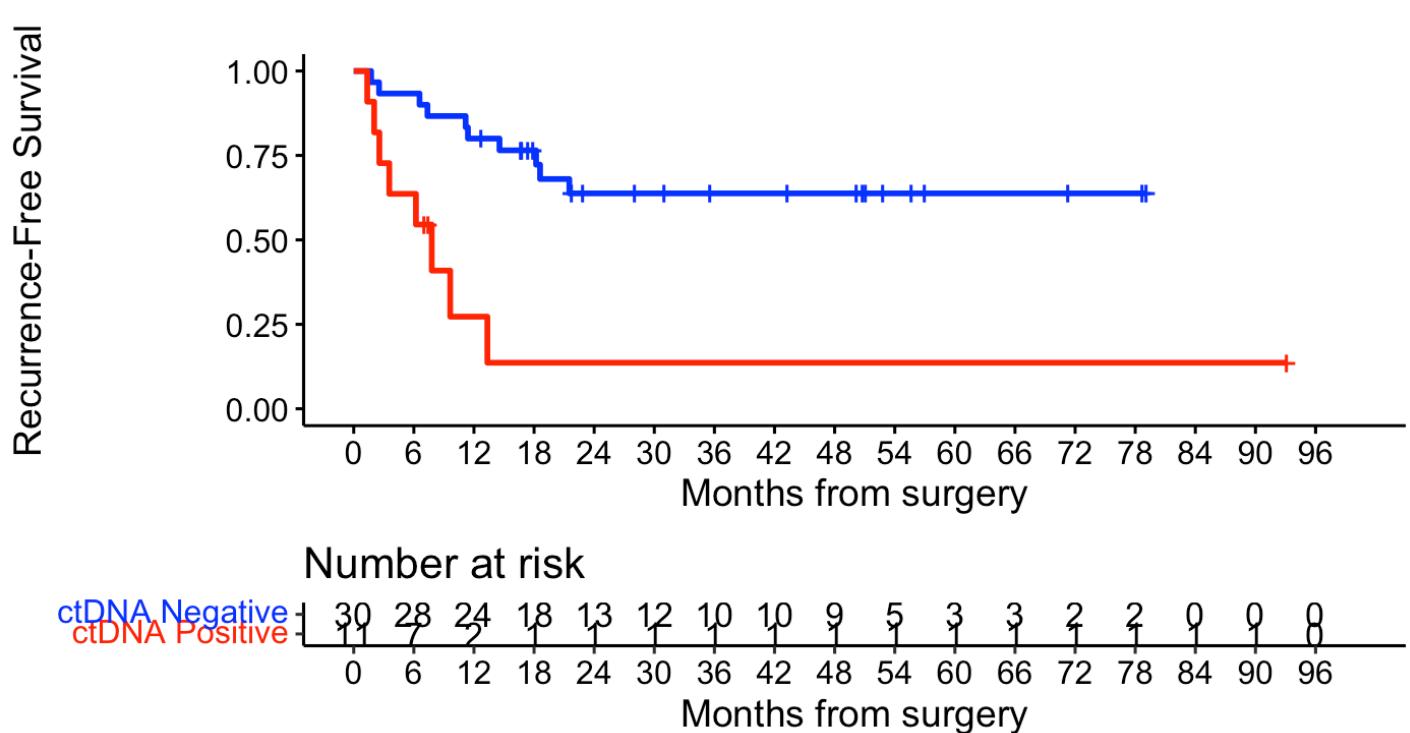
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.postNAC, data = circ_data)
```

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.postNAC=NEGATIVE	30	10	NA	21.52	NA
ctDNA.postNAC=POSITIVE	11	8	7.79	3.55	NA

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.postNAC, data = circ_data, conf.int=0.95, conf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE,
break.time.by=6, palette=c("blue","red"), title="RFS - ctDNA status post-NAT | All pts",
ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

RFS - ctDNA status post-NAT | All pts



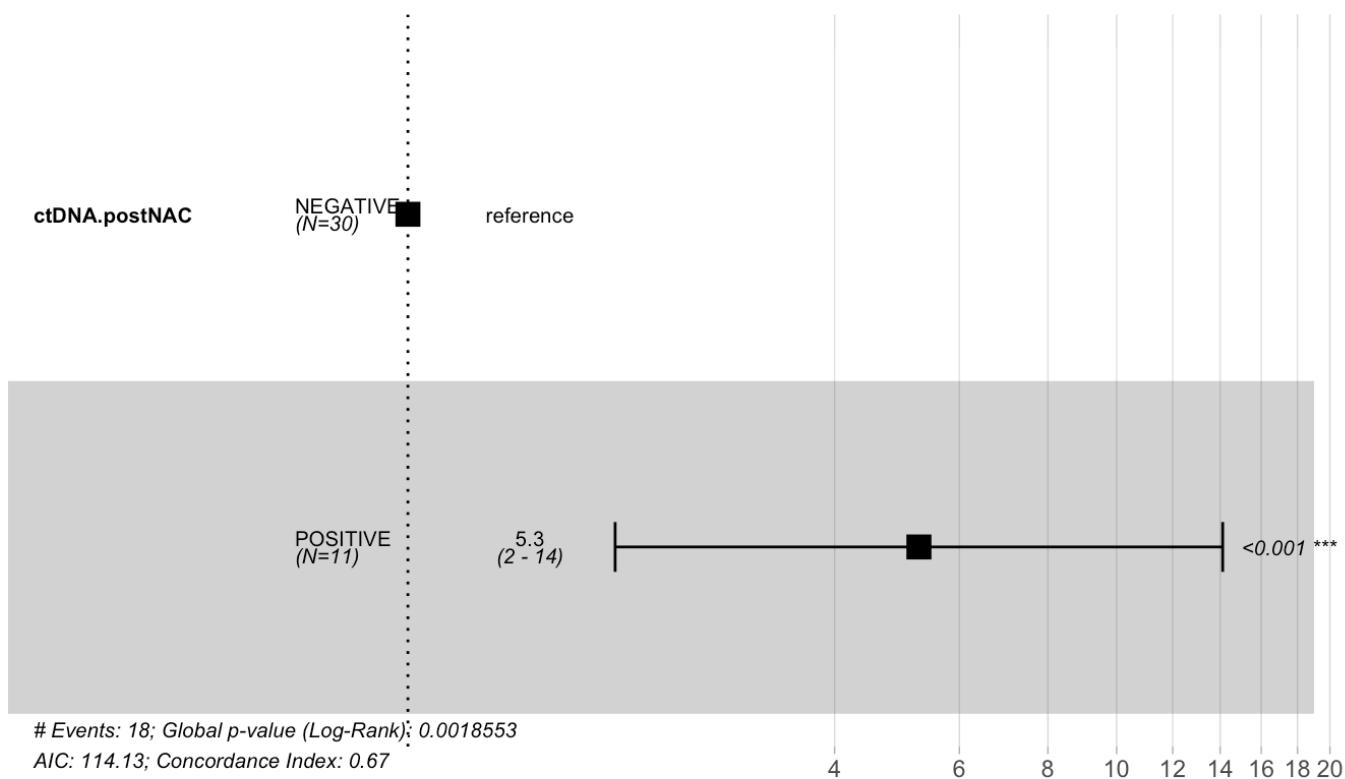
```
summary(KM_curve, times= c(24))
```

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

CI	time	ctDNA.postNAC=NEGATIVE		survival	std.err	95% CI	
		n.risk	n.event			lower	upper
83	24.0000	13.0000	10.0000	0.6377	0.0933	0.4269	0.78
CI	time	ctDNA.postNAC=POSITIVE		survival	std.err	95% CI	
		n.risk	n.event			lower	upper
63	24.00000	1.00000	8.00000	0.13636	0.12392	0.00767	0.442

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

Hazard ratio



Hide

```
summary(cox_fit)
```

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

`n= 41, number of events= 18`

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.postNACPOSITIVE	1.6603	5.2610	0.5036	3.297	0.000978 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.postNACPOSITIVE	5.261	0.1901	1.961	14.12

Concordance= 0.669 (se = 0.052)

Likelihood ratio test= 9.69 on 1 df, p=0.002

Wald test = 10.87 on 1 df, p=0.001

Score (logrank) test = 13.08 on 1 df, p=3e-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 = 5.26 (1.96-14.12); p = 0.001"
```

```
#OS by ctDNA post-NAT - all stages
```

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!=""]
circ_datadf <- as.data.frame(circ_data)

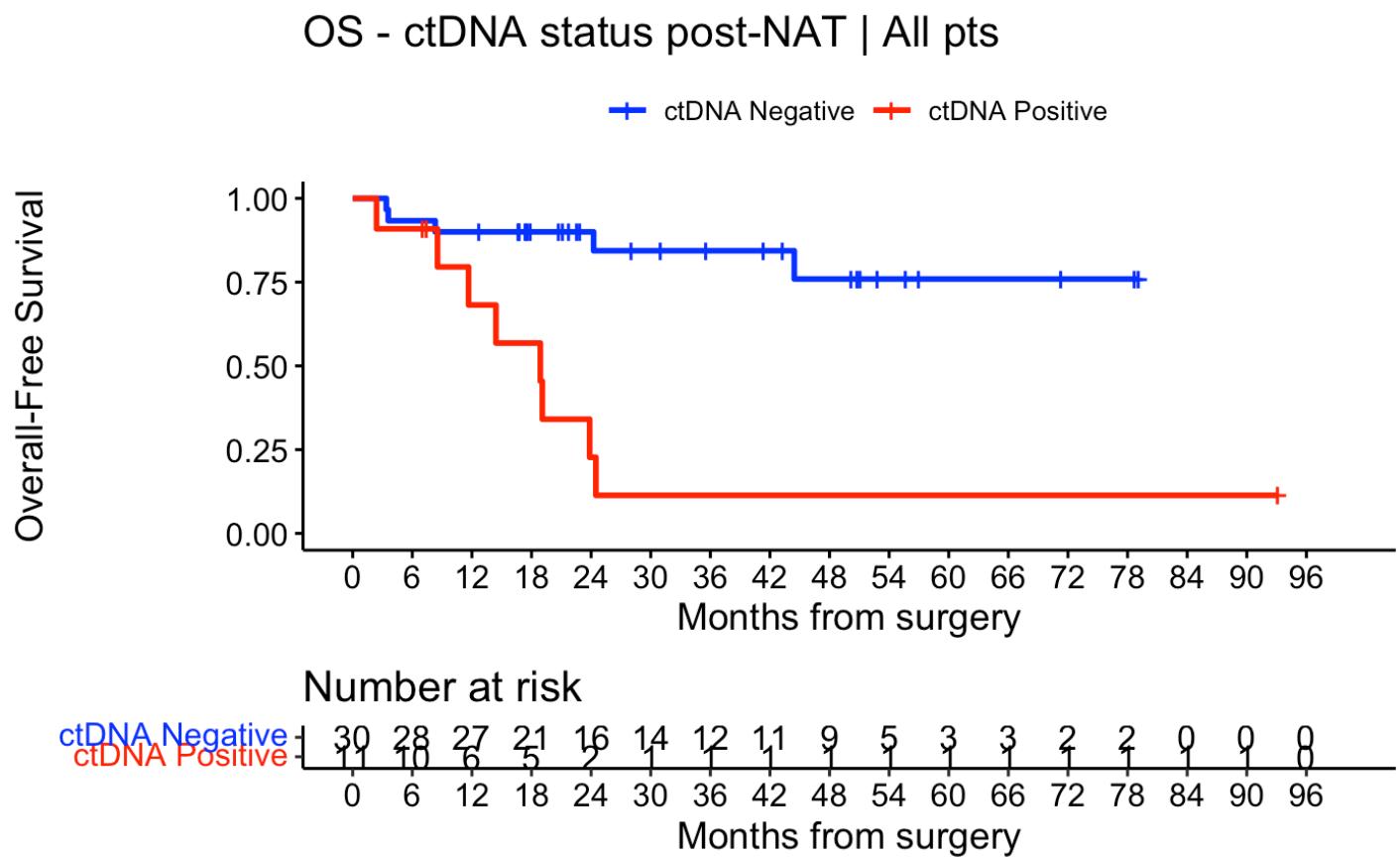
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.postNAC, data = circ_data)
```

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.postNAC=NEGATIVE	30	5	NA	NA	NA
ctDNA.postNAC=POSITIVE	11	8	18.9	11.7	NA

[Hide](#)

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



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

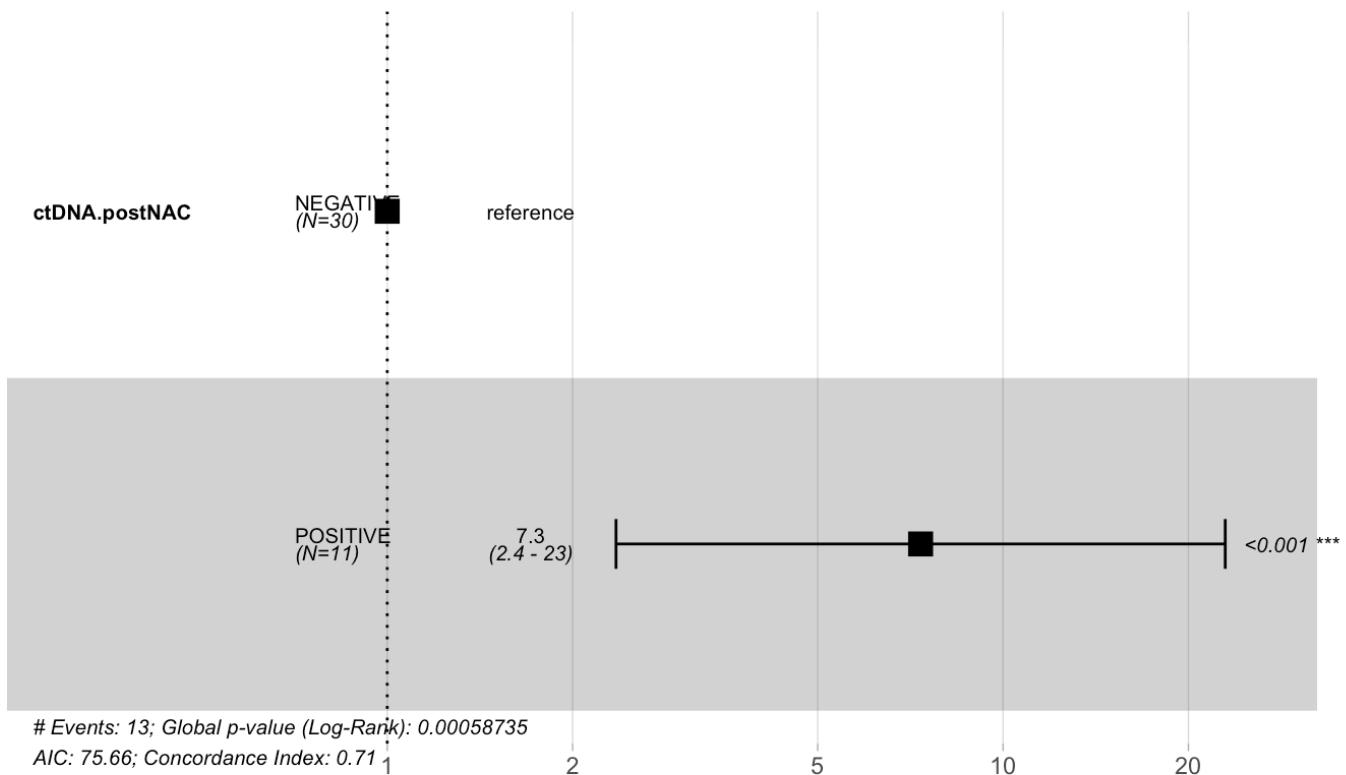
```
Call: survfit(formula = surv_object ~ ctDNA.postNAC, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.postNAC=NEGATIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
66	24.0000	16.0000	3.0000	0.9000	0.0548	0.7212		0.96	
ctDNA.postNAC=POSITIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
07	24.0000	2.0000	7.0000	0.2273	0.1409	0.0346		0.52	

Hide

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

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:

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

n= 41, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.postNACPOSITIVE	1.9944	7.3479	0.5811	3.432	0.000599 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.postNACPOSITIVE	7.348	0.1361	2.352	22.95

Concordance= 0.711 (se = 0.07)

Likelihood ratio test= 11.82 on 1 df, p=6e-04

Wald test = 11.78 on 1 df, p=6e-04

Score (logrank) test = 15.82 on 1 df, p=7e-05

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 = 7.35 (2.35-22.95); p = 0.001"

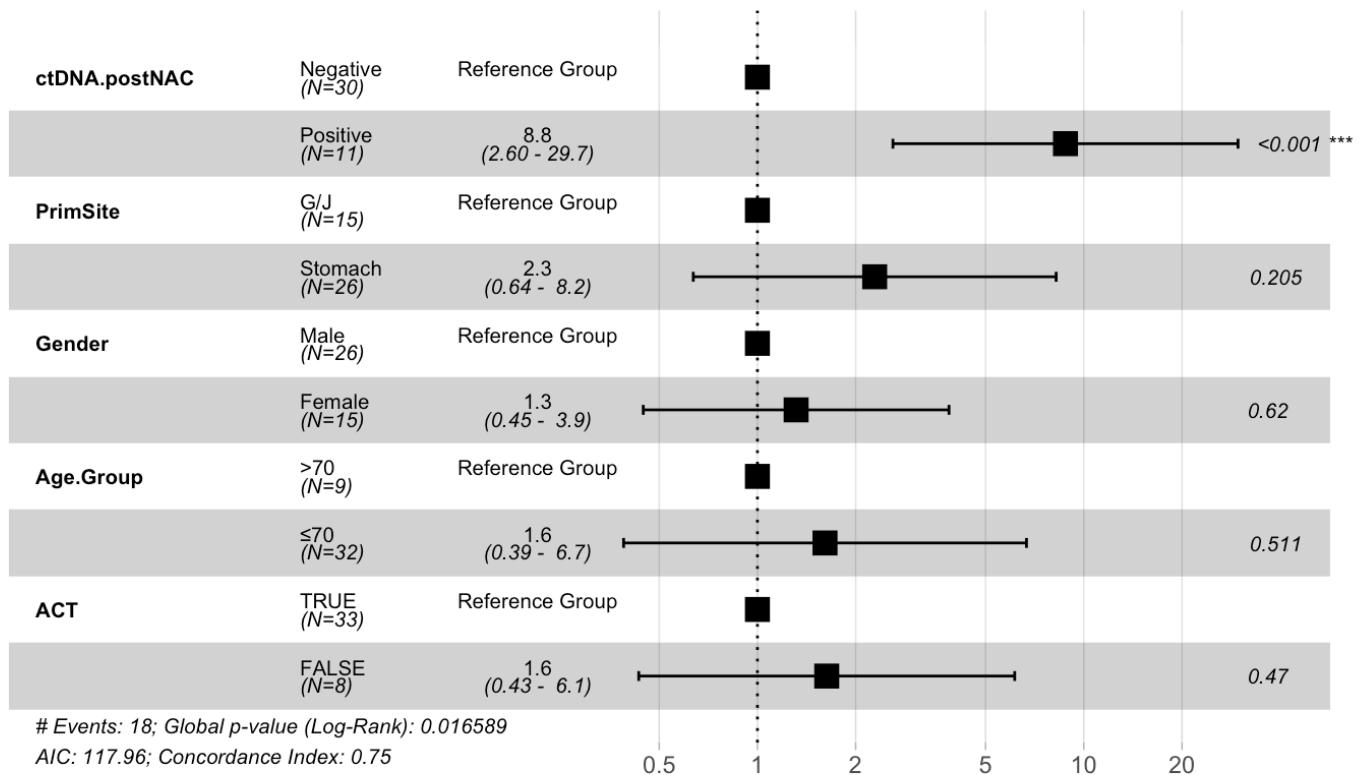
#Multivariate cox regression for RFS - ctDNA post-NAT

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.postNAC <- factor(circ_data$ctDNA.postNAC, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels=c("G/J","Stomach"))
circ_data$Gender <- factor(circ_data$Gender, levels=c("Male","Female"))
circ_data$Age.Group <- factor(circ_data$Age.Group, levels=c("2","1"), labels = c(">70", "<=70"))
circ_data$ypNStatus <- factor(circ_data$ypNStatus, levels=c("N0","N1-N3"))
circ_data$TRG <- factor(circ_data$TRG, levels=c("TRG1/2/3","TRG4/5"))
circ_data$ACT <- factor(circ_data$ACT, levels=c("TRUE","FALSE"))
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.postNAC + PrimSite + Gender + Age.Group + ACT, data =circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for RFS", refLabel = "Reference Group")
```

Multivariate Regression Model for RFS



Hide

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

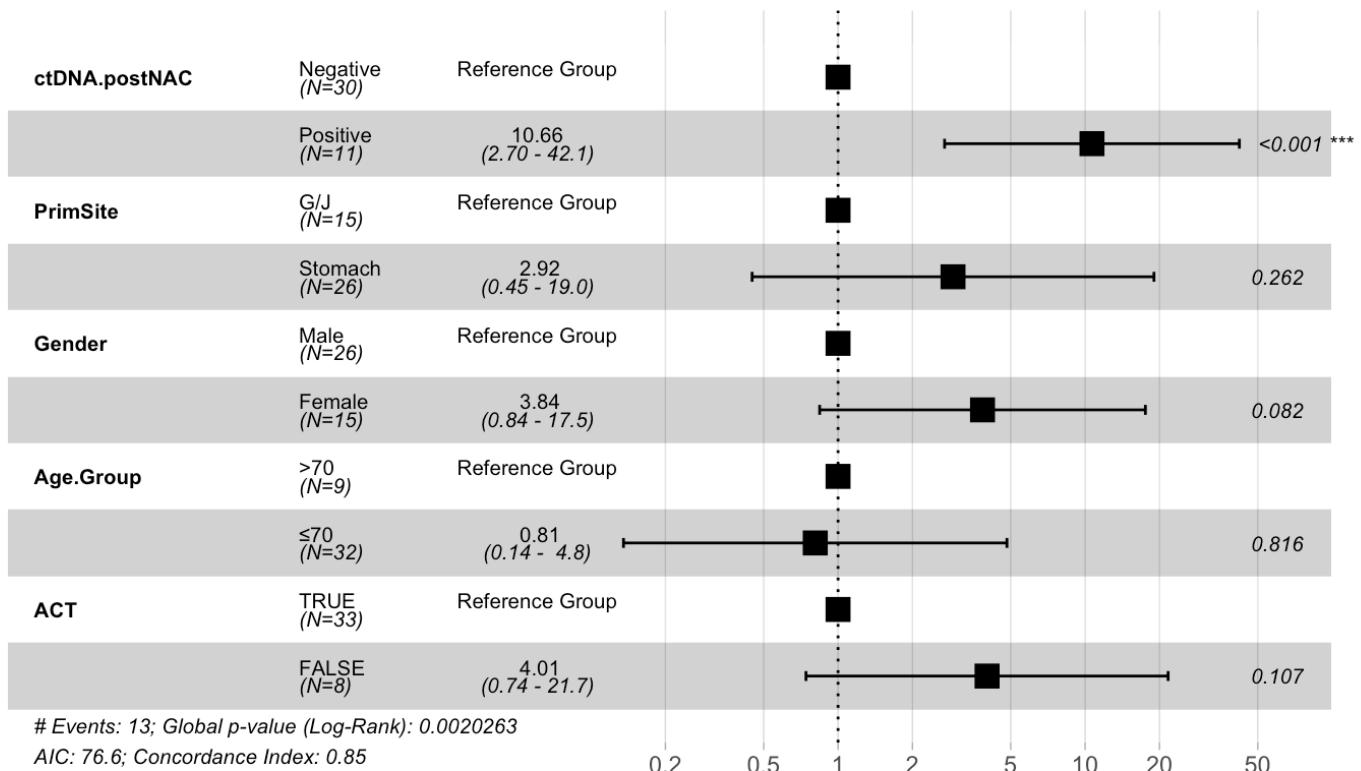
#Multivariate cox regression for OS - ctDNA post-NAT

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.postNAC <- factor(circ_data$ctDNA.postNAC, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels=c("G/J","Stomach"))
circ_data$Gender <- factor(circ_data$Gender, levels=c("Male","Female"))
circ_data$Age.Group <- factor(circ_data$Age.Group, levels=c("2","1"), labels = c(">70", "<=70"))
circ_data$ypNStatus <- factor(circ_data$ypNStatus, levels=c("N0","N1-N3"))
circ_data$TRG <- factor(circ_data$TRG, levels=c("TRG1/2/3","TRG4/5"))
circ_data$ACT <- factor(circ_data$ACT, levels=c("TRUE","FALSE"))
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.postNAC + PrimSite + Gender + Age.Group + ACT, data =circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for OS", refLabel = "Reference Group")
```

Multivariate Regression Model for OS



Hide

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

#RFS by ctDNA post-NAT & ypT

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TNM, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

