

Code ▾

AGEO PLAGAST_Clinical Final analysis_revisions 05042025

```
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			

#Neoadjuvant Treatment duration

[Hide](#)

```

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

circ_data$NAC.Start <- as.Date(circ_data$NAC.Start, format="%m/%d/%Y")
circ_data$NAC.End <- as.Date(circ_data$NAC.End, format="%m/%d/%Y")
circ_data$NAC.Duration <- as.numeric(difftime(circ_data$NAC.End, circ_data$NAC.Start, units="days"))
median_duration <- median(circ_data$NAC.Duration, na.rm=TRUE)
range_duration <- range(circ_data$NAC.Duration, na.rm=TRUE)
Q1 <- quantile(circ_data$NAC.Duration, 0.25, na.rm=TRUE)
Q3 <- quantile(circ_data$NAC.Duration, 0.75, na.rm=TRUE)
IQR_value <- IQR(circ_data$NAC.Duration, na.rm=TRUE)
cat("Median NAC Duration:", median_duration, "days\n")

```

Median NAC Duration: 43 days

[Hide](#)

```
cat("Range of NAC Duration:", range_duration[1], "-", range_duration[2], "days\n")
```

Range of NAC Duration: 28 – 242 days

[Hide](#)

```
cat("Q1 (25th percentile):", Q1, "days\n")
```

Q1 (25th percentile): 42 days

[Hide](#)

```
cat("Q3 (75th percentile):", Q3, "days\n")
```

Q3 (75th percentile): 56 days

[Hide](#)

```
cat("Interquartile Range (IQR):", IQR_value, "days\n")
```

Interquartile Range (IQR): 14 days

#Time of ctDNA draws from start and end of neodjuvant treatment

[Hide](#)

```

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

circ_data$NAC.Start <- as.Date(circ_data$NAC.Start, format="%m/%d/%Y")
circ_data$NAC.End <- as.Date(circ_data$NAC.End, format="%m/%d/%Y")
circ_data$ctDNA.C2D1.Date <- as.Date(circ_data$ctDNA.C2D1.Date, format="%m/%d/%Y")
circ_data$ctDNA.postNAC.Date <- as.Date(circ_data$ctDNA.postNAC.Date, format="%m/%d/%Y")

circ_data$Time_to_ctDNA_C2D1 <- as.numeric(difftime(circ_data$ctDNA.C2D1.Date, circ_data
$NAC.Start, units="days"))
circ_data$Time_to_ctDNA_postNAC <- as.numeric(difftime(circ_data$ctDNA.postNAC.Date, cir
c_data$NAC.End, units="days"))

median_ctDNA_C2D1 <- median(circ_data$Time_to_ctDNA_C2D1, na.rm=TRUE)
range_ctDNA_C2D1 <- range(circ_data$Time_to_ctDNA_C2D1, na.rm=TRUE)
Q1_ctDNA_C2D1 <- quantile(circ_data$Time_to_ctDNA_C2D1, 0.25, na.rm=TRUE)
Q3_ctDNA_C2D1 <- quantile(circ_data$Time_to_ctDNA_C2D1, 0.75, na.rm=TRUE)
IQR_ctDNA_C2D1 <- IQR(circ_data$Time_to_ctDNA_C2D1, na.rm=TRUE)
median_ctDNA_postNAC <- median(circ_data$Time_to_ctDNA_postNAC, na.rm=TRUE)
range_ctDNA_postNAC <- range(circ_data$Time_to_ctDNA_postNAC, na.rm=TRUE)
Q1_ctDNA_postNAC <- quantile(circ_data$Time_to_ctDNA_postNAC, 0.25, na.rm=TRUE)
Q3_ctDNA_postNAC <- quantile(circ_data$Time_to_ctDNA_postNAC, 0.75, na.rm=TRUE)
IQR_ctDNA_postNAC <- IQR(circ_data$Time_to_ctDNA_postNAC, na.rm=TRUE)
cat("Median time from NAC start to ctDNA draw during NAC:", median_ctDNA_C2D1, "days\n")

```

Median time from NAC start to ctDNA draw during NAC: 14 days

[Hide](#)

```
cat("Range:", range_ctDNA_C2D1[1], "-", range_ctDNA_C2D1[2], "days\n")
```

Range: 10 – 111 days

[Hide](#)

```
cat("Q1 (25th percentile):", Q1_ctDNA_C2D1, "days\n")
```

Q1 (25th percentile): 14 days

[Hide](#)

```
cat("Q3 (75th percentile):", Q3_ctDNA_C2D1, "days\n")
```

Q3 (75th percentile): 28 days

[Hide](#)

```
cat("Interquartile Range (IQR):", IQR_ctDNA_C2D1, "days\n\n")
```

Interquartile Range (IQR): 14 days

[Hide](#)

```
cat("Median time from NAC end to ctDNA draw post-NAC:", median_ctDNA_postNAC, "days\n")
```

Median time from NAC end to ctDNA draw post-NAC: 0 days

[Hide](#)

```
cat("Range:", range_ctDNA_postNAC[1], "-", range_ctDNA_postNAC[2], "days\n")
```

Range: -4 – 64 days

[Hide](#)

```
cat("Q1 (25th percentile):", Q1_ctDNA_postNAC, "days\n")
```

Q1 (25th percentile): 0 days

[Hide](#)

```
cat("Q3 (75th percentile):", Q3_ctDNA_postNAC, "days\n")
```

Q3 (75th percentile): 12 days

[Hide](#)

```
cat("Interquartile Range (IQR):", IQR_ctDNA_postNAC, "days\n")
```

Interquartile Range (IQR): 12 days

#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 (Min - Max); n (%)	

Characteristic	N = 62 ¹
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 (Min - Max); n (%)	

Characteristic **N = 62¹**

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 (Min - Max); n (%)[Hide](#)

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

Characteristic **N = 62¹**

Age	66 (34 - 86)
------------	--------------

Gender

Male	39 (63%)
Female	23 (37%)

PrimSite

Stomach	38 (61%)
---------	----------

¹Median (Min - Max); n (%)

Characteristic	N = 62¹
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%)
Stage	
0	3 (4.8%)
I	16 (26%)
II	26 (42%)
III	17 (27%)

¹Median (Min - Max); n (%)

Characteristic	N = 62 ¹
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	
Deceased	19 (31%)
Alive	43 (69%)
FU.months	
	29 (2 - 93)

¹Median (Min - Max); n (%)

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

#Demographics Table for those that included vs excluded from the study

```

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

circ_data_subset1 <- 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", "TRG
5")),
    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", "Ob
servation")),
    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", "Ali
ve")),
    FU.months = as.numeric(FU.months))

circ_data1 <- read.csv("PLAGAST_Clinical Data.csv")

circ_data_subset2 <- 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,
Included) %>%
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", "TRG
5")),
  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", "Ob
servation")),
  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", "Ali
ve")),
  FU.months = as.numeric(FU.months),
  Included = factor(Included, levels = c("TRUE", "FALSE"), labels = c("Included", "Ex
cluded")))
Overall <- circ_data_subset1 %>%
 tbl_summary(
  statistic = list(
    all_continuous() ~ "{median} ({min} - {max})",
    all_categorical() ~ "{n} ({p}%)") %>%
  bold_labels()
Overall

```

Characteristic	N = 82 ¹
Age	67 (34 - 86)
Gender	
Male	51 (62%)
Female	31 (38%)
PrimSite	
Stomach	54 (66%)
G/J	28 (34%)
Oesophagus	0 (0%)
NAC	
Neoadjuvant Therapy	74 (90%)
Upfront Surgery	8 (9.8%)
NAC.Regimen	
	8 (9.8%)
Chemoimmunotherapy	8 (9.8%)
Chemotherapy	65 (79%)
Radiotherapy	1 (1.2%)
TRG.Mandard	
TRG1	4 (5.6%)
TRG2	14 (19%)
TRG3	26 (36%)
TRG4	19 (26%)
TRG5	9 (13%)
Unknown	10
TNM	
T0-TisN0M0	3 (3.7%)
T1-T2N0	25 (30%)

¹ Median (Min - Max); n (%)

Characteristic	N = 82 ¹
T2-T3N0-N1	21 (26%)
T2N1-N2	4 (4.9%)
T3N2-N3	9 (11%)
T4N0-N1	12 (15%)
T4N2-N3	8 (9.8%)
Stage	
0	3 (3.7%)
I	25 (30%)
II	33 (40%)
III	21 (26%)
Grade	
G1	14 (26%)
G2	27 (50%)
G3	13 (24%)
Unknown	28
Lauren Class	
	8 (9.8%)
Diffuse	21 (26%)
Intestinal	45 (55%)
Mixed	8 (9.8%)
Margins	
R0	78 (95%)
R1	4 (4.9%)
ACT	
Adjuvant Treatment	68 (83%)
Observation	14 (17%)

¹ Median (Min - Max); n (%)

Characteristic	N = 82 ¹
ACT.Regimen	
	19 (23%)
Chemoimmunotherapy	4 (4.9%)
Chemotherapy	58 (71%)
Immunotherapy	1 (1.2%)
DFS.Event	
Recurrence	37 (45%)
No Recurrence	45 (55%)
OS.Event	
Deceased	25 (30%)
Alive	57 (70%)
FU.months	
	32 (2 - 93)

¹ Median (Min - Max); n (%)

[Hide](#)

```
ByctDNA_MRD <- circ_data_subset2 %>%
 tbl_summary(
  by = Included, # Subgroup by Included
  statistic = list(
    all_continuous() ~ "{median} ({min} - {max})",
    all_categorical() ~ "{n} ({p}%)") %>%
  add_p(test = all_categorical() ~ "fisher.test", pvalue_fun = ~style_pvalue(p.adjust(.x, method = "bonferroni"))) %>%
  bold_labels()
ByctDNA_MRD
```

Characteristic	Included N = 62 ¹	Excluded N = 20 ¹	p-value ²
Age	66 (34 - 86)	70 (44 - 82)	0.8
Gender			>0.9
Male	39 (63%)	12 (60%)	

¹ Median (Min - Max); n (%)

² Wilcoxon rank sum test; Fisher's exact test

Characteristic	Included N = 62¹	Excluded N = 20¹	p-value²
Female	23 (37%)	8 (40%)	
PrimSite			>0.9
Stomach	38 (61%)	16 (80%)	
G/J	24 (39%)	4 (20%)	
Oesophagus	0 (0%)	0 (0%)	
NAC			>0.9
Neoadjuvant Therapy	55 (89%)	19 (95%)	
Upfront Surgery	7 (11%)	1 (5.0%)	
NAC.Regimen			>0.9
	7 (11%)	1 (5.0%)	
Chemoimmunotherapy	7 (11%)	1 (5.0%)	
Chemotherapy	47 (76%)	18 (90%)	
Radiotherapy	1 (1.6%)	0 (0%)	
TRG.Mandard			>0.9
TRG1	3 (5.5%)	1 (5.9%)	
TRG2	8 (15%)	6 (35%)	
TRG3	21 (38%)	5 (29%)	
TRG4	16 (29%)	3 (18%)	
TRG5	7 (13%)	2 (12%)	
Unknown	7	3	
TNM			>0.9
T0-TisN0M0	3 (4.8%)	0 (0%)	
T1-T2N0	16 (26%)	9 (45%)	
T2-T3N0-N1	17 (27%)	4 (20%)	
T2N1-N2	2 (3.2%)	2 (10%)	

¹ Median (Min - Max); n (%)² Wilcoxon rank sum test; Fisher's exact test

Characteristic	Included N = 62¹	Excluded N = 20¹	p-value²			
T3N2-N3	8 (13%)	1 (5.0%)				
T4N0-N1	9 (15%)	3 (15%)				
T4N2-N3	7 (11%)	1 (5.0%)				
Stage	>0.9					
0	3 (4.8%)	0 (0%)				
I	16 (26%)	9 (45%)				
II	26 (42%)	7 (35%)				
III	17 (27%)	4 (20%)				
Grade	>0.9					
G1	11 (28%)	3 (21%)				
G2	21 (53%)	6 (43%)				
G3	8 (20%)	5 (36%)				
Unknown	22	6				
Lauren.Class	>0.9					
	5 (8.1%)	3 (15%)				
Diffuse	18 (29%)	3 (15%)				
Intestinal	34 (55%)	11 (55%)				
Mixed	5 (8.1%)	3 (15%)				
Margins	0.7					
R0	61 (98%)	17 (85%)				
R1	1 (1.6%)	3 (15%)				
ACT	>0.9					
Adjuvant Treatment	53 (85%)	15 (75%)				
Observation	9 (15%)	5 (25%)				
ACT.Regimen	0.2					
¹ Median (Min - Max); n (%)						
² Wilcoxon rank sum test; Fisher's exact test						

Characteristic	Included N = 62 ¹	Excluded N = 20 ¹	p-value ²
	9 (15%)	10 (50%)	
Chemoimmunotherapy	4 (6.5%)	0 (0%)	
Chemotherapy	48 (77%)	10 (50%)	
Immunotherapy	1 (1.6%)	0 (0%)	
DFS.Event			>0.9
Recurrence	29 (47%)	8 (40%)	
No Recurrence	33 (53%)	12 (60%)	
OS.Event			>0.9
Deceased	19 (31%)	6 (30%)	
Alive	43 (69%)	14 (70%)	
FU.months	29 (2 - 93)	34 (7 - 68)	>0.9

¹ Median (Min - Max); n (%)² Wilcoxon rank sum test; Fisher's exact test

```
merged_table <-tbl_merge(tbls=list(Overall, ByctDNA_MR))  
merged_table
```

Characteristic	Table 1		Table 2	
	N = 82 ¹	Included N = 62 ¹	Excluded N = 20 ¹	p-value ²
Age	67 (34 - 86)	66 (34 - 86)	70 (44 - 82)	0.8
Gender				>0.9
Male	51 (62%)	39 (63%)	12 (60%)	
Female	31 (38%)	23 (37%)	8 (40%)	
PrimSite				>0.9
Stomach	54 (66%)	38 (61%)	16 (80%)	

¹ Median (Min - Max); n (%)² Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1	Table 2		
	N = 82¹	Included N = 62¹	Excluded N = 20¹	p-value²
G/J	28 (34%)	24 (39%)	4 (20%)	
Oesophagus	0 (0%)	0 (0%)	0 (0%)	
NAC				>0.9
Neoadjuvant Therapy	74 (90%)	55 (89%)	19 (95%)	
Upfront Surgery	8 (9.8%)	7 (11%)	1 (5.0%)	
NAC.Regimen				>0.9
	8 (9.8%)	7 (11%)	1 (5.0%)	
Chemoimmunotherapy	8 (9.8%)	7 (11%)	1 (5.0%)	
Chemotherapy	65 (79%)	47 (76%)	18 (90%)	
Radiotherapy	1 (1.2%)	1 (1.6%)	0 (0%)	
TRG.Mandard				>0.9
TRG1	4 (5.6%)	3 (5.5%)	1 (5.9%)	
TRG2	14 (19%)	8 (15%)	6 (35%)	
TRG3	26 (36%)	21 (38%)	5 (29%)	
TRG4	19 (26%)	16 (29%)	3 (18%)	
TRG5	9 (13%)	7 (13%)	2 (12%)	
Unknown	10	7	3	
TNM				>0.9
T0-TisN0M0	3 (3.7%)	3 (4.8%)	0 (0%)	
T1-T2N0	25 (30%)	16 (26%)	9 (45%)	
T2-T3N0-N1	21 (26%)	17 (27%)	4 (20%)	
T2N1-N2	4 (4.9%)	2 (3.2%)	2 (10%)	
T3N2-N3	9 (11%)	8 (13%)	1 (5.0%)	
T4N0-N1	12 (15%)	9 (15%)	3 (15%)	

¹ Median (Min - Max); n (%)² Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1	Table 2		
	N = 82¹	Included N = 62¹	Excluded N = 20¹	p-value²
T4N2-N3	8 (9.8%)	7 (11%)	1 (5.0%)	
Stage				>0.9
0	3 (3.7%)	3 (4.8%)	0 (0%)	
I	25 (30%)	16 (26%)	9 (45%)	
II	33 (40%)	26 (42%)	7 (35%)	
III	21 (26%)	17 (27%)	4 (20%)	
Grade				>0.9
G1	14 (26%)	11 (28%)	3 (21%)	
G2	27 (50%)	21 (53%)	6 (43%)	
G3	13 (24%)	8 (20%)	5 (36%)	
Unknown	28	22	6	
Lauren.Class				>0.9
	8 (9.8%)	5 (8.1%)	3 (15%)	
Diffuse	21 (26%)	18 (29%)	3 (15%)	
Intestinal	45 (55%)	34 (55%)	11 (55%)	
Mixed	8 (9.8%)	5 (8.1%)	3 (15%)	
Margins				0.7
R0	78 (95%)	61 (98%)	17 (85%)	
R1	4 (4.9%)	1 (1.6%)	3 (15%)	
ACT				>0.9
Adjuvant Treatment	68 (83%)	53 (85%)	15 (75%)	
Observation	14 (17%)	9 (15%)	5 (25%)	
ACT.Regimen				0.2
	19 (23%)	9 (15%)	10 (50%)	

¹ Median (Min - Max); n (%)² Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1		Table 2	
	N = 82¹	Included N = 62¹	Excluded N = 20¹	p-value²
Chemoimmunotherapy	4 (4.9%)	4 (6.5%)	0 (0%)	
Chemotherapy	58 (71%)	48 (77%)	10 (50%)	
Immunotherapy	1 (1.2%)	1 (1.6%)	0 (0%)	
DFS.Event				>0.9
Recurrence	37 (45%)	29 (47%)	8 (40%)	
No Recurrence	45 (55%)	33 (53%)	12 (60%)	
OS.Event				>0.9
Deceased	25 (30%)	19 (31%)	6 (30%)	
Alive	57 (70%)	43 (69%)	14 (70%)	
FU.months	32 (2 - 93)	29 (2 - 93)	34 (7 - 68)	>0.9

¹ Median (Min - Max); n (%)

² Wilcoxon rank sum test; Fisher's exact test

Hide

```
fit1 <- as_flex_table(
  merged_table,
  include = everything(),
  return_calls = FALSE)
fit1
```

Characteristic	Table 1		Table 2	
	N = 82¹	Included N = 62¹	Excluded N = 20¹	p-value²
Age	67 (34 - 86)	66 (34 - 86)	70 (44 - 82)	0.8
Gender				>0.9
Male	51 (62%)	39 (63%)	12 (60%)	
Female	31 (38%)	23 (37%)	8 (40%)	
PrimSite				>0.9
Stomach	54 (66%)	38 (61%)	16 (80%)	

¹Median (Min - Max); n (%)

²Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1	Table 2		p-value²
	N = 82¹	Included N = 62¹	Excluded N = 20¹	
G/J	28 (34%)	24 (39%)	4 (20%)	
Oesophagus	0 (0%)	0 (0%)	0 (0%)	
NAC				>0.9
Neoadjuvant Therapy	74 (90%)	55 (89%)	19 (95%)	
Upfront Surgery	8 (9.8%)	7 (11%)	1 (5.0%)	
NAC.Regimen				>0.9
	8 (9.8%)	7 (11%)	1 (5.0%)	
Chemoimmunotherapy	8 (9.8%)	7 (11%)	1 (5.0%)	
Chemotherapy	65 (79%)	47 (76%)	18 (90%)	
Radiotherapy	1 (1.2%)	1 (1.6%)	0 (0%)	
TRG.Mandard				>0.9
TRG1	4 (5.6%)	3 (5.5%)	1 (5.9%)	
TRG2	14 (19%)	8 (15%)	6 (35%)	
TRG3	26 (36%)	21 (38%)	5 (29%)	
TRG4	19 (26%)	16 (29%)	3 (18%)	
TRG5	9 (13%)	7 (13%)	2 (12%)	
Unknown	10	7	3	
TNM				>0.9
T0-TisN0M0	3 (3.7%)	3 (4.8%)	0 (0%)	
T1-T2N0	25 (30%)	16 (26%)	9 (45%)	
T2-T3N0-N1	21 (26%)	17 (27%)	4 (20%)	
T2N1-N2	4 (4.9%)	2 (3.2%)	2 (10%)	
T3N2-N3	9 (11%)	8 (13%)	1 (5.0%)	
T4N0-N1	12 (15%)	9 (15%)	3 (15%)	
T4N2-N3	8 (9.8%)	7 (11%)	1 (5.0%)	
Stage				>0.9
0	3 (3.7%)	3 (4.8%)	0 (0%)	
I	25 (30%)	16 (26%)	9 (45%)	

¹Median (Min - Max); n (%)²Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1	Table 2		p-value²
	N = 82¹	Included N = 62¹	Excluded N = 20¹	
II	33 (40%)	26 (42%)	7 (35%)	
III	21 (26%)	17 (27%)	4 (20%)	
Grade				>0.9
G1	14 (26%)	11 (28%)	3 (21%)	
G2	27 (50%)	21 (53%)	6 (43%)	
G3	13 (24%)	8 (20%)	5 (36%)	
Unknown	28	22	6	
Lauren.Class				>0.9
	8 (9.8%)	5 (8.1%)	3 (15%)	
Diffuse	21 (26%)	18 (29%)	3 (15%)	
Intestinal	45 (55%)	34 (55%)	11 (55%)	
Mixed	8 (9.8%)	5 (8.1%)	3 (15%)	
Margins				0.7
R0	78 (95%)	61 (98%)	17 (85%)	
R1	4 (4.9%)	1 (1.6%)	3 (15%)	
ACT				>0.9
Adjuvant Treatment	68 (83%)	53 (85%)	15 (75%)	
Observation	14 (17%)	9 (15%)	5 (25%)	
ACT.Regimen				0.2
	19 (23%)	9 (15%)	10 (50%)	
Chemoimmunotherapy	4 (4.9%)	4 (6.5%)	0 (0%)	
Chemotherapy	58 (71%)	48 (77%)	10 (50%)	
Immunotherapy	1 (1.2%)	1 (1.6%)	0 (0%)	
DFS.Event				>0.9
Recurrence	37 (45%)	29 (47%)	8 (40%)	
No Recurrence	45 (55%)	33 (53%)	12 (60%)	
OS.Event				>0.9
Deceased	25 (30%)	19 (31%)	6 (30%)	

¹Median (Min - Max); n (%)²Wilcoxon rank sum test; Fisher's exact test

Characteristic	Table 1		Table 2	
	N = 82 ¹	Included N = 62 ¹	Excluded N = 20 ¹	p-value ²
Alive	57 (70%)	43 (69%)	14 (70%)	
FU.months	32 (2 - 93)	29 (2 - 93)	34 (7 - 68)	>0.9

¹Median (Min - Max); n (%)²Wilcoxon rank sum test; Fisher's exact test

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

#RFS by inclusion status

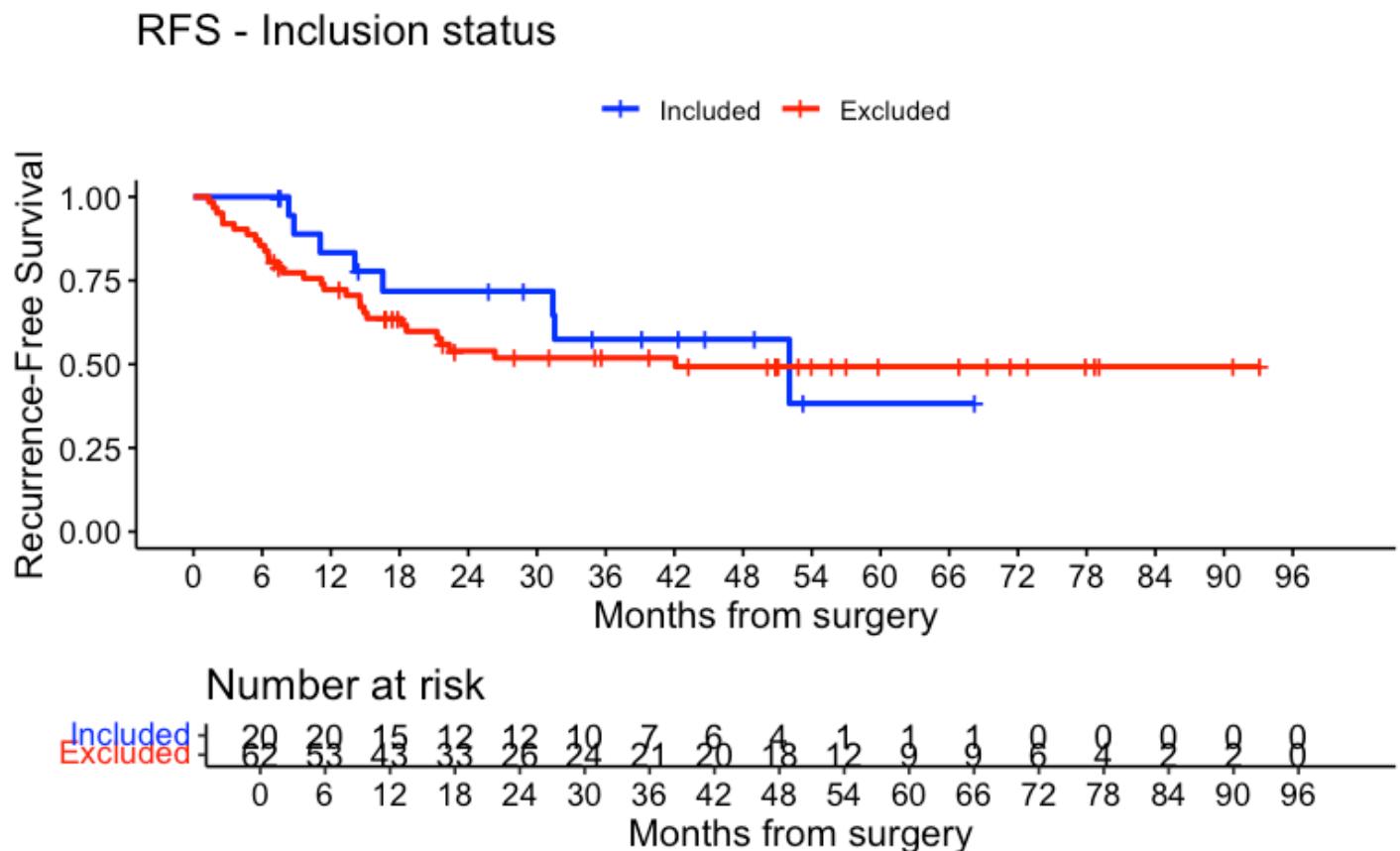
```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("PLAGAST_Clinical Data.csv")