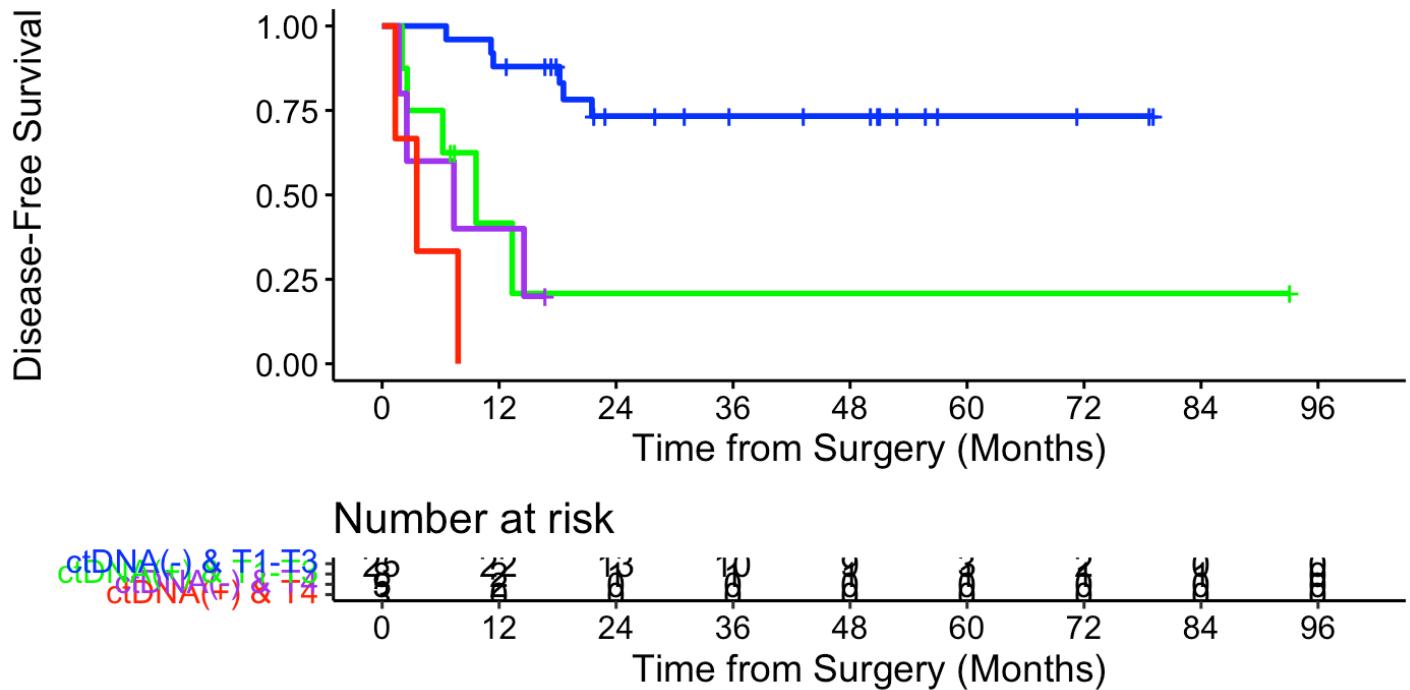
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	25	6	NA	NA	NA
ctDNA.Stage.II.TNM=2	8	5	9.63	6.21	NA
ctDNA.Stage.II.TNM=3	5	4	7.36	2.53	NA
ctDNA.Stage.II.TNM=4	3	3	3.55	1.35	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con-
f.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 post-
NAT & ypTN", ylab= "Disease-Free Survival", xlab="Time from Surgery (Months)", legend.la-
bs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.t-
itle="")
```

DFS - ctDNA post-NAT & ypTN

— ctDNA(-) & T1-T3 — ctDNA(+) & T1-T3 — ctDNA(-) & T4 — ctDNA(+) & T4



Hide

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

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

ctDNA.Stage.II.TNM=1

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	25	0	1.000	0.0000	1.000		1.000	
24	13	6	0.733	0.0944	0.497		0.872	

ctDNA.Stage.II.TNM=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	8	0	1.000	0.000	1.00		1.000	
24	1	5	0.208	0.179	0.01		0.586	

ctDNA.Stage.II.TNM=3

CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
1	0	5	0	1	0		1		

ctDNA.Stage.II.TNM=4

CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
1	0	3	0	1	0		1		

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4",
"ctDNA(+) & T4"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 18

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	1.9291	6.8836	0.6303	3.061	0.00221 **
ctDNA.Stage.II.TNMctDNA(-) & T4	2.2440	9.4313	0.6932	3.237	0.00121 **
ctDNA.Stage.II.TNMctDNA(+) & T4	3.2728	26.3857	0.8116	4.032	5.52e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	6.884	0.1453	2.001	23.68
ctDNA.Stage.II.TNMctDNA(-) & T4	9.431	0.1060	2.424	36.70
ctDNA.Stage.II.TNMctDNA(+) & T4	26.386	0.0379	5.376	129.49

Concordance= 0.775 (se = 0.043)

Likelihood ratio test= 20.62 on 3 df, p=1e-04

Wald test = 19.79 on 3 df, p=2e-04

Score (logrank) test = 29.52 on 3 df, p=2e-06

Hide

```
#Repeat analysis to compare ctDNA post-NAT (+) T1-T3 vs T4
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TNM, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	25	6	NA	NA	NA
ctDNA.Stage.II.TNM=2	8	5	9.63	6.21	NA
ctDNA.Stage.II.TNM=3	5	4	7.36	2.53	NA
ctDNA.Stage.II.TNM=4	3	3	3.55	1.35	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("2","4","1","3"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)
```

n= 41, number of events= 18

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	1.3437	3.8331	0.7683	1.749	0.08031 .
ctDNA.Stage.II.TNM1	-1.9291	0.1453	0.6303	-3.061	0.00221 **
ctDNA.Stage.II.TNM3	0.3149	1.3701	0.6866	0.459	0.64652

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	3.8331	0.2609	0.85028	17.2799
ctDNA.Stage.II.TNM1	0.1453	6.8836	0.04224	0.4997
ctDNA.Stage.II.TNM3	1.3701	0.7299	0.35669	5.2629

Concordance= 0.775 (se = 0.043)

Likelihood ratio test= 20.62 on 3 df, p=1e-04

Wald test = 19.79 on 3 df, p=2e-04

Score (logrank) test = 29.52 on 3 df, p=2e-06

#OS by ctDNA post-NAT & ypT

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
)

circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)

```

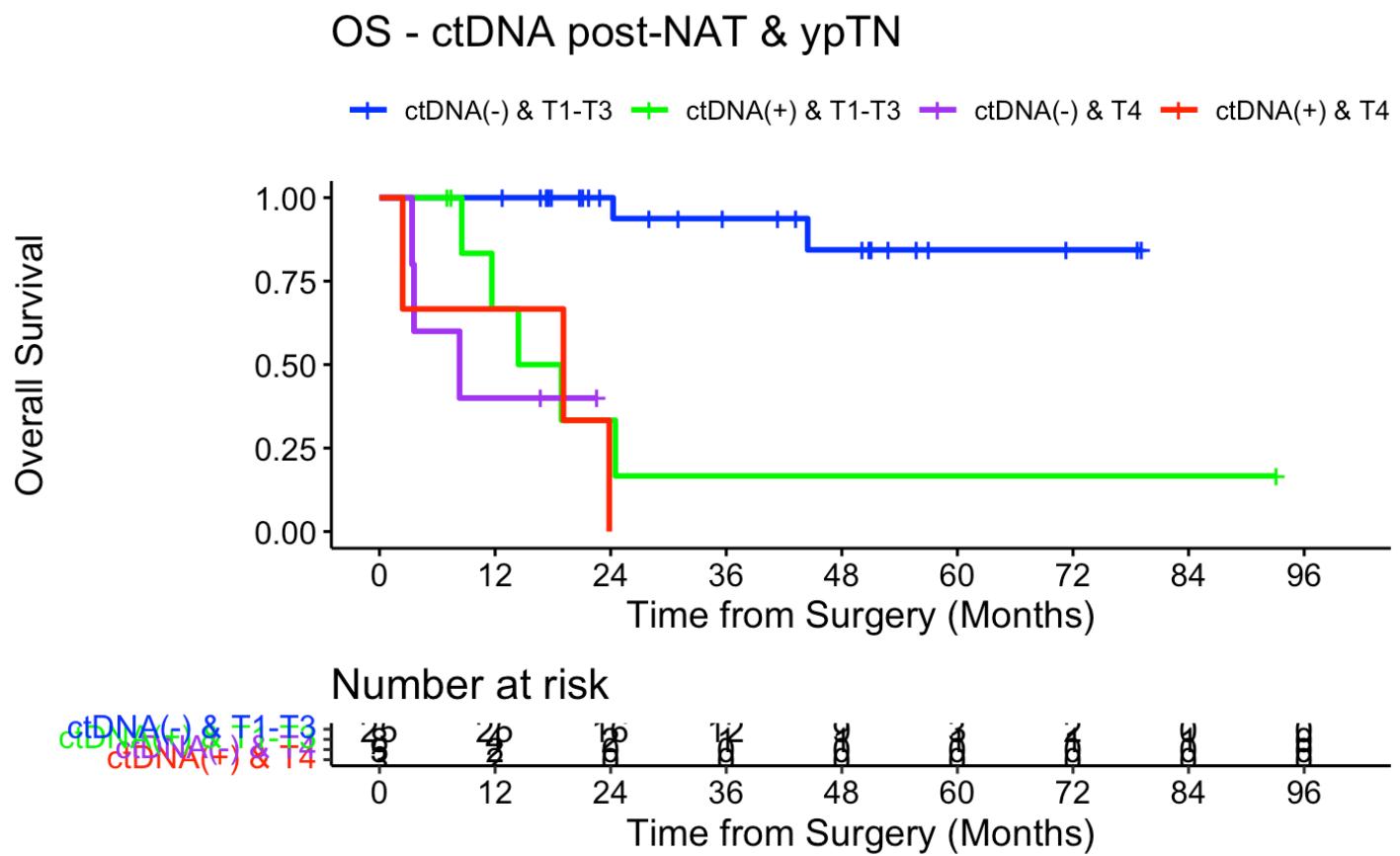
Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~
ctDNA.Stage.II.TNM, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	25	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	8	5	16.66	11.66	NA
ctDNA.Stage.II.TNM=3	5	3	8.31	3.58	NA
ctDNA.Stage.II.TNM=4	3	3	19.09	2.40	NA

Hide

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.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="OS - ctDNA post-N
AT & ypTN", ylab= "Overall Survival", xlab="Time from Surgery (Months)", legend.labs=c
("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.title
="")
```



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

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

ctDNA.Stage.II.TNM=1

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	25	0	1	0	1		1	
24	16	0	1	0	NA		NA	

ctDNA.Stage.II.TNM=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	8	0	1.000	0.000	1.0000		1.000	
24	2	4	0.333	0.192	0.0461		0.676	

ctDNA.Stage.II.TNM=3

CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
1	0	5	0	1	0		1		

ctDNA.Stage.II.TNM=4

CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
1	0	3	0	1	0		1		

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4",
"ctDNA(+) & T4"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	2.7833	16.1727	0.8451	3.293	0.000990 ***
ctDNA.Stage.II.TNMctDNA(-) & T4	3.6483	38.4083	1.0196	3.578	0.000346 ***
ctDNA.Stage.II.TNMctDNA(+) & T4	3.6058	36.8115	0.9977	3.614	0.000302 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	16.17	0.06183	3.086	84.75
ctDNA.Stage.II.TNMctDNA(-) & T4	38.41	0.02604	5.207	283.33
ctDNA.Stage.II.TNMctDNA(+) & T4	36.81	0.02717	5.208	260.17

Concordance= 0.863 (se = 0.035)

Likelihood ratio test= 24.39 on 3 df, p=2e-05

Wald test = 16.23 on 3 df, p=0.001

Score (logrank) test = 30.45 on 3 df, p=1e-06

Hide

```
#Repeat analysis to compare ctDNA post-NAT (+) T1-T3 vs T4
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.postNAC == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)
```

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	25	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	8	5	16.66	11.66	NA
ctDNA.Stage.II.TNM=3	5	3	8.31	3.58	NA
ctDNA.Stage.II.TNM=4	3	3	19.09	2.40	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("2","4","1","3"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:
 coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	0.82248	2.27615	0.79900	1.029	0.30329
ctDNA.Stage.II.TNM1	-2.78333	0.06183	0.84513	-3.293	0.00099 ***
ctDNA.Stage.II.TNM3	0.86495	2.37489	0.80811	1.070	0.28447

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	2.27615	0.4393	0.4754	10.897
ctDNA.Stage.II.TNM1	0.06183	16.1727	0.0118	0.324
ctDNA.Stage.II.TNM3	2.37489	0.4211	0.4873	11.575

Concordance= 0.863 (se = 0.035)
 Likelihood ratio test= 24.39 on 3 df, p=2e-05
 Wald test = 16.23 on 3 df, p=0.001
 Score (logrank) test = 30.45 on 3 df, p=1e-06

#RFS by ctDNA post-NAT & ypN

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TN
M, data = circ_data)

```

Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.Stage.II.TNM, data = circ_data)

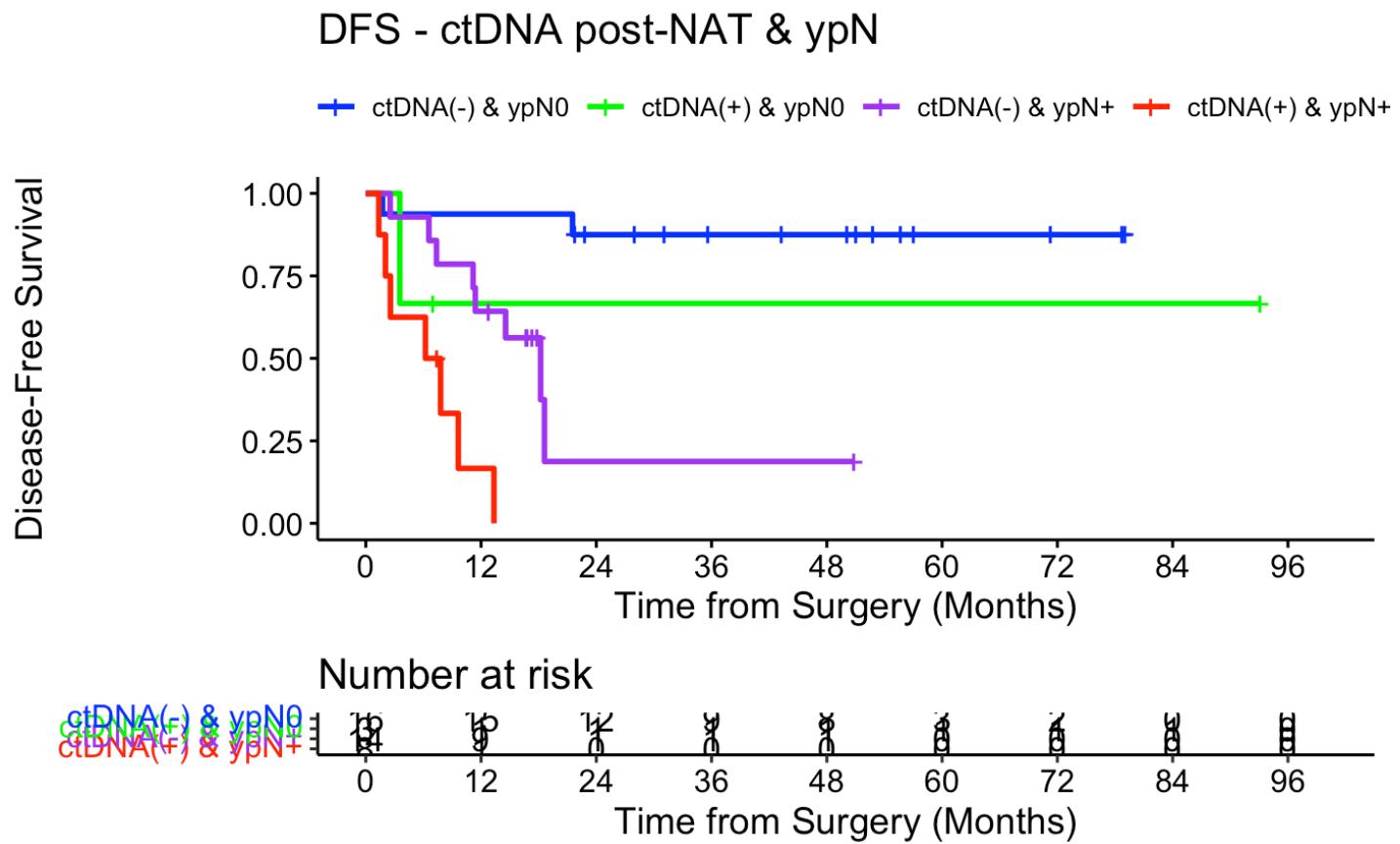
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	16	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	3	1	NA	3.55	NA
ctDNA.Stage.II.TNM=3	14	8	18.2	11.40	NA
ctDNA.Stage.II.TNM=4	8	7	7.0	2.56	NA

Hide

```

surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.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 post-
NAT & ypN", ylab= "Disease-Free Survival", xlab="Time from Surgery (Months)", legend.lab
s=c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"), legend.
title="")

```



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

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

  ctDNA.Stage.II.TNM=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      16      0     1.000  0.0000      1.000      1.000
   24      12      2     0.875  0.0827      0.586      0.967

  ctDNA.Stage.II.TNM=2
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      3      0     1.000  0.000      1.0000      1.000
   24      1      1     0.667  0.272      0.0541      0.945

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      14      0     1.000  0.00      1.0000      1.000
   24      1      8     0.188  0.16      0.0107      0.539

  ctDNA.Stage.II.TNM=4
  time      n.risk      n.event      survival      std.err      lower 95% CI      upper 95% CI
  0          8          0            1            0            1
```

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
  ("1","2","3","4"), labels = c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 18

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & ypN0	1.8009	6.0550	1.2294	1.465	0.14295
ctDNA.Stage.II.TNMctDNA(-) & ypN+	2.3395	10.3756	0.8452	2.768	0.00564 **
ctDNA.Stage.II.TNMctDNA(+) & ypN+	3.7725	43.4875	0.9323	4.046	5.2e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & ypN0	6.055	0.16515	0.5441	67.38
ctDNA.Stage.II.TNMctDNA(-) & ypN+	10.376	0.09638	1.9794	54.39
ctDNA.Stage.II.TNMctDNA(+) & ypN+	43.488	0.02300	6.9948	270.37

Concordance= 0.773 (se = 0.058)

Likelihood ratio test= 23.01 on 3 df, p=4e-05

Wald test = 17.09 on 3 df, p=7e-04

Score (logrank) test = 26.38 on 3 df, p=8e-06

Hide

```
#Repeat analysis to compare ctDNA post-NAT (-) vs (+) in ypN+
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TNM, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	16	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	3	1	NA	3.55	NA
ctDNA.Stage.II.TNM=3	14	8	18.2	11.40	NA
ctDNA.Stage.II.TNM=4	8	7	7.0	2.56	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("3","4","1","2"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)
```

n= 41, number of events= 18

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	1.43302	4.19133	0.57130	2.508	0.01213 *
ctDNA.Stage.II.TNM1	-2.33945	0.09638	0.84525	-2.768	0.00564 **
ctDNA.Stage.II.TNM2	-0.53857	0.58358	1.08425	-0.497	0.61939

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	4.19133	0.2386	1.36791	12.8424
ctDNA.Stage.II.TNM1	0.09638	10.3756	0.01839	0.5052
ctDNA.Stage.II.TNM2	0.58358	1.7136	0.06969	4.8867

Concordance= 0.773 (se = 0.058)

Likelihood ratio test= 23.01 on 3 df, p=4e-05

Wald test = 17.09 on 3 df, p=7e-04

Score (logrank) test = 26.38 on 3 df, p=8e-06

#OS by ctDNA post-NAT & ypN

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
)

circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)

```

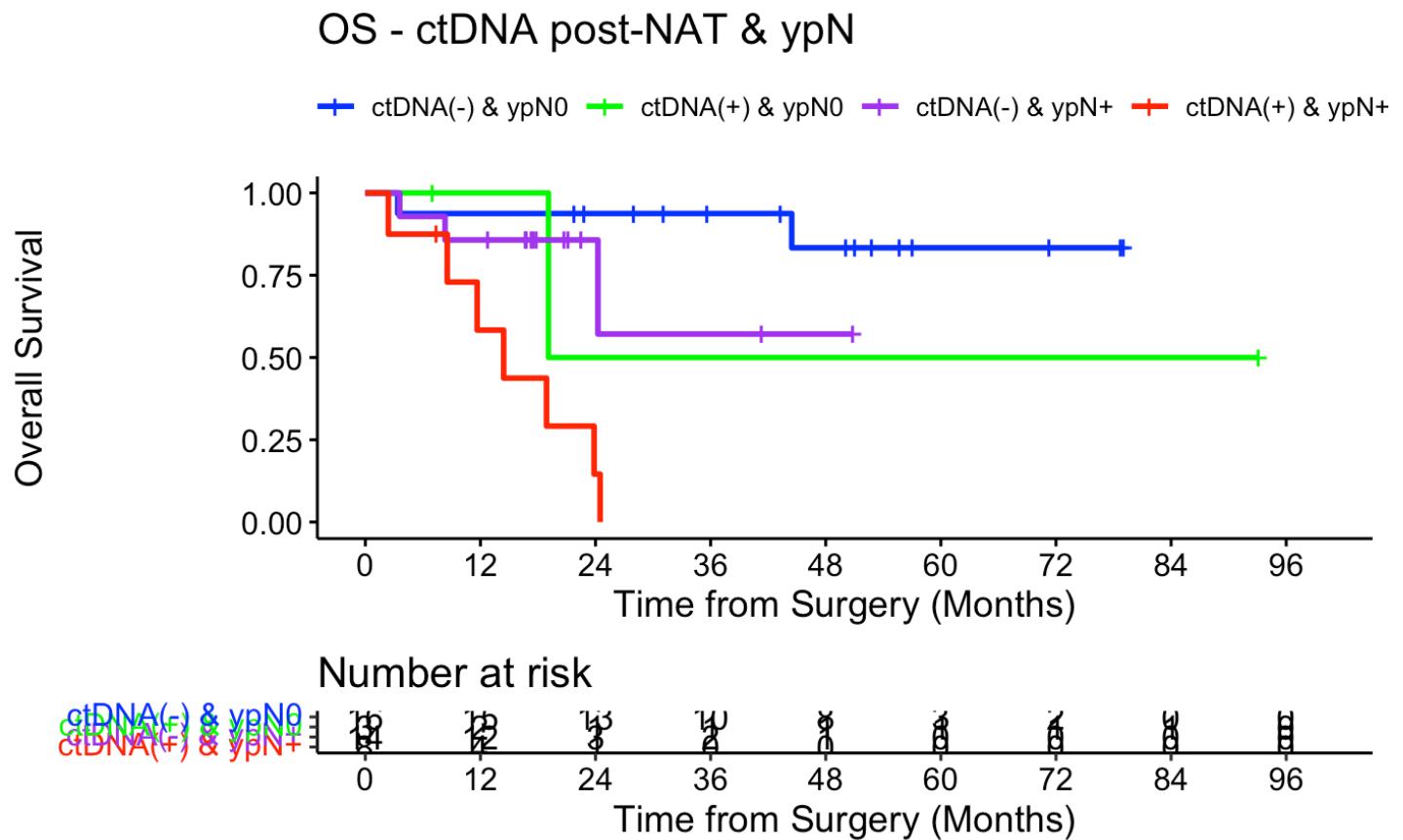
Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~
ctDNA.Stage.II.TNM, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	16	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	3	1	19.1	19.09	NA
ctDNA.Stage.II.TNM=3	14	3	NA	24.25	NA
ctDNA.Stage.II.TNM=4	8	7	14.4	8.54	NA

Hide

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.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="OS - ctDNA post-N
AT & ypN", ylab= "Overall Survival", xlab="Time from Surgery (Months)", legend.labs=c("c
tDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"), legend.title
="")
```



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

Hide

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

  ctDNA.Stage.II.TNM=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     16      0     1.000  0.0000      1.000      1.000
  24     13      1     0.938  0.0605      0.632      0.991

  ctDNA.Stage.II.TNM=2
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      3      0     1.0    0.000      1.00000      1.00
  24      1      1     0.5    0.354      0.00598      0.91

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     14      0     1.000  0.0000      1.000      1.000
  24      3      2     0.857  0.0935      0.539      0.962

  ctDNA.Stage.II.TNM=4
  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
  24      1      6     0.146  0.135      0.00726      0.471
```

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & ypN0	1.4824	4.4036	1.2274	1.208	0.227138
ctDNA.Stage.II.TNMctDNA(-) & ypN+	1.4115	4.1022	0.9495	1.487	0.137123
ctDNA.Stage.II.TNMctDNA(+) & ypN+	3.1167	22.5721	0.8727	3.571	0.000355 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & ypN0	4.404	0.2271	0.3972	48.82
ctDNA.Stage.II.TNMctDNA(-) & ypN+	4.102	0.2438	0.6380	26.38
ctDNA.Stage.II.TNMctDNA(+) & ypN+	22.572	0.0443	4.0804	124.87

Concordance= 0.765 (se = 0.076)

Likelihood ratio test= 16.79 on 3 df, p=8e-04

Wald test = 15.08 on 3 df, p=0.002

Score (logrank) test = 23.63 on 3 df, p=3e-05

Hide

```
#Repeat analysis to compare ctDNA post-NAT (+) ypN0 vs ypN+
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.postNAC!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.postNAC == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.postNAC == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c("2","4","1","3"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 41, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	1.63430	5.12585	1.11292	1.468	0.142
ctDNA.Stage.II.TNM1	-1.48242	0.22709	1.22741	-1.208	0.227
ctDNA.Stage.II.TNM3	-0.07091	0.93155	1.17530	-0.060	0.952

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	5.1259	0.1951	0.57869	45.403
ctDNA.Stage.II.TNM1	0.2271	4.4036	0.02048	2.517
ctDNA.Stage.II.TNM3	0.9315	1.0735	0.09306	9.325

Concordance= 0.765 (se = 0.076)

Likelihood ratio test= 16.79 on 3 df, p=8e-04

Wald test = 15.08 on 3 df, p=0.002

Score (logrank) test = 23.63 on 3 df, p=3e-05

#DFS by ctDNA Clearance post-NAT

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.C2D1.Clearance <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C2D1.Clearance = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~ "TRUE",
    ctDNA.Base == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~ "FALSE",
    ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.C2D1.Clearance),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.C2D1.Clearance, data = circ_data)

```

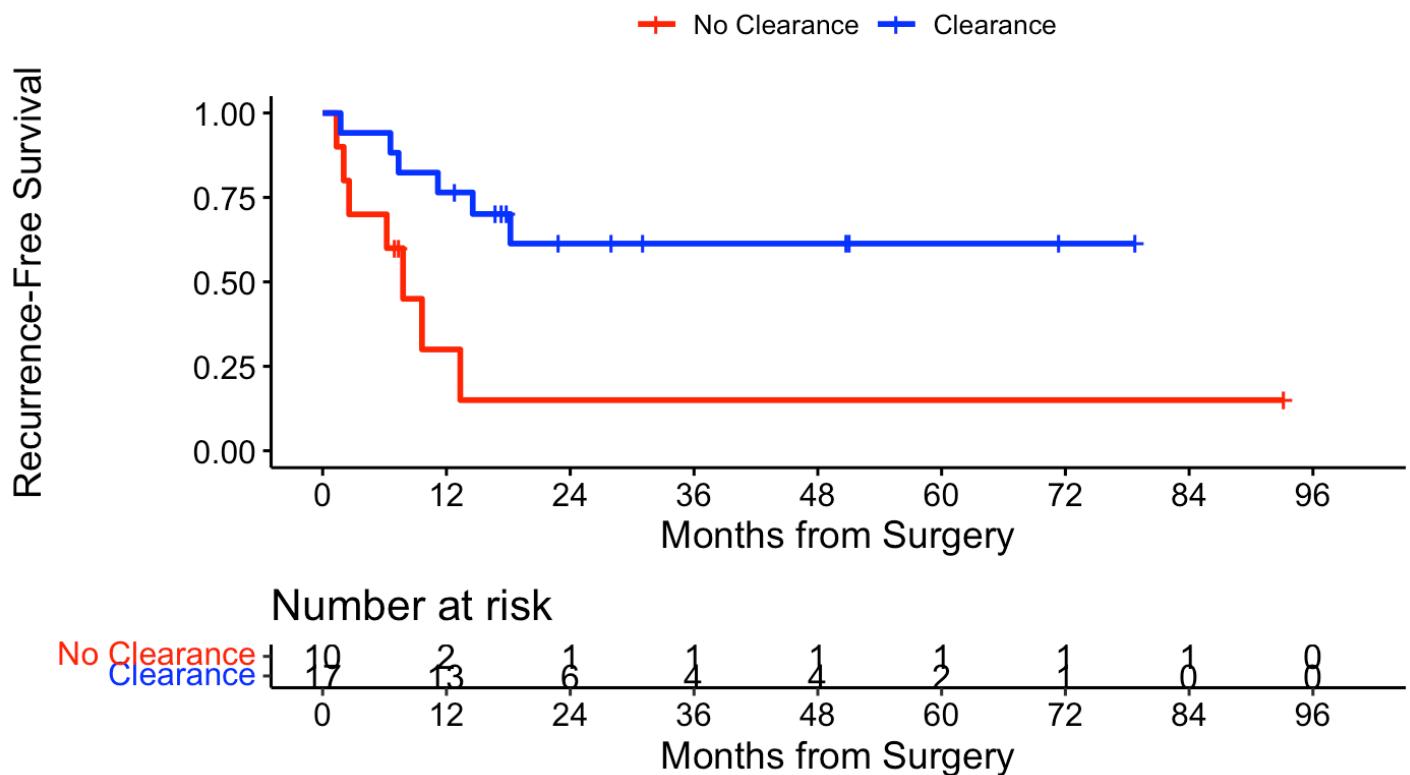
Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.C2D1.Clearance, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1.Clearance=FALSE	10	7	7.79	2.56	NA
ctDNA.C2D1.Clearance=TRUE	17	6	NA	18.20	NA

Hide

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1.Clearance, data = circ_data, conf.int=0.95, c
onf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=12, palette=c("red","blue"), title="RFS - ctDNA clearance post-NAT", yl
ab= "Recurrence-Free Survival", xlab="Months from Surgery", legend.labs=c("No Clearanc
e", "Clearance"), legend.title="")
```

RFS - ctDNA clearance post-NAT



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

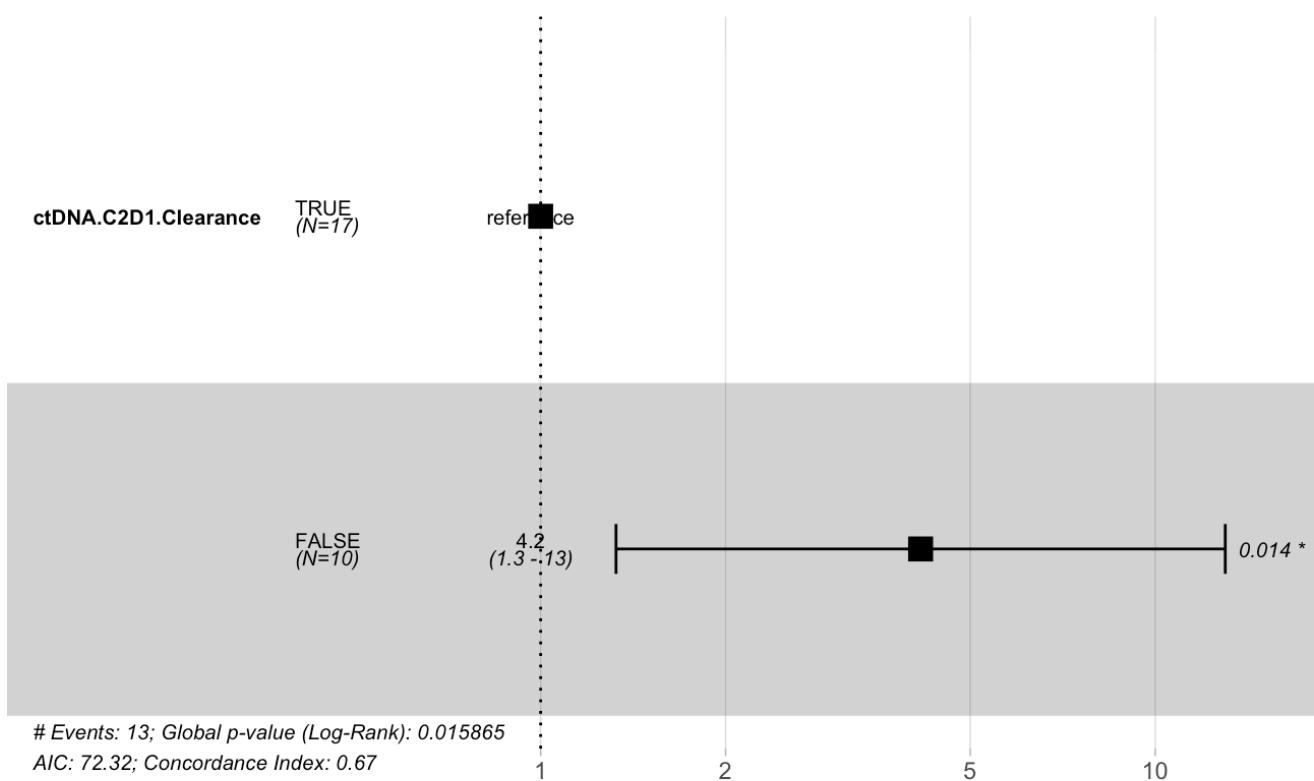
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

CI	ctDNA.C2D1.Clearance=FALSE						
	time	n.risk	n.event	survival	std.err	lower	95% CI upper
24.00000	1.00000	7.00000	0.15000	0.13555	0.00802	0.474	

CI	ctDNA.C2D1.Clearance=TRUE						
	time	n.risk	n.event	survival	std.err	lower	95% CI upper
24.00	6.000	6.000	0.613	0.128	0.324	0.8	

```
circ_data$ctDNA.C2D1.Clearance <- factor(circ_data$ctDNA.C2D1.Clearance, levels=c("TRUE","FALSE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1.Clearance, data=circ_data)
ggforest(cox_fit, data = circ_data)
```

Hazard ratio



```
summary(cox_fit)
```

Call:
`coxph(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data)`

n= 27, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C2D1.Clearance	FALSE	1.4237	4.1523	0.5823	2.445 0.0145 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C2D1.Clearance	FALSE	4.152	0.2408	1.326 13

Concordance= 0.668 (se = 0.064)
 Likelihood ratio test= 5.82 on 1 df, p=0.02
 Wald test = 5.98 on 1 df, p=0.01
 Score (logrank) test = 6.85 on 1 df, p=0.009

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 4.15 (1.33-13); p = 0.014"
```

#OS by ctDNA Clearance post-NAT

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

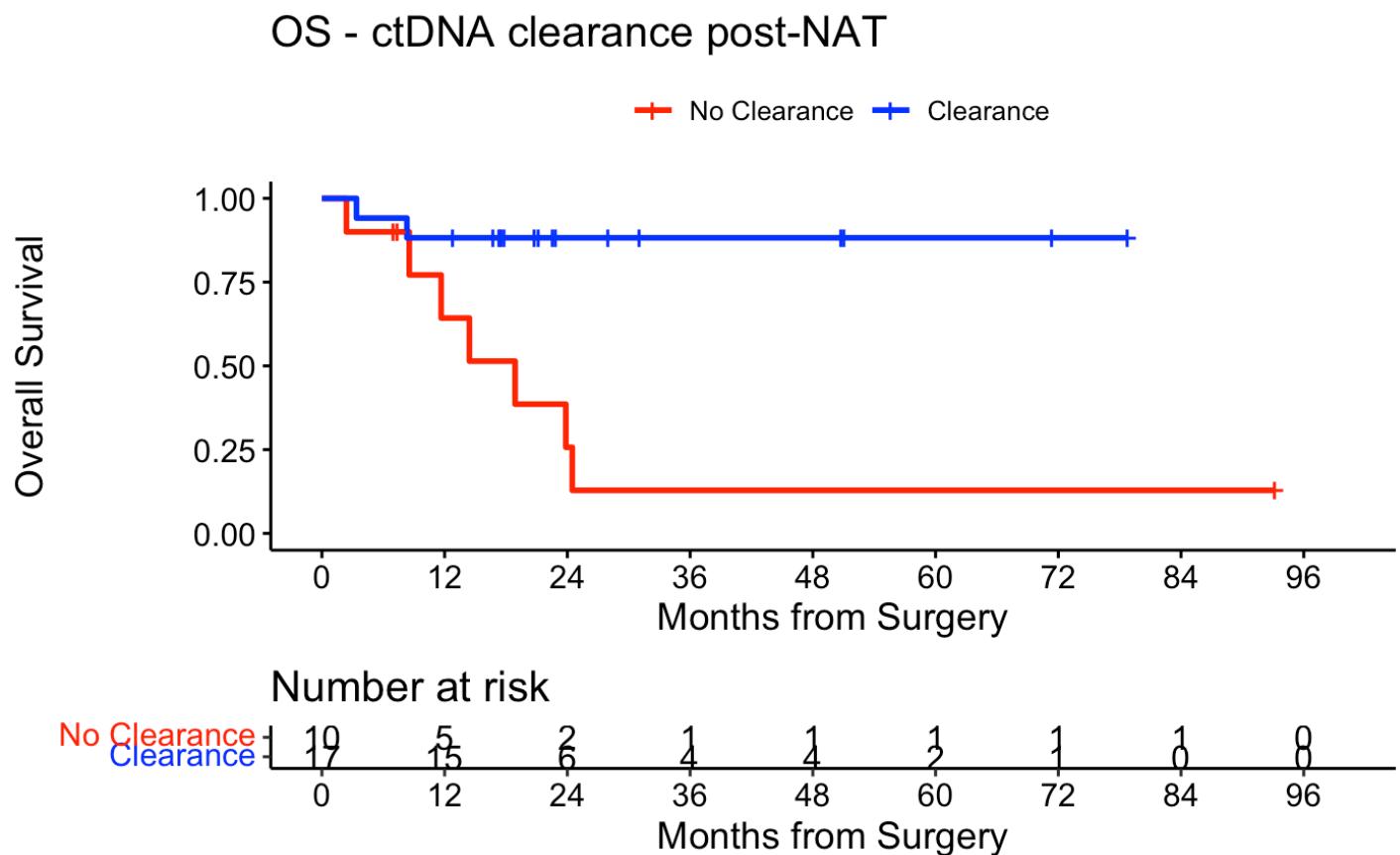
circ_data$ctDNA.C2D1.Clearance <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.C2D1.Clearance = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~ "TRUE",
    ctDNA.Base == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~ "FALSE",
    ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.C2D1.Clearance),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.C2D1.Clearance, data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~ ctDNA.C2D1.Clearance, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.C2D1.Clearance=FALSE	10	7	18.9	11.7	NA
ctDNA.C2D1.Clearance=TRUE	17	2	NA	NA	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.C2D1.Clearance, 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("red","blue"), title="OS - ctDNA clearance post-NAT", ylab= "Overall Survival", xlab="Months from Surgery", legend.labs=c("No Clearance", "Clearance"), legend.title="")
```



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

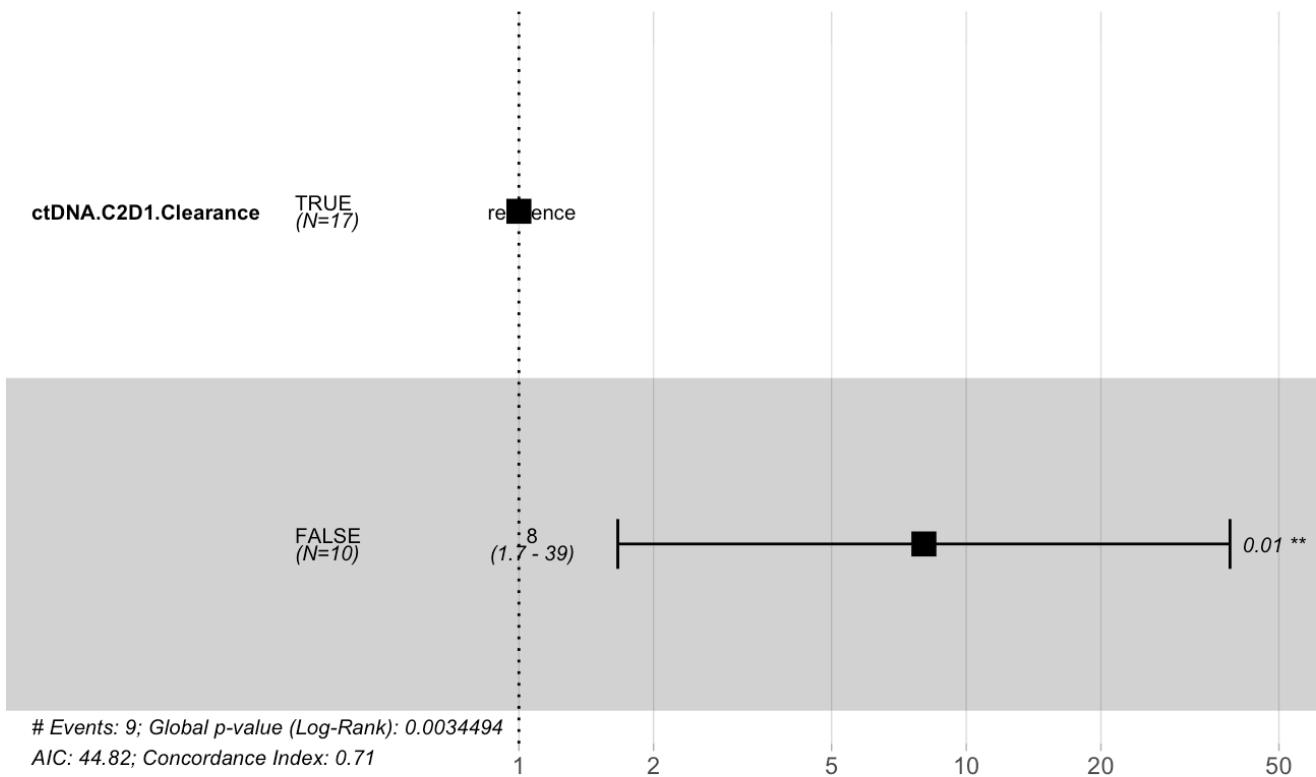
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

CI	ctDNA.C2D1.Clearance=FALSE							
	time	n.risk	n.event	survival	std.err	lower	95% CI upper	95%
79	24.0000	2.0000	6.0000	0.2571	0.1560	0.0384	0.56	
CI	ctDNA.C2D1.Clearance=TRUE							
	time	n.risk	n.event	survival	std.err	lower	95% CI upper	95%
92	24.0000	6.0000	2.0000	0.8824	0.0781	0.6060	0.96	

Hide

```
circ_data$ctDNA.C2D1.Clearance <- factor(circ_data$ctDNA.C2D1.Clearance, levels=c("TRUE", "FALSE"))
cox_fit <- coxph(surv_object ~ ctDNA.C2D1.Clearance, data=circ_data)
ggforest(cox_fit, data = circ_data)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

Call:

```
coxph(formula = surv_object ~ ctDNA.C2D1.Clearance, data = circ_data)
```

n= 27, number of events= 9

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.C2D1.Clearance	2.0846	8.0411	0.8039	2.593	0.00951 **

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.C2D1.Clearance	8.041	0.1244	1.664	38.87

Concordance= 0.709 (se = 0.089)

Likelihood ratio test= 8.55 on 1 df, p=0.003

Wald test = 6.72 on 1 df, p=0.01

Score (logrank) test = 9.48 on 1 df, p=0.002

[Hide](#)

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

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

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

[1] "HR = 8.04 (1.66-38.87); p = 0.01"

#Association of ctDNA Dynamics post-NAT and Response

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/--",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~
    "ctDNA +/+/+"
  ))
#Vertical Fisher plot for ctDNA clearance post-NAT and Rec Status
circ_data <- circ_data[circ_data$ctDNA.Dynamics!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Dynamics, data = circ_data)

```

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

40 observations deleted due to missingness

	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=ctDNA +/--	5	1	NA	18.20	NA
ctDNA.Dynamics=ctDNA +/+-	10	5	14.55	7.36	NA
ctDNA.Dynamics=ctDNA +/+/+	7	6	6.21	2.04	NA

Hide

```

circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("ctDNA +/--", "ctDNA +/+-", "ctDNA +/+/+"))
circ_data$TRG <- factor(circ_data$TRG, levels = c("TRG1/2/3", "TRG4/5"))
contingency_table <- table(circ_data$ctDNA.Dynamics, circ_data$TRG)
fisher_exact_test <- fisher.test(contingency_table)
chi_square_test <- chisq.test(contingency_table)

```

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

Hide

```
print(chi_square_test)
```

```
Pearson's Chi-squared test
```

```
data: contingency_table
X-squared = 11.378, df = 2, p-value = 0.003383
```

[Hide](#)

```
print(fisher_exact_test)
```

```
Fisher's Exact Test for Count Data
```

```
data: contingency_table
p-value = 0.002557
alternative hypothesis: two.sided
```

[Hide](#)

```
print(contingency_table)
```

	TRG1/2/3	TRG4/5
ctDNA +/-/-	4	0
ctDNA +/+/-	6	4
ctDNA +/+/+	0	7

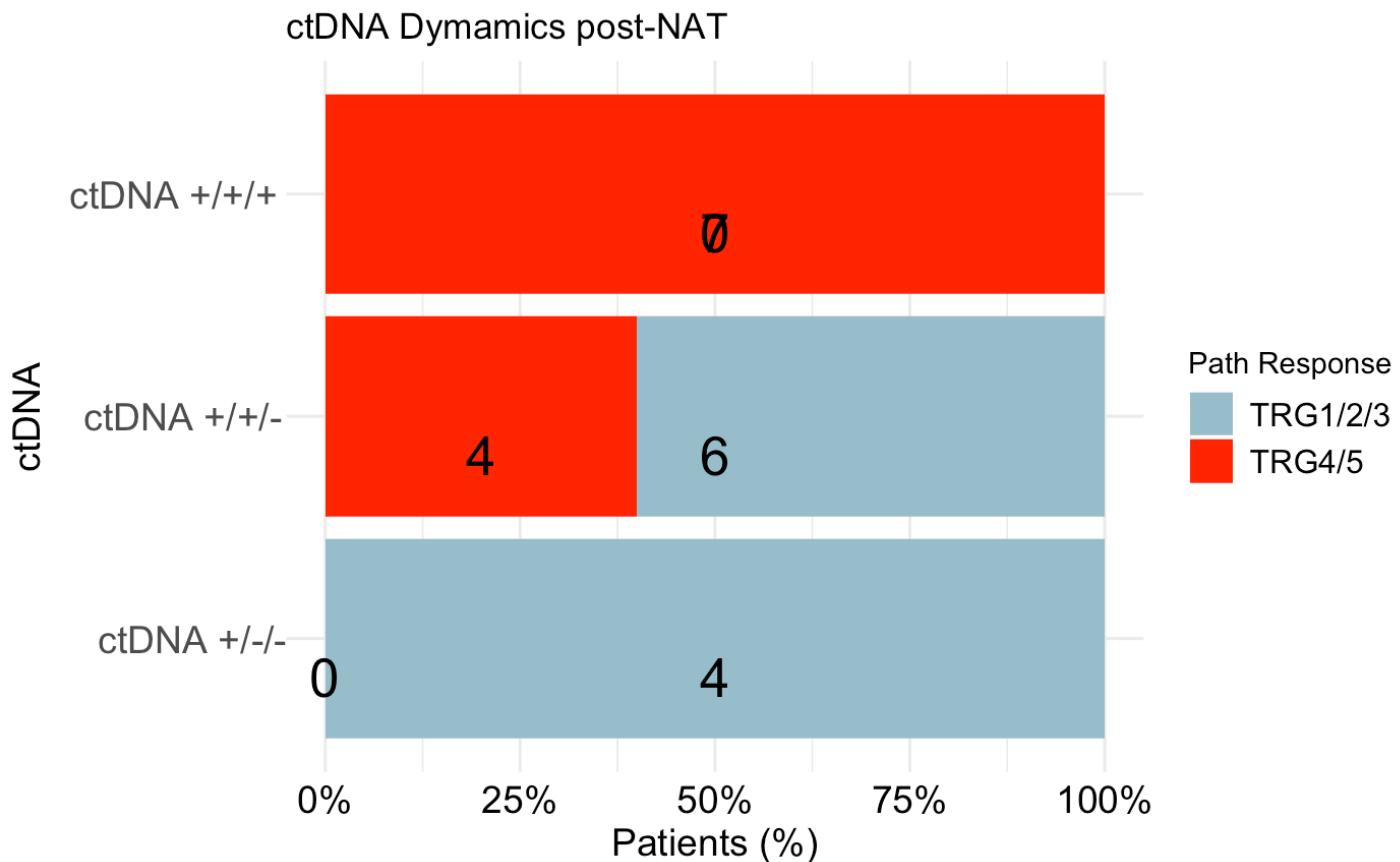
[Hide](#)

```

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

# Swapping x and y in ggplot function to make bar plot vertical
ggplot(table_df, aes(y = Var1, x = Percentage, fill = Var2)) +
  geom_bar(stat = "identity") +
  geom_text(aes(x = MiddlePercentage, label = Freq), position = "stack", color = "black", vjust = 1.5, size = 7) +
  theme_minimal() +
  labs(title = "ctDNA Dynamics post-NAT", y = "ctDNA", x = "Patients (%)", fill = "Path Response") +
  scale_x_continuous(labels = scales::percent_format()) +
  scale_fill_manual(values = c("TRG1/2/3" = "lightblue3", "TRG4/5" = "red")) + # define custom colors
  theme(axis.text.y = element_text(angle = 0, hjust = 1.5, size = 14), # increase y-axis text size
        axis.text.x = element_text(size = 14, color = "black"), # increase x-axis text size
        axis.title.y = element_text(size = 14, color = "black"), # increase y-axis label size
        axis.title.x = element_text(size = 14, color = "black"), # increase x-axis label size
        legend.text = element_text(size = 12, color = "black")) # increase Recurrence label size

```



Hide

```
#Calculating p-value with Fisher exact test for the ctDNA +/+/+ vs ctDNA +/+-/
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~
    "ctDNA +/+/",
    ))
#Vertical Fisher plot for ctDNA clearance post-NAT and Rec Status
circ_data <- circ_data[circ_data$ctDNA.Dynamics!="",]
circ_data <- circ_data[circ_data$ctDNA.Dynamics!="ctDNA +/+-",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Dynamics, data = circ_data)
```

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

```
40 observations deleted due to missingness
      n  events median 0.95LCL 0.95UCL
ctDNA.Dynamics=ctDNA +/+- 10      5  14.55    7.36     NA
ctDNA.Dynamics=ctDNA +/+/+  7      6   6.21    2.04     NA
```

Hide

```
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("ctDNA +/+-", "ctDNA +/+/"))
circ_data$TRG <- factor(circ_data$TRG, levels = c("TRG1/2/3", "TRG4/5"))
contingency_table <- table(circ_data$ctDNA.Dynamics, circ_data$TRG)
fisher_exact_test <- fisher.test(contingency_table)
chi_square_test <- chisq.test(contingency_table)
```

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

Hide

```
print(chi_square_test)
```

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

```
data: contingency_table
X-squared = 4.1295, df = 1, p-value = 0.04214
```

[Hide](#)

```
print(fisher_exact_test)
```

```
Fisher's Exact Test for Count Data
```

```
data: contingency_table
p-value = 0.0345
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.176016      Inf
sample estimates:
odds ratio
      Inf
```

[Hide](#)

```
print(contingency_table)
```

```
TRG1/2/3 TRG4/5
```

ctDNA +/+-	6	4
ctDNA +/+/+	0	7

```
#DFS by ctDNA Dynamics post-NAT
```

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/--",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~
    "ctDNA +/++"
  ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.Dynamics),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Dynamics, data = circ_data)