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

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

	n	events	median	0.95LCL	0.95UCL
Included=FALSE	20	8	52.0	31.4	NA
Included=TRUE	62	29	42.1	18.6	NA

```
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ Included, 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 - Inclusion status", ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("Included", "Excluded"), legend.title="")
```



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

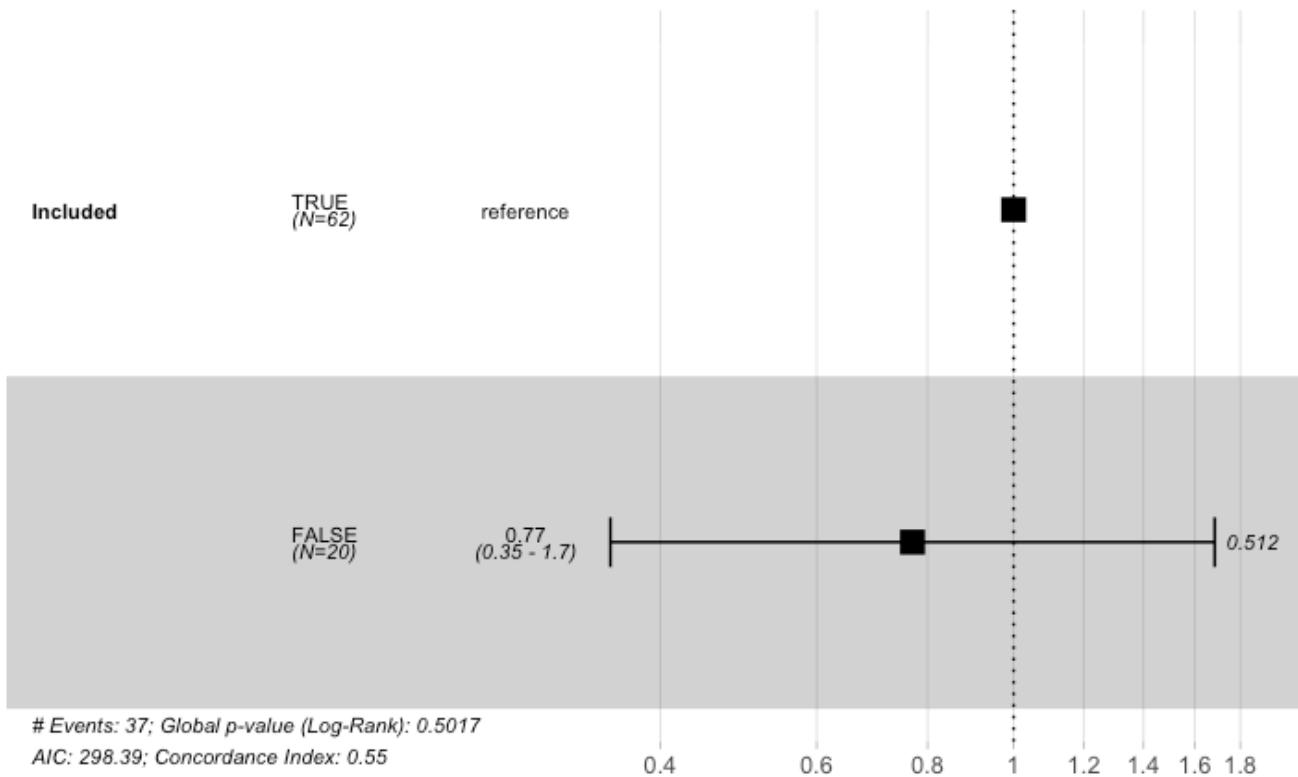
```
Call: survfit(formula = surv_object ~ Included, data = circ_data, conf.int = 0.95,
  conf.type = "log-log")
```

Included=FALSE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
72	24.000	12.000	5.000	0.718	0.107	0.449		0.8	
Included=TRUE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
82	24.0000	26.0000	27.0000	0.5396	0.0661	0.4024		0.65	

[Hide](#)

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

Hazard ratio



[Hide](#)

```
summary(cox_fit)
```

Call:

coxph(formula = surv_object ~ Included, data = circ_data)

n= 82, number of events= 37

	coef	exp(coef)	se(coef)	z	Pr(> z)
Included=FALSE	-0.2622	0.7693	0.3998	-0.656	0.512

	exp(coef)	exp(-coef)	lower .95	upper .95
Included=FALSE	0.7693	1.3	0.3514	1.684

Concordance= 0.546 (se = 0.031)

Likelihood ratio test= 0.45 on 1 df, p=0.5

Wald test = 0.43 on 1 df, p=0.5

Score (logrank) test = 0.43 on 1 df, p=0.5

Hide

cox_fit_summary <- summary(cox_fit)

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

[1] "HR = 0.77 (0.35-1.68); p = 0.512"

#OS by inclusion status

Hide

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

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

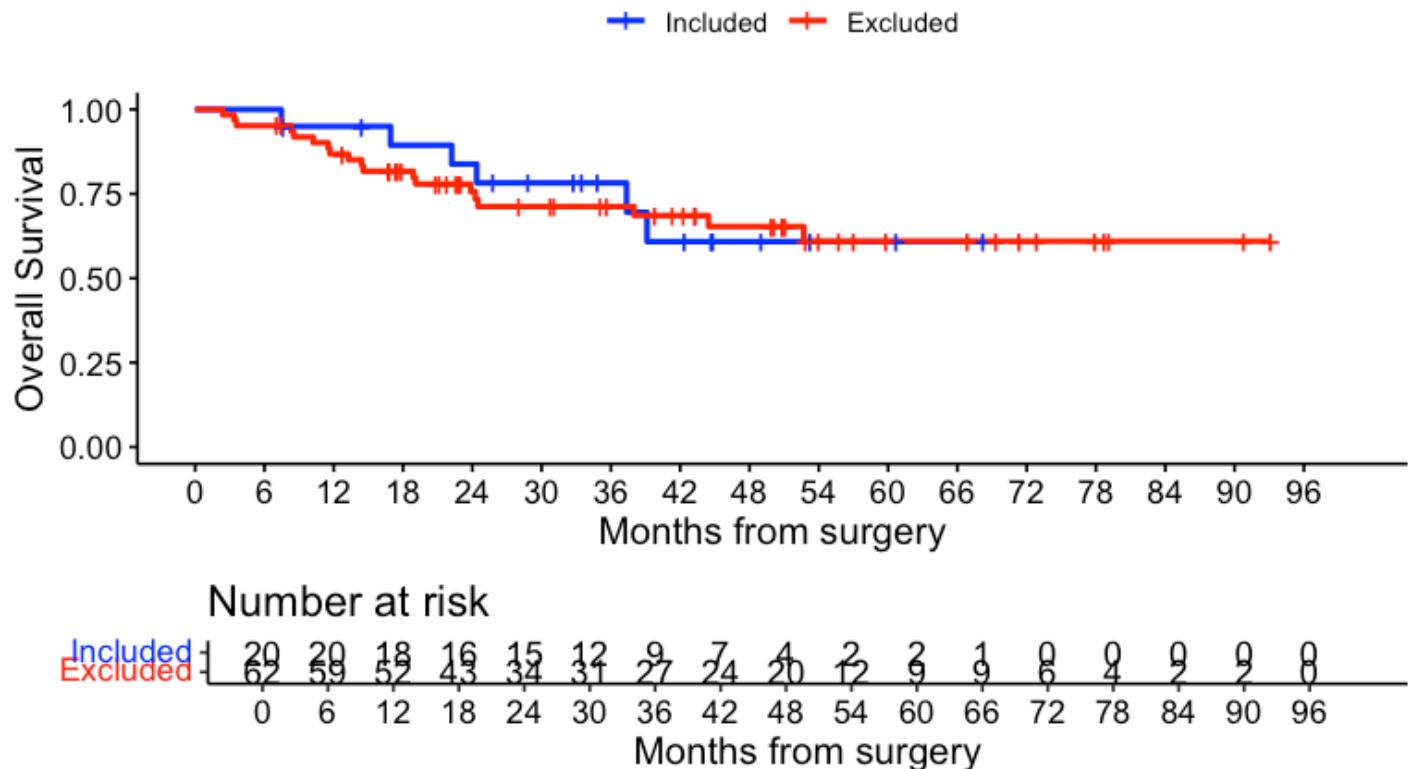
Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~ Included, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
Included=FALSE	20	6	NA	37.4	NA
Included=TRUE	62	19	NA	52.7	NA

Hide

```
surv_object <- Surv(time = circ_data$FU.months, event = circ_data$OS.Event)
KM_curve <- survfit(surv_object ~ Included, 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 - Inclusion status", ylab= "Overall Survival", xlab="Months from surgery", legend.labs=c("Included", "Excluded"), legend.title="")
```

OS - Inclusion status



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

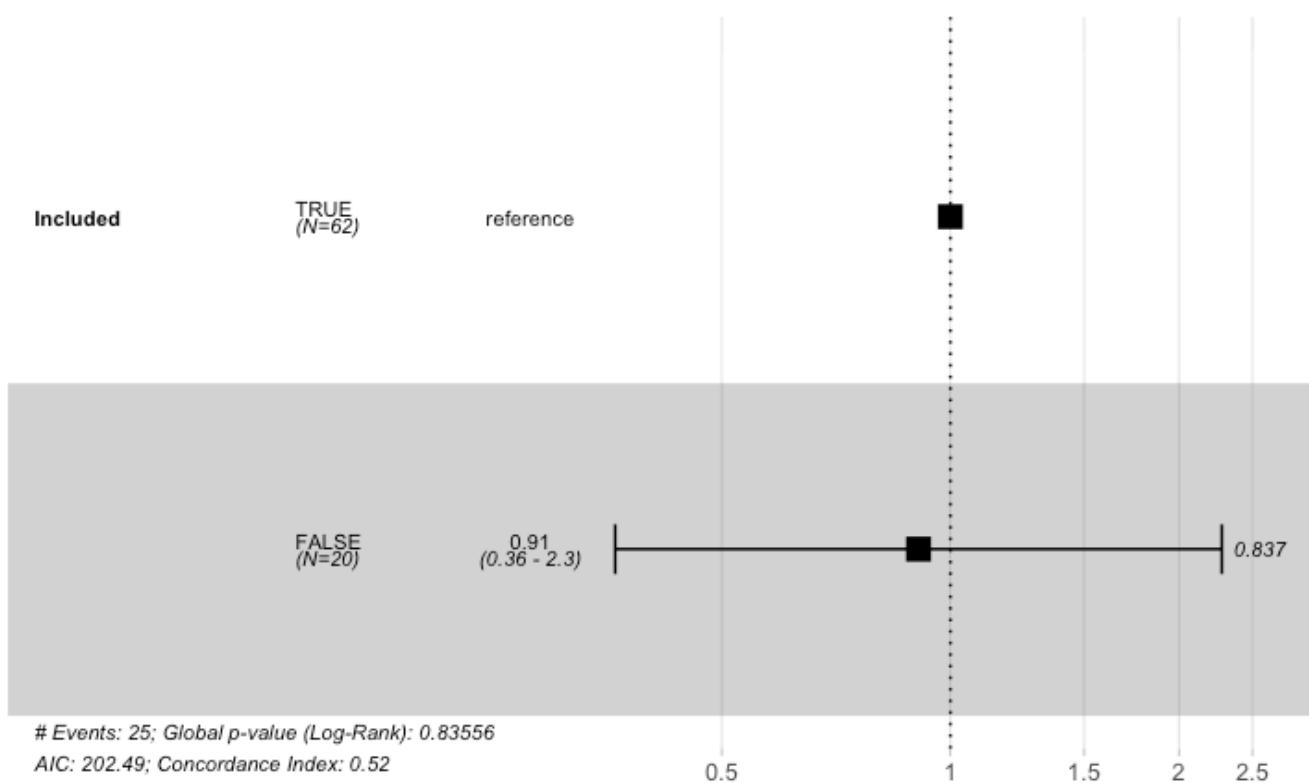
```
Call: survfit(formula = surv_object ~ Included, data = circ_data, conf.int = 0.95,
  conf.type = "log-log")
```

CI	Included=FALSE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.0000	15.0000	3.0000	0.8382	0.0858	0.5770	0.94	

CI	Included=TRUE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.0000	34.0000	14.0000	0.7568	0.0572	0.6225	0.84	

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

Hazard ratio



```
summary(cox_fit)
```

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

n= 82, number of events= 25

	coef	exp(coef)	se(coef)	z	Pr(> z)
IncludedFALSE	-0.09663	0.90789	0.46937	-0.206	0.837

	exp(coef)	exp(-coef)	lower .95	upper .95
IncludedFALSE	0.9079	1.101	0.3618	2.278

Concordance= 0.524 (se = 0.041)
 Likelihood ratio test= 0.04 on 1 df, p=0.8
 Wald test = 0.04 on 1 df, p=0.8
 Score (logrank) test = 0.04 on 1 df, p=0.8

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 0.91 (0.36-2.28); p = 0.837"
```

#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()

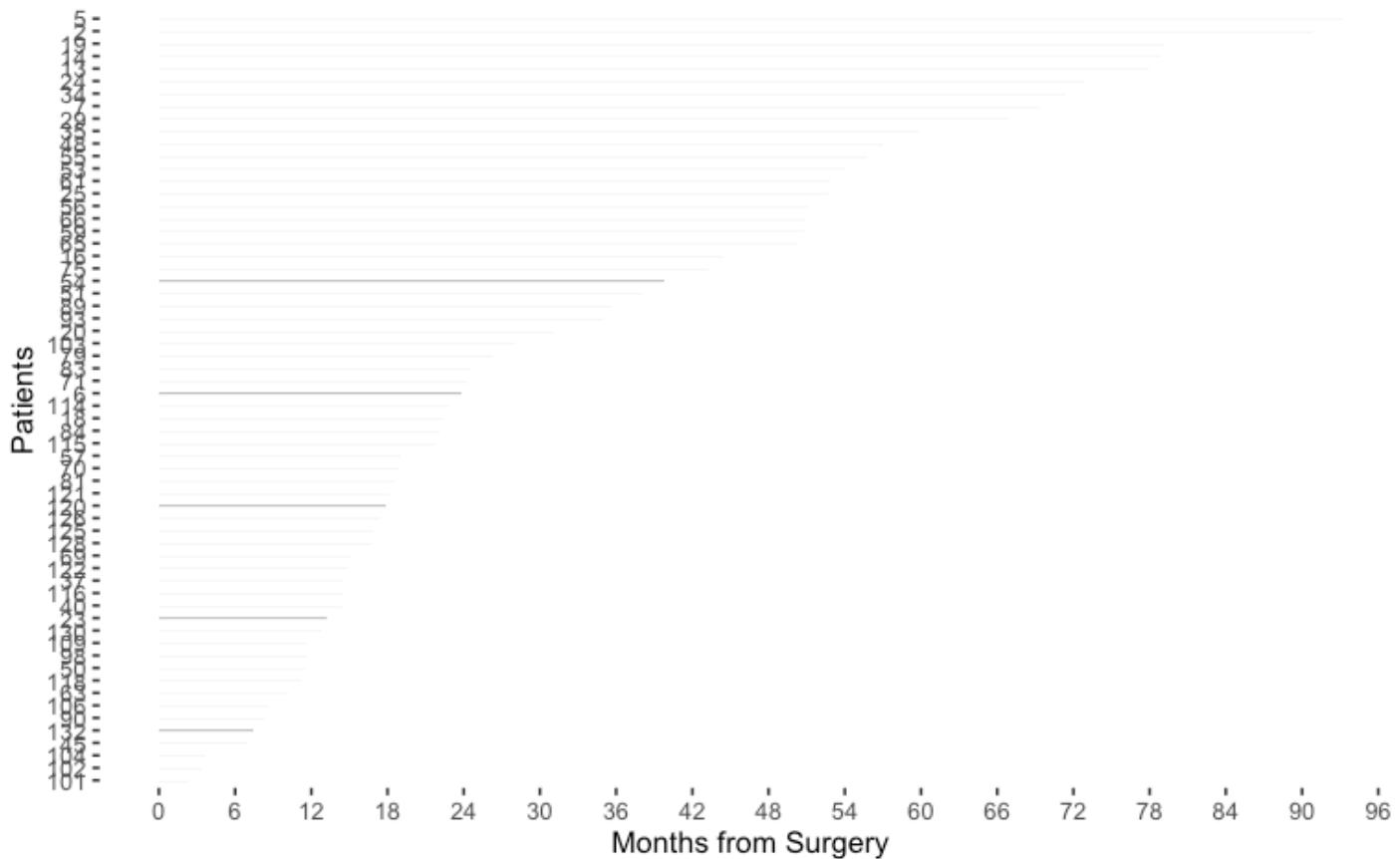
```

#Overview Plot

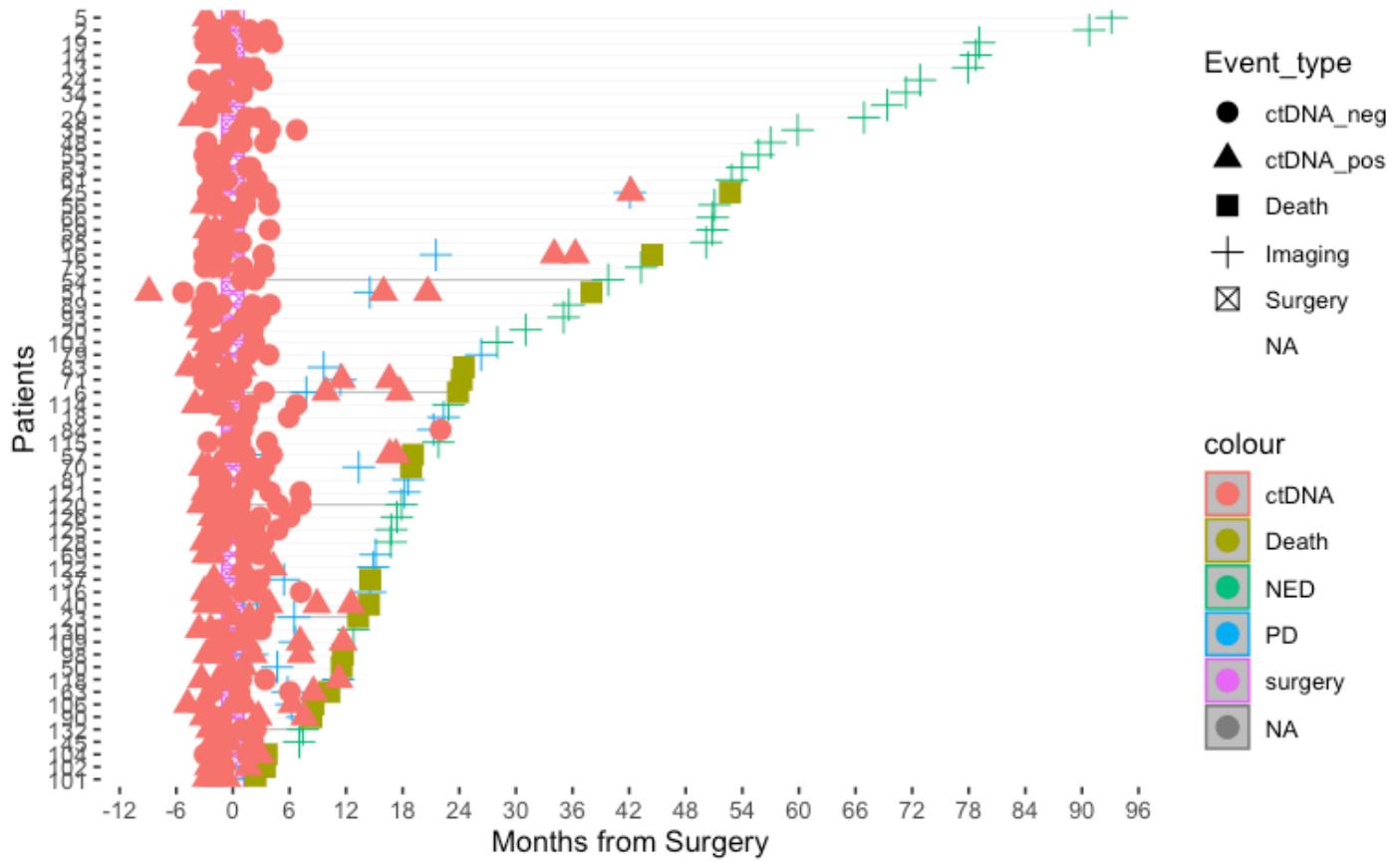
```
setwd("~/Downloads")
clinstage<- read.csv("PLAGAST_OP.csv")
clinstage_df<- as.data.frame(clinstage)

#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](#)

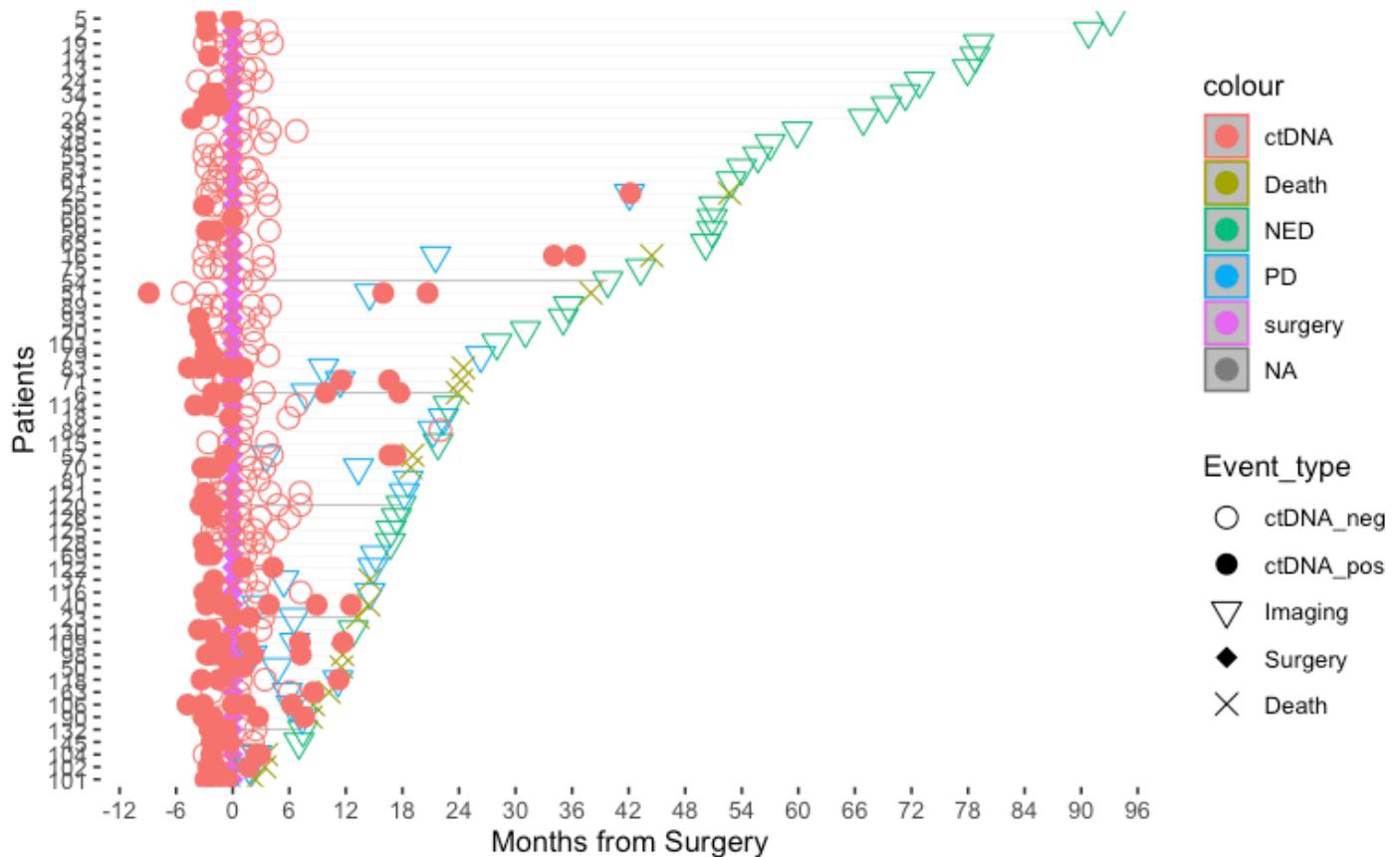
```
##plot events
oplot_ev1 <- oplot + swimmer_points(df_points=clinstage_df,
                                       id='PatientName',
                                       time='date.diff.months',
                                       name_shape ='Event_type',
                                    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



```
#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'))

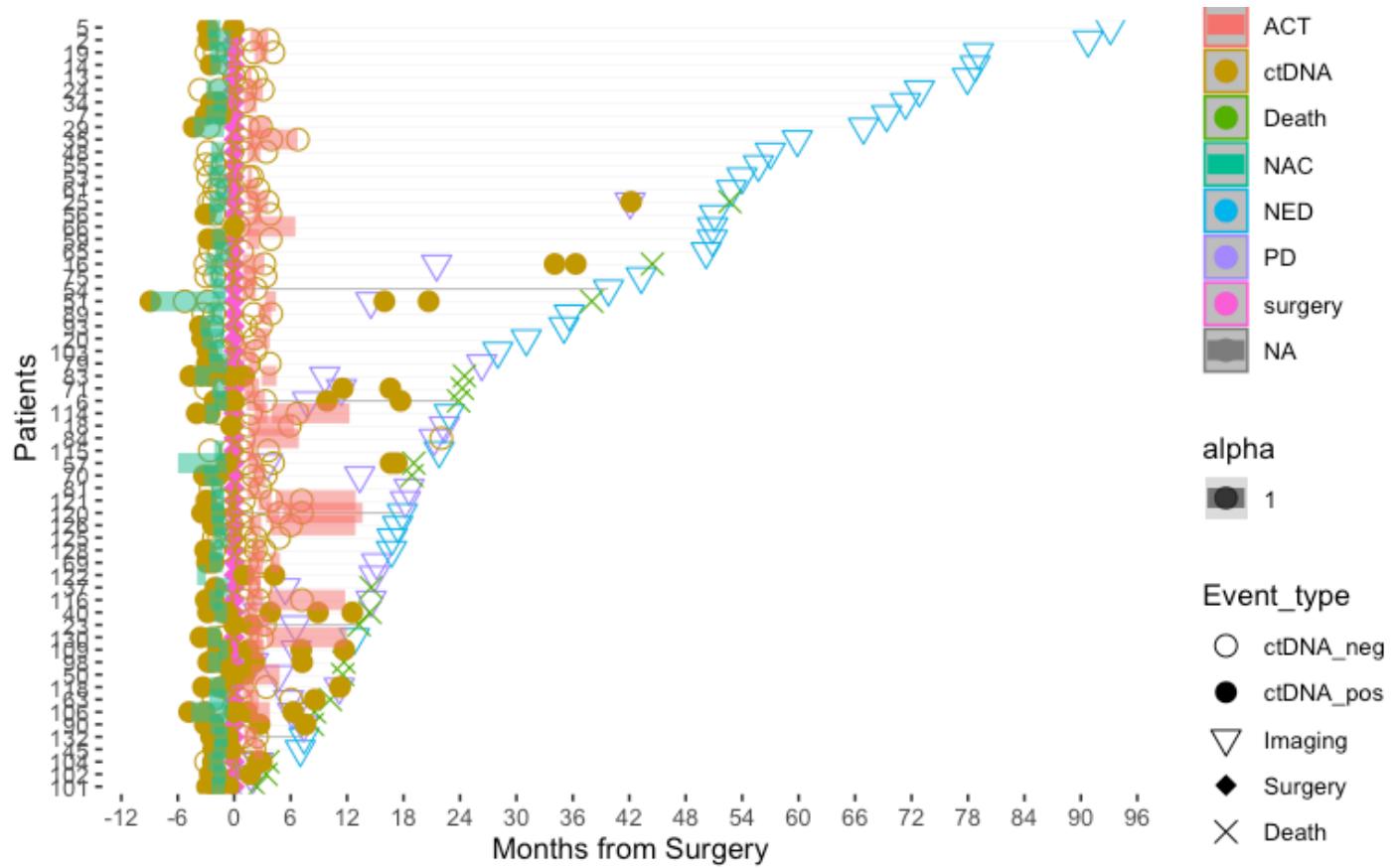
oplot_ev1.1
```



Hide

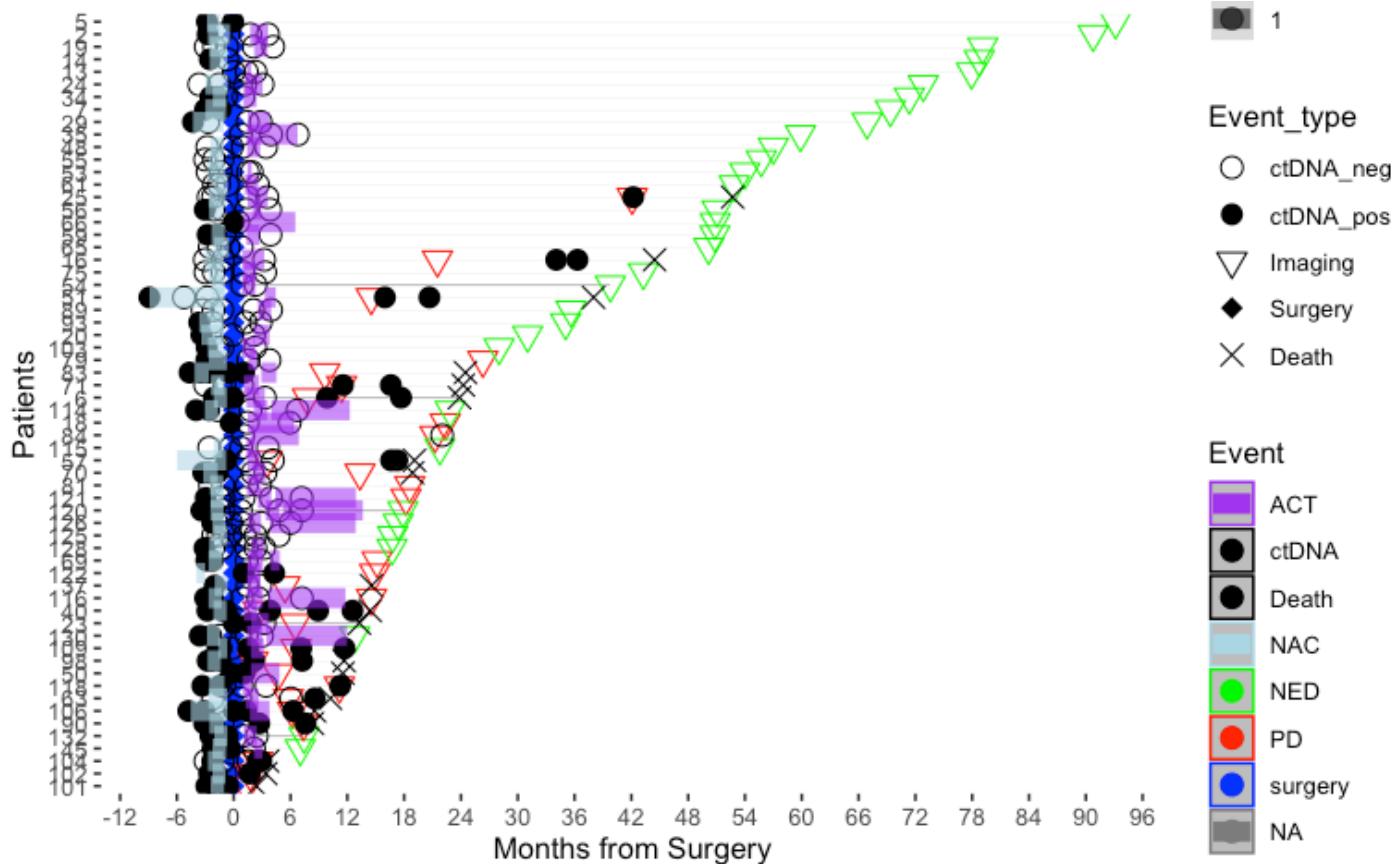
```
#plot treatment

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
```



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
```



#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)
```

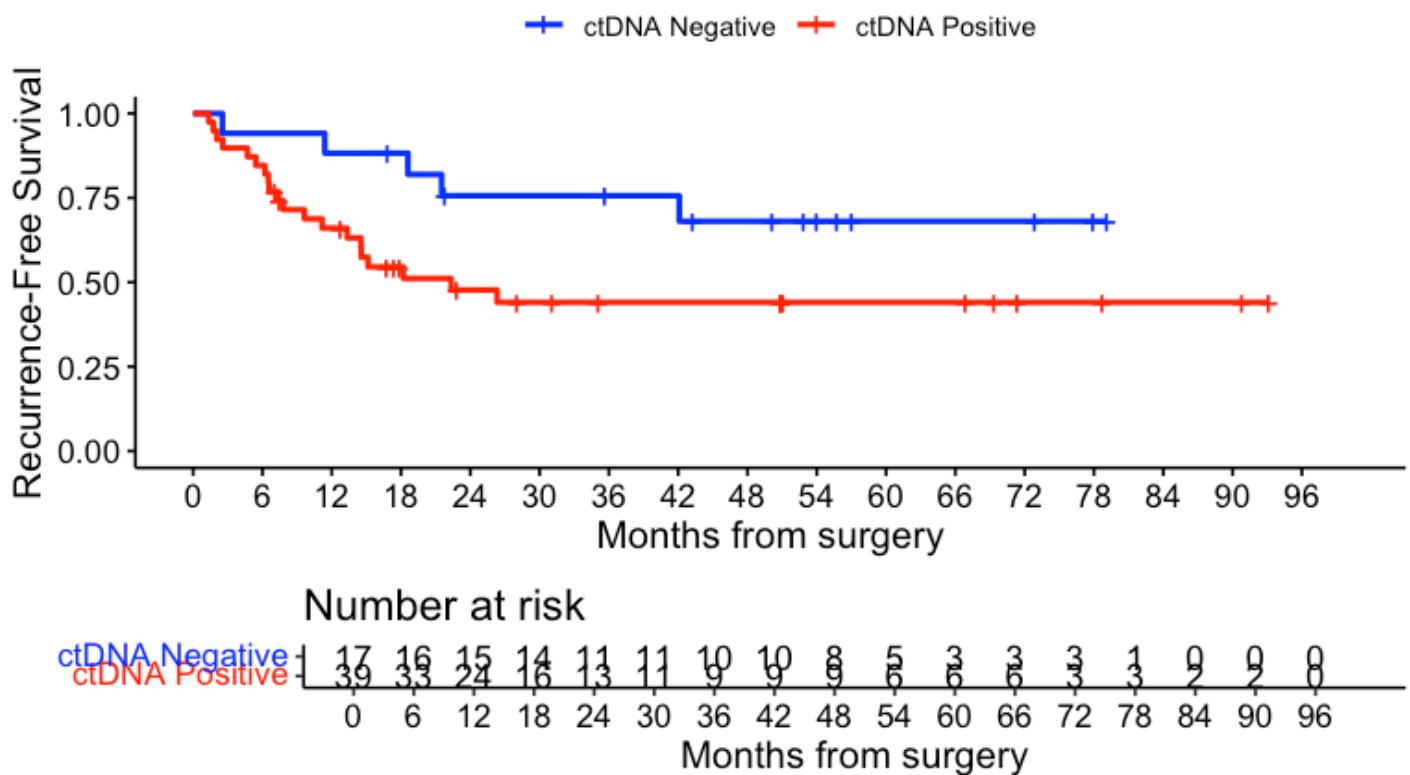
Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.Base, data = circ_data)

	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



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

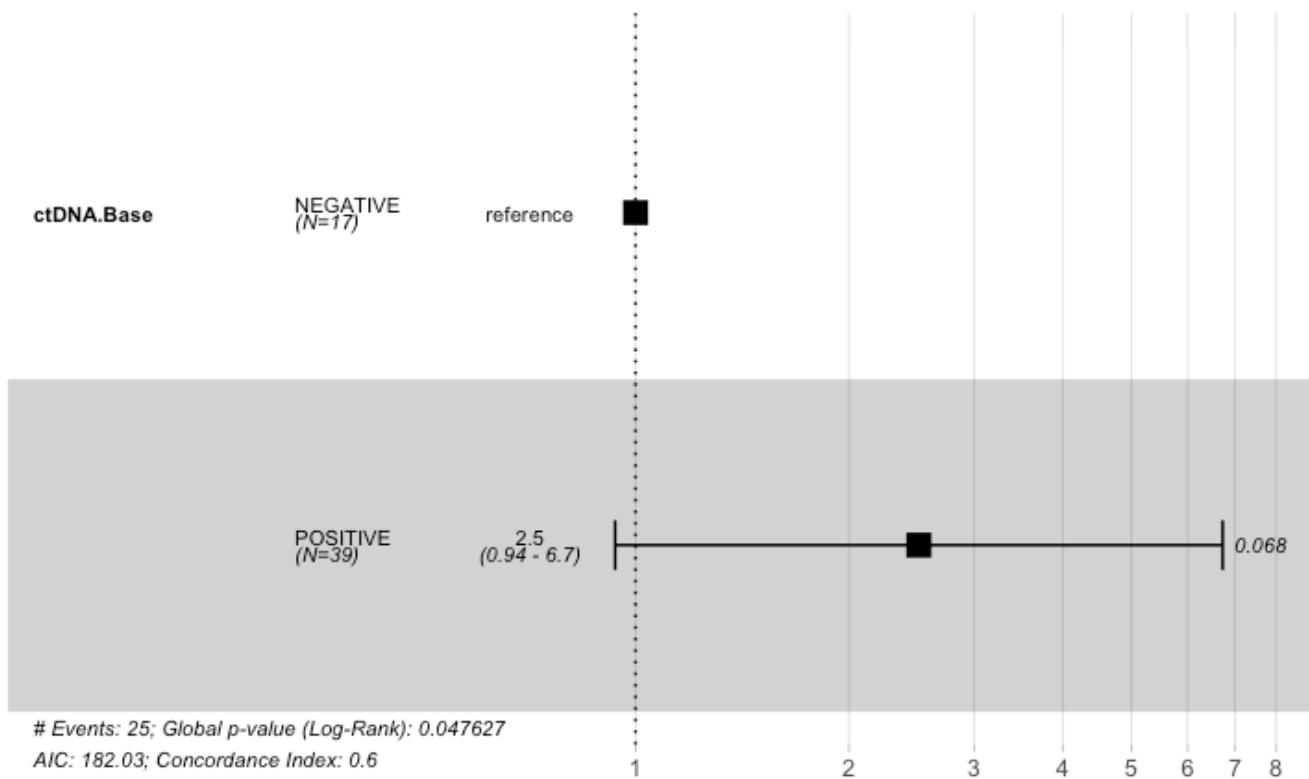
```
Call: survfit(formula = surv_object ~ ctDNA.Base, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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

ctDNA.Base=POSITIVE						
CI	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

```
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



```
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

```
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)
```

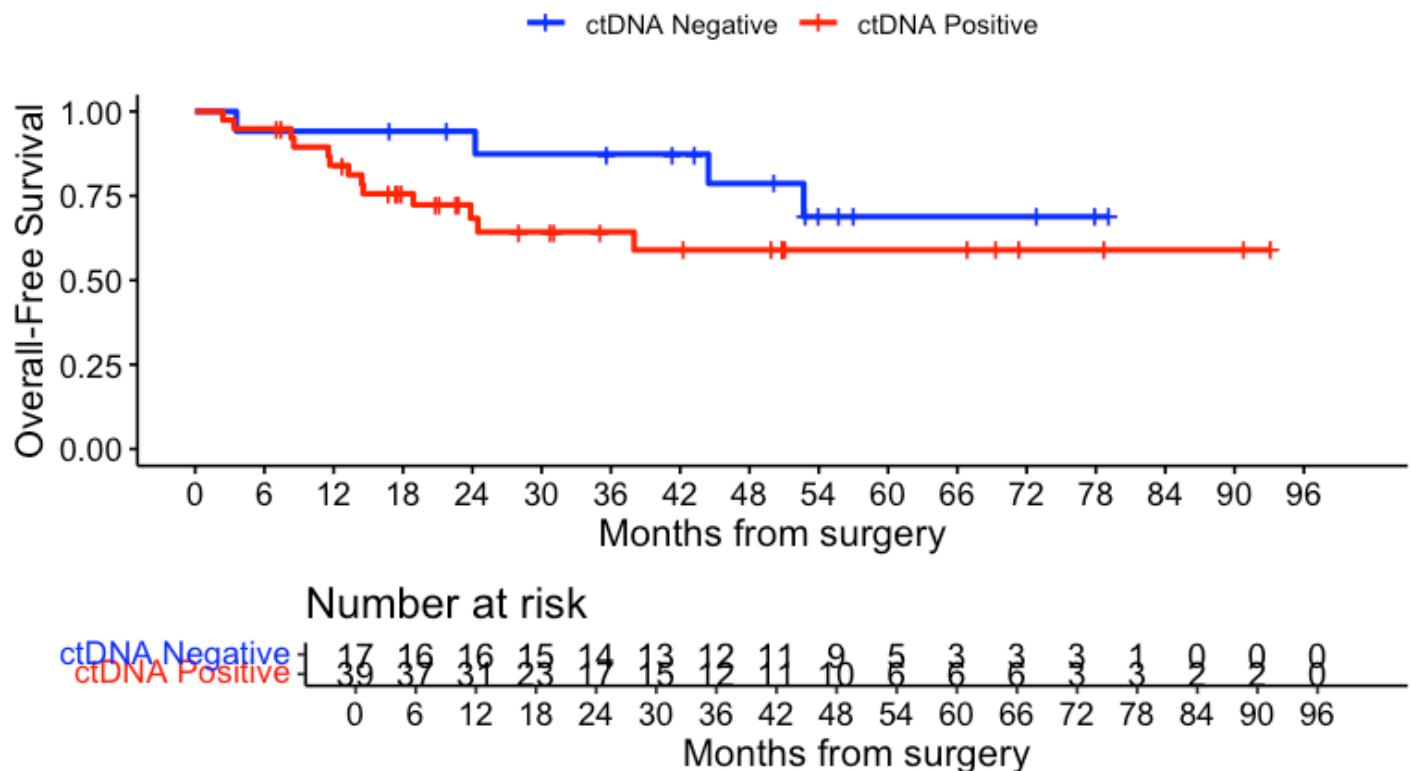
Call: survfit(formula = Surv(time = circ_data\$FU.months, event = circ_data\$OS.Event) ~ ctDNA.Base, data = circ_data)

	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="")
```