```

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

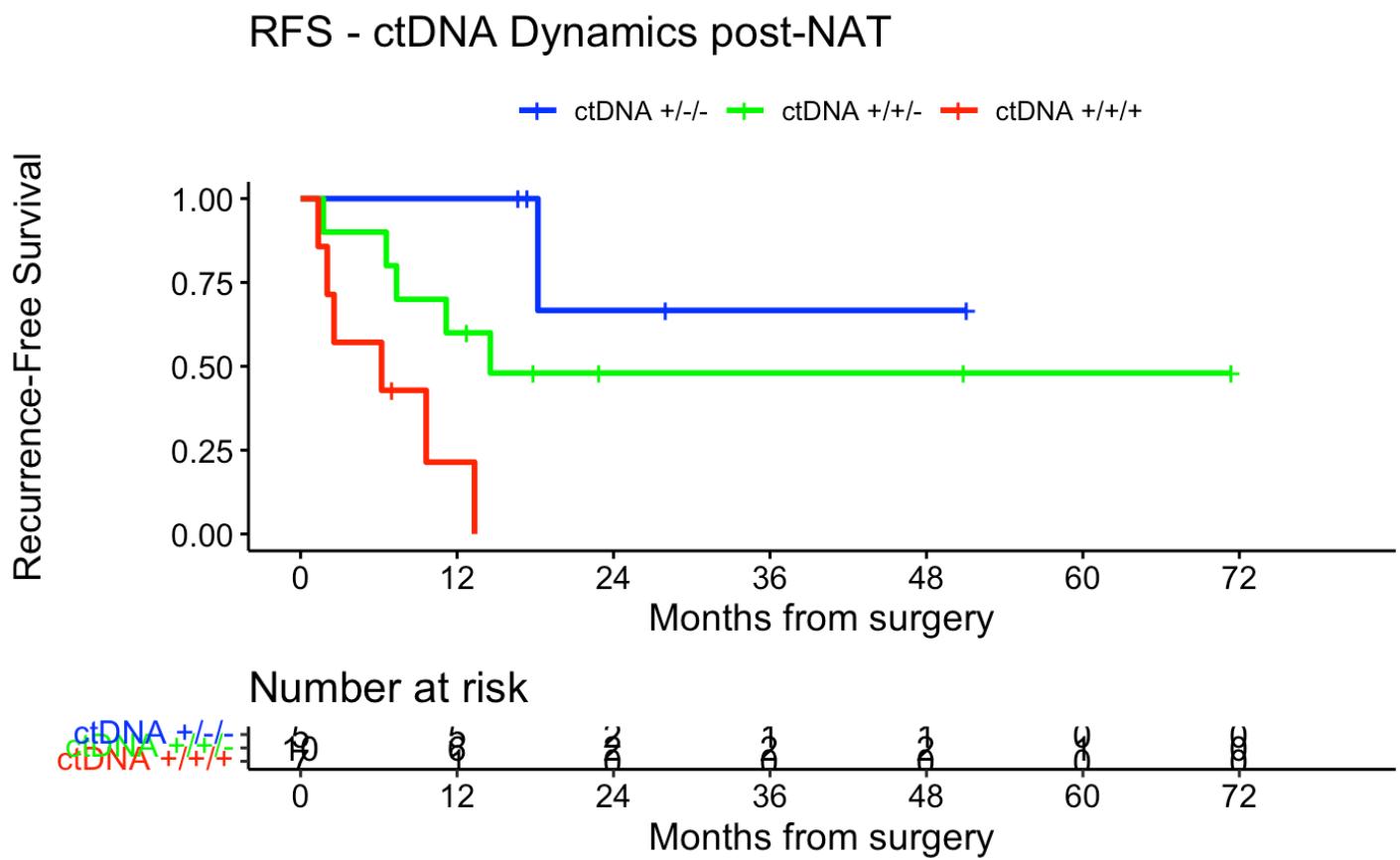
	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=ctDNA +/--	5	1	NA	18.20	NA
ctDNA.Dynamics=ctDNA +/+-	10	5	14.55	7.36	NA
ctDNA.Dynamics=ctDNA +/++	7	6	6.21	2.04	NA

Hide

```

surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.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=12, palette=c("blue","green","red"), title="RFS - ctDNA Dynamics post-NAT",
ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA +/--","ctDNA +/+-", "ctDNA +/++"), legend.title="")

```



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

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

```
ctDNA.Dynamics=ctDNA +/--
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0      5      0    1.000  0.000      1.0000    1.000
 24      2      1    0.667  0.272      0.0541    0.945
```

```
ctDNA.Dynamics=ctDNA +/+-
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     10      0    1.00  0.000      1.000    1.000
 24      2      5    0.48  0.164      0.161    0.745
```

```
ctDNA.Dynamics=ctDNA +/+/+
time      n.risk      n.event      survival      std.err      lower 95% CI      upper 95% CI
  0          7          0          1          0          1
```

[Hide](#)

```
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("ctDNA +/-/-", "ctDNA +/+-", "ctDNA +/+/+"))
cox_fit <- coxph(surv_object ~ ctDNA.Dynamics, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 22, number of events= 12

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.DynamicsctDNA +/--	1.343	3.831	1.100	1.221	0.2219
ctDNA.DynamicsctDNA +/+/+	2.921	18.568	1.157	2.526	0.0116 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.DynamicsctDNA +/--	3.831	0.26100	0.4439	33.07
ctDNA.DynamicsctDNA +/+/+	18.568	0.05386	1.9237	179.22

Concordance= 0.756 (se = 0.053)

Likelihood ratio test= 10.01 on 2 df, p=0.007

Wald test = 8.49 on 2 df, p=0.01

Score (logrank) test = 11.47 on 2 df, p=0.003

#OS by ctDNA Dynamics post-NAT

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/-/-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~
    "ctDNA +/+/+"
  ))
  
circ_data <- circ_data[!is.na(circ_data$ctDNA.Dynamics),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Dynamics, data = circ_data)

```

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

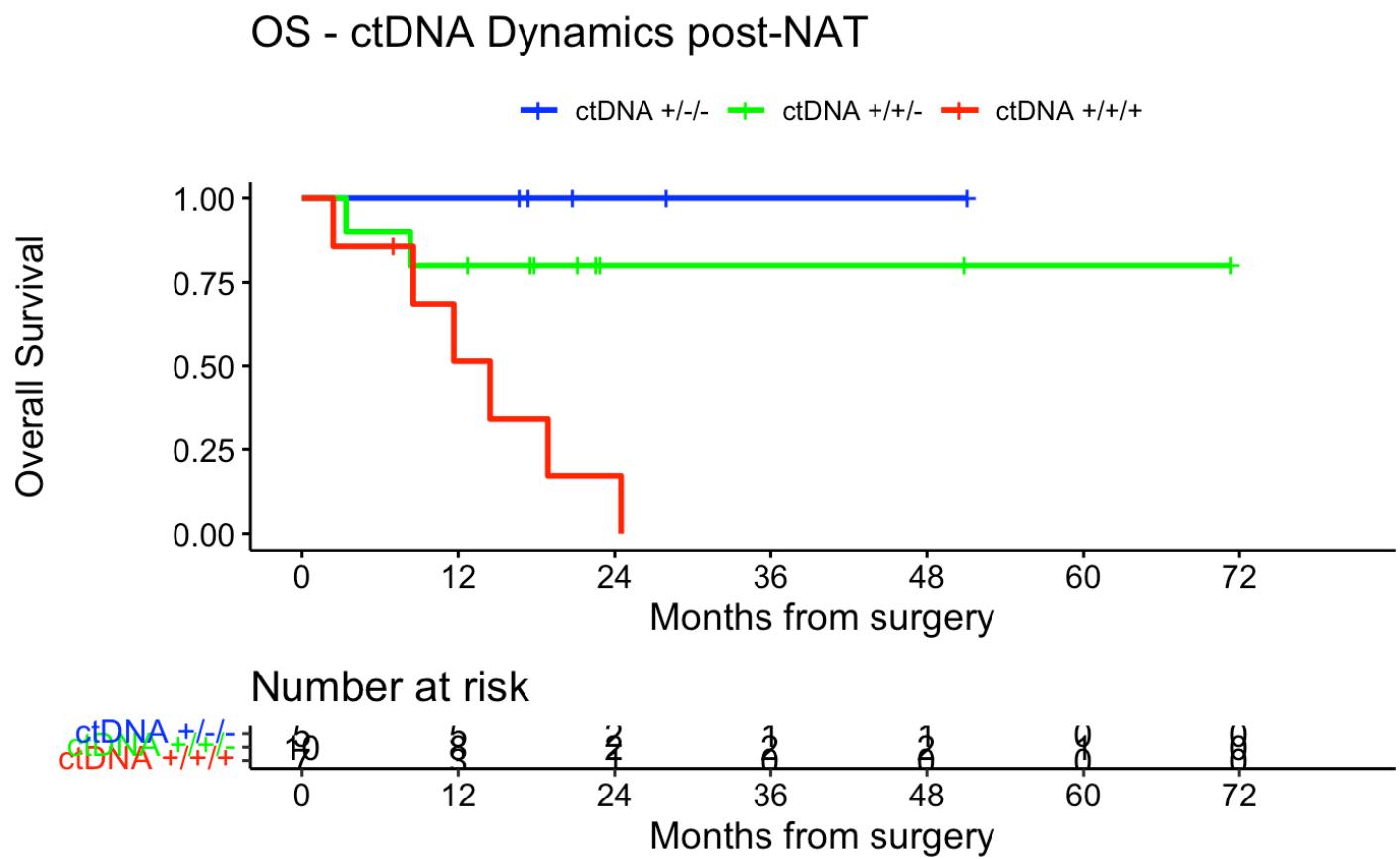
	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=ctDNA +/-/-	5	0	NA	NA	NA
ctDNA.Dynamics=ctDNA +/+-	10	2	NA	NA	NA
ctDNA.Dynamics=ctDNA +/+/+	7	6	14.4	8.54	NA

Hide

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$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=12, palette=c("blue","green","red"), title="OS - ctDNA Dynamics post-NA T",
ylab= "Overall Survival", xlab="Months from surgery", legend.labs=c("ctDNA +/-/-", "ctDNA +/+-", "ctDNA +/+/+"), legend.title="")

```



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

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

  ctDNA.Dynamics=ctDNA +/--
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0      5      0      1      0      1      1
  24     2      0      1      0      NA      NA

  ctDNA.Dynamics=ctDNA +/+-
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     10      0     1.0  0.000  1.000  1.000
  24     2      2     0.8  0.126  0.409  0.946

  ctDNA.Dynamics=ctDNA +/+/+
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0      7      0     1.000  0.000  1.00000  1.000
  24     1      5     0.171  0.156  0.00794  0.526
```

[Hide](#)

```
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("ctDNA +/-/-", "ctDNA +/+-", "ctDNA +/+/+"))
cox_fit <- coxphf(surv_object ~ ctDNA.Dynamics, data=circ_data)
summary(cox_fit)
```

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

Model fitted by Penalized ML

Confidence intervals and p-values by Profile Likelihood

	coef	se(coef)	exp(coef)	lower 0.95	upper 0.95	Chisq
p						
ctDNA.Dynamics	ctDNA +/- 1.178268	1.696304	3.248742	0.2627946	448.914	0.7281482 0.
	393484095					
ctDNA.Dynamics	ctDNA +/+/+ 2.776185	1.609915	16.057642	1.8668569	2102.174	7.2774206 0.
	006982673					

Likelihood ratio test=8.847995 on 2 df, p=0.01198622, n=22

Wald test = 5.979249 on 2 df, p = 0.05030632

Covariance-Matrix:

	ctDNA.Dynamics	ctDNA +/-	ctDNA.Dynamics	ctDNA +/+/+
ctDNA.Dynamics		2.877449		2.403504
ctDNA +/-			2.403504	2.591825

#Levels of Baseline MTM/mL in ctDNA Dynamics post-NAT categories

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.Base!=""]
circ_data <- as.data.frame(circ_data)

circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/--",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/+-",
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "POSITIVE" & ctDNA.postNAC == "POSITIVE" ~
    "ctDNA +/++"
  ))
  
# Transform ctDNA.Base.MTM with log10
circ_data <- subset(circ_data, !is.na(ctDNA.Dynamics))
circ_data$ctDNA.Base.MTM <- as.numeric(as.character(circ_data$ctDNA.Base.MTM))
circ_data$ctDNA.Dynamics <- factor(circ_data$ctDNA.Dynamics, levels=c("ctDNA +/--","ctDNA +/+-", "ctDNA +/++"))
median_Base_MTM <- aggregate(ctDNA.Base.MTM ~ ctDNA.Dynamics, data = circ_data, FUN = median)
print(median_Base_MTM)

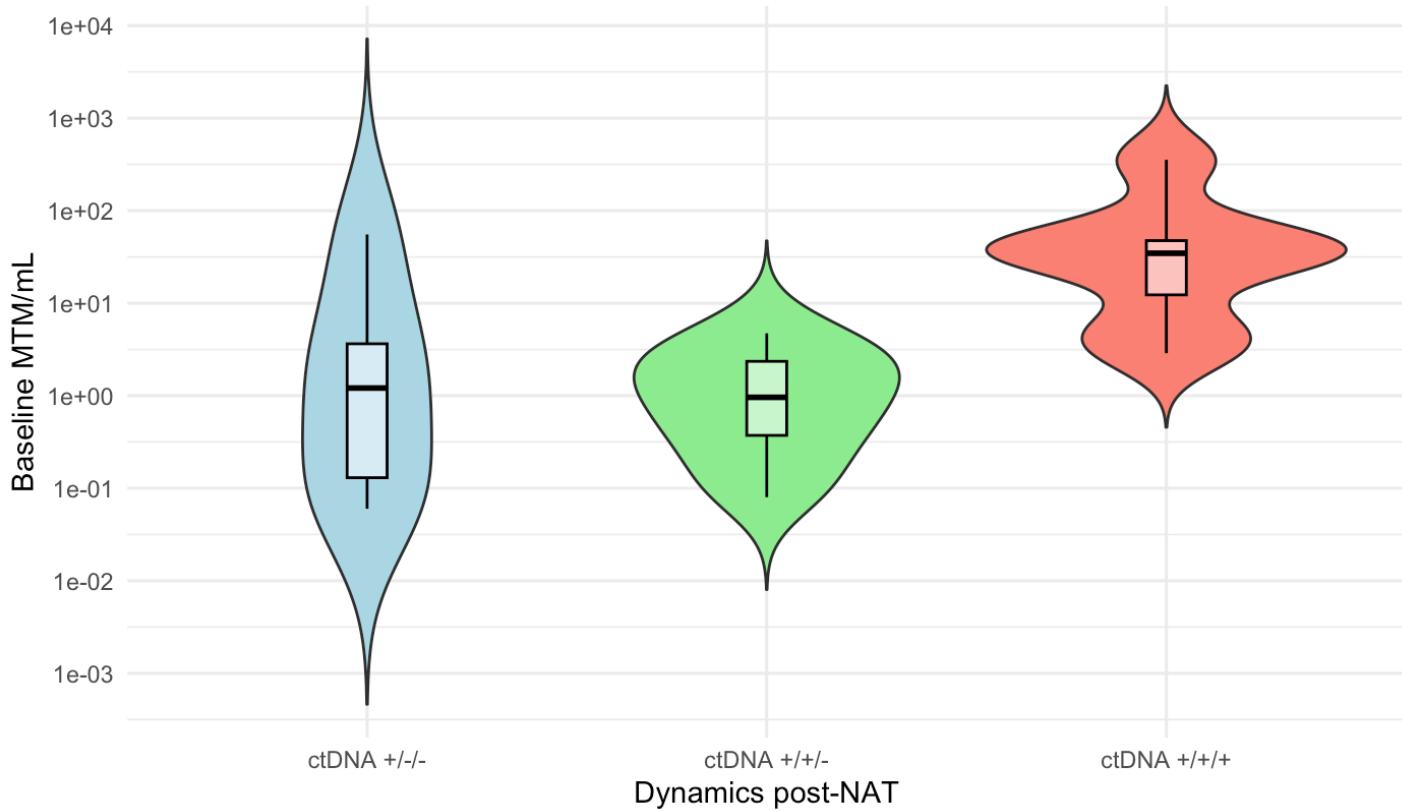
```

ctDNA.Dynamics	ctDNA.Base.MTM
<fctr>	<dbl>
ctDNA +/--	1.21
ctDNA +/+-	0.96
ctDNA +/++	34.63
3 rows	

Hide

```
# Create violin plot with log10 scale on y-axis
ggplot(circ_data, aes(x=ctDNA.Dynamics, y=ctDNA.Base.MTM, fill=ctDNA.Dynamics)) +
  geom_violin(trim=FALSE) +
  scale_fill_manual(values=c("ctDNA +/−/"="lightblue", "ctDNA +/+/−"="lightgreen", "ctDNA +/+/+"="salmon")) +
  geom_boxplot(width=0.1, fill="white", colour="black", alpha=0.5) +
  scale_y_log10(breaks=c(0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000)) +
  labs(title="Baseline MTM/mL | Dynamics post-NAT", x="Dynamics post-NAT", y="Baseline MTM/mL") +
  theme_minimal() +
  theme(legend.position="none")
```

Baseline MTM/mL | Dynamics post-NAT



Hide

```
m3_1v2 <- wilcox.test(ctDNA.Base.MTM ~ ctDNA.Dynamics,
                        data = circ_data[circ_data$ctDNA.Dynamics %in% c("ctDNA +/−/", "ctDNA +/+/−")], ,
                        na.rm = TRUE)
```

Warning in wilcox.test.default(x = DATA[[1L]], y = DATA[[2L]], ...):
cannot compute exact p-value with ties

Hide

```
print(m3_1v2)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by ctDNA.Dynamics
W = 26, p-value = 0.9511
alternative hypothesis: true location shift is not equal to 0
```

[Hide](#)

```
m3_1v3 <- wilcox.test(ctDNA.Base.MTM ~ ctDNA.Dynamics,
                        data = circ_data[circ_data$ctDNA.Dynamics %in% c("ctDNA +/−/−", "ctDNA +/+/+"), ],
                        na.rm = TRUE)
print(m3_1v3)
```

Wilcoxon rank sum exact test

```
data: ctDNA.Base.MTM by ctDNA.Dynamics
W = 7, p-value = 0.1061
alternative hypothesis: true location shift is not equal to 0
```

[Hide](#)

```
m3_2v3 <- wilcox.test(ctDNA.Base.MTM ~ ctDNA.Dynamics,
                        data = circ_data[circ_data$ctDNA.Dynamics %in% c("ctDNA +/+/−", "ctDNA +/+/+"), ],
                        na.rm = TRUE)
```

Warning in wilcox.test.default(x = DATA[[1L]], y = DATA[[2L]], ...) :
cannot compute exact p-value with ties

[Hide](#)

```
print(m3_2v3)
```

Wilcoxon rank sum test with continuity correction

```
data: ctDNA.Base.MTM by ctDNA.Dynamics
W = 2, p-value = 0.001506
alternative hypothesis: true location shift is not equal to 0
```

#DFS by ctDNA post-NAT & TRG combination

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.pCR <- NA #first we create the variable for the ctDNA & NAC combination,
and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.postNAC == "NEGATIVE" & TRG == "TRG1/2/3" ~ "1",
    ctDNA.postNAC == "NEGATIVE" & TRG == "TRG4/5" ~ "2",
    ctDNA.postNAC == "POSITIVE" & TRG == "TRG4/5" ~ "3"
  ))
  ))

circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.pCR, data =
circ_data)

```

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

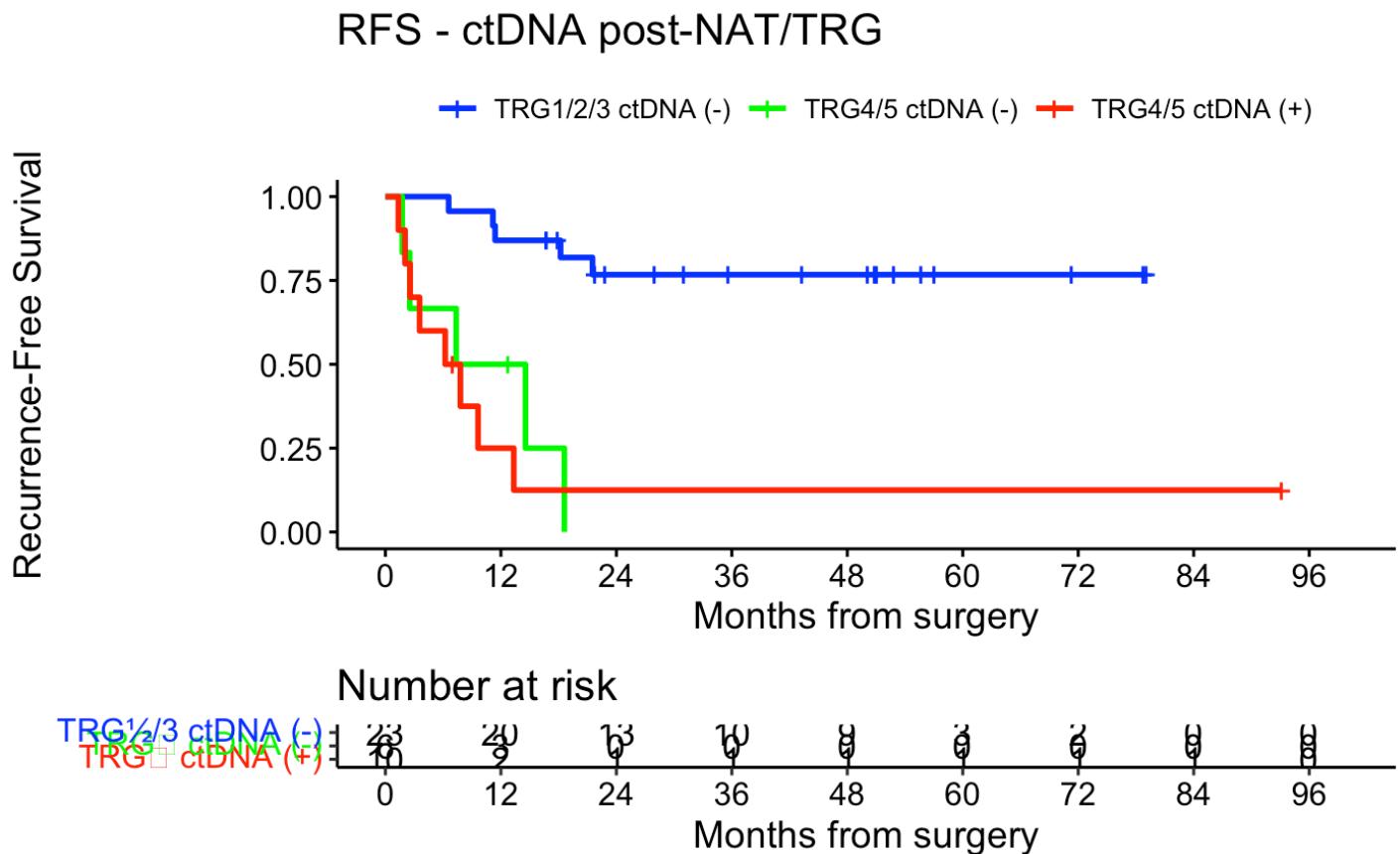
	n	events	median	0.95LCL	0.95UCL
ctDNA.pCR=1	23	5	NA	NA	NA
ctDNA.pCR=2	6	5	11	2.53	NA
ctDNA.pCR=3	10	8	7	2.56	NA

Hide

```

surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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","red"), title="RFS - ctDNA post-NAT/TRG",
ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("TRG1/2/3 ct
DNA (-)","TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"), legend.title="")

```



Hide

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

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

```
ctDNA.pCR=1
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     23      0     1.000  0.000      1.000      1.000
 24     13      5     0.767  0.092      0.526      0.897

ctDNA.pCR=2
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0       6      0       1       0       1
 1

ctDNA.pCR=3
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     10      0     1.000  0.000      1.00000      1.000
 24      1      8     0.125  0.115      0.00702      0.418
```

[Hide](#)

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2","3"), labels=c("TRG
1/2/3 ctDNA (-)", "TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"))
cox_fit <- coxph(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

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

n= 39, number of events= 18

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.pCR	2.1228	8.3545	0.6529	3.251	0.00115 **
ctDNA.pCR	2.3546	10.5336	0.5969	3.945	7.99e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.pCR	8.355	0.11970	2.323	30.04
ctDNA.pCR	10.534	0.09493	3.270	33.94

Concordance= 0.765 (se = 0.044)

Likelihood ratio test= 19.53 on 2 df, p=6e-05

Wald test = 17.01 on 2 df, p=2e-04

Score (logrank) test = 23.41 on 2 df, p=8e-06

#OS by ctDNA post-NAT & TRG combination

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]

circ_data$ctDNA.pCR <- NA #first we create the variable for the ctDNA & NAC combination,
and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.postNAC == "NEGATIVE" & TRG == "TRG1/2/3" ~ "1",
    ctDNA.postNAC == "NEGATIVE" & TRG == "TRG4/5" ~ "2",
    ctDNA.postNAC == "POSITIVE" & TRG == "TRG4/5" ~ "3"
  ))
circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.pCR, data = circ_data)

```

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

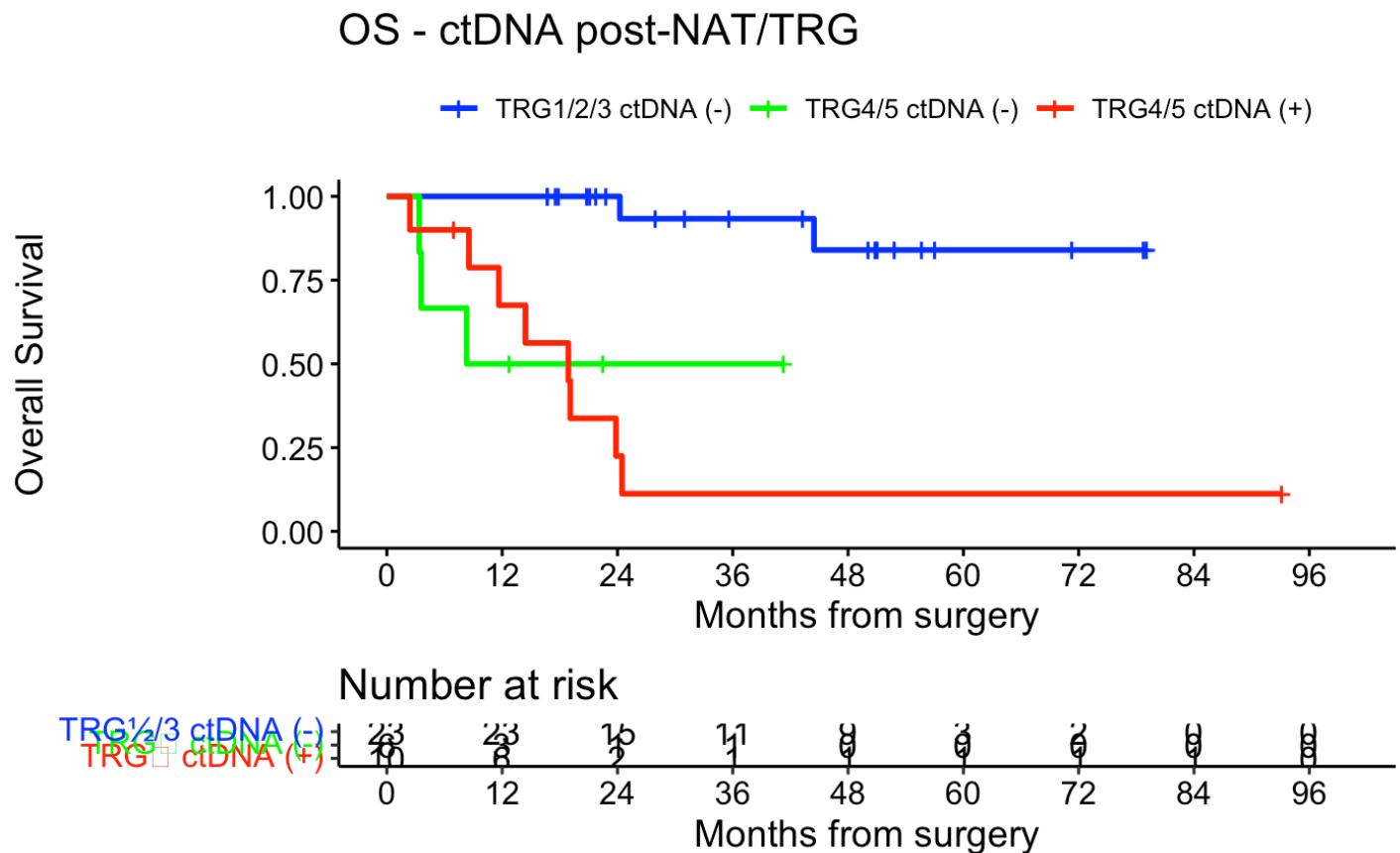
	n	events	median	0.95LCL	0.95UCL
ctDNA.pCR=1	23	2	NA	NA	NA
ctDNA.pCR=2	6	3	8.31	3.58	NA
ctDNA.pCR=3	10	8	18.89	11.66	NA

[Hide](#)

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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","red"), title="OS - ctDNA post-NAT/TRG", ylab= "Overall Survival", xlab="Months from surgery", legend.labs=c("TRG1/2/3 ctDNA (-)", "TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)", legend.title="")