OS - ctDNA Baseline | All pts



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

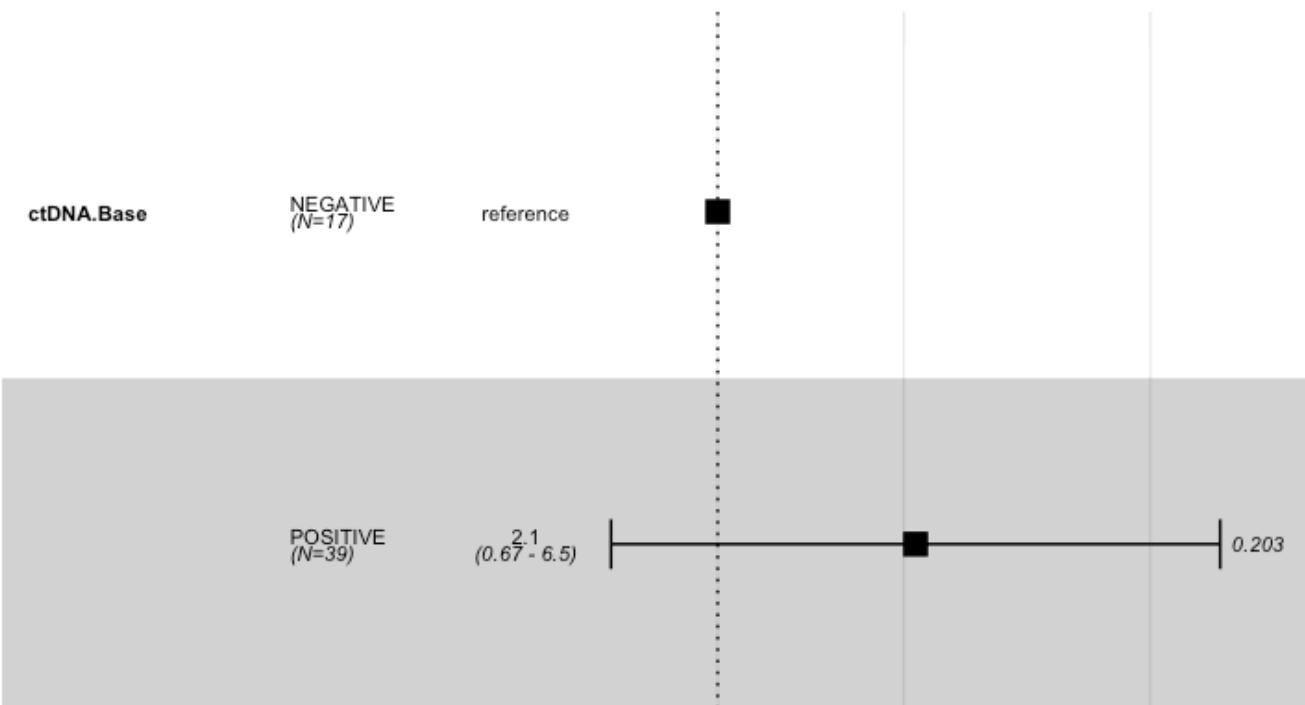
```
Call: survfit(formula = surv_object ~ ctDNA.Base, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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



[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

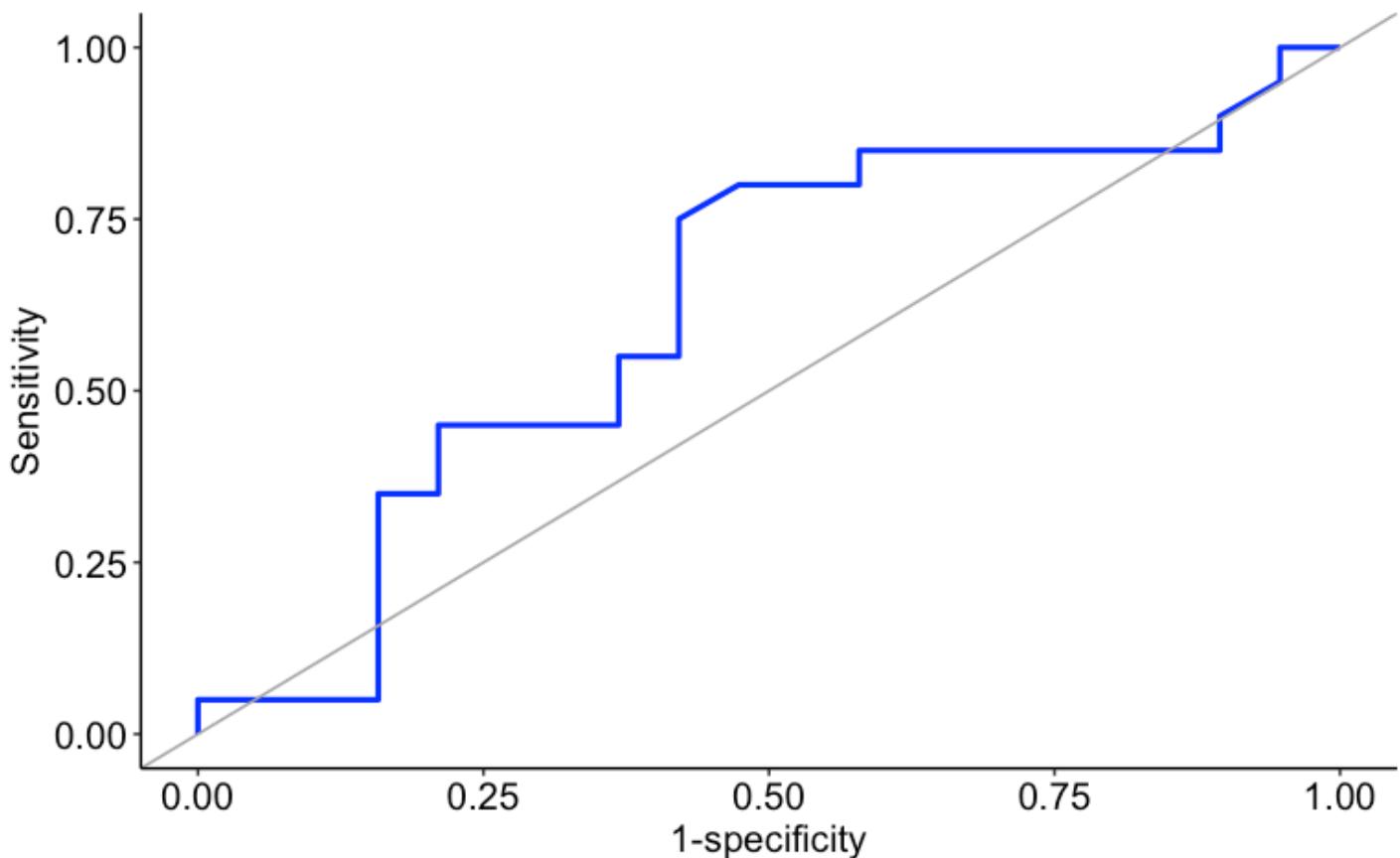
Hide

```
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: Ignoring unknown parameters: `aes`
Warning: Using `size` aesthetic for lines was
deprecated in ggplot2 3.4.0.
Please use `linewidth` instead.

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)

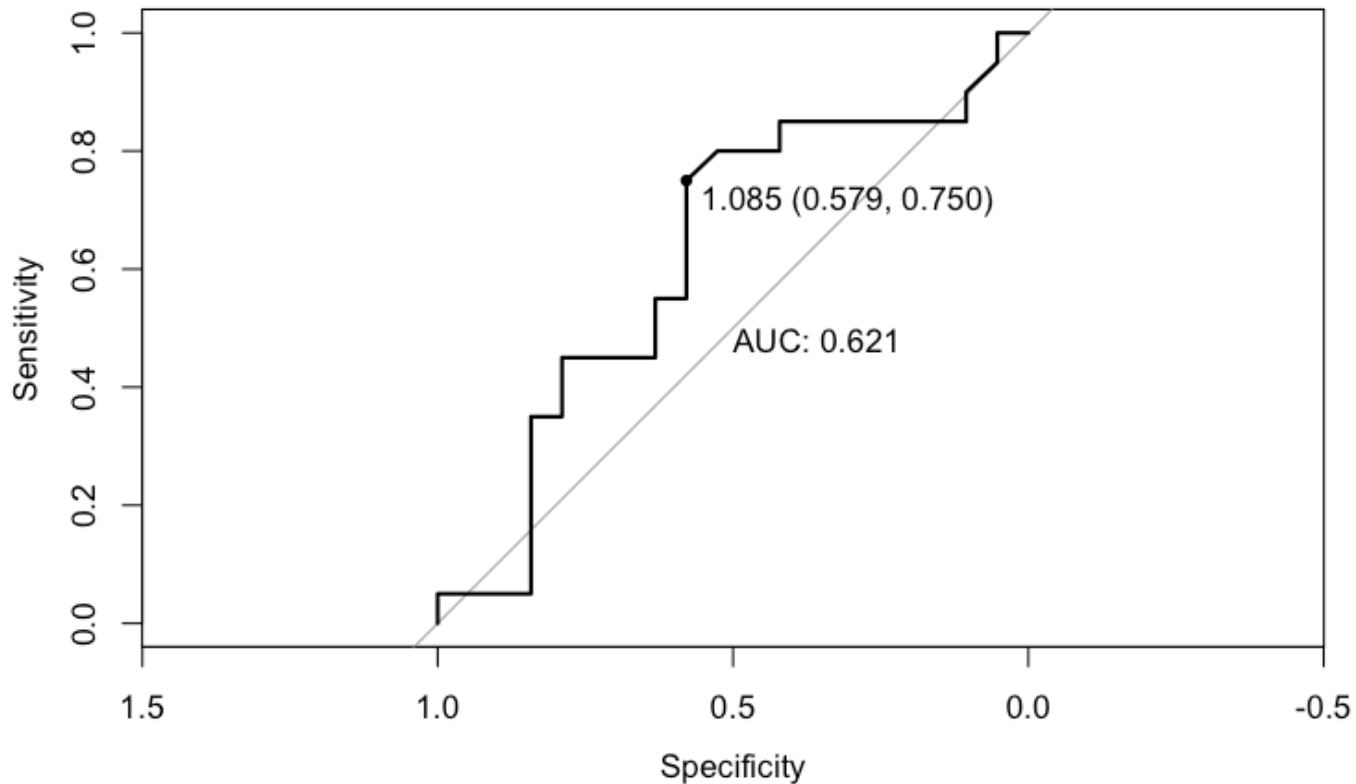
Hide

```
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

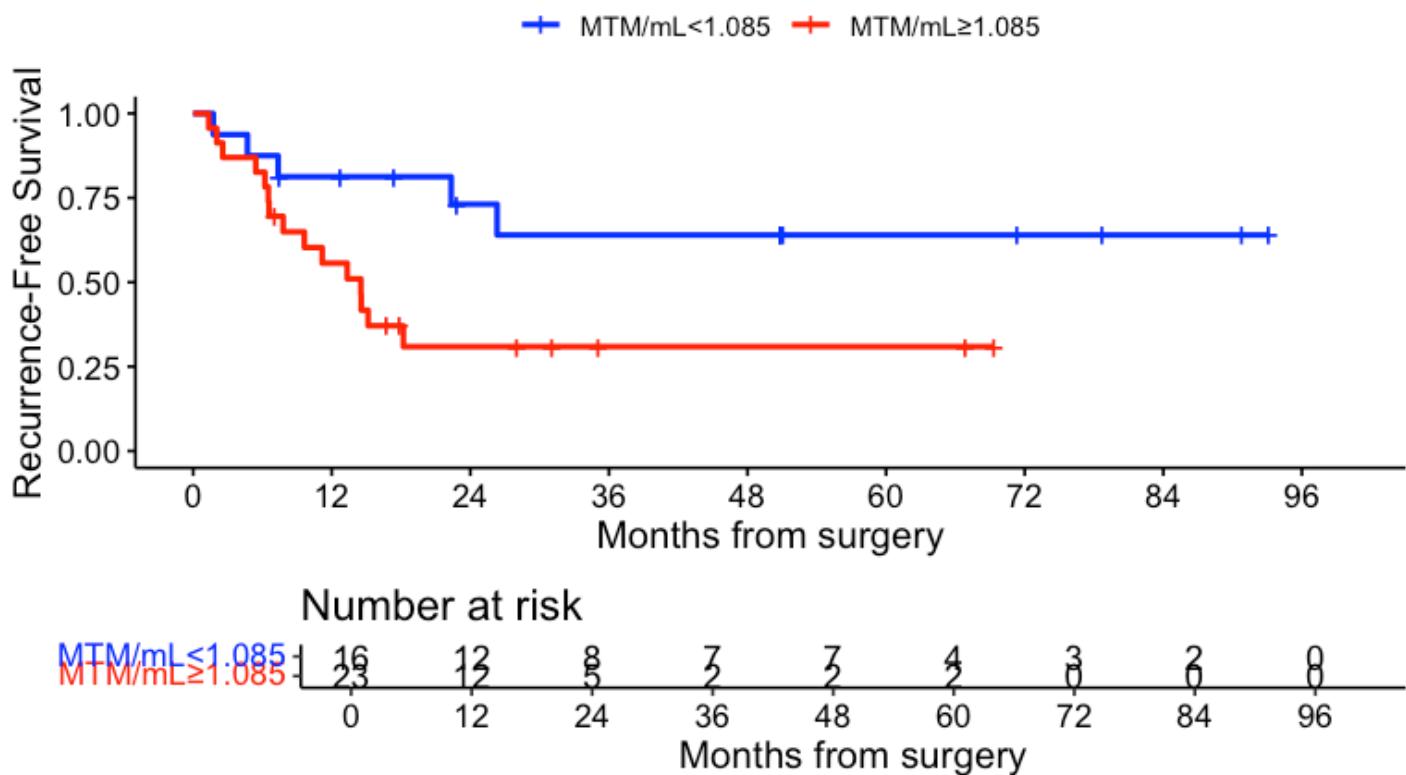
Hide

```

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="RFS - ctDNA MTM/mL groups at Baseline", ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("MTM/mL <1.085", "MTM/mL≥1.085"), legend.title="")

```

RFS - ctDNA MTM/mL groups at Baseline



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

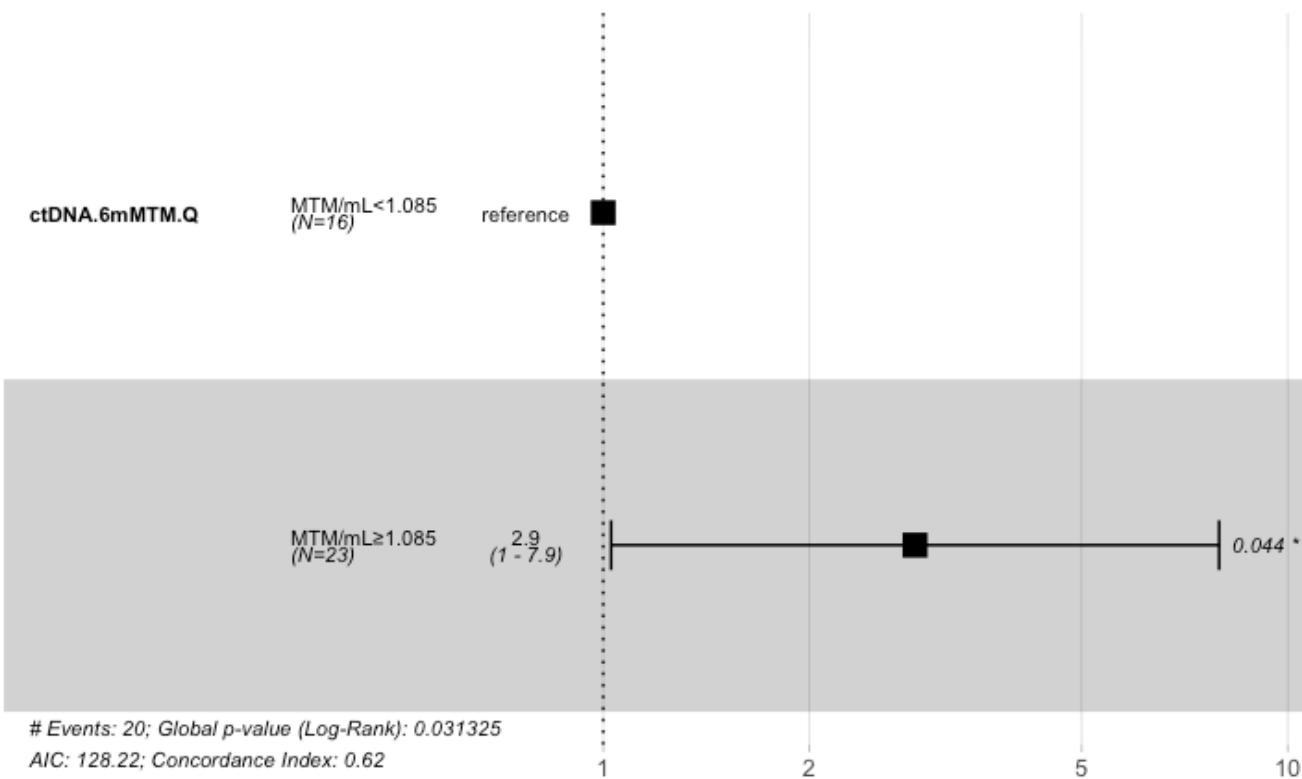
```
Call: survfit(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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	

```
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)
```

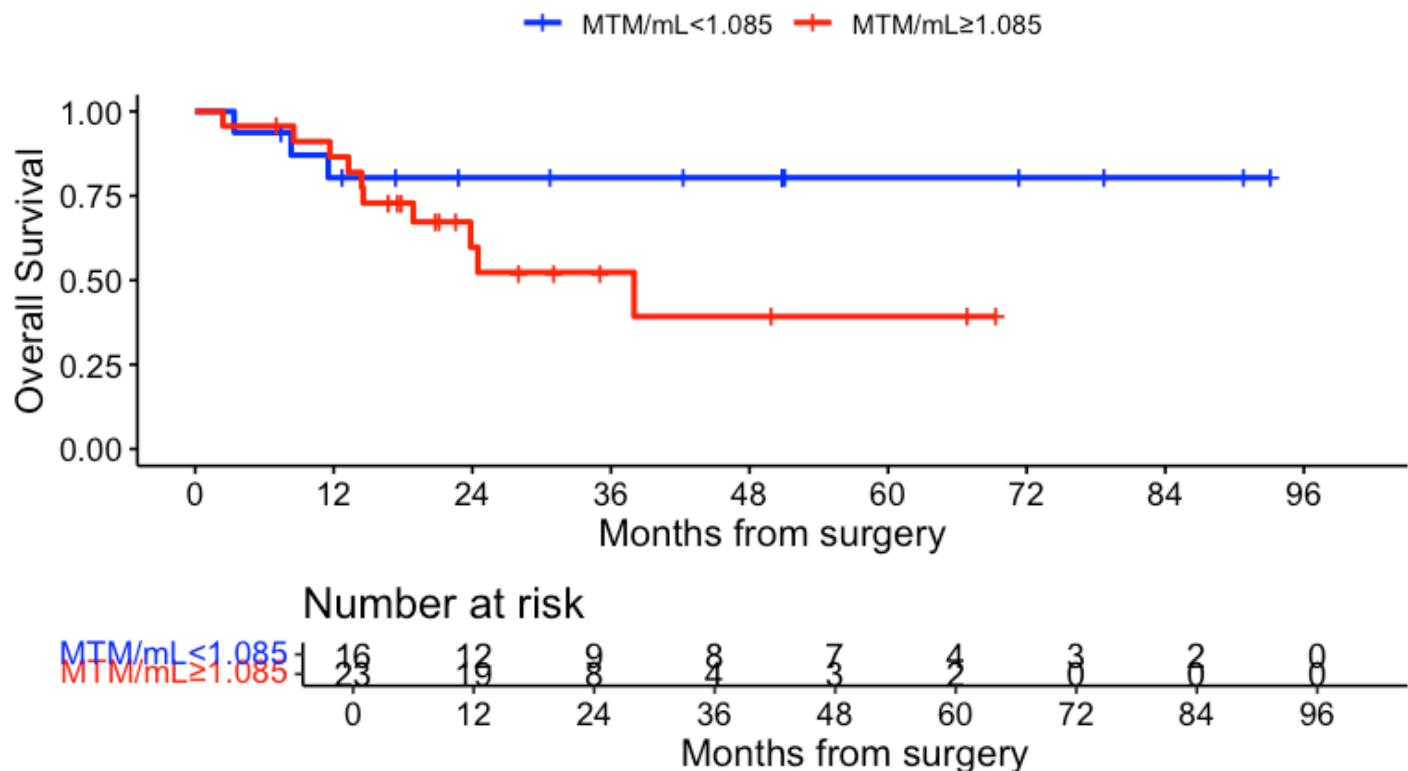
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.6mMTM.Q, data = circ_data)
```

	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



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

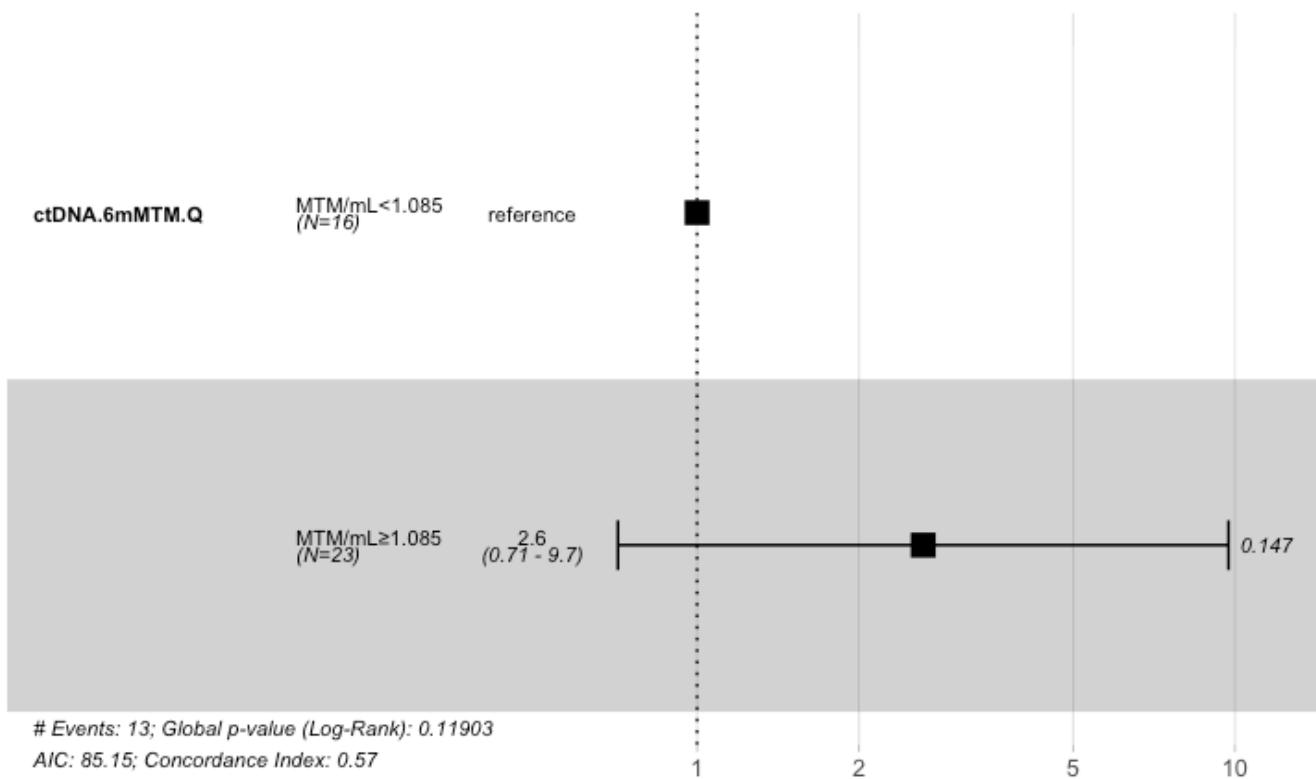
```
Call: survfit(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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

```
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 at Baseline - excluding pts that did not receive NAC

[Hide](#)

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

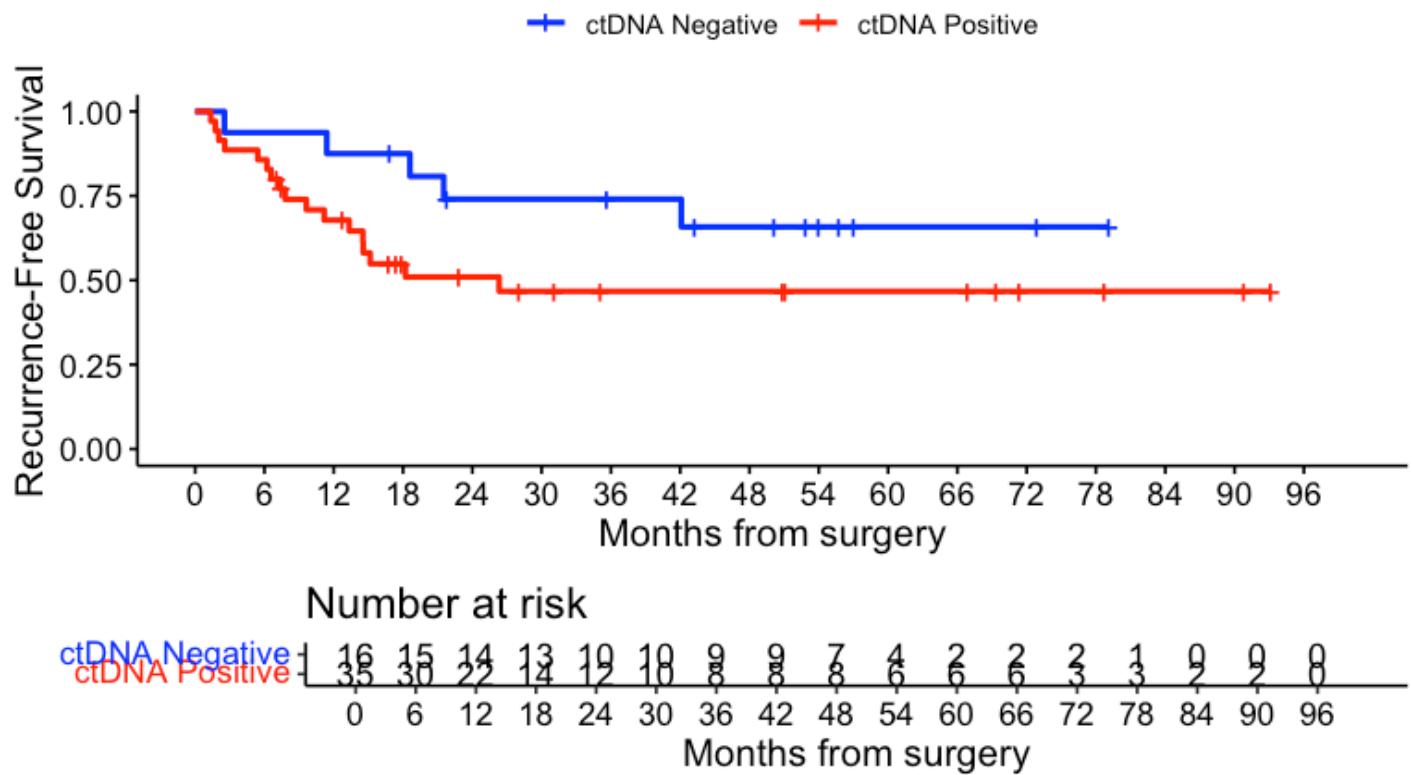
Call: survfit(formula = Surv(time = circ_data\$DFS.months, event = circ_data\$DFS.Event) ~ ctDNA.Base, data = circ_data)

	n	events	median	0.95LCL	0.95UCL
ctDNA.Base=NEGATIVE	16	5	NA	42.1	NA
ctDNA.Base=POSITIVE	35	17	26.3	14.5	NA

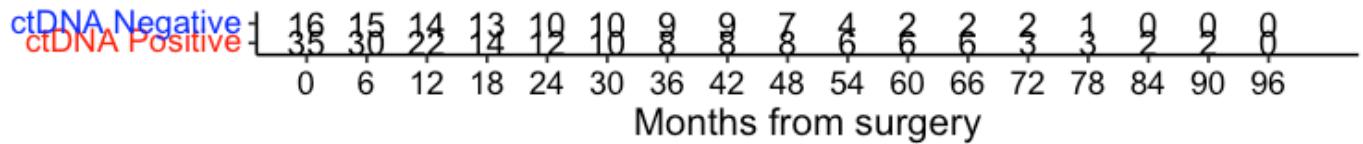
[Hide](#)

```
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 | NAC treated only", ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

RFS - ctDNA Baseline | NAC treated only



Number at risk



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

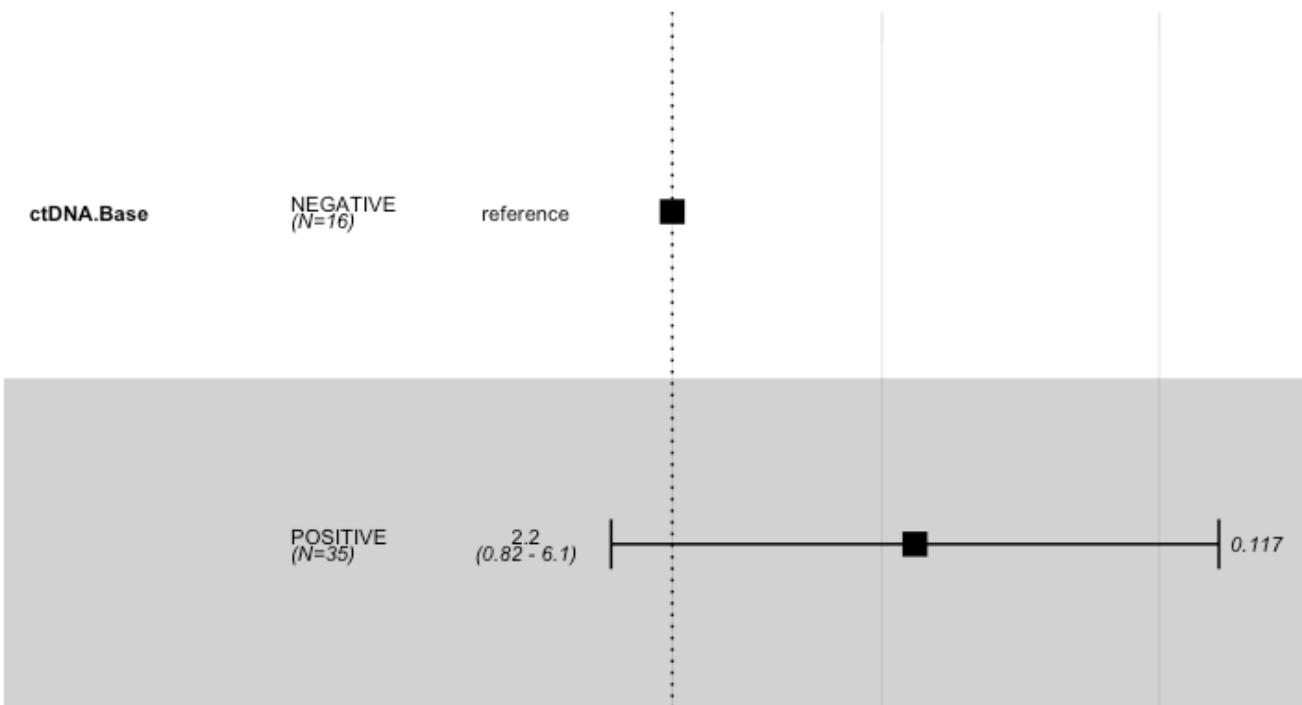
```
Call: survfit(formula = surv_object ~ ctDNA.Base, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.Base=NEGATIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
94	24.000	10.000	4.000	0.740	0.112	0.446		0.8	
ctDNA.Base=POSITIVE									
CI	time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95%
70	24.0000	12.0000	16.0000	0.5096	0.0893	0.3256		0.66	

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: 22; Global p-value (Log-Rank): 0.095453
 AIC: 157.13; Concordance Index: 0.6

Hide

```
summary(cox_fit)
```

Call:

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

n= 51, number of events= 22

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.BasePOSITIVE	0.8020	2.2300	0.5121	1.566	0.117

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.BasePOSITIVE	2.23	0.4484	0.8173	6.085

Concordance= 0.599 (se = 0.044)

Likelihood ratio test= 2.78 on 1 df, p=0.1

Wald test = 2.45 on 1 df, p=0.1

Score (logrank) test = 2.58 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 = 2.23 (0.82-6.08); p = 0.117"

#OS by ctDNA at Baseline - excluding pts that did not receive NAC

[Hide](#)

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

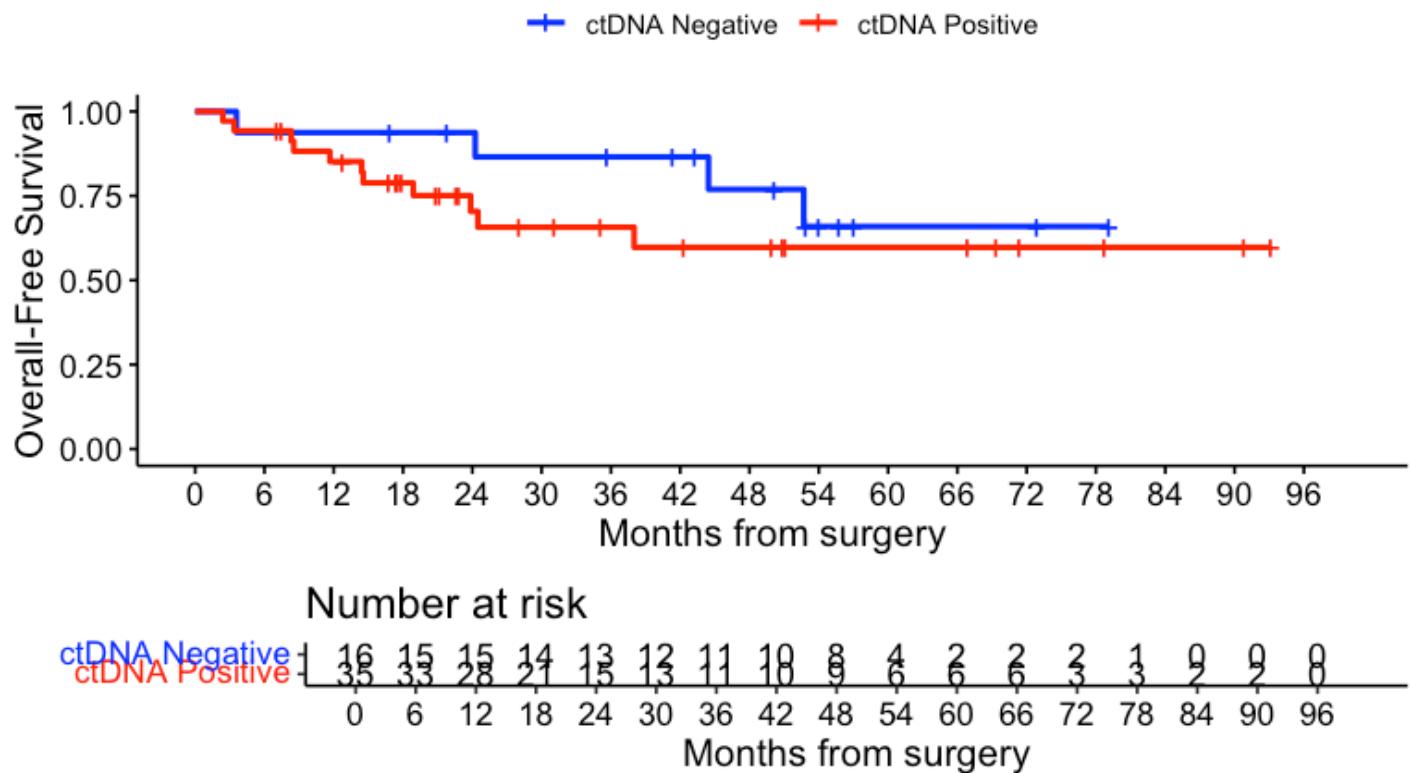
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.Base, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.Base=NEGATIVE	16	4	NA	52.7	NA
ctDNA.Base=POSITIVE	35	11	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 | NAC treated only",
ylab= "Overall-Free Survival", xlab="Months from surgery", legend.labs=c("ctDNA Negative",
"ctDNA Positive"), legend.title="")
```

OS - ctDNA Baseline | NAC treated only



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")
```

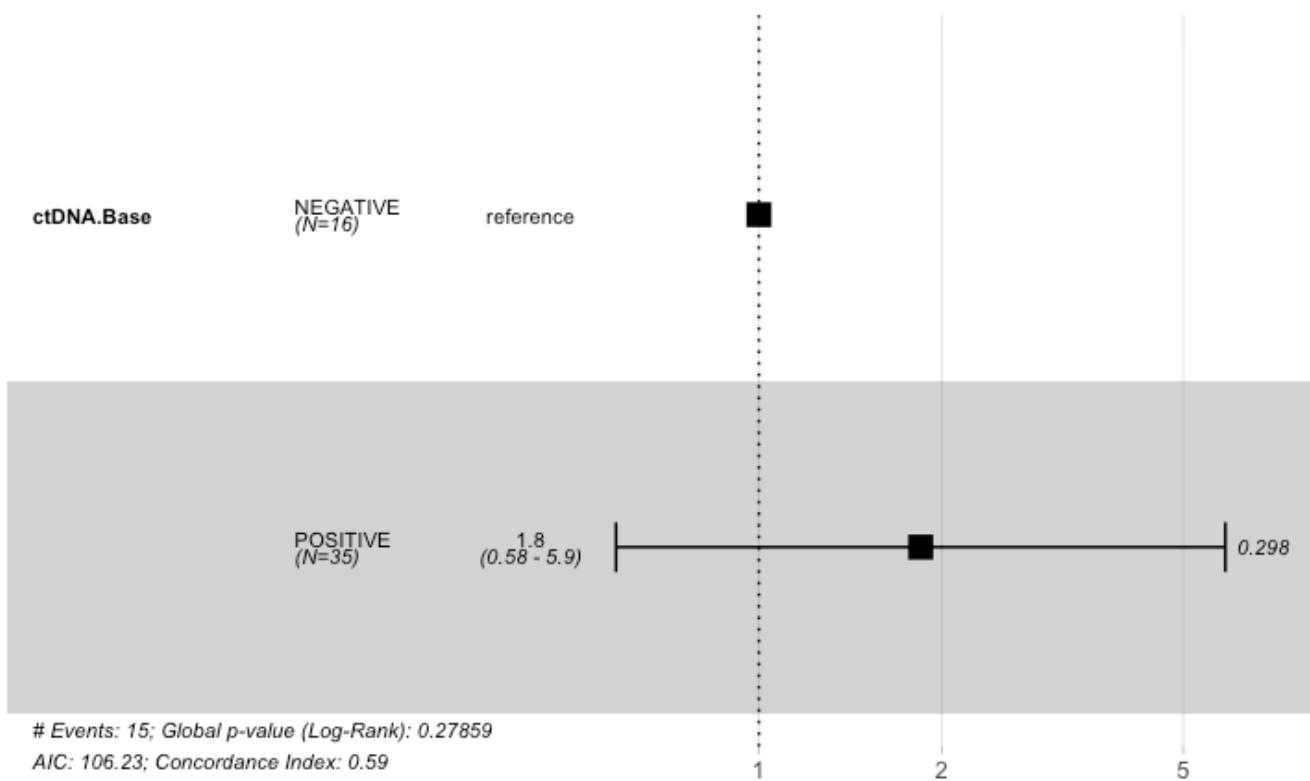
CI	ctDNA.Base=NEGATIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.0000	13.0000	1.0000	0.9375	0.0605	0.6323	0.99	

CI	ctDNA.Base=POSITIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.0000	15.0000	9.0000	0.7040	0.0854	0.5012	0.83	

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= 51, number of events= 15

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.BasePOSITIVE	0.6136	1.8470	0.5897	1.041	0.298

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.BasePOSITIVE	1.847	0.5414	0.5815	5.867

Concordance= 0.59 (se = 0.055)
 Likelihood ratio test= 1.17 on 1 df, p=0.3
 Wald test = 1.08 on 1 df, p=0.3
 Score (logrank) test = 1.12 on 1 df, p=0.3

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 1.85 (0.58-5.87); p = 0.298"
```

#RFS by ctDNA levels at Baseline based on AUC optimal MTM/ml level - ctDNA Baseline positive excluding pts that did not receive NAC

[Hide](#)

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

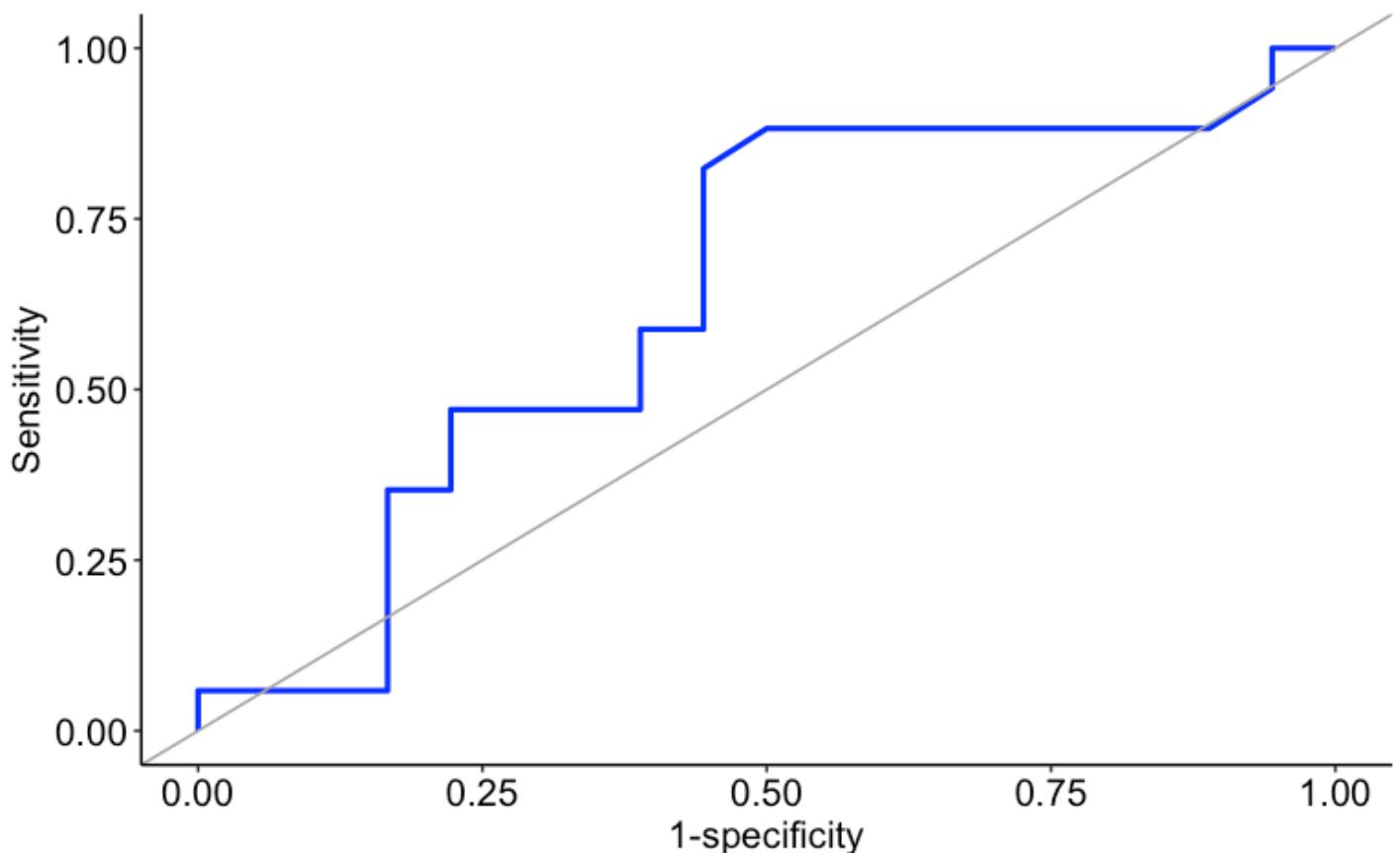
[Hide](#)

```
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_blank(), legend.text = element_text(size=14))
```