```



Hide

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

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

  ctDNA.pCR=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     23      0       1       0           1           1
  24     15      0       1       0           NA          NA

  ctDNA.pCR=2
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      6      0     1.0  0.000      1.000      1.000
  24      1      3     0.5  0.204      0.111      0.804

  ctDNA.pCR=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     10      0     1.000  0.00       1.0000     1.000
  24      2      7     0.225  0.14       0.0342     0.517
```

[Hide](#)

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2","3"), labels=c("TRG
1/2/3 ctDNA (-)", "TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"))
cox_fit <- coxph(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

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

n= 39, number of events= 13

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.pCRTRG4/5 ctDNA (-)	2.7343	15.3983	0.9416	2.904	0.003685 **
ctDNA.pCRTRG4/5 ctDNA (+)	2.8771	17.7633	0.8035	3.581	0.000343 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.pCRTRG4/5 ctDNA (-)	15.40	0.06494	2.432	97.49
ctDNA.pCRTRG4/5 ctDNA (+)	17.76	0.05630	3.678	85.80

Concordance= 0.811 (se = 0.04)
 Likelihood ratio test= 19.74 on 2 df, p=5e-05
 Wald test = 13.12 on 2 df, p=0.001
 Score (logrank) test = 22.59 on 2 df, p=1e-05

#DFS by ctDNA at the MRD Window - all stages

[Hide](#)

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

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

```

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

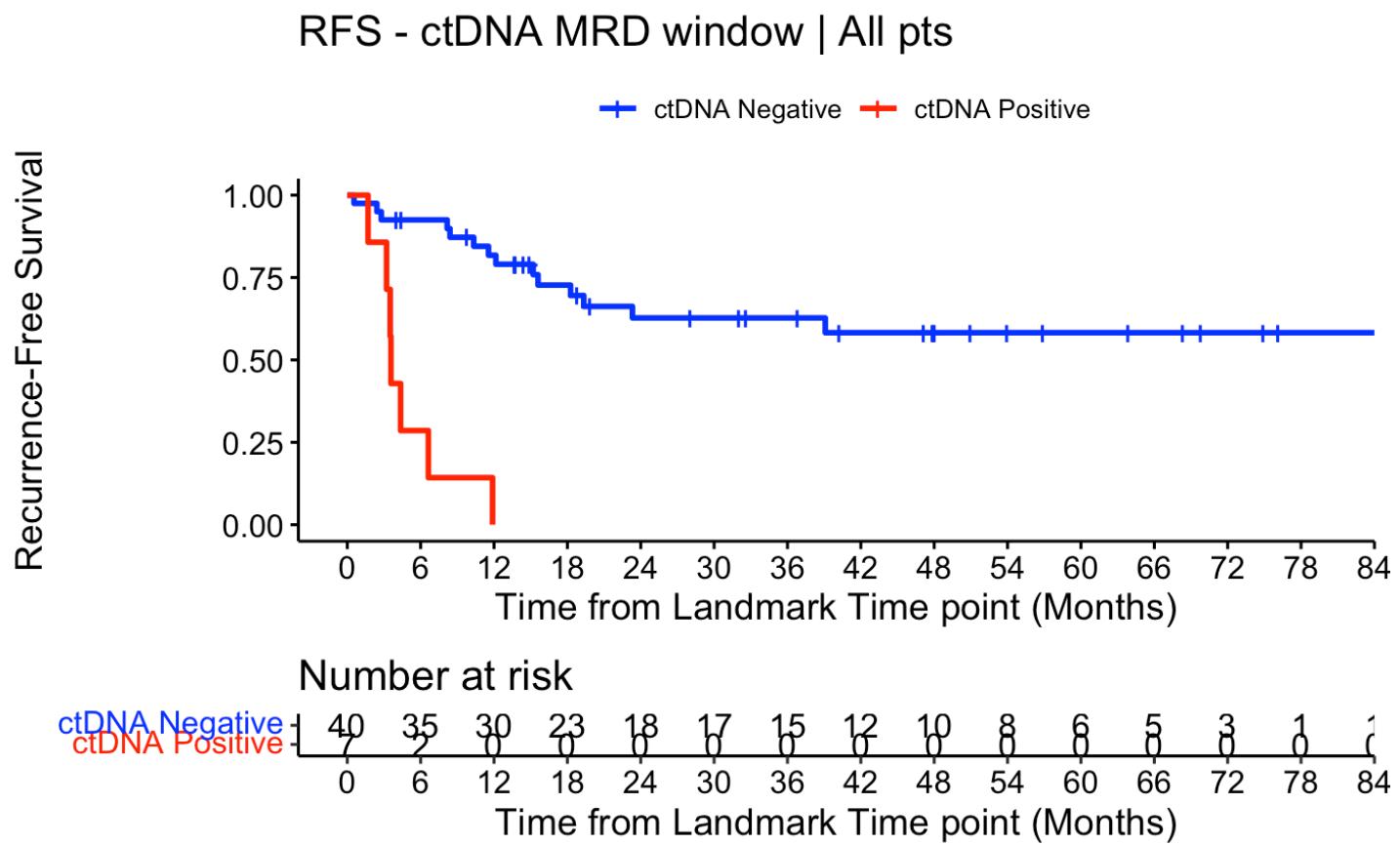
	n	events	median	0.95LCL	0.95UCL
ctDNA.MRD=NEGATIVE	40	14	NA	23.32	NA
ctDNA.MRD=POSITIVE	7	7	3.57	3.21	NA

[Hide](#)

```

surv_object <-Surv(time = circ_data$DFS.months, event = circ_data$DFS.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=6, palette=c("blue","red"), title="RFS - ctDNA MRD window | All pts", ylab= "Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")

```



```
summary(KM_curve, times= c(0, 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
0	40	0	1.000	0.0000	1.000		1.000	
24	18	13	0.628	0.0836	0.442		0.766	

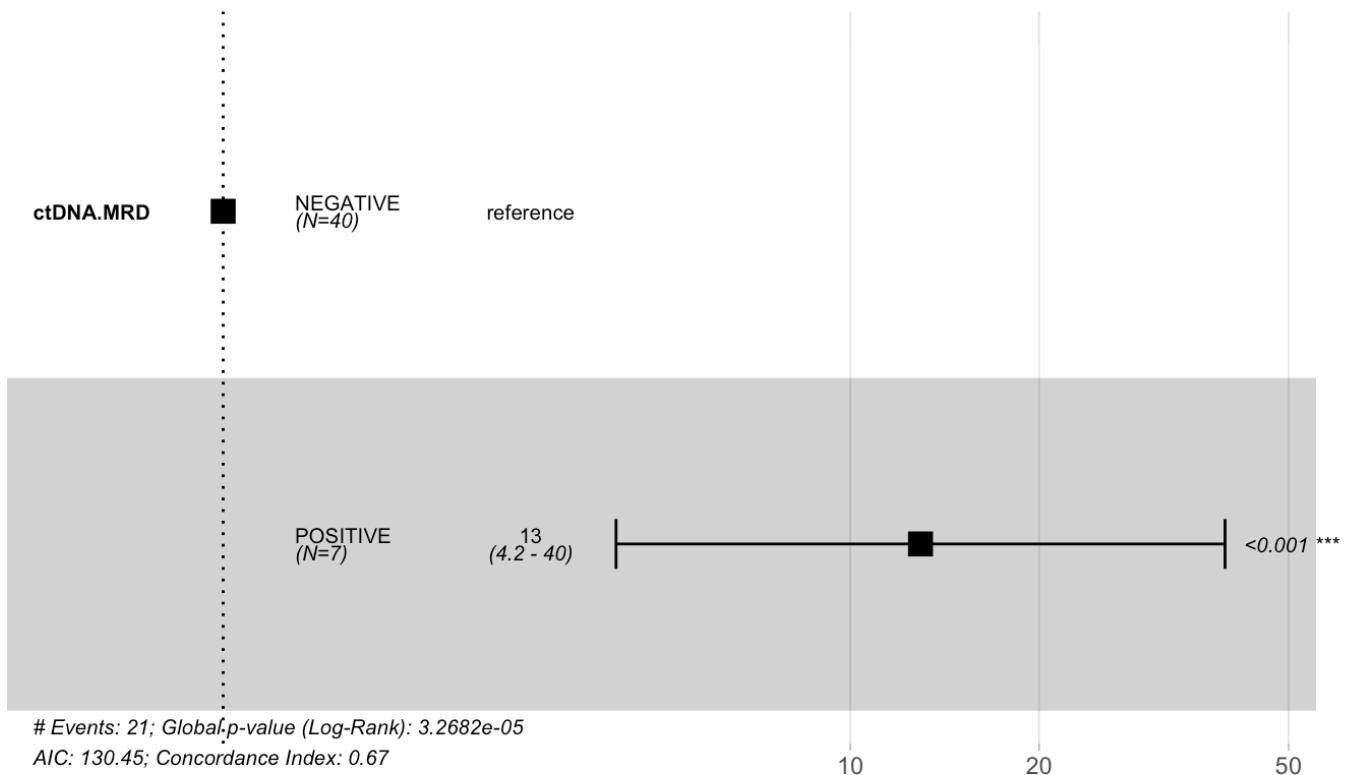
ctDNA.MRD=POSITIVE

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	7	0	1	0	0	1	1	
1								

[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)
```

Hazard ratio



[Hide](#)

```
summary(cox_fit)
```

Call:

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

n= 47, number of events= 21

            coef exp(coef)  se(coef)      z Pr(>|z|)    
ctDNA.MRDPOSITIVE 2.5604    12.9415   0.5704  4.489 7.17e-06 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

            exp(coef) exp(-coef) lower .95 upper .95    
ctDNA.MRDPOSITIVE     12.94      0.07727     4.231     39.59 

Concordance= 0.666  (se = 0.047 ) 
Likelihood ratio test= 17.26  on 1 df,  p=3e-05 
Wald test             = 20.15  on 1 df,  p=7e-06 
Score (logrank) test = 31.36  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_CI, 2), "); p = ", round(p_value, 3))
print(label_text)
```

```
[1] "HR = 12.94 (4.23-39.59); p = 0"
```

#OS by ctDNA at the MRD Window - all stages

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]
circ_datadf <- as.data.frame(circ_data)

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

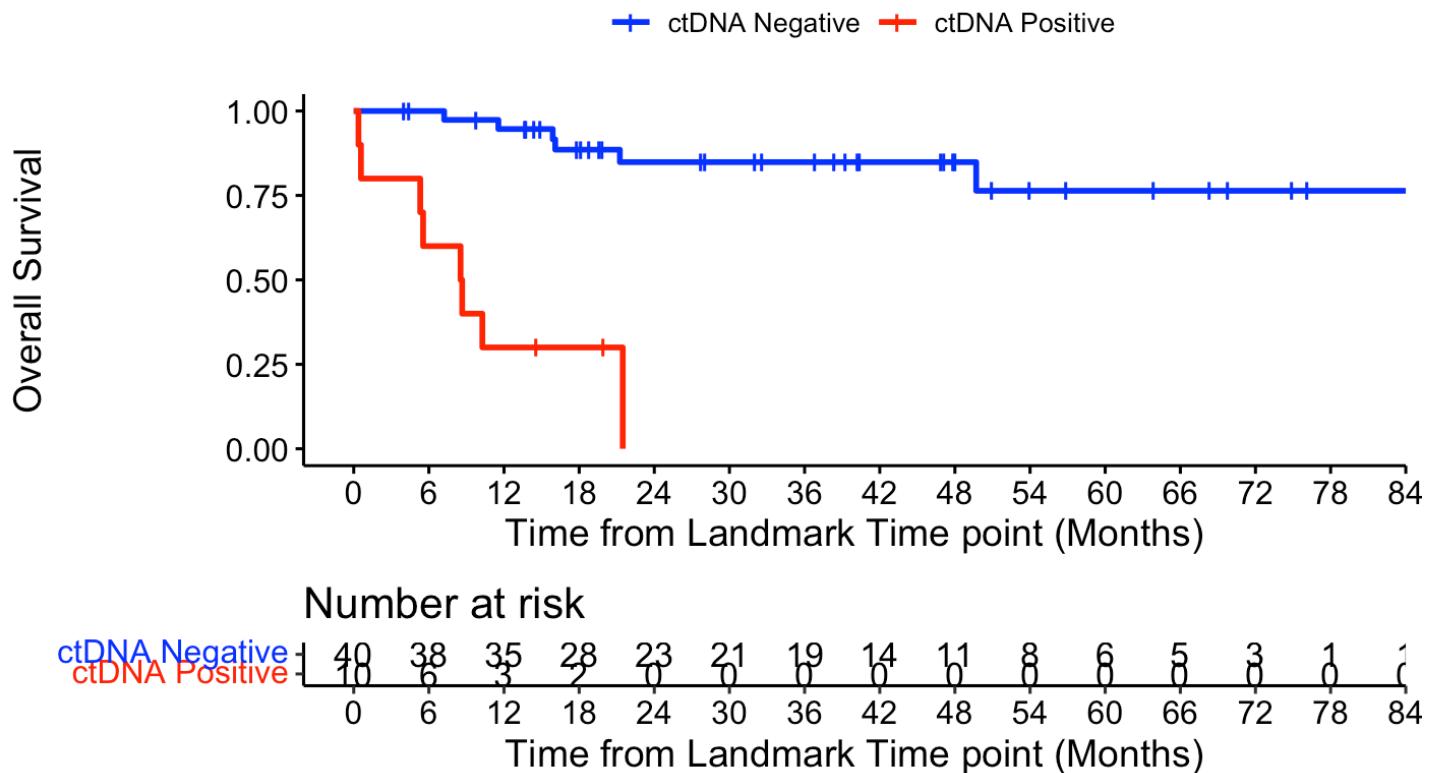
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.MRD, data = circ_data)

      n  events  median  0.95LCL  0.95UCL
ctDNA.MRD=NEGATIVE 40      6      NA      NA
ctDNA.MRD=POSITIVE 10      8     8.59     5.31      NA
```

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$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=6, palette=c("blue","red"), title="OS - ctDNA MRD window | All pts", yl
ab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ct
DNA Negative", "ctDNA Positive"), legend.title="")
```

OS - ctDNA MRD window | All pts



Hide

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

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

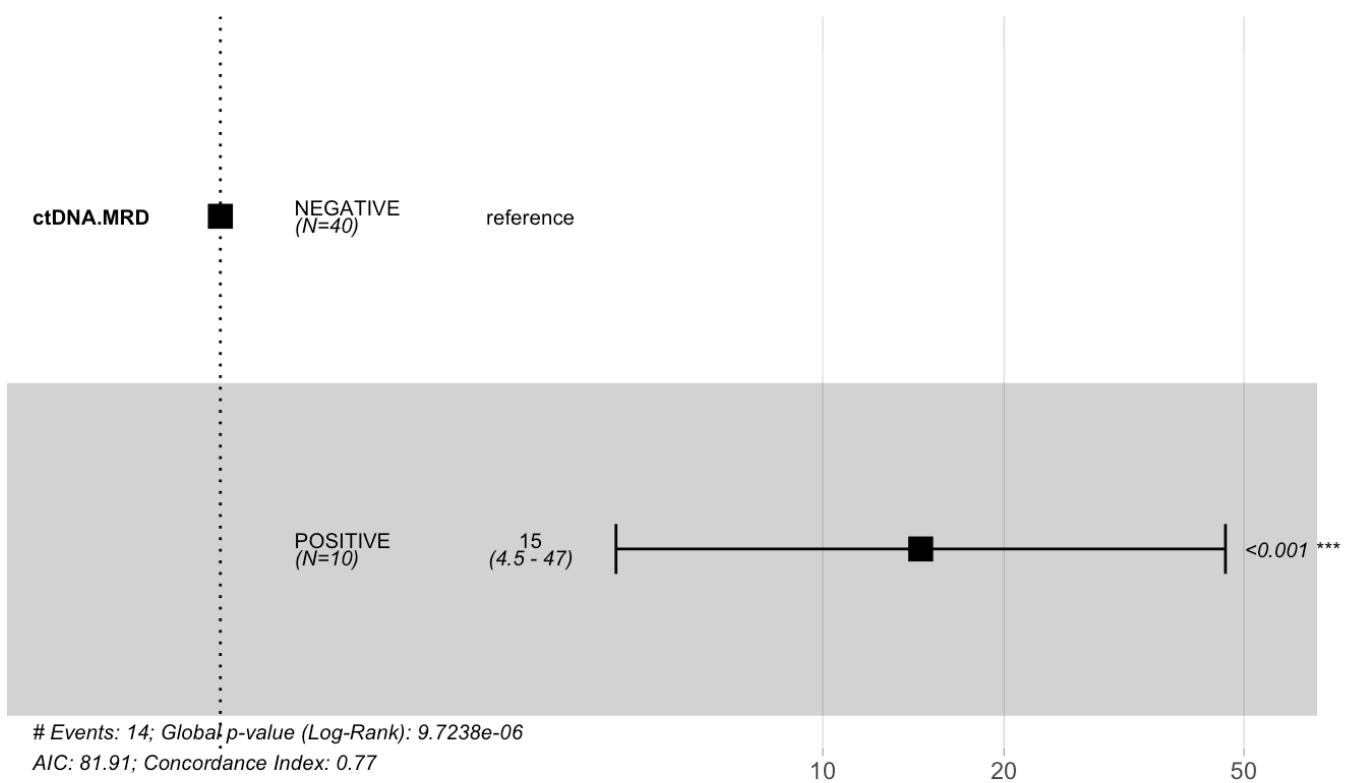
```
ctDNA.MRD=NEGATIVE
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     40      0    1.000  0.0000    1.000    1.000
 24     23      5    0.849  0.0632    0.671    0.935
```

```
ctDNA.MRD=POSITIVE
time      n.risk      n.event      survival      std.err      lower 95% CI      upper 95% CI
  0          10          0            1            0            1
```

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)
```

Hazard ratio



Hide

```
summary(cox_fit)
```

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

n= 50, number of events= 14

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.MRDPOSITIVE	2.6769	14.5406	0.5942	4.505	6.64e-06 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.MRDPOSITIVE	14.54	0.06877	4.537	46.6

Concordance= 0.771 (se = 0.055)
 Likelihood ratio test= 19.56 on 1 df, p=1e-05
 Wald test = 20.29 on 1 df, p=7e-06
 Score (logrank) test = 33.05 on 1 df, p=9e-09

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 = 14.54 (4.54-46.6); p = 0"
```

#RFS by ctDNA at the MRD Window & ypT

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TNM, data = circ_data)
```

Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.Stage.II.TNM, data = circ_data)

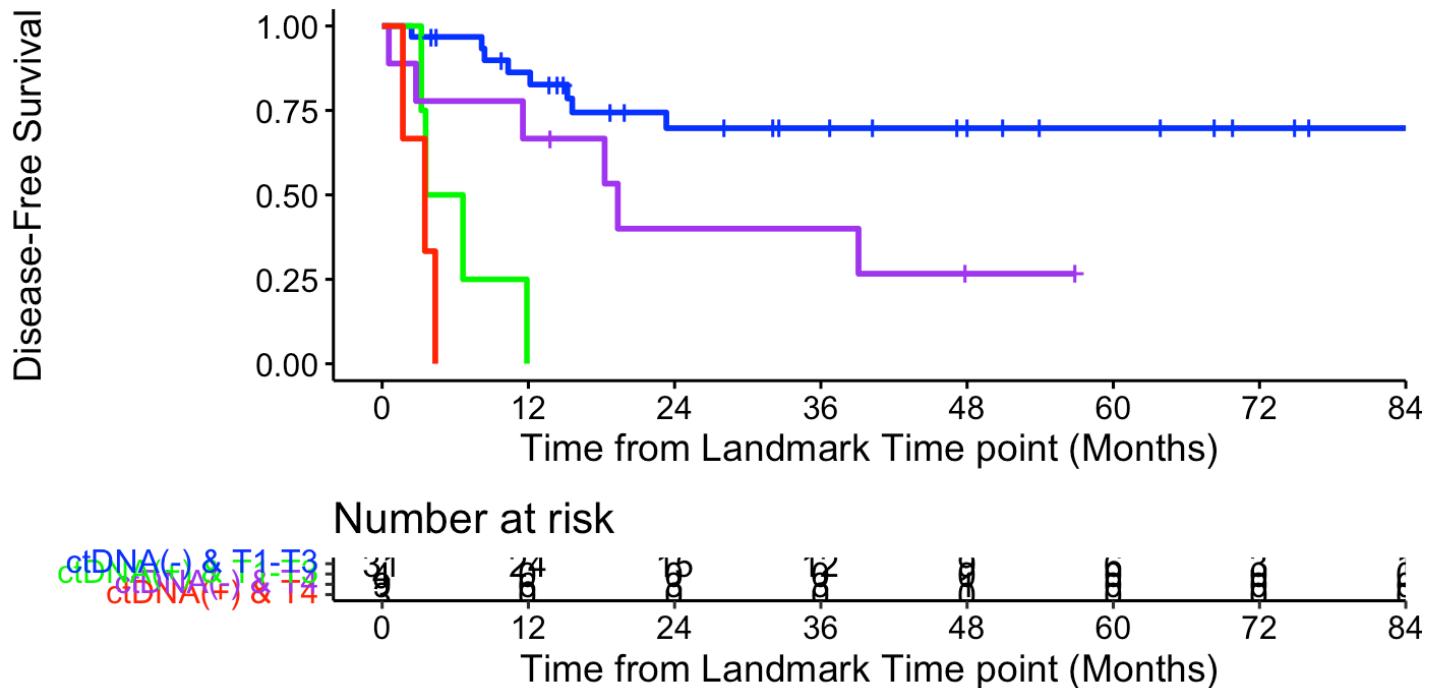
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	31	8	NA	NA	NA
ctDNA.Stage.II.TNM=2	4	4	5.10	3.21	NA
ctDNA.Stage.II.TNM=3	9	6	19.34	11.55	NA
ctDNA.Stage.II.TNM=4	3	3	3.51	1.70	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, 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 MRD & ypTN",
ylab= "Disease-Free Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.title="")
```

DFS - ctDNA MRD & ypTN

— ctDNA(-) & T1-T3 — ctDNA(+) & T1-T3 — ctDNA(-) & T4 — ctDNA(+) & T4



Hide

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

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

  ctDNA.Stage.II.TNM=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
   24     15      8     0.698  0.091      0.481      0.838

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           4           0           1           0           1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     9      0     1.0  0.000      1.0000      1.000
   24     3      5     0.4  0.174      0.0981      0.697

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           3           0           1           0           1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4",
"ctDNA(+) & T4"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 47, number of events= 21

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	2.6818	14.6110	0.6876	3.900	9.62e-05 ***
ctDNA.Stage.II.TNMctDNA(-) & T4	1.0893	2.9722	0.5407	2.015	0.0439 *
ctDNA.Stage.II.TNMctDNA(+) & T4	3.4934	32.8977	0.8353	4.182	2.89e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	14.611	0.06844	3.796	56.233
ctDNA.Stage.II.TNMctDNA(-) & T4	2.972	0.33645	1.030	8.576
ctDNA.Stage.II.TNMctDNA(+) & T4	32.898	0.03040	6.399	169.126

Concordance= 0.731 (se = 0.051)

Likelihood ratio test= 21.92 on 3 df, p=7e-05

Wald test = 22.7 on 3 df, p=5e-05

Score (logrank) test = 37.96 on 3 df, p=3e-08

Hide

```
#Repeat analysis to compare ctDNA MRD (-) vs (+) in T4
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TNM, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