Warning: Ignoring unknown parameters: `aes`

[Hide](#)

p

[Hide](#)

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

Area under the curve: 0.6373

[Hide](#)

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

95% CI: 0.4424–0.8321 (DeLong)

[Hide](#)

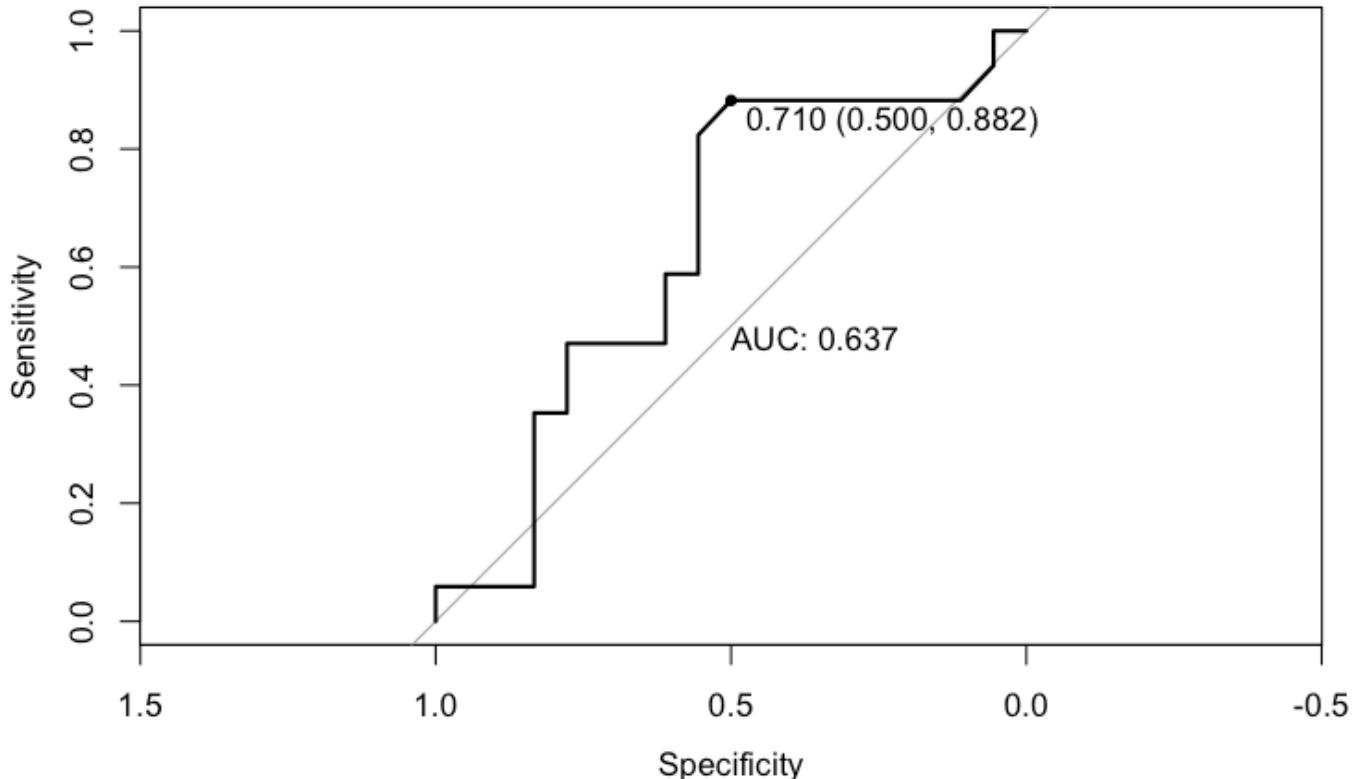
```
res.roc <- roc(circ_data$DFS.Event, circ_data$ctDNA.Base.MTM)
```

```
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$Included==TRUE,]
circ_data <- circ_data[circ_data$NAC==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 < 0.71 ~ 1,
    ctDNA.Base.MTM >= 0.71 ~ 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	11	2	NA	NA	NA
ctDNA.6mMTM.Q=2	24	15	14.5	7.79	NA

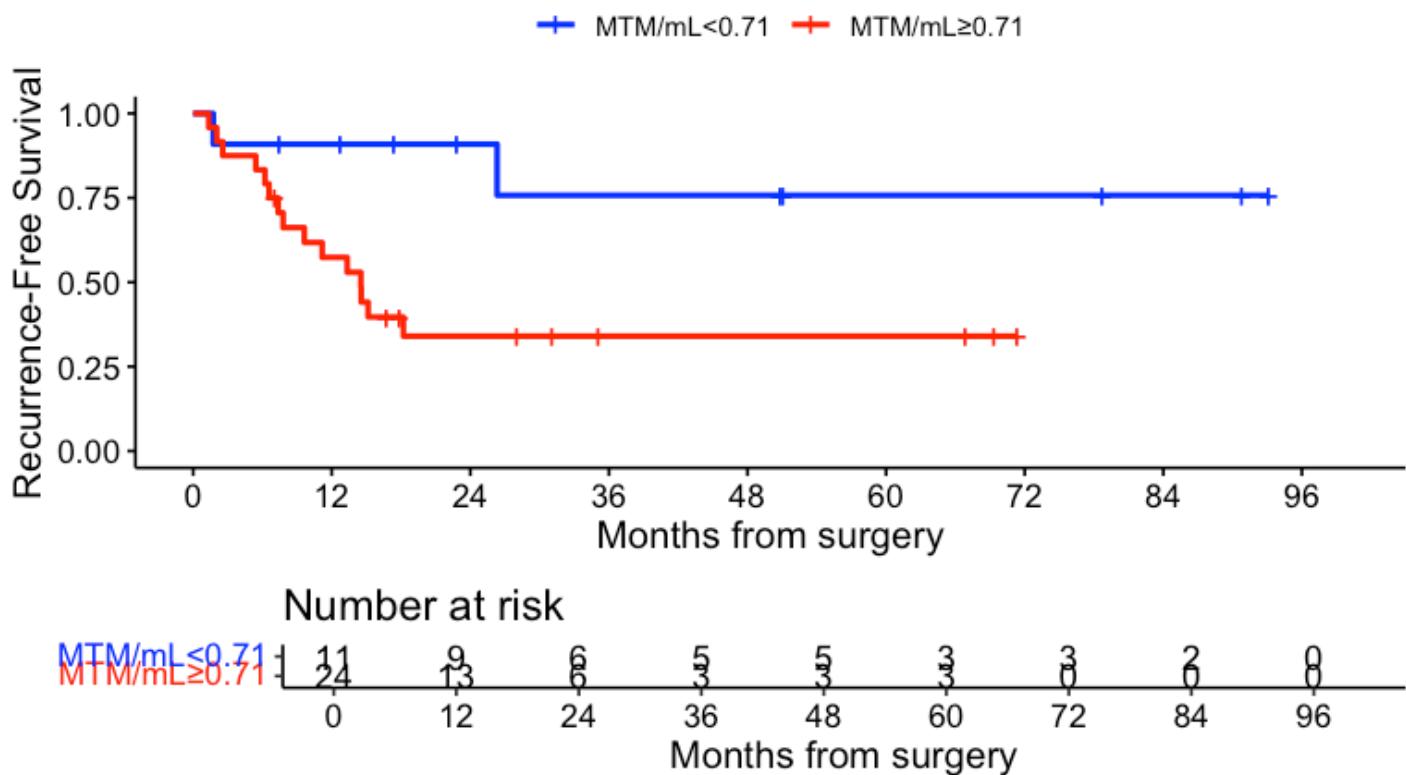
[Hide](#)

```

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="RFS - ctDNA MTM/mL groups at Baseline",
ylab= "Recurrence-Free Survival", xlab="Months from surgery", legend.labs=c("MTM/mL <0.71", "MTM/mL≥0.71"), legend.title="")

```

RFS - ctDNA MTM/mL groups at Baseline



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

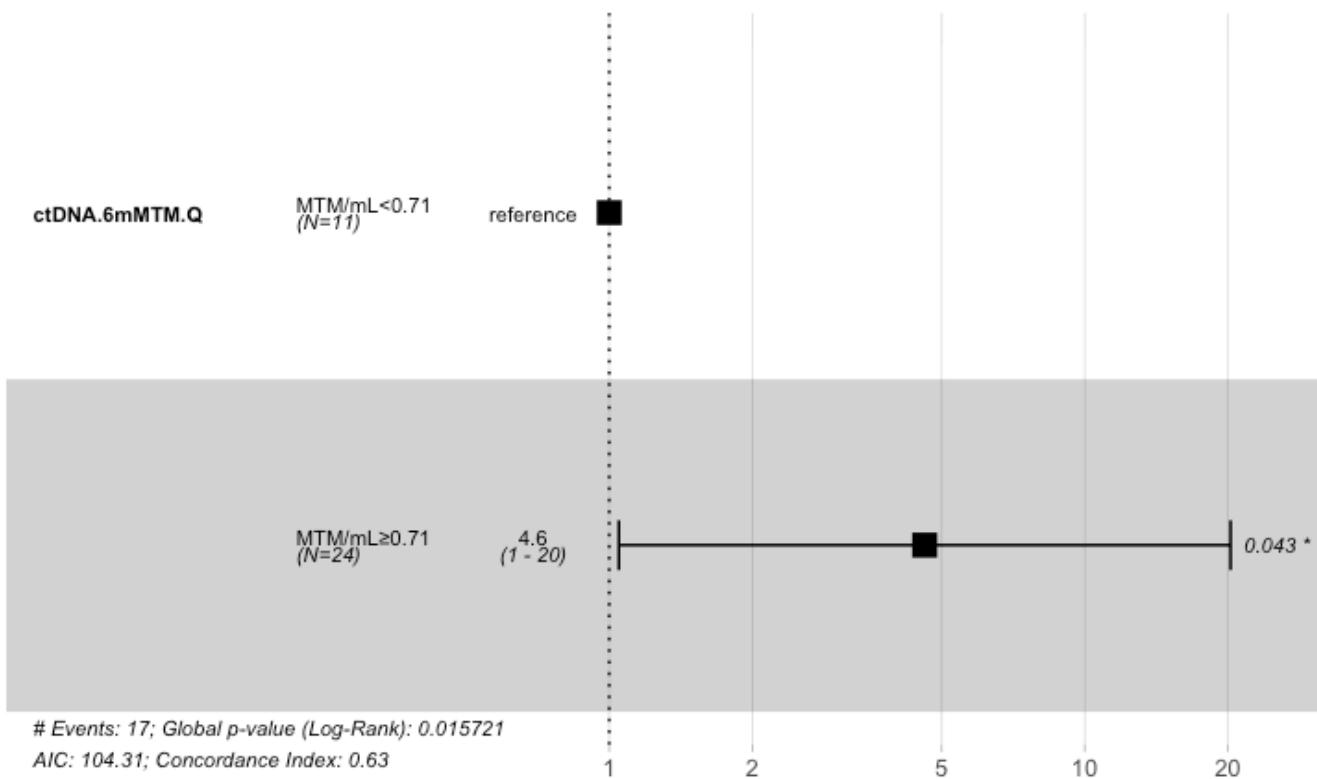
```
Call: survfit(formula = surv_object ~ ctDNA.6mMTM.Q, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

CI	ctDNA.6mMTM.Q=1						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
67	24.0000	6.0000	1.0000	0.9091	0.0867	0.5081	0.98

CI	ctDNA.6mMTM.Q=2						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
35	24.000	6.000	15.000	0.340	0.102	0.156	0.5

```
circ_data$ctDNA.6mMTM.Q <- factor(circ_data$ctDNA.6mMTM.Q, levels=c("1","2"), labels = c
("MTM/mL<0.71", "MTM/mL≥0.71"))
cox_fit <- coxph(surv_object ~ ctDNA.6mMTM.Q, data=circ_data)
ggforest(cox_fit,data = circ_data)
```

Hazard ratio



```
summary(cox_fit)
```

Call:

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

n= 35, number of events= 17

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.6mMTM.QMTM/mL ≥ 0.71	1.5274	4.6061	0.7556	2.021	0.0432 *

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.6mMTM.QMTM/mL ≥ 0.71	4.606	0.2171	1.048	20.25

Concordance= 0.632 (se = 0.056)

Likelihood ratio test= 5.83 on 1 df, p=0.02

Wald test = 4.09 on 1 df, p=0.04

Score (logrank) test = 4.92 on 1 df, p=0.03

```
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 = 4.61 (1.05-20.25); p = 0.043"

#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)

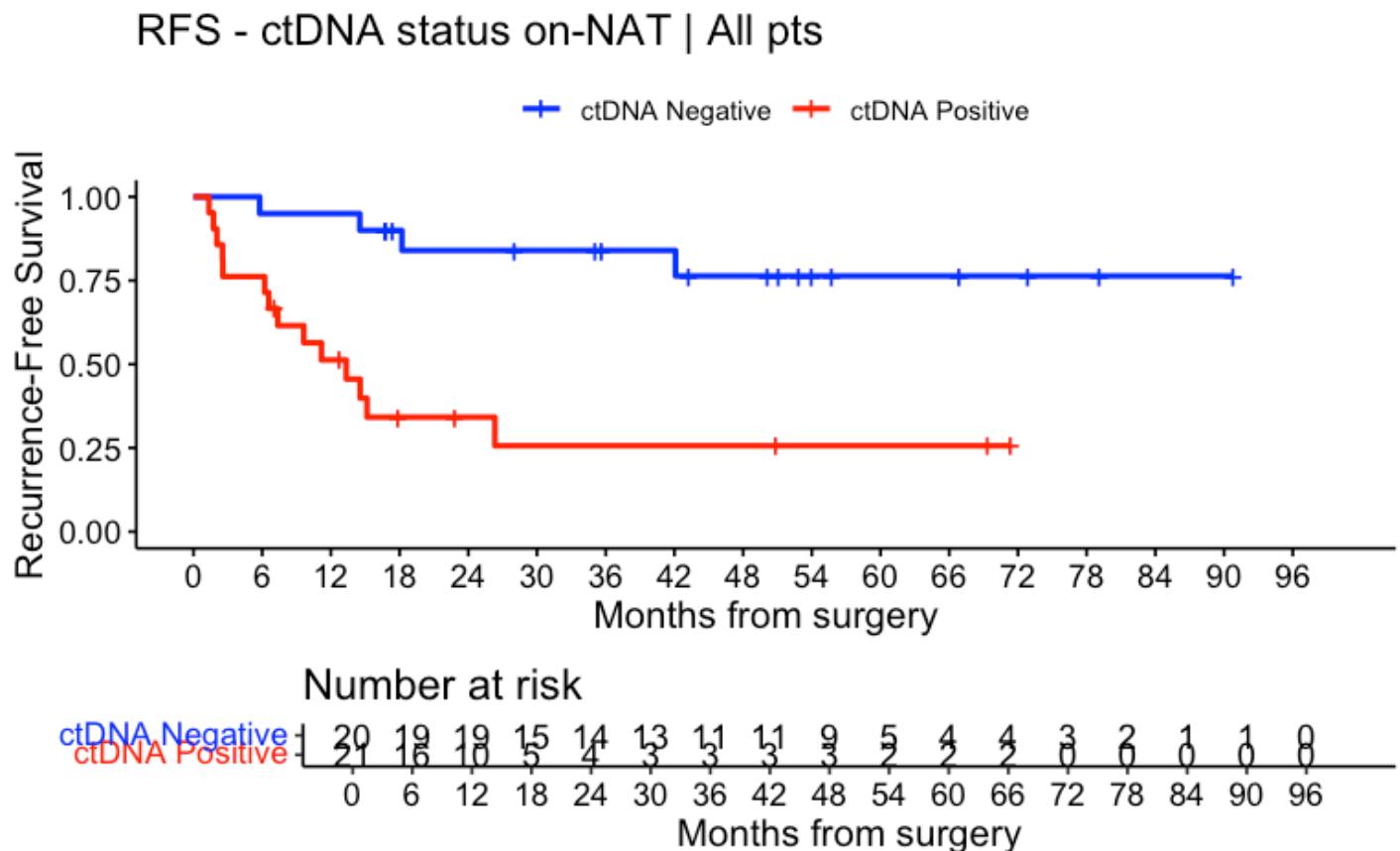
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.C2D1, data = circ_data)
```

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="")
```



Hide

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

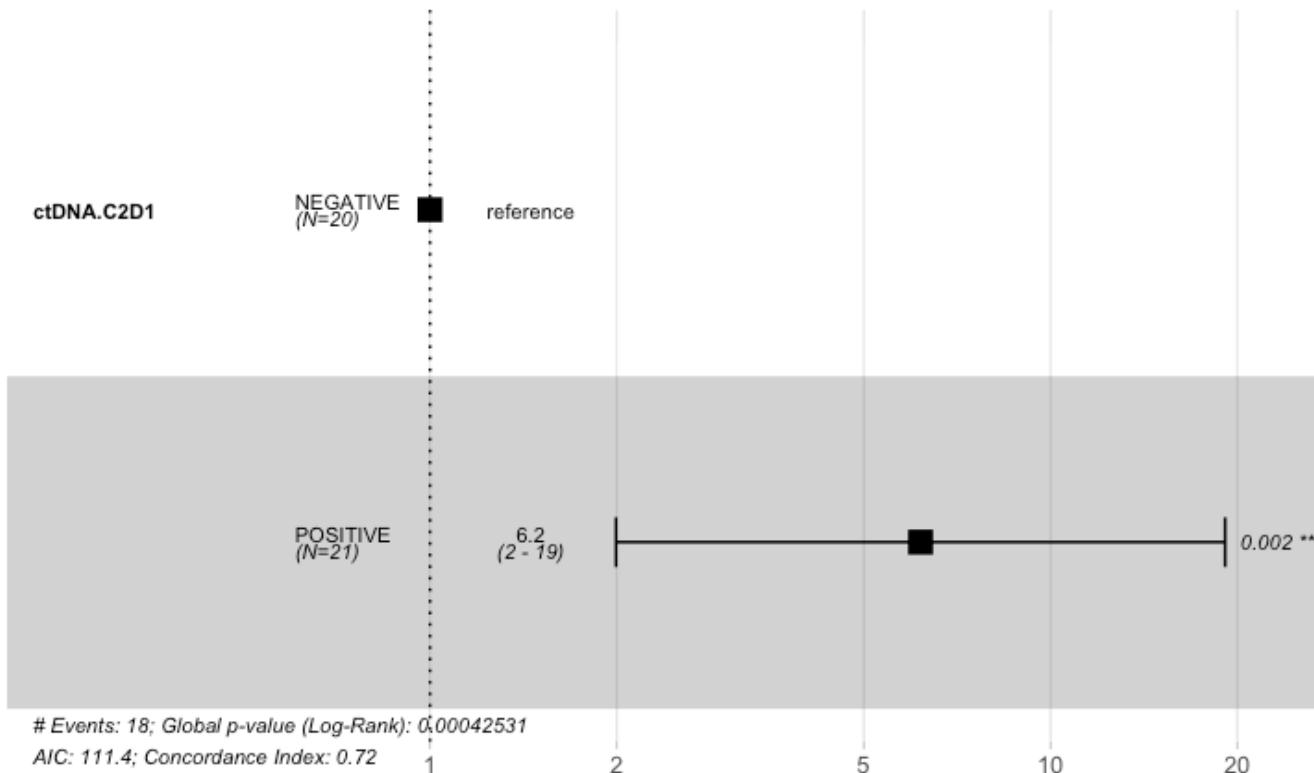
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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	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	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)
```

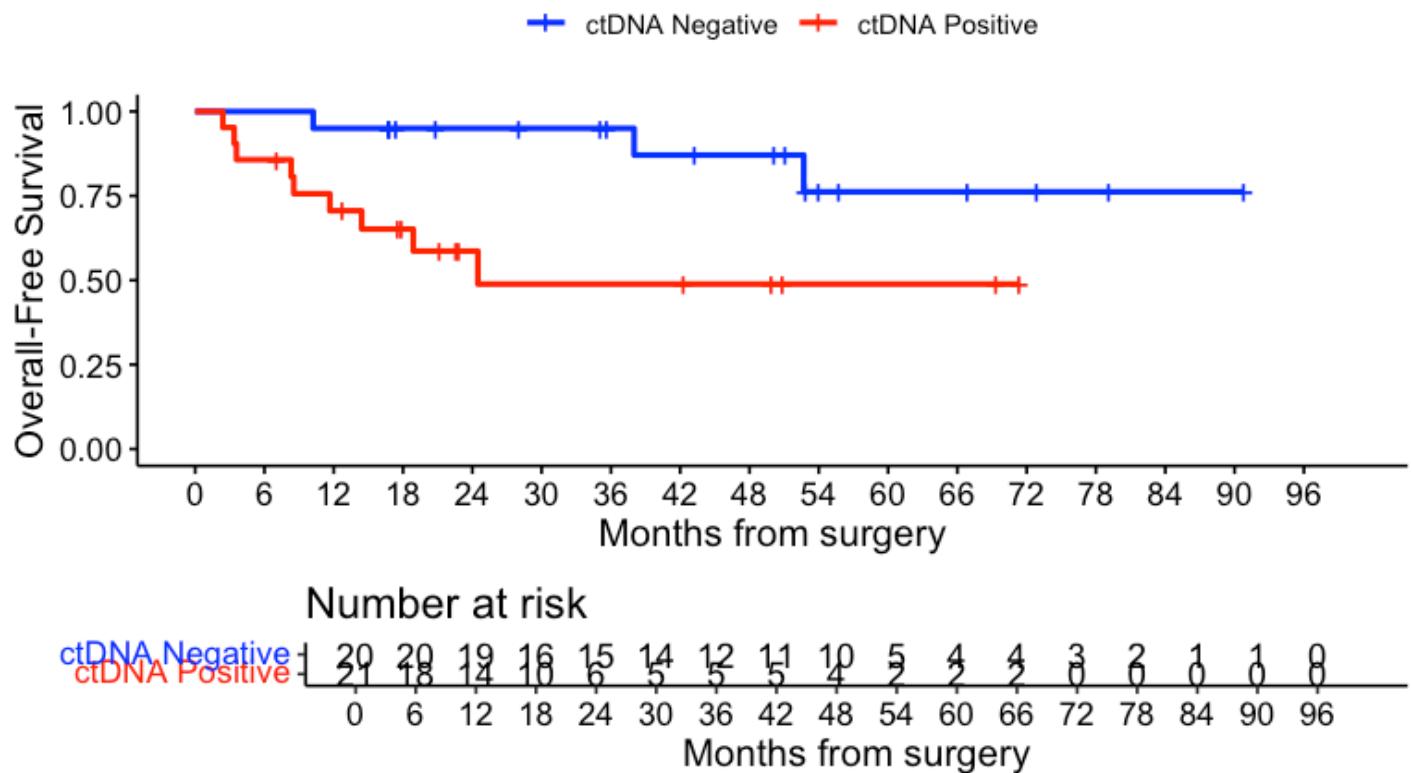
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.C2D1, data = circ_data)
```

	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



Hide

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

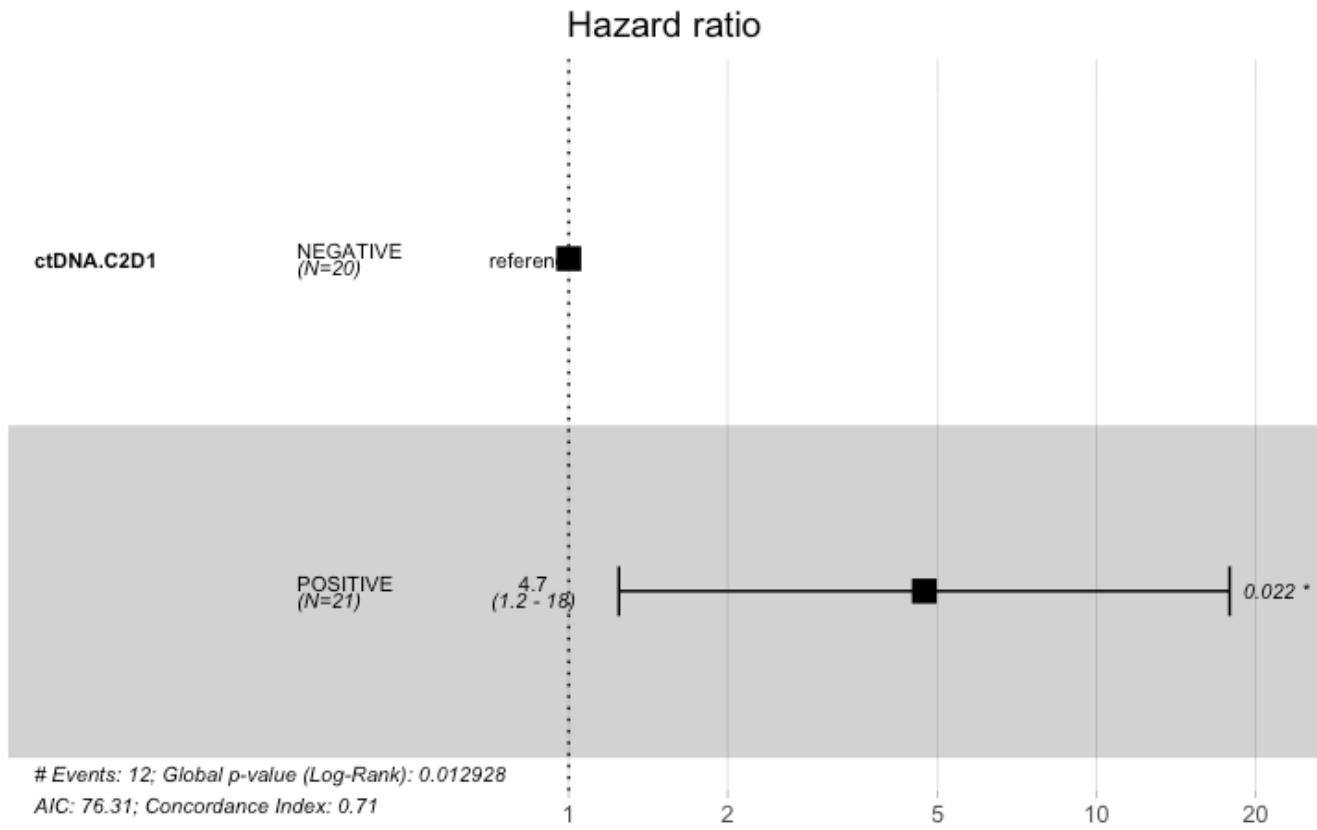
```
Call: survfit(formula = surv_object ~ ctDNA.C2D1, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.C2D1=NEGATIVE						
CI	time	n.risk	n.event	survival	std.err	lower 95% CI upper 95%
28	24.0000	15.0000	1.0000	0.9500	0.0487	0.6947 0.99

ctDNA.C2D1=POSITIVE						
CI	time	n.risk	n.event	survival	std.err	lower 95% CI upper 95%
71	24.000	6.000	8.000	0.586	0.114	0.335 0.7

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)
```



```
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

```
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"
```

#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)

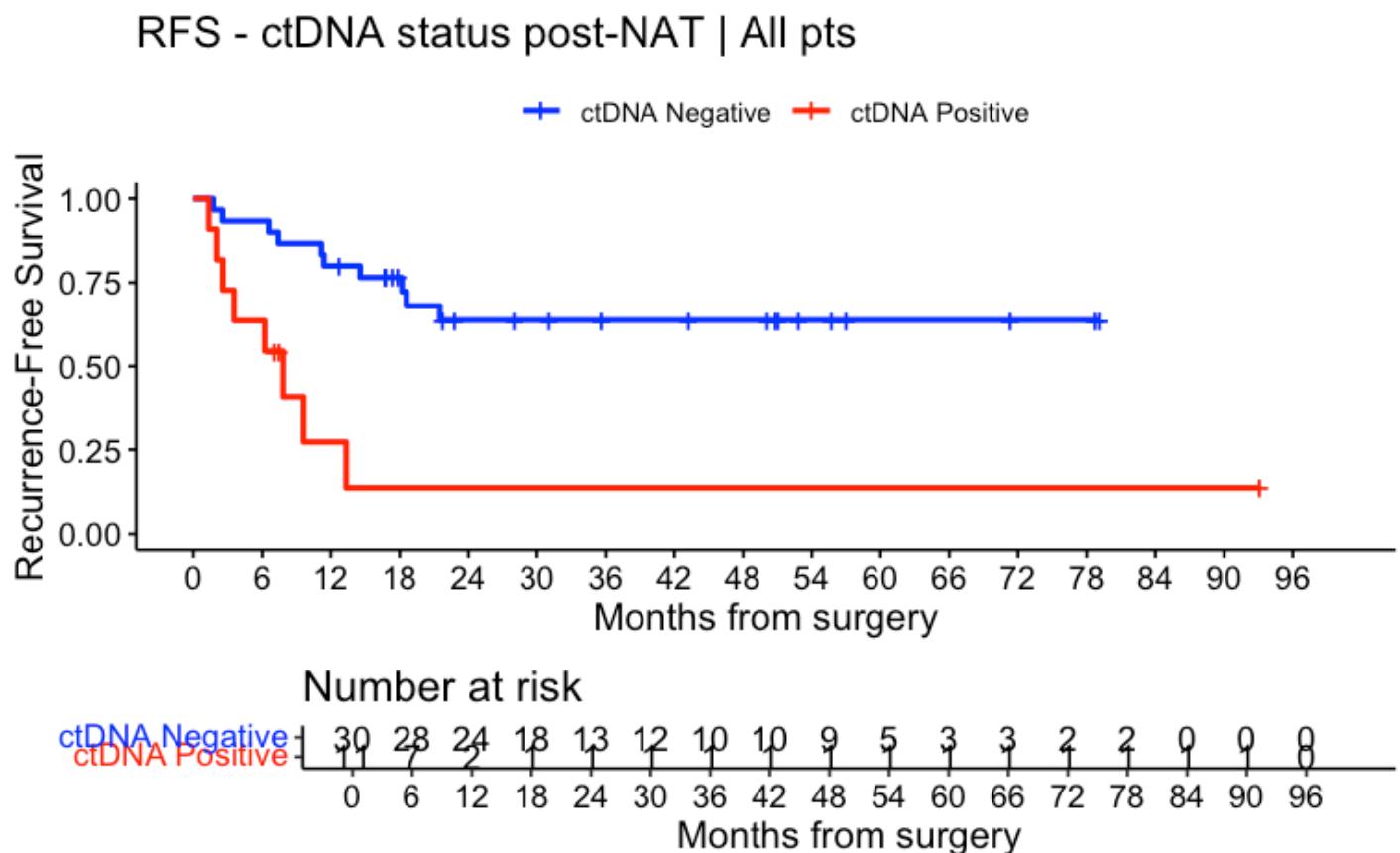
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.postNAC, data = circ_data)
```

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

[Hide](#)

```
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="")
```



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

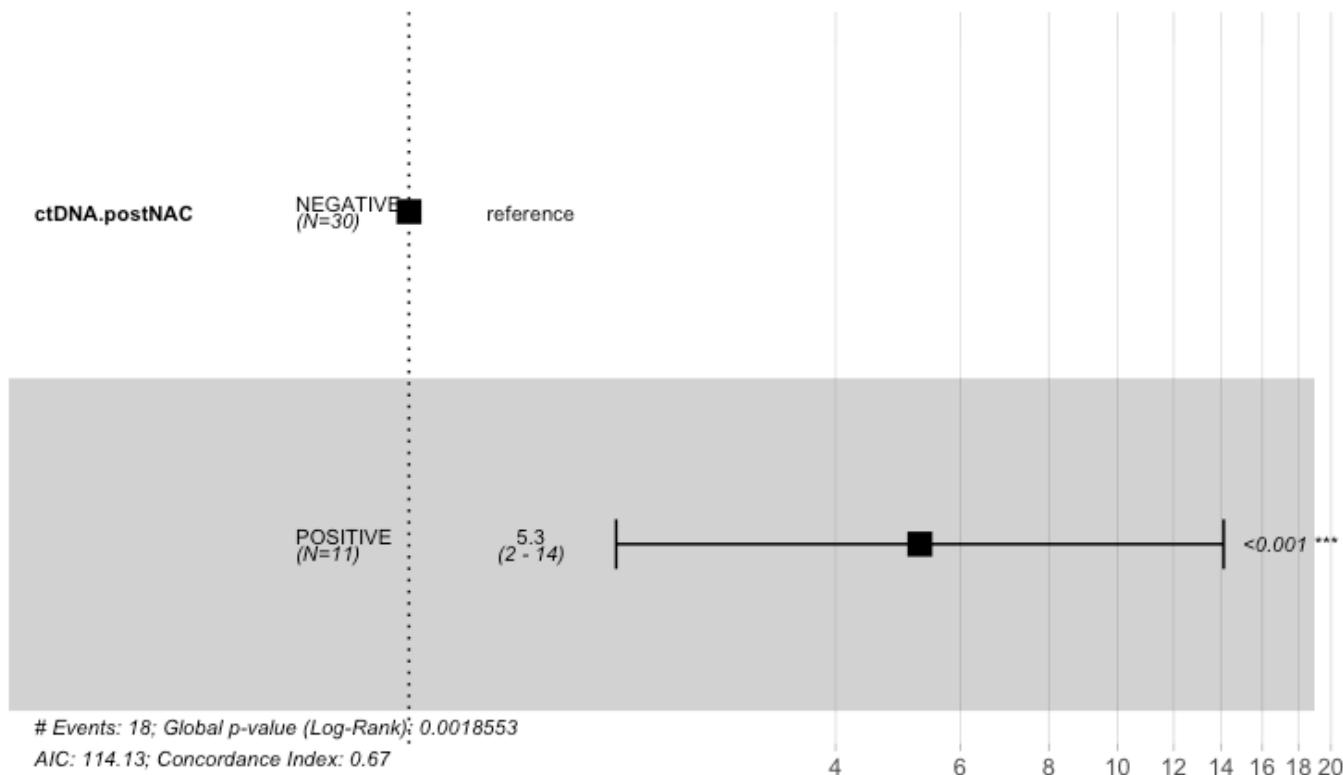
```
Call: survfit(formula = surv_object ~ ctDNA.postNAC, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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

[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= 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)
```

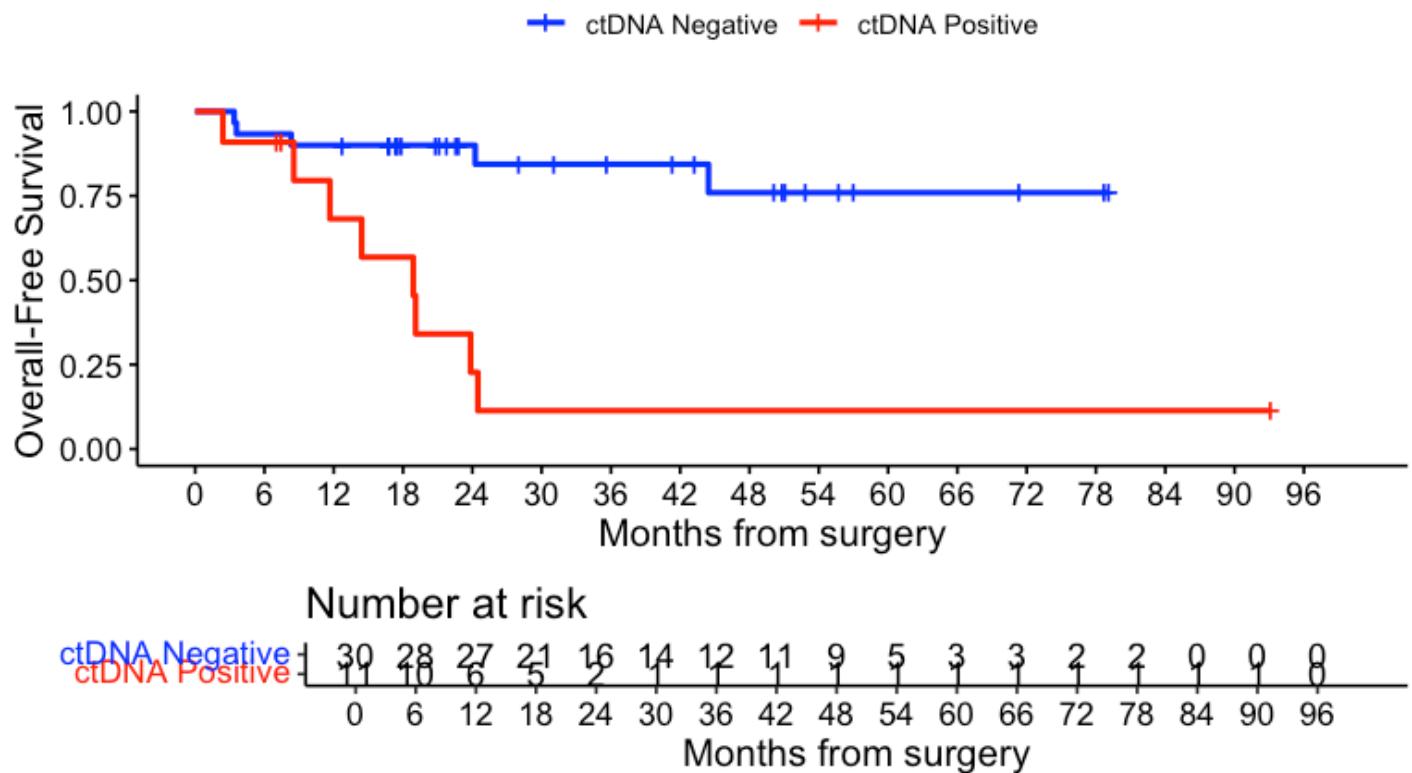
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.postNAC, data = circ_data)
```

	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="")
```

OS - ctDNA status post-NAT | All pts



Hide

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

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

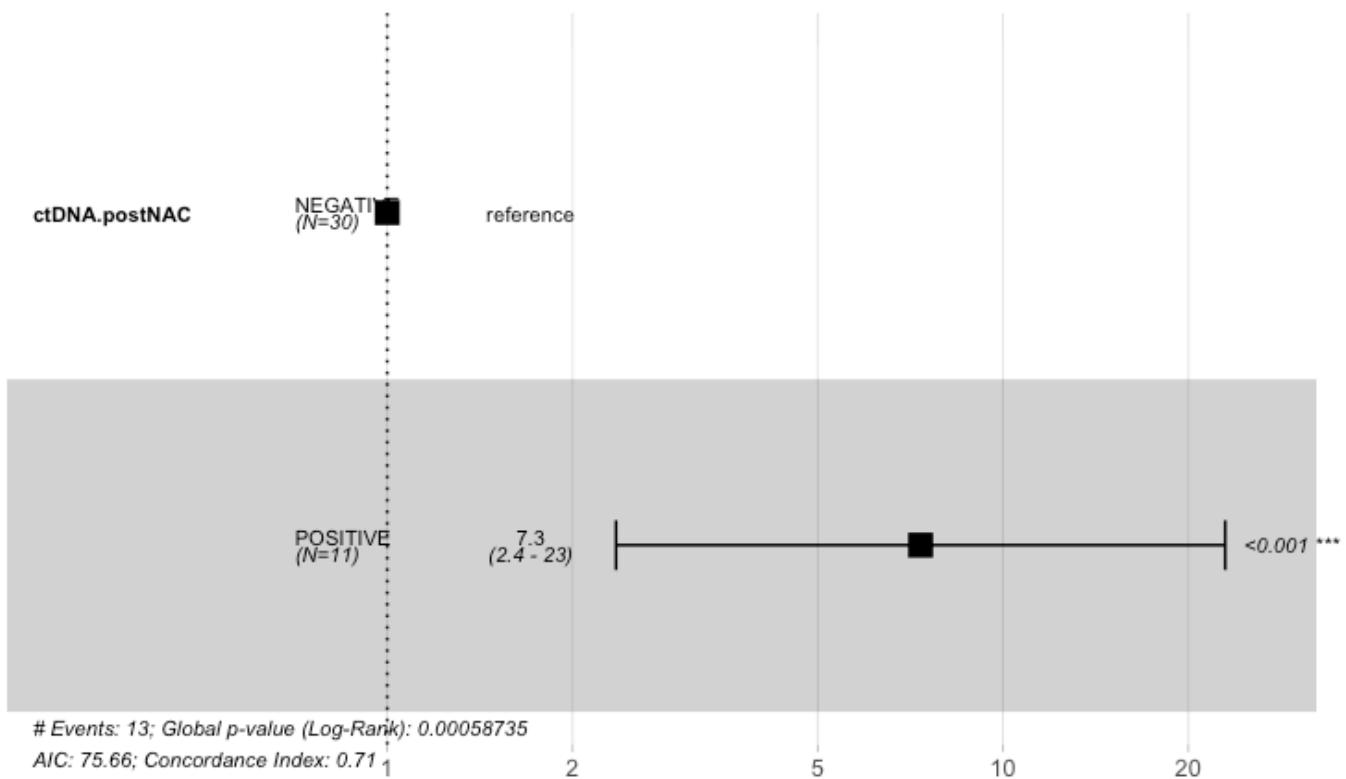
CI	ctDNA.postNAC=NEGATIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
24.0000	16.0000	3.0000	0.9000	0.0548	0.7212	0.96	

CI	ctDNA.postNAC=POSITIVE						
	time	n.risk	n.event	survival	std.err	lower 95%	CI upper 95%
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)
```

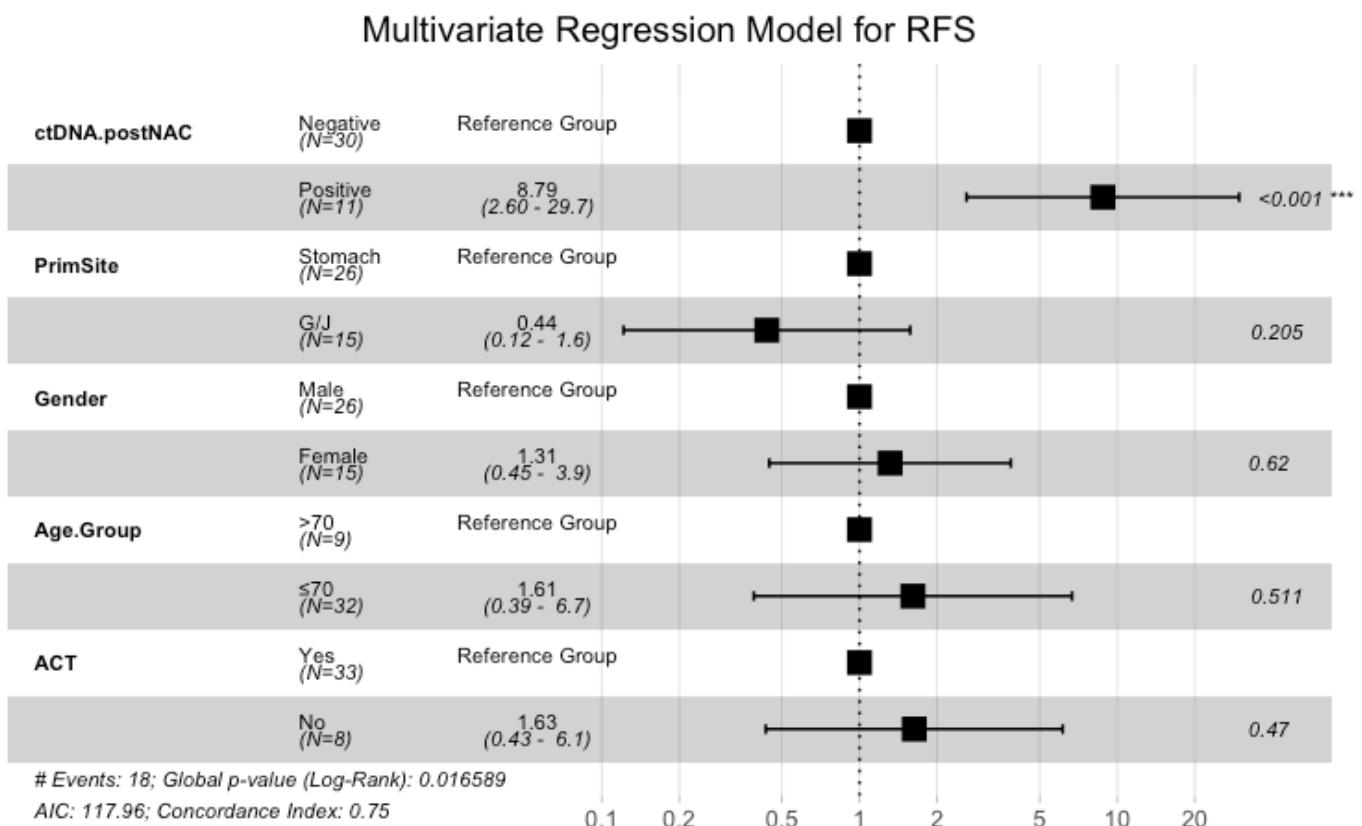
```
[1] "HR = 7.35 (2.35-22.95); p = 0.001"
```