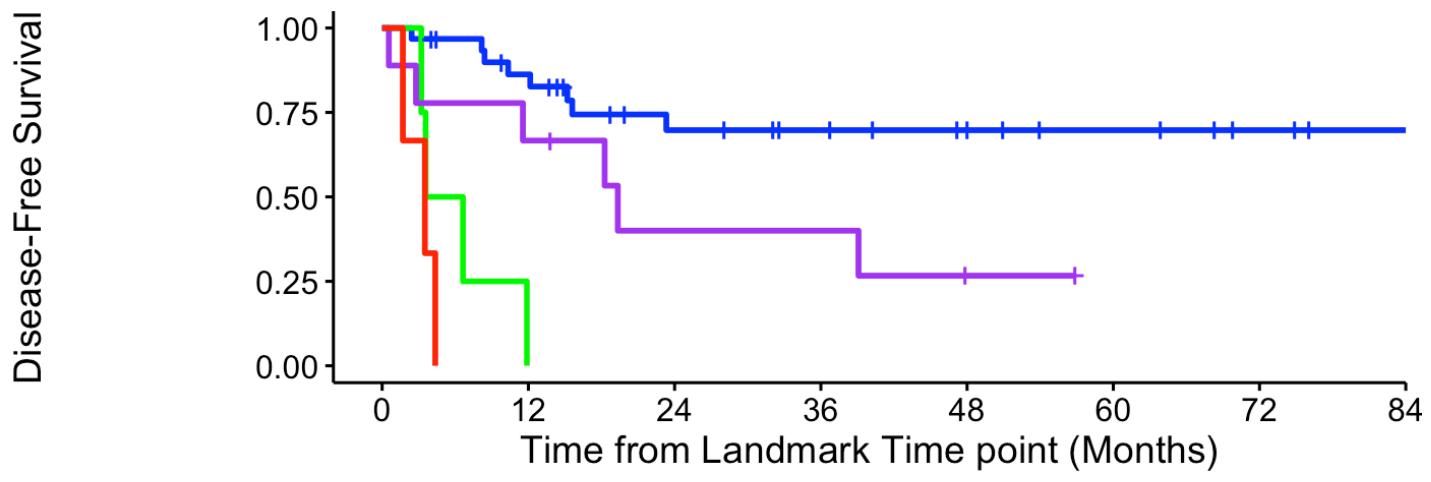
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	31	8	NA	NA	NA
ctDNA.Stage.II.TNM=2	4	4	5.10	3.21	NA
ctDNA.Stage.II.TNM=3	9	6	19.34	11.55	NA
ctDNA.Stage.II.TNM=4	3	3	3.51	1.70	NA

Hide

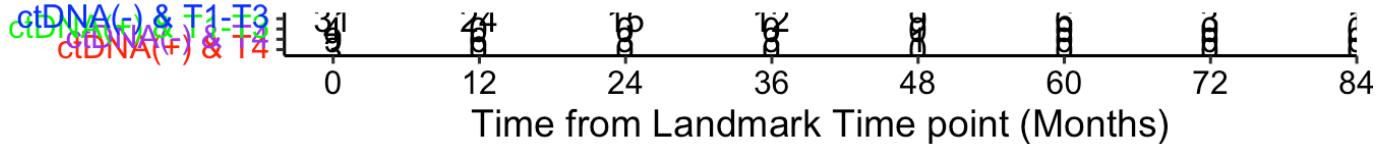
```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=12, palette=c("blue","green","purple", "red"), title="DFS - ctDNA MRD &
ypTN", ylab= "Disease-Free Survival", xlab="Time from Landmark Time point (Months)", leg
end.labs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), le
gend.title="")
```

DFS - ctDNA MRD & ypTN

— ctDNA(-) & T1-T3 — ctDNA(+) & T1-T3 — ctDNA(-) & T4 — ctDNA(+) & T4



Number at risk



Hide

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

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

  ctDNA.Stage.II.TNM=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
   24     15      8     0.698  0.091      0.481      0.838

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           4           0           1           0           1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     9      0     1.0  0.000      1.0000      1.000
   24     3      5     0.4  0.174      0.0981      0.697

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           3           0           1           0           1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("2","4","1","3"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 47, number of events= 21

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	0.81163	2.25158	0.80842	1.004	0.3154
ctDNA.Stage.II.TNM1	-2.68177	0.06844	0.68763	-3.900	9.62e-05 ***
ctDNA.Stage.II.TNM3	-1.59246	0.20342	0.71596	-2.224	0.0261 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	2.25158	0.4441	0.46170	10.9804
ctDNA.Stage.II.TNM1	0.06844	14.6110	0.01778	0.2634
ctDNA.Stage.II.TNM3	0.20342	4.9158	0.05000	0.8276

Concordance= 0.731 (se = 0.051)

Likelihood ratio test= 21.92 on 3 df, p=7e-05

Wald test = 22.7 on 3 df, p=5e-05

Score (logrank) test = 37.96 on 3 df, p=3e-08

#OS by ctDNA at the MRD Window & ypT

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))

circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)

```

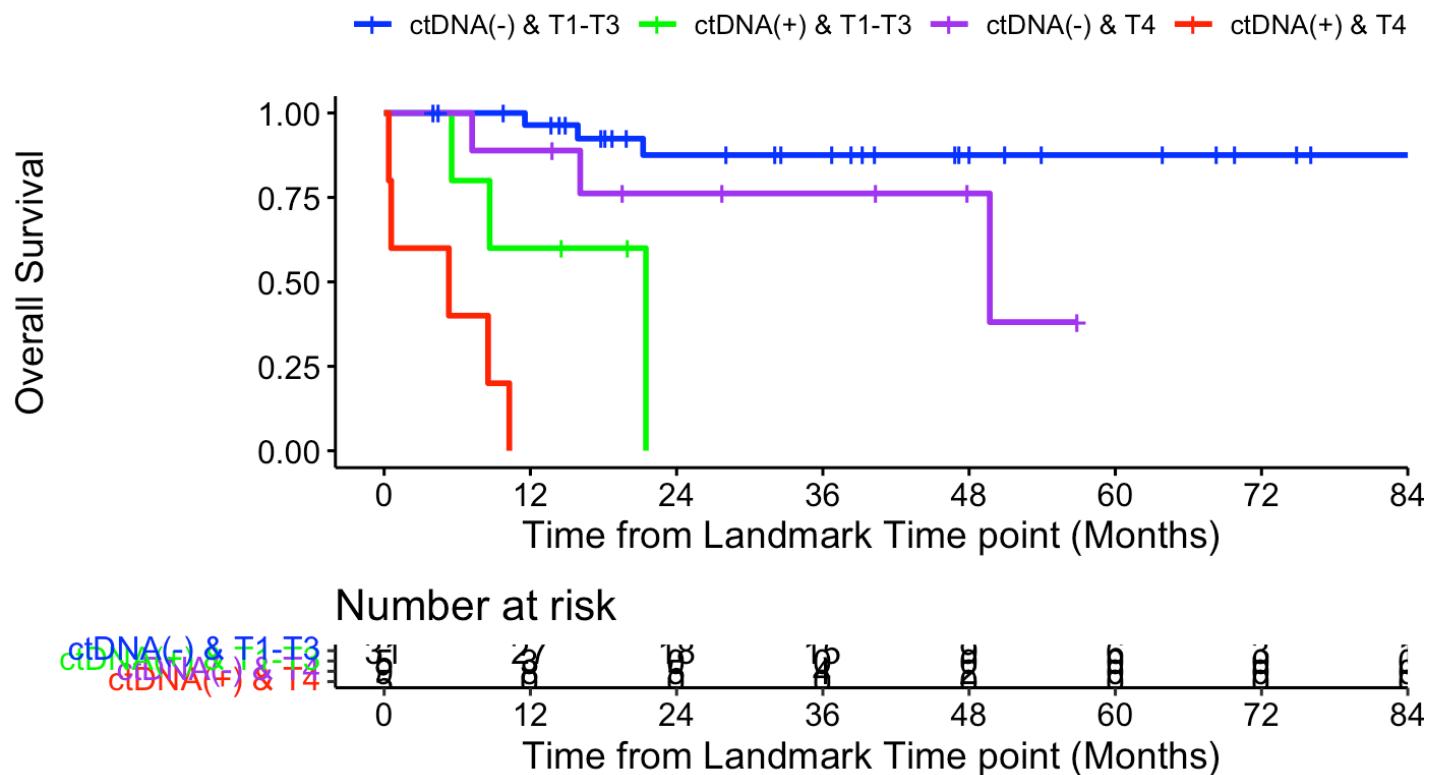
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	31	3	NA	NA	NA
ctDNA.Stage.II.TNM=2	5	3	21.48	8.66	NA
ctDNA.Stage.II.TNM=3	9	3	49.70	49.70	NA
ctDNA.Stage.II.TNM=4	5	5	5.31	0.58	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.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="OS - ctDNA MRD &
ypTN", ylab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.l
abs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.
title="")
```

OS - ctDNA MRD & ypTN



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

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

  ctDNA.Stage.II.TNM=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
   24     18      3     0.875  0.0681    0.658    0.959

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1          0          5          0          1          0          1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     9      0     1.000  0.000    1.000    1.000
   24     5      2     0.762  0.148    0.332    0.935

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1          0          5          0          1          0          1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4",
"ctDNA(+) & T4"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 50, number of events= 14

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	2.5315	12.5725	0.8498	2.979	0.00289 **
ctDNA.Stage.II.TNMctDNA(-) & T4	1.2740	3.5751	0.8169	1.560	0.11886
ctDNA.Stage.II.TNMctDNA(+) & T4	4.5082	90.7580	0.9252	4.873	1.1e-06 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & T1-T3	12.573	0.07954	2.377	66.49
ctDNA.Stage.II.TNMctDNA(-) & T4	3.575	0.27971	0.721	17.73
ctDNA.Stage.II.TNMctDNA(+) & T4	90.758	0.01102	14.804	556.40

Concordance= 0.829 (se = 0.056)

Likelihood ratio test= 27.86 on 3 df, p=4e-06

Wald test = 24.63 on 3 df, p=2e-05

Score (logrank) test = 59.15 on 3 df, p=9e-13

Hide

```
#Repeat analysis to compare ctDNA post-NAT (-) vs (+) in T4
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T1-T3" ~ 1,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T1-T3" ~ 2,
    ctDNA.MRD == "NEGATIVE" & pT.Stage == "T4" ~ 3,
    ctDNA.MRD == "POSITIVE" & pT.Stage == "T4" ~ 4
  ))

circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)
```

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

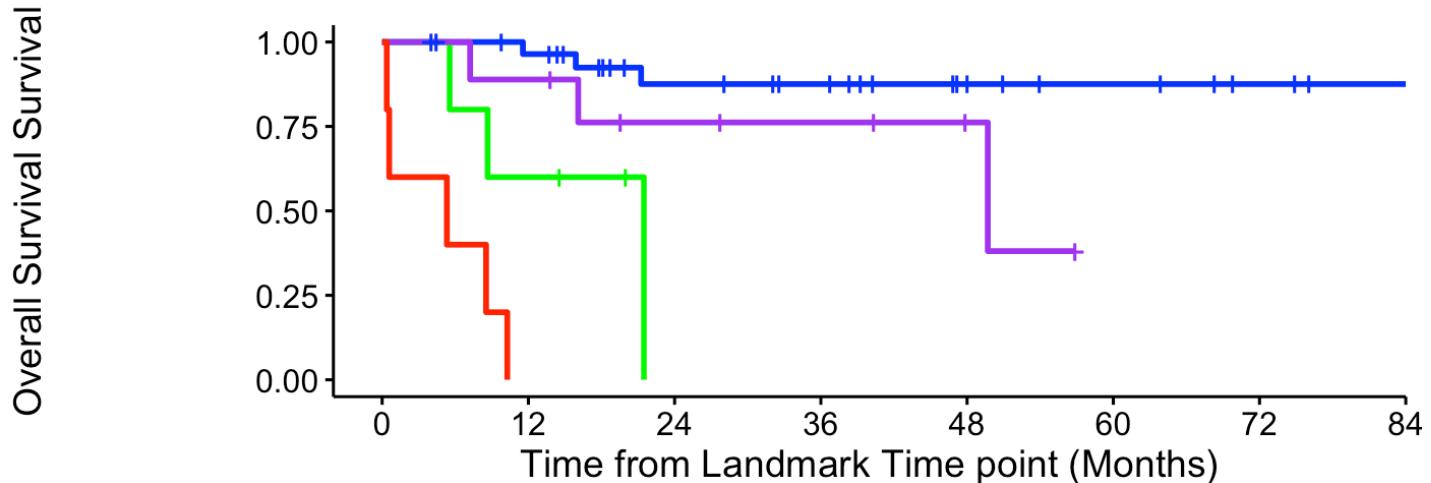
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	31	3	NA	NA	NA
ctDNA.Stage.II.TNM=2	5	3	21.48	8.66	NA
ctDNA.Stage.II.TNM=3	9	3	49.70	49.70	NA
ctDNA.Stage.II.TNM=4	5	5	5.31	0.58	NA

Hide

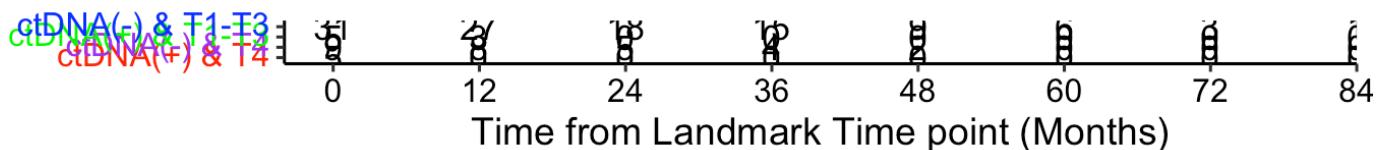
```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=12, palette=c("blue","green","purple", "red"), title="OS - ctDNA MRD &
ypTN", ylab= "Overall Survival Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.title="")
```

OS - ctDNA MRD & ypTN

— ctDNA(-) & T1-T3 — ctDNA(+) & T1-T3 — ctDNA(-) & T4 — ctDNA(+) & T4



Number at risk



Hide

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

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

  ctDNA.Stage.II.TNM=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
   24     18      3     0.875  0.0681    0.658    0.959

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           5           0           1           0           1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     9      0     1.000  0.000    1.000    1.000
   24     5      2     0.762  0.148    0.332    0.935

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           5           0           1           0           1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("2","4","1","3"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 50, number of events= 14

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNM4	1.97669	7.21877	0.84482	2.340	0.01930 *
ctDNA.Stage.II.TNM1	-2.53151	0.07954	0.84981	-2.979	0.00289 **
ctDNA.Stage.II.TNM3	-1.25752	0.28436	0.84413	-1.490	0.13629

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNM4	7.21877	0.1385	1.37833	37.8071
ctDNA.Stage.II.TNM1	0.07954	12.5725	0.01504	0.4207
ctDNA.Stage.II.TNM3	0.28436	3.5167	0.05437	1.4873

Concordance= 0.829 (se = 0.056)

Likelihood ratio test= 27.86 on 3 df, p=4e-06

Wald test = 24.63 on 3 df, p=2e-05

Score (logrank) test = 59.15 on 3 df, p=9e-13

#RFS by ctDNA at the MRD Window & ypN

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.MRD == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.MRD == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.MRD == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.Stage.II.TN
M, data = circ_data)

```

```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

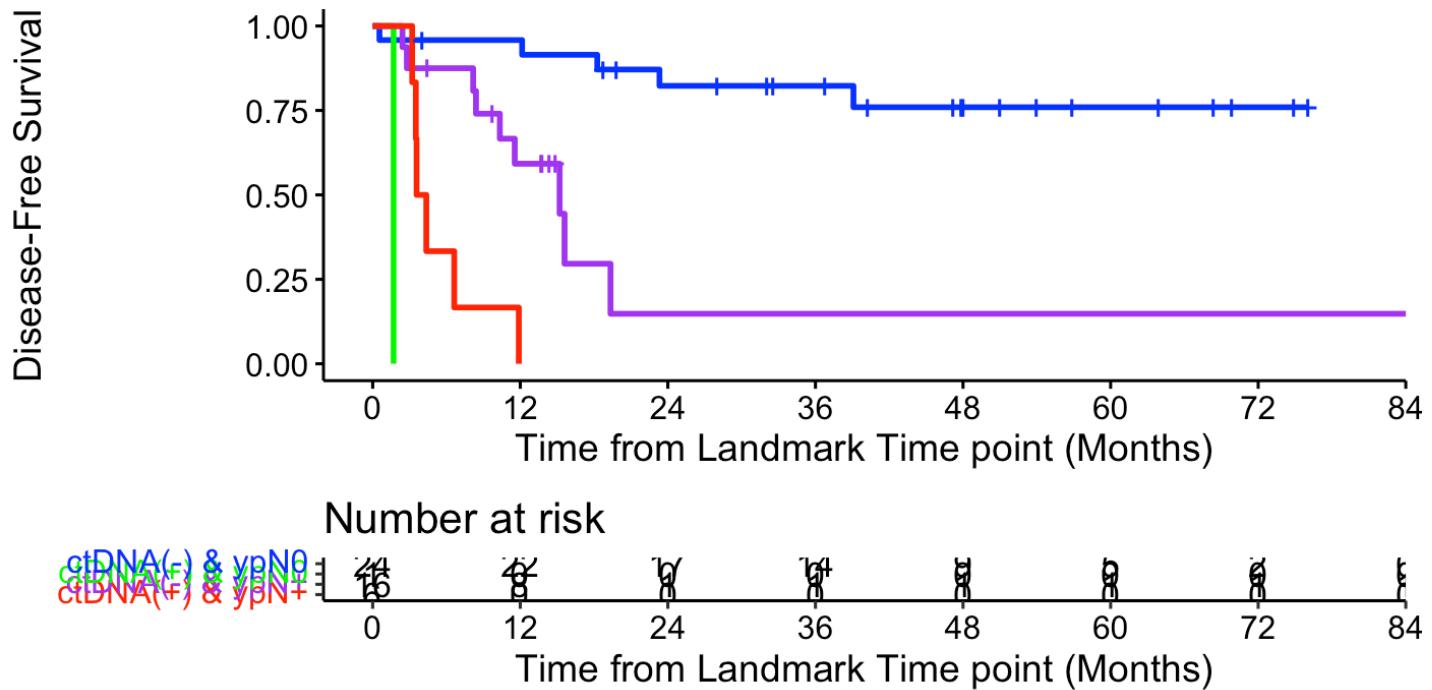
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	24	5	NA	NA	NA
ctDNA.Stage.II.TNM=2	1	1	1.70	NA	NA
ctDNA.Stage.II.TNM=3	16	9	15.20	10.34	NA
ctDNA.Stage.II.TNM=4	6	6	3.97	3.51	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con-
f.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 MRD &
ypN", ylab= "Disease-Free Survival", xlab="Time from Landmark Time point (Months)", lege-
nd.labs=c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"), l-
egend.title="")
```

DFS - ctDNA MRD & ypN

— ctDNA(-) & ypN0 — ctDNA(+) & ypN0 — ctDNA(-) & ypN+ — ctDNA(+) & ypN+



Hide

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

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

  ctDNA.Stage.II.TNM=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     24      0     1.000  0.0000    1.000      1.00
   24     17      4     0.823  0.0807    0.593      0.93

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           1           0           1           0           1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     16      0     1.000  0.000    1.00000    1.000
   24     1       9     0.148  0.132    0.00842    0.466

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           6           0           1           0           1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
("1","2","3","4"), labels = c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 47, number of events= 21

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & ypN0	5.9532	384.9884	1.5428	3.859	0.000114 ***
ctDNA.Stage.II.TNMctDNA(-) & ypN+	2.0297	7.6117	0.6062	3.348	0.000814 ***
ctDNA.Stage.II.TNMctDNA(+) & ypN+	3.6876	39.9503	0.7639	4.827	1.38e-06 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & ypN0	384.988	0.002597	18.715	7919.53
ctDNA.Stage.II.TNMctDNA(-) & ypN+	7.612	0.131377	2.320	24.97
ctDNA.Stage.II.TNMctDNA(+) & ypN+	39.950	0.025031	8.939	178.56

Concordance= 0.792 (se = 0.048)

Likelihood ratio test= 31.15 on 3 df, p=8e-07

Wald test = 27.95 on 3 df, p=4e-06

Score (logrank) test = 54.44 on 3 df, p=9e-12

#OS by ctDNA at the MRD Window & ypN

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.Stage.II.Risk <- NA #first we create the variable for the ctDNA & NAC co
mbination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Stage.II.TNM = case_when(
    ctDNA.MRD == "NEGATIVE" & ypNStatus == "N0" ~ 1,
    ctDNA.MRD == "POSITIVE" & ypNStatus == "N0" ~ 2,
    ctDNA.MRD == "NEGATIVE" & ypNStatus == "N1-N3" ~ 3,
    ctDNA.MRD == "POSITIVE" & ypNStatus == "N1-N3" ~ 4
  ))
  
circ_data <- circ_data[circ_data$ctDNA.Stage.II.TNM!="",]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.Stage.II.TNM,
data = circ_data)

```

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

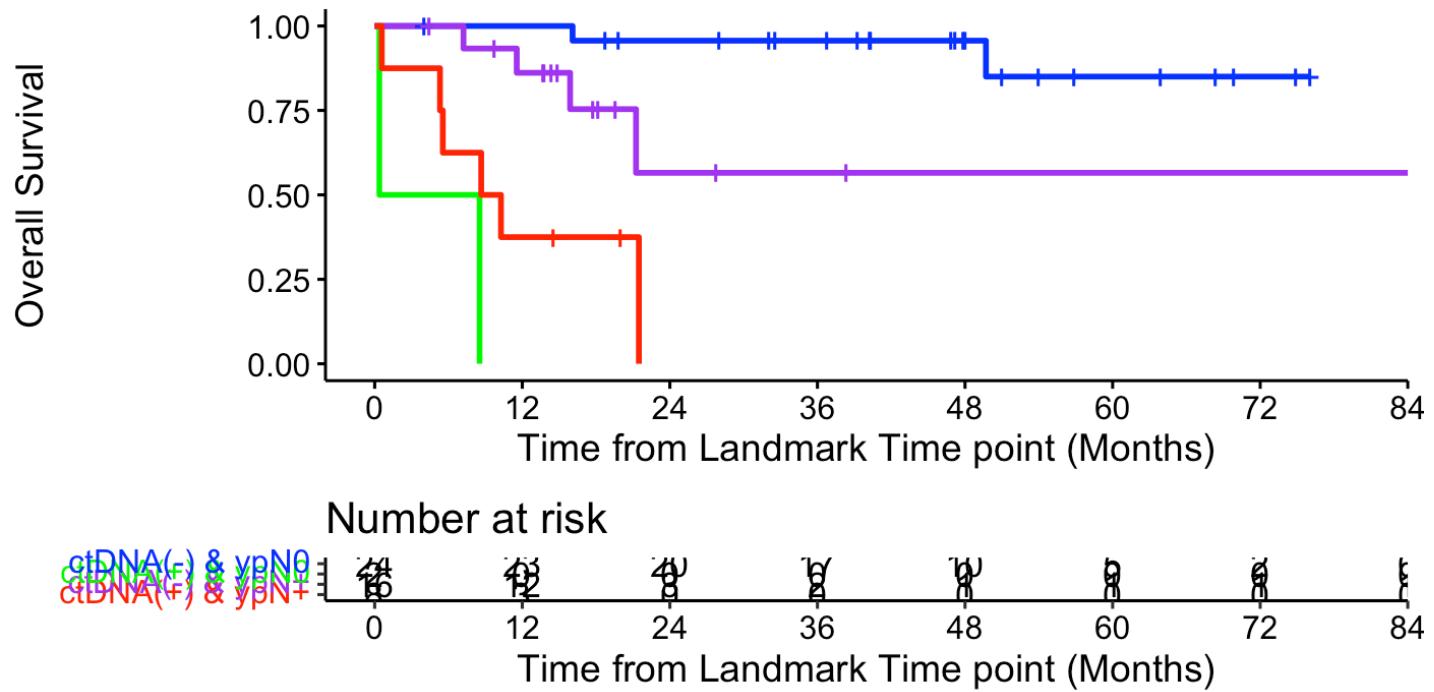
	n	events	median	0.95LCL	0.95UCL
ctDNA.Stage.II.TNM=1	24	2	NA	NA	NA
ctDNA.Stage.II.TNM=2	2	2	4.46	0.38	NA
ctDNA.Stage.II.TNM=3	16	4	NA	21.25	NA
ctDNA.Stage.II.TNM=4	8	6	9.46	5.54	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.Stage.II.TNM, data = circ_data, conf.int=0.95, con
f.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="OS - ctDNA MRD &
ypN", ylab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.la
bs=c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"), legen
d.title="")
```

OS - ctDNA MRD & ypN

— ctDNA(-) & ypN0 — ctDNA(+) & ypN0 — ctDNA(-) & ypN+ — ctDNA(+) & ypN+



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

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

  ctDNA.Stage.II.TNM=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     24      0     1.000  0.0000      1.000      1.000
   24     20      1     0.957  0.0425      0.729      0.994

  ctDNA.Stage.II.TNM=2
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           2           0           1           0           1

  ctDNA.Stage.II.TNM=3
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     16      0     1.000  0.00      1.000      1.000
   24     3       4     0.565  0.19      0.165      0.835

  ctDNA.Stage.II.TNM=4
  time     n.risk     n.event     survival     std.err lower 95% CI upper 95% CI
  1       0           8           0           1           0           1
```

Hide

```
circ_data$ctDNA.Stage.II.TNM <- factor(circ_data$ctDNA.Stage.II.TNM, levels=c
  ("1","2","3","4"), labels = c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN+"))
cox_fit <- coxph(surv_object ~ ctDNA.Stage.II.TNM, data=circ_data)
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ ctDNA.Stage.II.TNM, data = circ_data)

n= 50, number of events= 14

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.Stage.II.TNMctDNA(+) & ypN0	5.0035	148.9381	1.1735	4.264	2.01e-05 ***
ctDNA.Stage.II.TNMctDNA(-) & ypN+	1.8810	6.5597	0.9016	2.086	0.037 *
ctDNA.Stage.II.TNMctDNA(+) & ypN+	3.5092	33.4225	0.8993	3.902	9.53e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.Stage.II.TNMctDNA(+) & ypN0	148.94	0.006714	14.932	1485.5
ctDNA.Stage.II.TNMctDNA(-) & ypN+	6.56	0.152445	1.120	38.4
ctDNA.Stage.II.TNMctDNA(+) & ypN+	33.42	0.029920	5.736	194.8

Concordance= 0.86 (se = 0.04)

Likelihood ratio test= 26.4 on 3 df, p=8e-06

Wald test = 23 on 3 df, p=4e-05

Score (logrank) test = 46.69 on 3 df, p=4e-10

#DFS by ctDNA at the MRD Window & TRG combination

Hide

```

rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]

circ_data$ctDNA.pCR <- NA #first we create the variable for the ctDNA & NAC combination,
and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.MRD == "NEGATIVE" & TRG == "TRG1/2/3" ~ "1",
    ctDNA.MRD == "NEGATIVE" & TRG == "TRG4/5" ~ "2",
    ctDNA.MRD == "POSITIVE" & TRG == "TRG4/5" ~ "3"
  ))

circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.pCR, data =
circ_data)

```

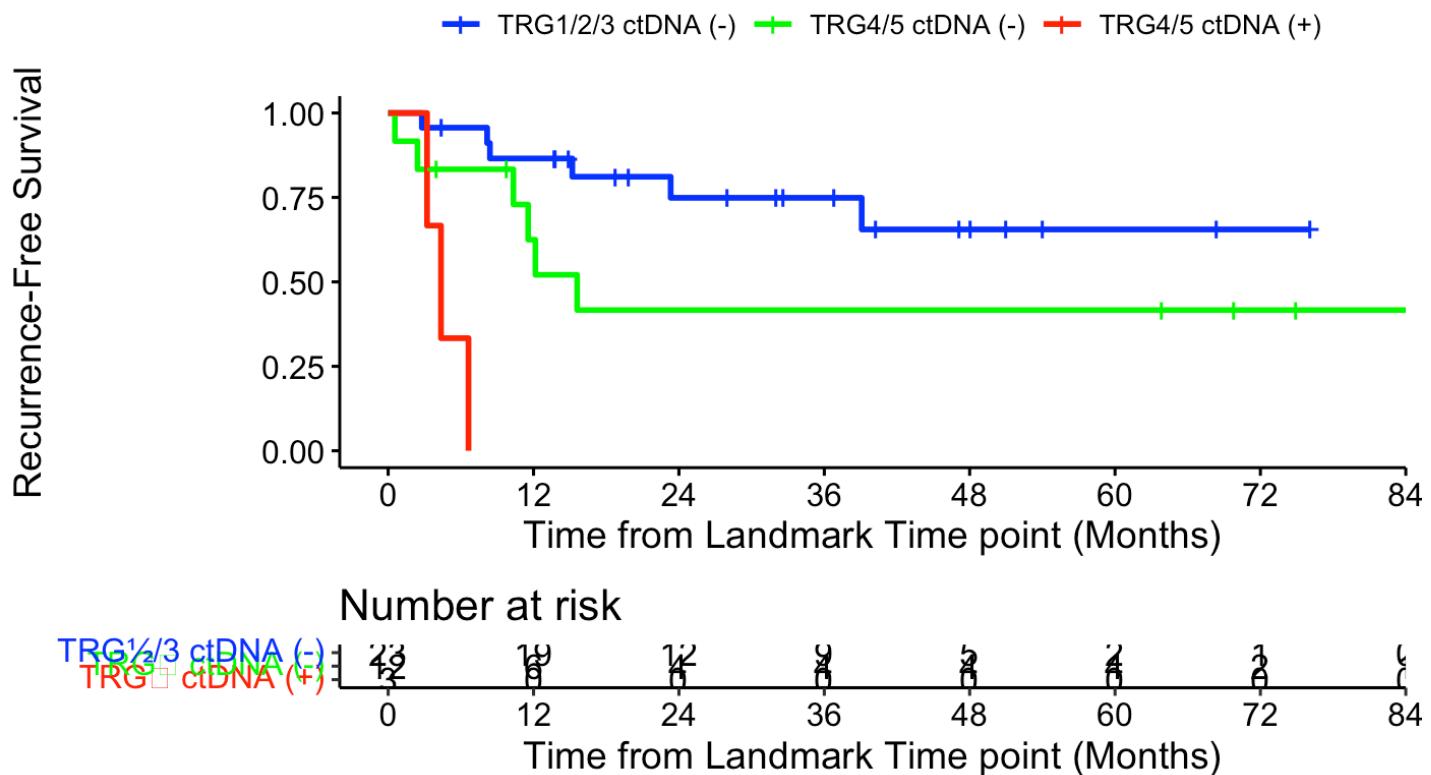
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.pCR, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.pCR=1	23	6	NA	39.09	NA
ctDNA.pCR=2	12	6	15.60	11.55	NA
ctDNA.pCR=3	3	3	4.36	3.21	NA

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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","red"), title="RFS - ctDNA MRD/TRG", ylab="Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("TRG1/2/3 ctDNA (-)", "TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)", legend.title="")
```

RFS - ctDNA MRD/TRG



Hide

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

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

ctDNA.pCR=1

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	23	0	1.000	0.0000	1.000	1.000	1.000	1.000
24	12	5	0.749	0.0993	0.492	0.492	0.889	0.889

ctDNA.pCR=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	12	0	1.000	0.000	1.000	1.000	1.000	1.000
24	4	6	0.417	0.157	0.131	0.131	0.686	0.686

ctDNA.pCR=3

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	3	0	1	0	0	1	1	1

Hide

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2","3"), labels=c("TRG 1/2/3 ctDNA (-)","TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"))
cox_fit <- coxph(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 38, number of events= 15

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.pCRTRG4/5 ctDNA (-)	0.9323	2.5404	0.5794	1.609	0.107610
ctDNA.pCRTRG4/5 ctDNA (+)	3.0190	20.4704	0.8850	3.411	0.000647 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.pCRTRG4/5 ctDNA (-)	2.54	0.39364	0.816	7.909
ctDNA.pCRTRG4/5 ctDNA (+)	20.47	0.04885	3.612	116.004

Concordance= 0.7 (se = 0.064)

Likelihood ratio test= 10.66 on 2 df, p=0.005

Wald test = 11.83 on 2 df, p=0.003

Score (logrank) test = 19.3 on 2 df, p=6e-05

#OS by ctDNA at the MRD Window & TRG combination

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]

circ_data$ctDNA.pCR <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.MRD == "NEGATIVE" & TRG == "TRG1/2/3" ~ "1",
    ctDNA.MRD == "NEGATIVE" & TRG == "TRG4/5" ~ "2",
    ctDNA.MRD == "POSITIVE" & TRG == "TRG4/5" ~ "3"
  ))

circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.pCR, data = circ_data)
```

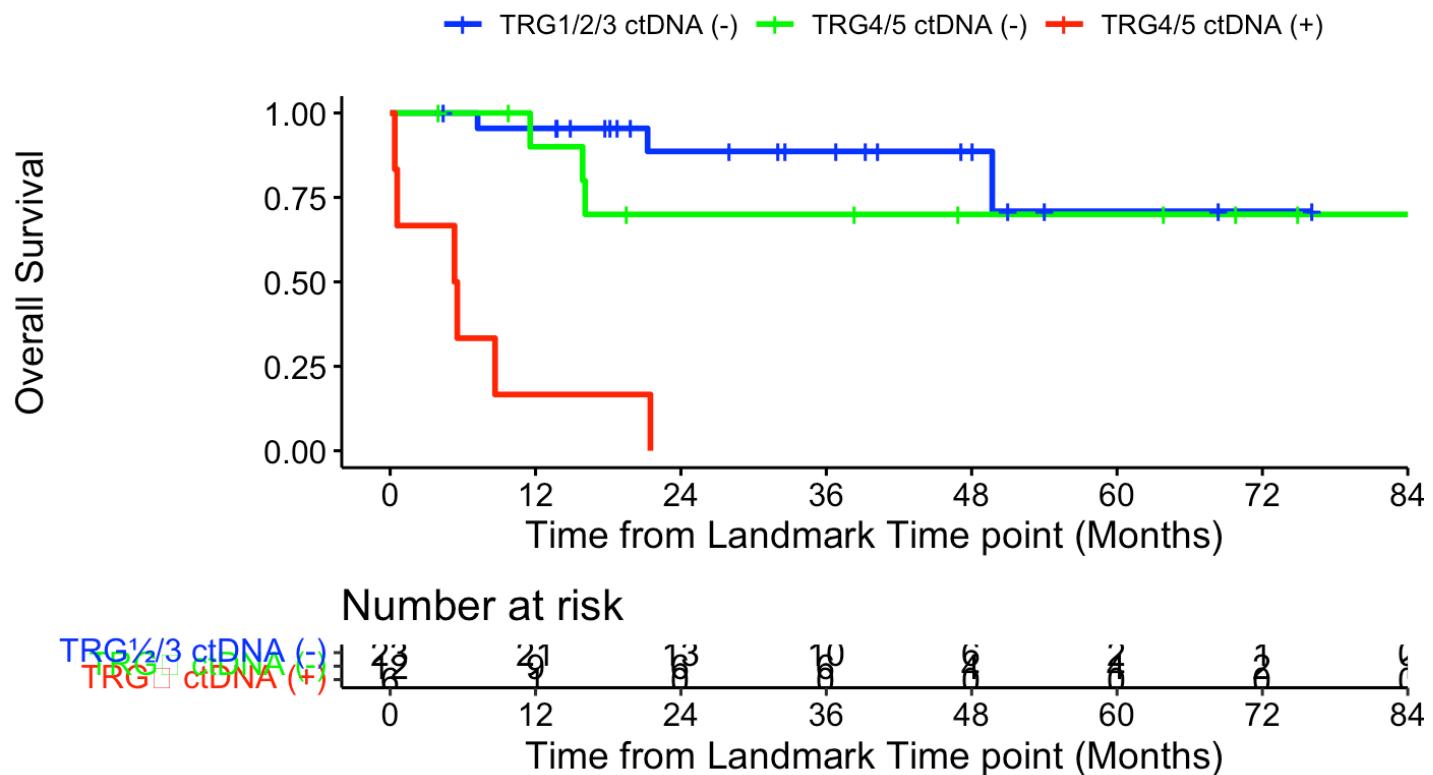
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.pCR, data = circ_data)

      n events median 0.95LCL 0.95UCL
ctDNA.pCR=1 23      3      NA  49.70      NA
ctDNA.pCR=2 12      3      NA  16.09      NA
ctDNA.pCR=3  6      6  5.42    0.58      NA
```

[Hide](#)

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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","red"), title="OS - ctDNA MRD/TRG", ylab="Overall Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("TRG1/2/3 ctDNA (-)","TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"), legend.title="")
```

OS - ctDNA MRD/TRG



Hide

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

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

ctDNA.pCR=1

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	23	0	1.000	0.0000	1.000	1.000	1.000	1.000
24	13	2	0.886	0.0776	0.607	0.971		

ctDNA.pCR=2

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	12	0	1.0	0.000	1.000	1.000	1.000	1.000
24	6	3	0.7	0.145	0.329	0.892		

ctDNA.pCR=3

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	6	0	1	0	0	1	1	1

Hide

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2","3"), labels=c("TRG 1/2/3 ctDNA (-)","TRG4/5 ctDNA (-)", "TRG4/5 ctDNA (+)"))
cox_fit <- coxph(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 41, number of events= 12

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.pCRTRG4/5 ctDNA (-)	0.6450	1.9061	0.8202	0.786	0.432
ctDNA.pCRTRG4/5 ctDNA (+)	3.1621	23.6201	0.7401	4.273	1.93e-05 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.pCRTRG4/5 ctDNA (-)	1.906	0.52464	0.3819	9.513
ctDNA.pCRTRG4/5 ctDNA (+)	23.620	0.04234	5.5374	100.753

Concordance= 0.812 (se = 0.065)

Likelihood ratio test= 18.73 on 2 df, p=9e-05

Wald test = 21.36 on 2 df, p=2e-05

Score (logrank) test = 38.75 on 2 df, p=4e-09

#DFS by ctDNA at the MRD Window & ypTN Characteristics

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included == TRUE,]
circ_data$DFS.months <- circ_data$DFS.months - 3
circ_data <- circ_data[circ_data$DFS.months >= 0,]
circ_data$ctDNA.pCR <- NA

# Define new categories
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage=="T1-T3" & ypNStatus == "N0" ~ "1",
    ctDNA.MRD == "POSITIVE" | pT.Stage == "T4" | ypNStatus == "N1-N3" ~ "2"
  ))

circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.pCR, data = circ_data)
```

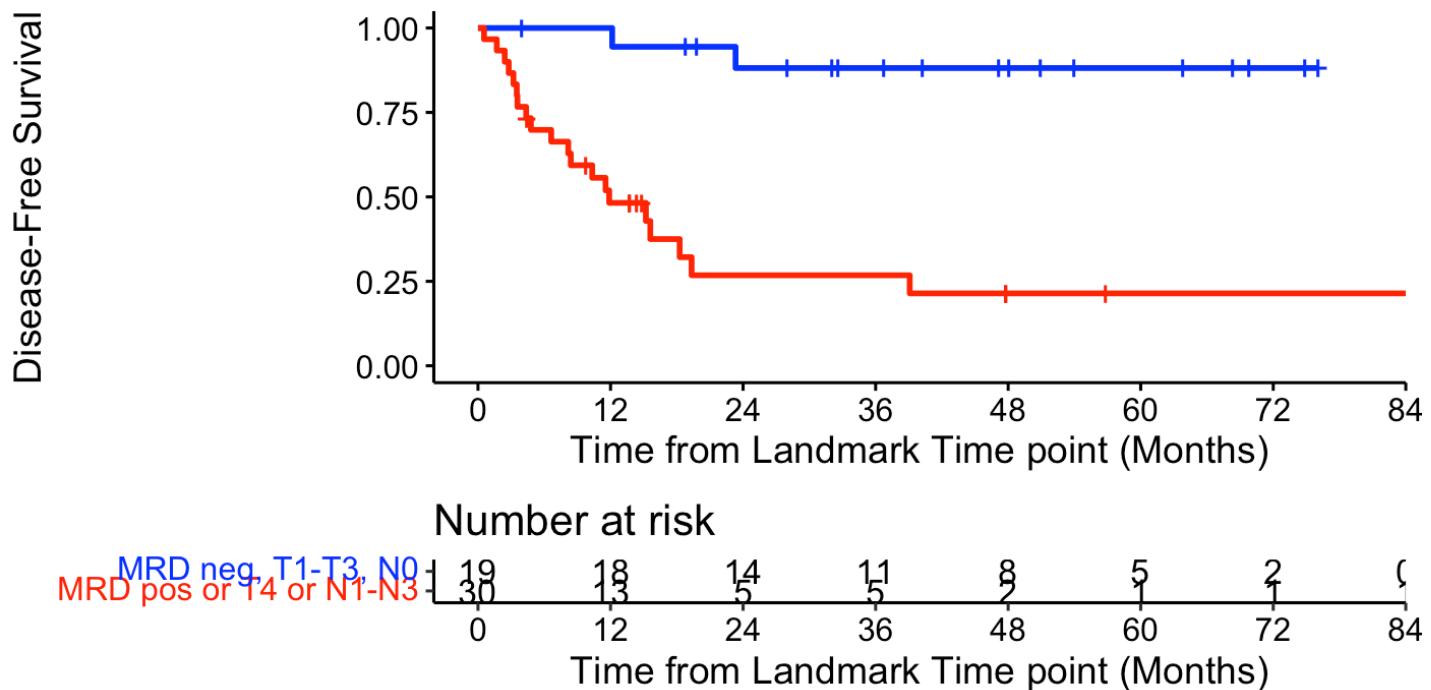
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.pCR, data = circ_data)

      n events median 0.95LCL 0.95UCL
ctDNA.pCR=1 19      2      NA      NA
ctDNA.pCR=2 30     20    11.9    8.17    39.1
```

[Hide](#)

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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 MRD/Clinical Characteristics",
  ylab = "Disease-Free Survival", xlab = "Time from Landmark Time point (Months)",
  legend.labs = c("MRD neg, T1-T3, N0", "MRD pos or T4 or N1-N3"),
  legend.title = "")
```

DFS - ctDNA MRD/Clinical Characteristics



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

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

```
ctDNA.pCR=1
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     19      0     1.000  0.000     1.000     1.000
 24     14      2     0.881  0.079     0.602     0.969
```

```
ctDNA.pCR=2
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     30      0     1.000  0.0000     1.000     1.000
 24      5     19     0.268  0.0954     0.107     0.461
```

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2"), labels=c("MRD neg,
T1-T3, N0","MRD pos or T4 or N1-N3"))
cox_fit <- coxph(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

Call:

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

n= 49, number of events= 22

            coef  exp(coef)  se(coef)      z Pr(>|z|)    
ctDNA.pCRM RD pos or T4 or N1-N3  2.5366   12.6367   0.7503 3.381 0.000723 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

            exp(coef)  exp(-coef) lower .95 upper .95    
ctDNA.pCRM RD pos or T4 or N1-N3      12.64     0.07913     2.904     54.99

Concordance= 0.723  (se = 0.036 )
Likelihood ratio test= 20.4  on 1 df,  p=6e-06
Wald test             = 11.43  on 1 df,  p=7e-04
Score (logrank) test = 18  on 1 df,  p=2e-05
```

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 = 12.64 (2.9-54.99); p = 0.001"
```

#OS by ctDNA at the MRD Window & ypTN Characteristics

Hide

```

# Clear workspace and set working directory
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included == TRUE,]
circ_data$FU.months <- circ_data$FU.months - 3
circ_data <- circ_data[circ_data$FU.months >= 0,]
circ_data$ctDNA.pCR <- NA

# Define new categories
circ_data <- circ_data %>%
  mutate(ctDNA.pCR = case_when(
    ctDNA.MRD == "NEGATIVE" & pT.Stage=="T1-T3" & ypNStatus == "N0" ~ "1",
    ctDNA.MRD == "POSITIVE" | pT.Stage == "T4" | ypNStatus == "N1-N3" ~ "2"
  ))

circ_data <- circ_data[!is.na(circ_data$ctDNA.pCR),]
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.pCR, data = circ_data)

```

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

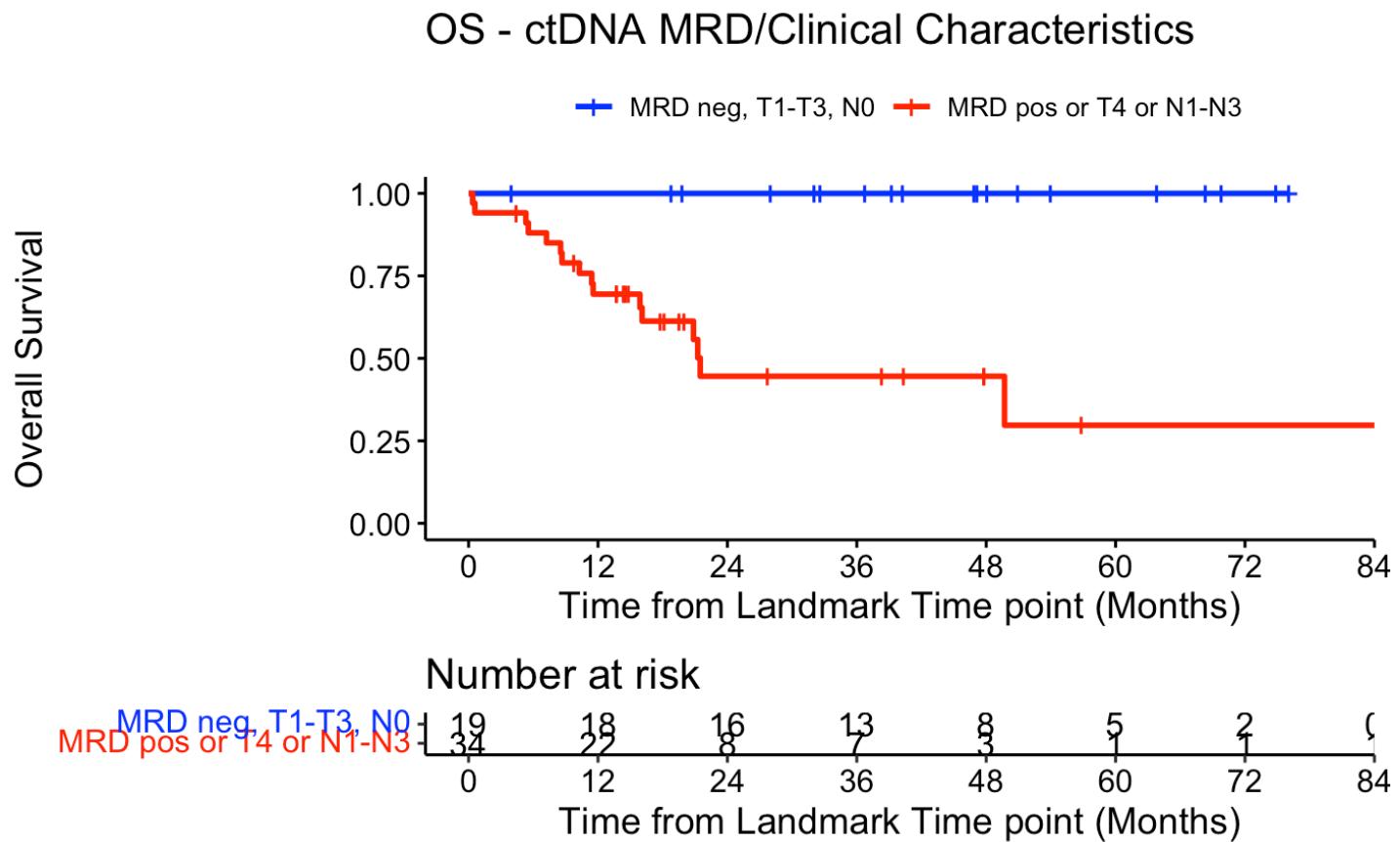
	n	events	median	0.95LCL	0.95UCL
ctDNA.pCR=1	19	0	NA	NA	NA
ctDNA.pCR=2	34	16	21.5	16.1	NA

[Hide](#)

```

surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, 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/Clinical Characteristics",
           ylab = "Overall Survival", xlab = "Time from Landmark Time point (Months)",
           legend.labs = c("MRD neg, T1-T3, N0", "MRD pos or T4 or N1-N3"),
           legend.title = "")

```



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

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

  ctDNA.pCR=1
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     19      0       1       0           1           1
  24     16      0       1       0           NA          NA

  ctDNA.pCR=2
  time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0     34      0     1.000  0.000       1.000      1.000
  24      8     15     0.446  0.105       0.239      0.634
```

```
circ_data$ctDNA.pCR <- factor(circ_data$ctDNA.pCR, levels=c("1","2"), labels=c("MRD neg",
T1-T3, N0","MRD pos or T4 or N1-N3"))
cox_fit <- coxphf(surv_object ~ ctDNA.pCR, data=circ_data)
summary(cox_fit)
```

```
coxphf(formula = surv_object ~ ctDNA.pCR, data = circ_data)
```

Model fitted by Penalized ML

Confidence intervals and p-values by Profile Likelihood

	coef	se(coef)	exp(coef)	lower 0.95	upper 0.95	Chi
sq	p					
ctDNA.pCRM	RD pos or T4 or N1-N3	3.493535	1.495302	32.90204	4.363659	4215.243 18.019
42	2.186631e-05					

Likelihood ratio test=18.01942 on 1 df, p=2.186631e-05, n=53

Wald test = 5.458488 on 1 df, p = 0.01947349

Covariance-Matrix:

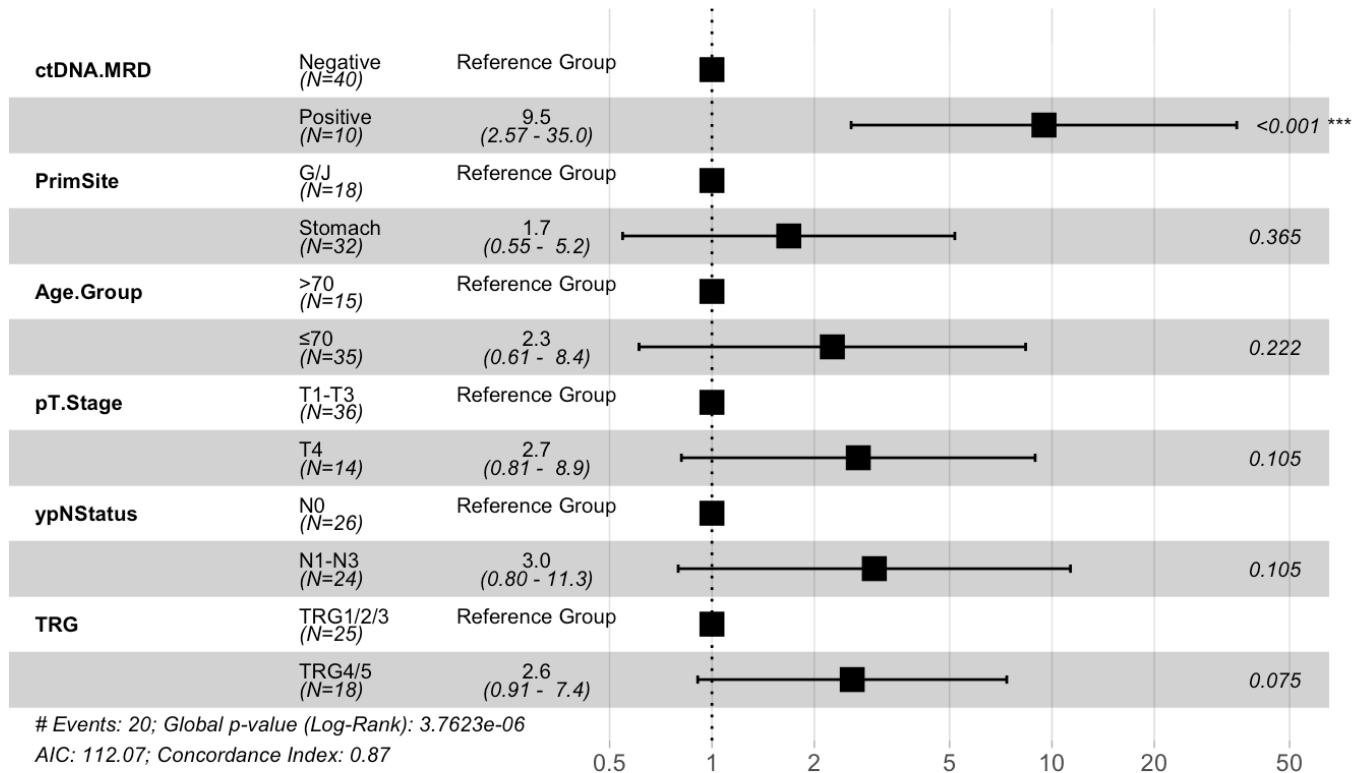
	ctDNA.pCRM	RD pos or T4 or N1-N3
ctDNA.pCRM	RD pos or T4 or N1-N3	2.235927
RD pos or T4 or N1-N3		

#Multivariate cox regression for RFS - ctDNA MRD Window

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels=c("G/J","Stomach"))
circ_data$Gender <- factor(circ_data$Gender, levels=c("Male","Female"))
circ_data$Age.Group <- factor(circ_data$Age.Group, levels=c("2","1"), labels = c(">70", "≤70"))
circ_data$pT.Stage <- factor(circ_data$pT.Stage, levels=c("T1-T3","T4"))
circ_data$ypNStatus <- factor(circ_data$ypNStatus, levels=c("N0","N1-N3"))
circ_data$TRG <- factor(circ_data$TRG, levels=c("TRG1/2/3","TRG4/5"))
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.MRD + PrimSite + Age.Group + pT.Stage + ypNStatus + TRG, data=circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for RFS", refLabel = "Reference Group")
```

Multivariate Regression Model for RFS



Hide

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

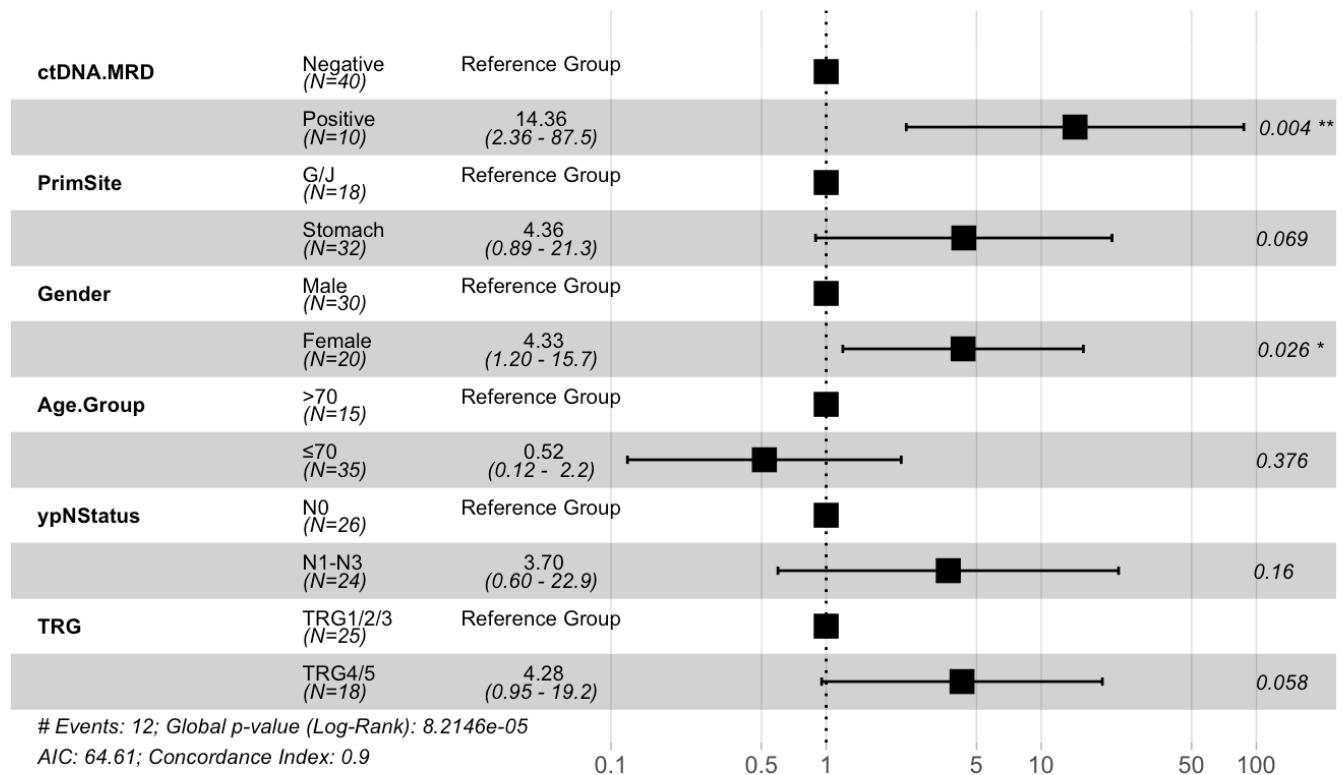
```
#Multivariate cox regression for OS - ctDNA MRD Window v1
```

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_dataadf <- as.data.frame(circ_data)

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels=c("G/J","Stomach"))
circ_data$Gender <- factor(circ_data$Gender, levels=c("Male","Female"))
circ_data$Age.Group <- factor(circ_data$Age.Group, levels=c("2","1"), labels = c(">70", "≤70"))
circ_data$ypNStatus <- factor(circ_data$ypNStatus, levels=c("N0","N1-N3"))
circ_data$TRG <- factor(circ_data$TRG, levels=c("TRG1/2/3","TRG4/5"))
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.MRD + PrimSite + Gender + Age.Group + ypNStatus + TRG, data=circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for OS", refLabel = "Reference Group")
```

Multivariate Regression Model for OS



[Hide](#)

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

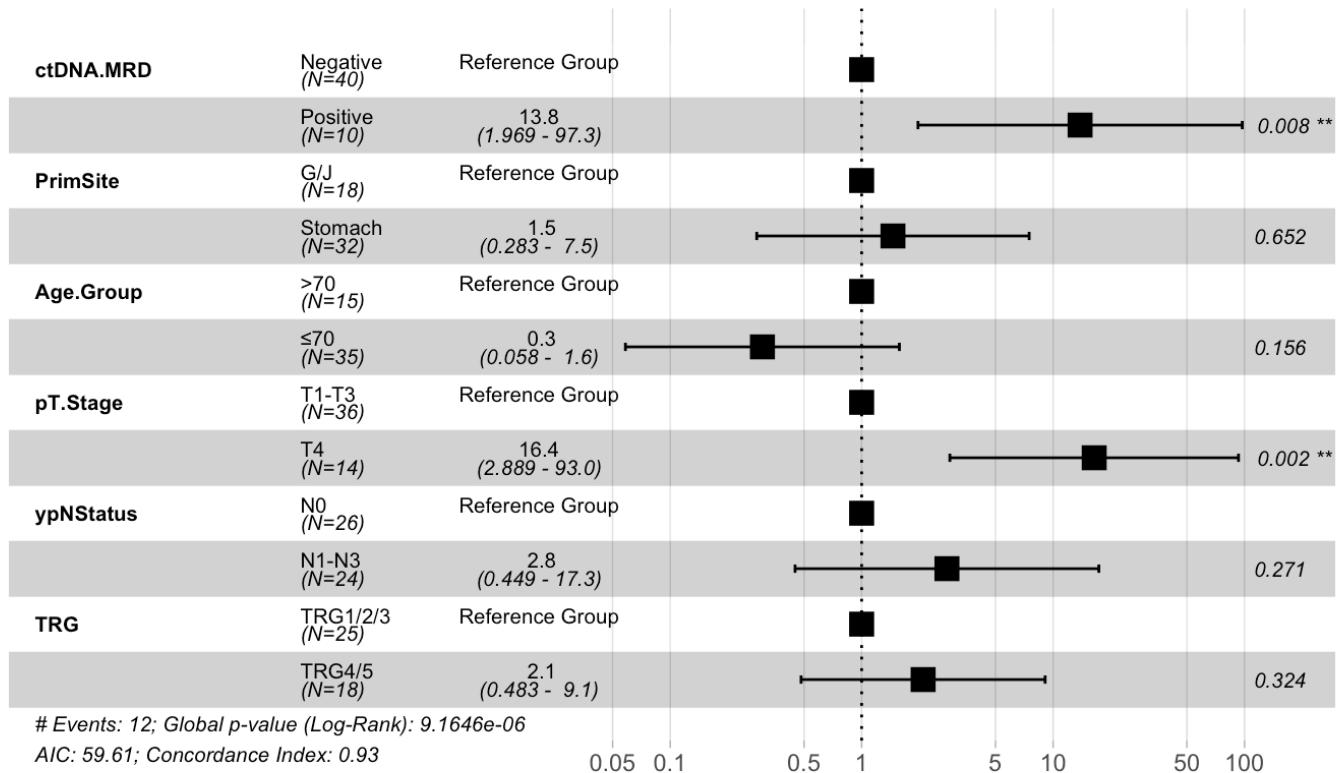
#Multivariate cox regression for OS - ctDNA MRD Window v2

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD!="",]
circ_datadf <- as.data.frame(circ_data)

circ_data$ctDNA.MRD <- factor(circ_data$ctDNA.MRD, levels=c("NEGATIVE","POSITIVE"), labels = c("Negative", "Positive"))
circ_data$PrimSite <- factor(circ_data$PrimSite, levels=c("G/J","Stomach"))
circ_data$Gender <- factor(circ_data$Gender, levels=c("Male","Female"))
circ_data$Age.Group <- factor(circ_data$Age.Group, levels=c("2","1"), labels = c(">70", "<70"))
circ_data$pT.Stage <- factor(circ_data$pT.Stage, levels=c("T1-T3","T4"))
circ_data$ypNStatus <- factor(circ_data$ypNStatus, levels=c("N0","N1-N3"))
circ_data$TRG <- factor(circ_data$TRG, levels=c("TRG1/2/3","TRG4/5"))
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
cox_fit <- coxph(surv_object ~ ctDNA.MRD + PrimSite + Age.Group + pT.Stage + ypNStatus + TRG, data=circ_data)
ggforest(cox_fit, data = circ_data, main = "Multivariate Regression Model for OS", refLabel = "Reference Group")
```

Multivariate Regression Model for OS



Hide

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

#DFS by ctDNA at the Surveillance Window - all stages

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.surveillance!="",]
circ_data$DFS.months=circ_data$DFS.months-3
circ_data <- circ_data[circ_data$DFS.months>=0,]
circ_datadf <- as.data.frame(circ_data)

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

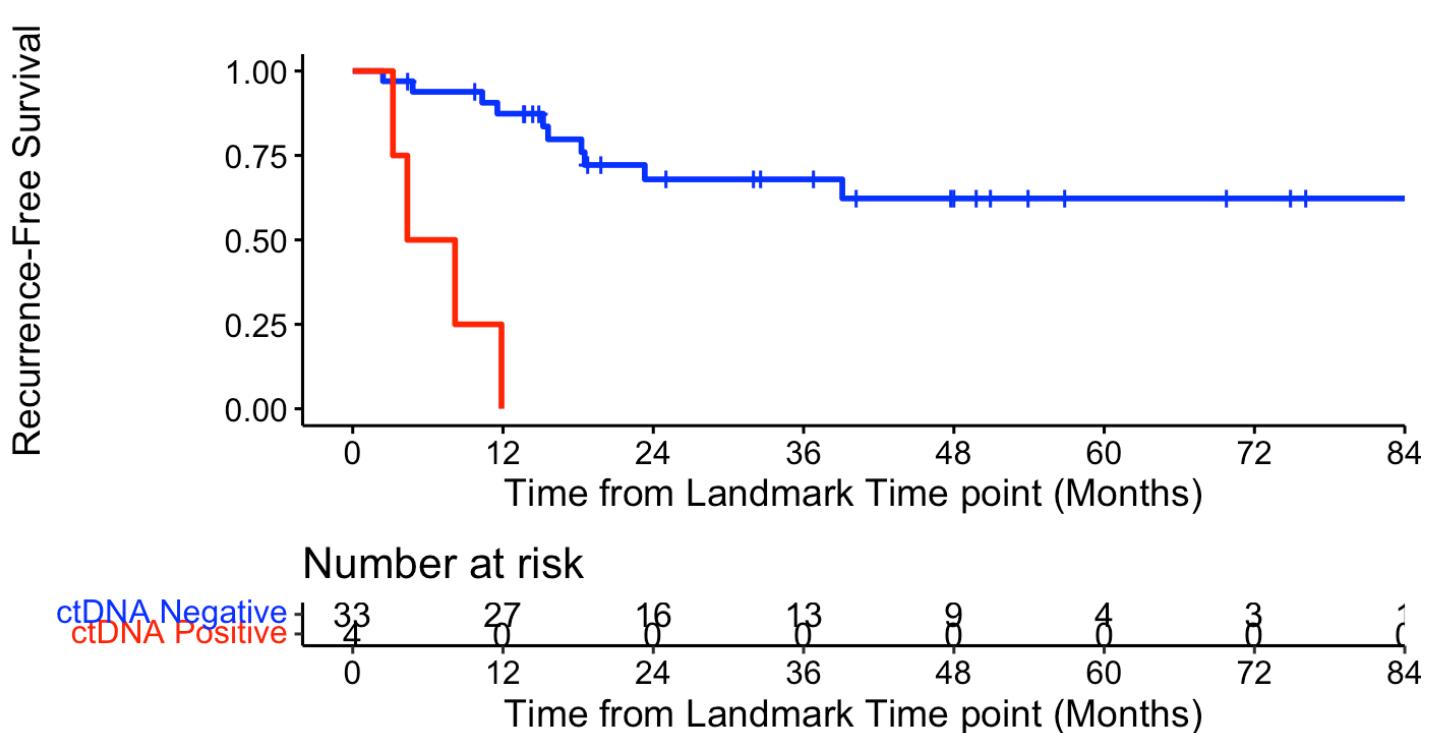
Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.surveillance, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.surveillance=NEGATIVE	33	10	NA	39.09	NA
ctDNA.surveillance=POSITIVE	4	4	6.27	3.21	NA

Hide

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.surveillance, 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="RFS - ctDNA Surveillance window | All pts",
ylab= "Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

RFS - ctDNA Surveillance window | All pts




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

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

ctDNA.surveillance=NEGATIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	33	0	1.000	0.0000	1.00	1.00
24	16	9	0.679	0.0896	0.47	0.82

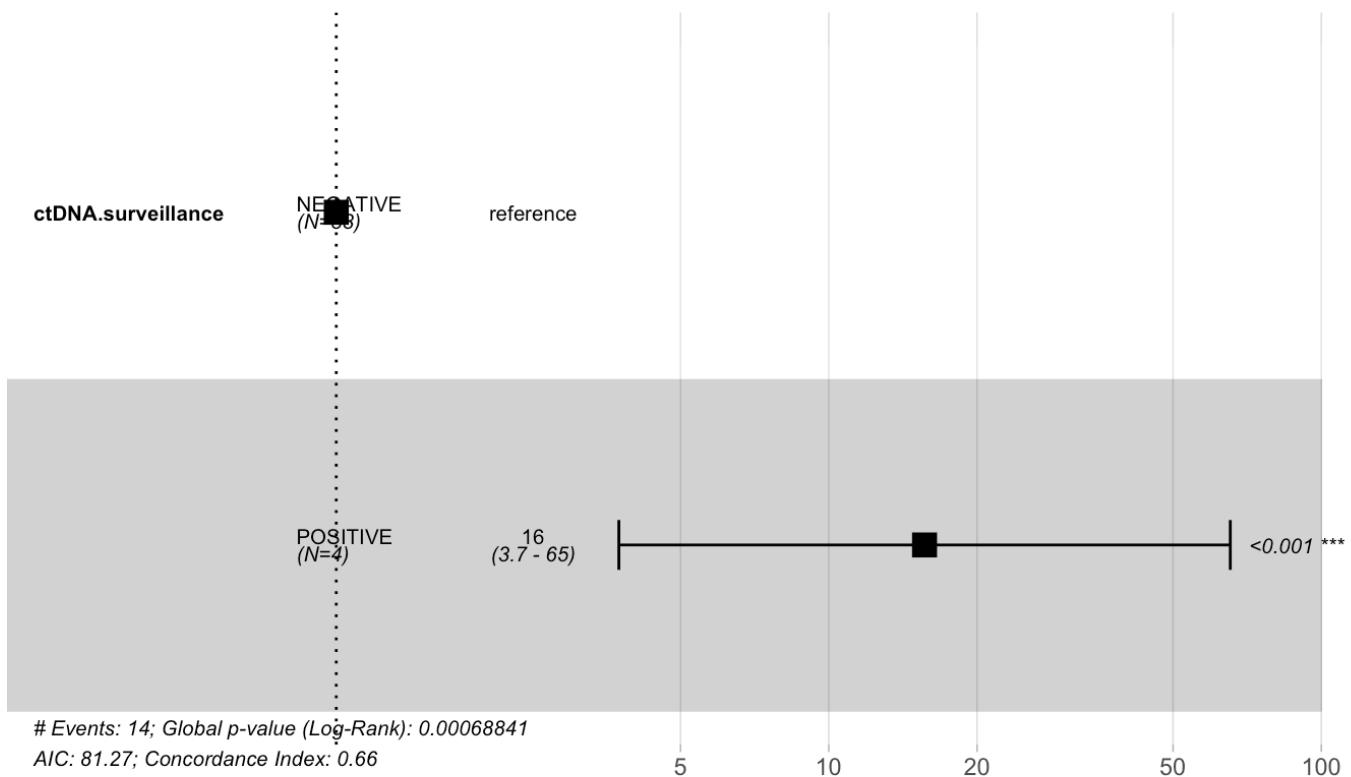
ctDNA.surveillance=POSITIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	4	0	1	0	0	1

```

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

```

Hazard ratio



Hide

```
summary(cox_fit)
```

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

n= 37, number of events= 14

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA_surveillancePOSITIVE	2.7504	15.6482	0.7292	3.772	0.000162 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA_surveillancePOSITIVE	15.65	0.06391	3.748	65.34

Concordance= 0.657 (se = 0.06)
 Likelihood ratio test= 11.52 on 1 df, p=7e-04
 Wald test = 14.23 on 1 df, p=2e-04
 Score (logrank) test = 24.6 on 1 df, p=7e-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 = 15.65 (3.75-65.34); p = 0"
```

#OS by ctDNA at the Surveillance Window - all stages

[Hide](#)

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")
circ_data <- circ_data[circ_data$Included==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.surveillance!="",]
circ_data$FU.months=circ_data$FU.months-3
circ_data <- circ_data[circ_data$FU.months>=0,]
circ_datadf <- as.data.frame(circ_data)

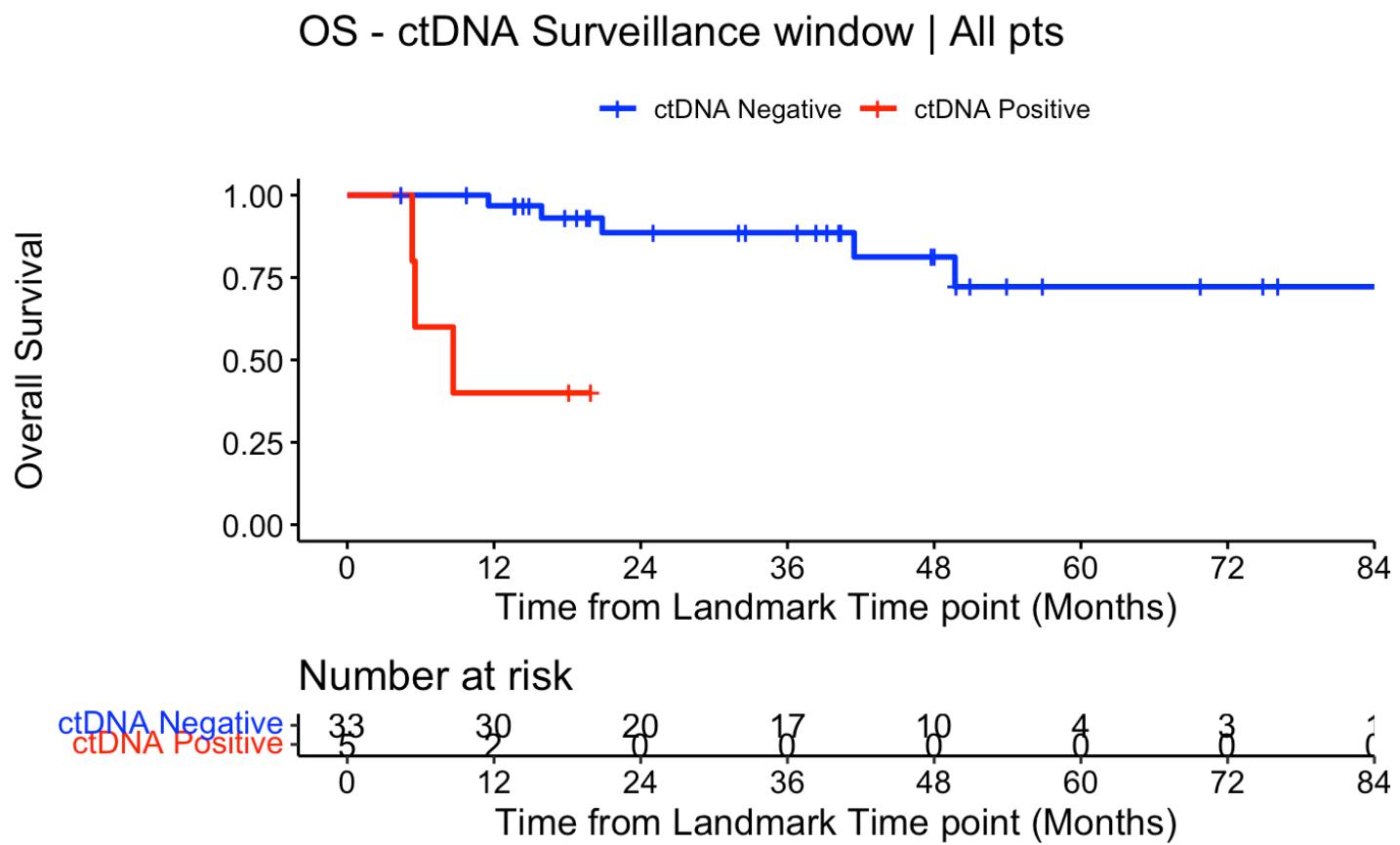
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.surveillance,
data = circ_data)
```

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

	n	events	median	0.95LCL	0.95UCL
ctDNA.surveillance=NEGATIVE	33	5	NA	NA	NA
ctDNA.surveillance=POSITIVE	5	3	8.66	5.54	NA

[Hide](#)

```
surv_object <-Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.surveillance, 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 Surveillance window | All pts",
ylab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



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

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

ctDNA.surveillance=NEGATIVE

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	33	0	1.000	0.0000	1.000		1.000	
24	20	3	0.886	0.0626	0.684		0.962	

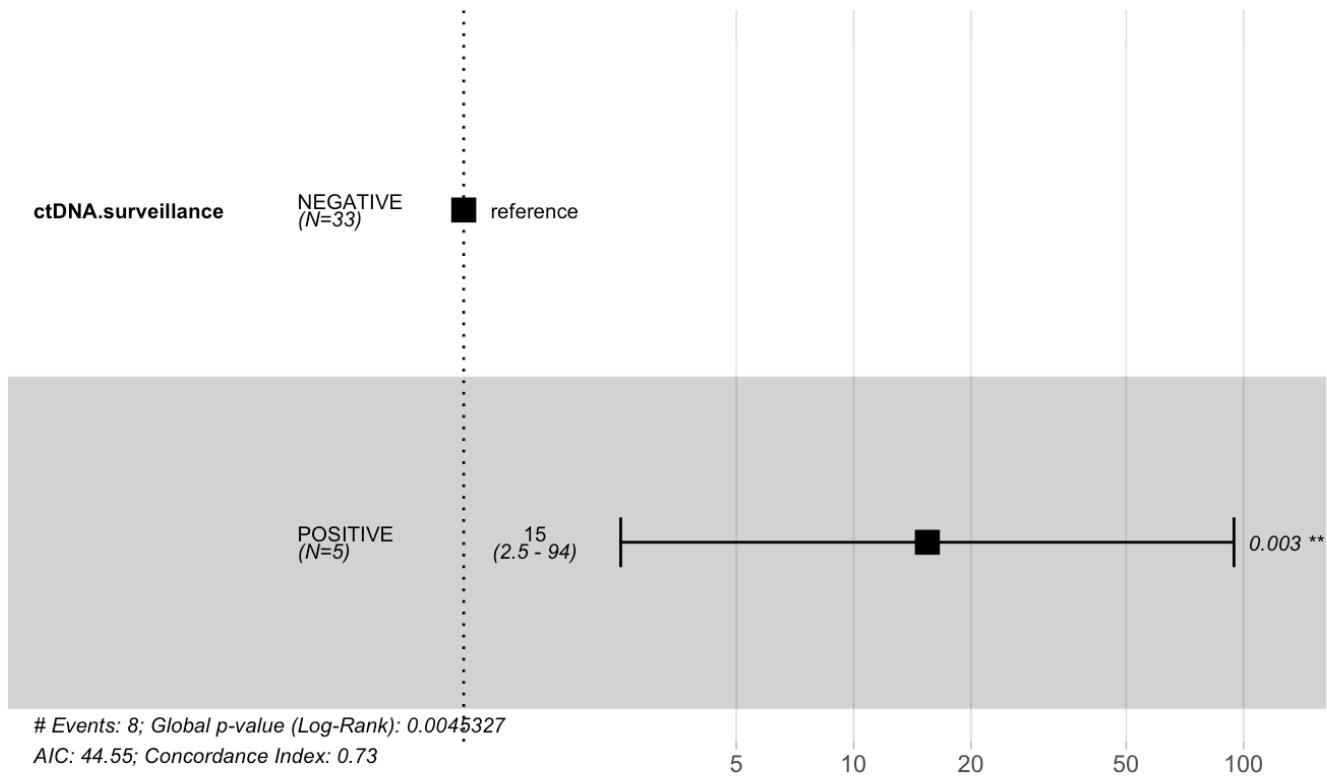
ctDNA.surveillance=POSITIVE

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	5	0	1	0	0		1	
1								

[Hide](#)

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

Hazard ratio



[Hide](#)

```
summary(cox_fit)
```

Call:

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

n= 38, number of events= 8

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.surveillancePOSITIVE	2.7378	15.4530	0.9239	2.963	0.00304 **

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.surveillancePOSITIVE	15.45	0.06471	2.527	94.5

Concordance= 0.727 (se = 0.088)

Likelihood ratio test= 8.06 on 1 df, p=0.005

Wald test = 8.78 on 1 df, p=0.003

Score (logrank) test = 15.41 on 1 df, p=9e-05

[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 = 15.45 (2.53-94.5); p = 0.003"