#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("Stomach","G/J"))
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"))
circ_data$ACT <- factor(circ_data$ACT, levels=c("TRUE","FALSE"), labels = c("Yes", "No"))
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")
```



Hide

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

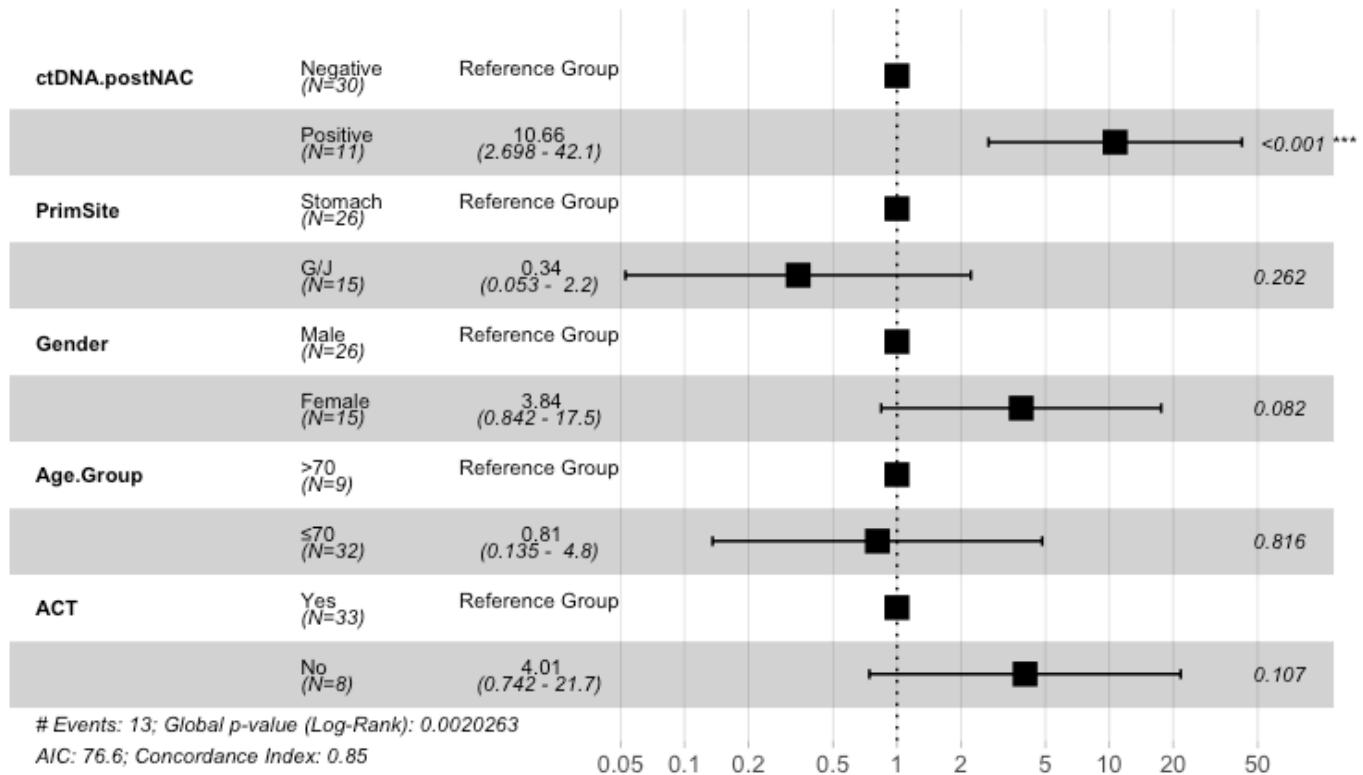
#Multivariate cox regression for OS - ctDNA post-NAT

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("Stomach","G/J"))
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"))
circ_data$ACT <- factor(circ_data$ACT, levels=c("TRUE","FALSE"), labels = c("Yes", "No"))
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)
```

#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: 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

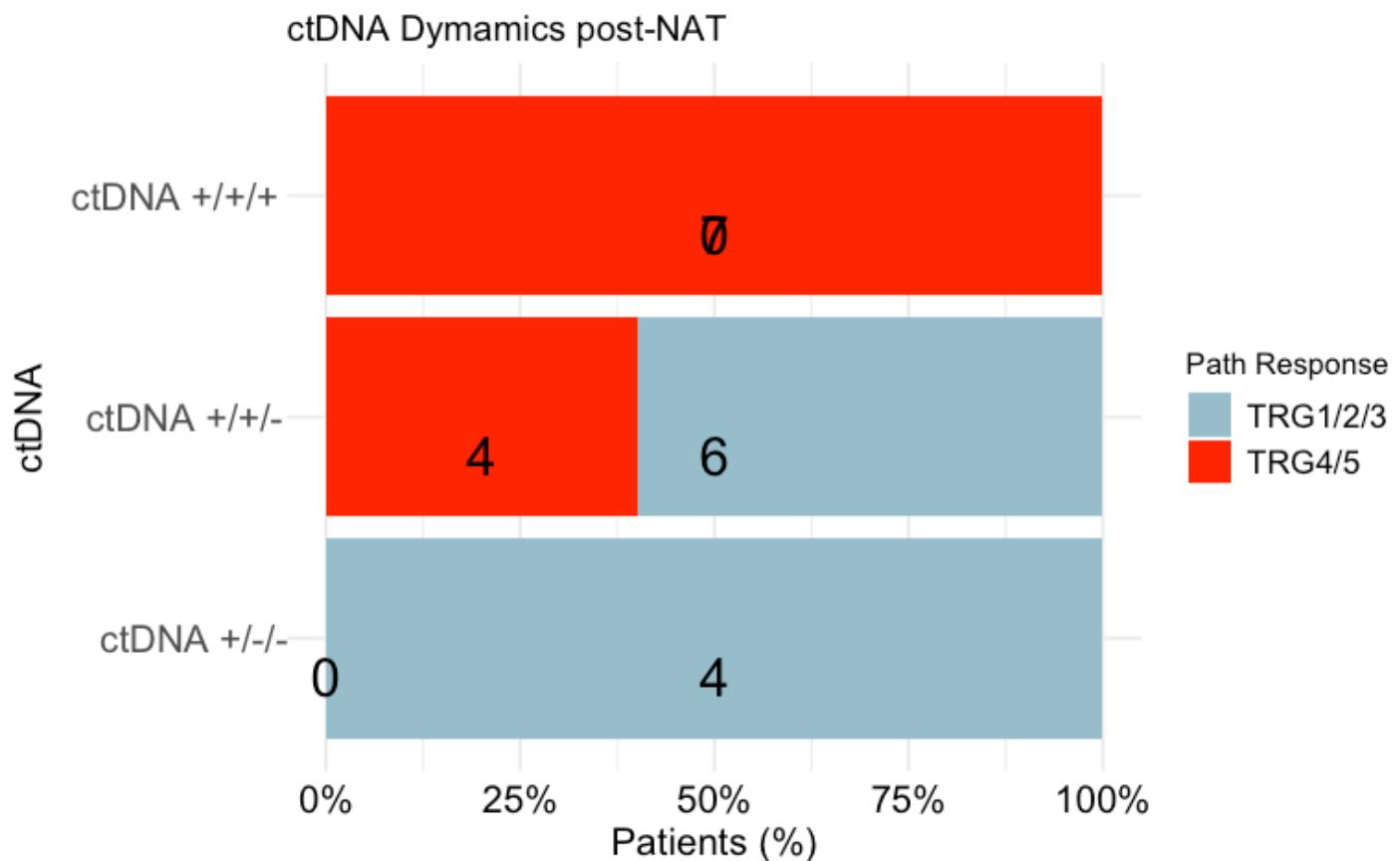
[Hide](#)

```

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: 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
```

```
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
```

```
print(contingency_table)
```

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

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

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

```

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)

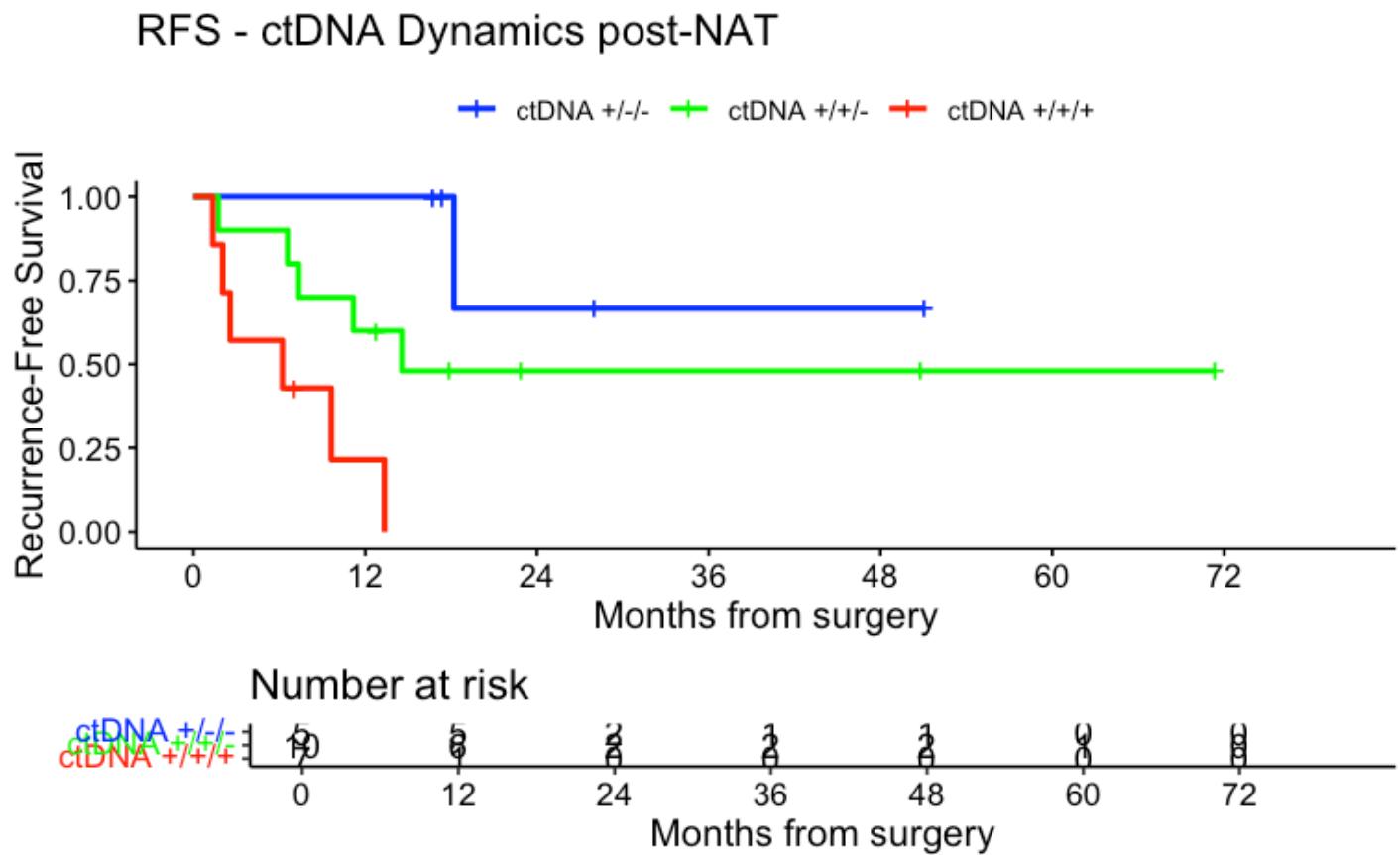
	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=ctDNA +/--	5	1	NA	18.20	NA
ctDNA.Dynamics=ctDNA +/+-	10	5	14.55	7.36	NA
ctDNA.Dynamics=ctDNA +/++	7	6	6.21	2.04	NA

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)

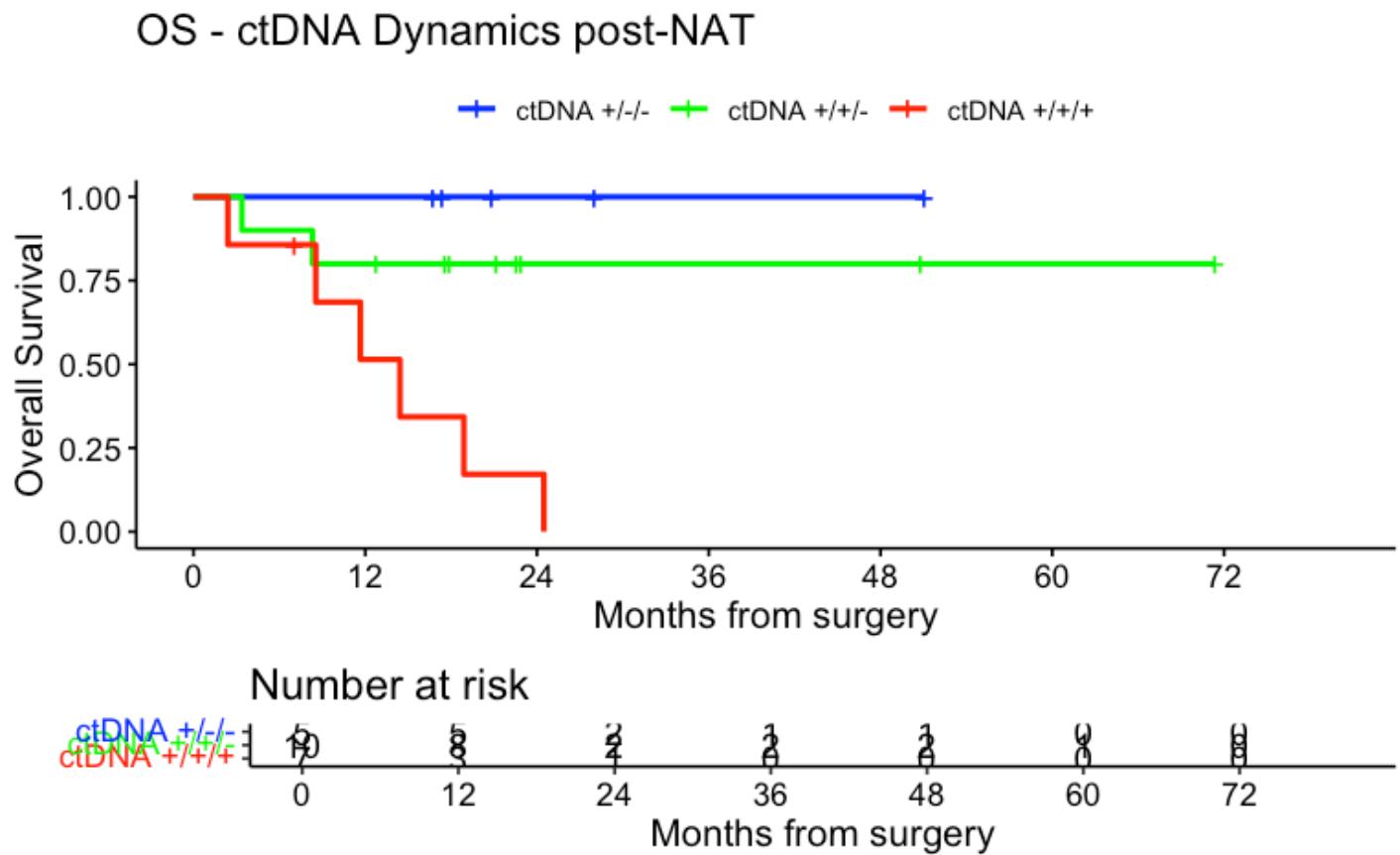
	n	events	median	0.95LCL	0.95UCL
ctDNA.Dynamics=ctDNA +/--	5	0	NA	NA	NA
ctDNA.Dynamics=ctDNA +/+-	10	2	NA	NA	NA
ctDNA.Dynamics=ctDNA +/++	7	6	14.4	8.54	NA

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
```

```
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

```

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)

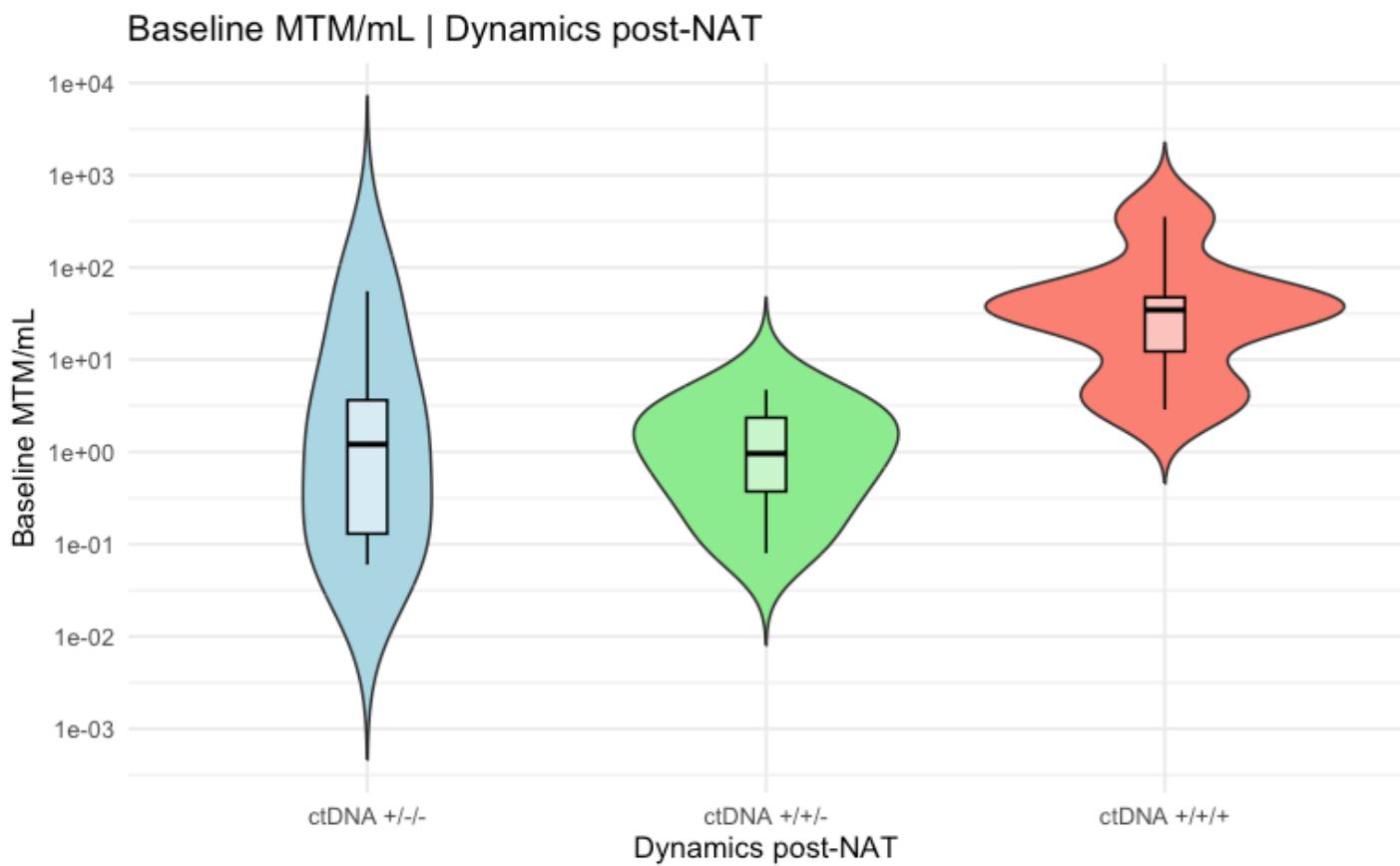
circ_data$ctDNA.Dynamics <- NA #first we create the variable for the ctDNA & NAC combination, and we assign values
circ_data <- circ_data %>%
  mutate(ctDNA.Dynamics = case_when(
    ctDNA.Base == "POSITIVE" & ctDNA.C2D1 == "NEGATIVE" & ctDNA.postNAC == "NEGATIVE" ~
    "ctDNA +/--",
    ctDNA.Base == "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")
```



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: 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: 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
```

#Time of ctDNA draws at the MRD window from surgery

[Hide](#)

```

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

circ_data$ctDNA.MRD.Date <- as.Date(circ_data$ctDNA.MRD.Date, format="%m/%d/%Y")
circ_data$Surgery.Date <- as.Date(circ_data$Surgery.Date, format="%m/%d/%Y")
circ_data$Time_to_ctDNA_MRD <- as.numeric(difftime(circ_data$ctDNA.MRD.Date, circ_data$Surgery.Date, units="days"))
median_ctDNA_MRД <- median(circ_data$Time_to_ctDNA_MRД, na.rm=TRUE)
range_ctDNA_MRД <- range(circ_data$Time_to_ctDNA_MRД, na.rm=TRUE)
Q1_ctDNA_MRД <- quantile(circ_data$Time_to_ctDNA_MRД, 0.25, na.rm=TRUE)
Q3_ctDNA_MRД <- quantile(circ_data$Time_to_ctDNA_MRД, 0.75, na.rm=TRUE)
IQR_ctDNA_MRД <- IQR(circ_data$Time_to_ctDNA_MRД, na.rm=TRUE)
cat("Median time from surgery to ctDNA MRД time point:", median_ctDNA_MRД, "days\n")

```

Median time from surgery to ctDNA MRД time point: 40.5 days

[Hide](#)

```
cat("Range:", range_ctDNA_MRД[1], "-", range_ctDNA_MRД[2], "days\n")
```

Range: 15 – 82 days

[Hide](#)

```
cat("Q1 (25th percentile):", Q1_ctDNA_MRД, "days\n")
```

Q1 (25th percentile): 30.25 days

[Hide](#)

```
cat("Q3 (75th percentile):", Q3_ctDNA_MRД, "days\n")
```

Q3 (75th percentile): 47 days

[Hide](#)

```
cat("Interquartile Range (IQR):", IQR_ctDNA_MRД, "days\n")
```

Interquartile Range (IQR): 16.75 days

#RFS by ctDNA at the MRД 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)

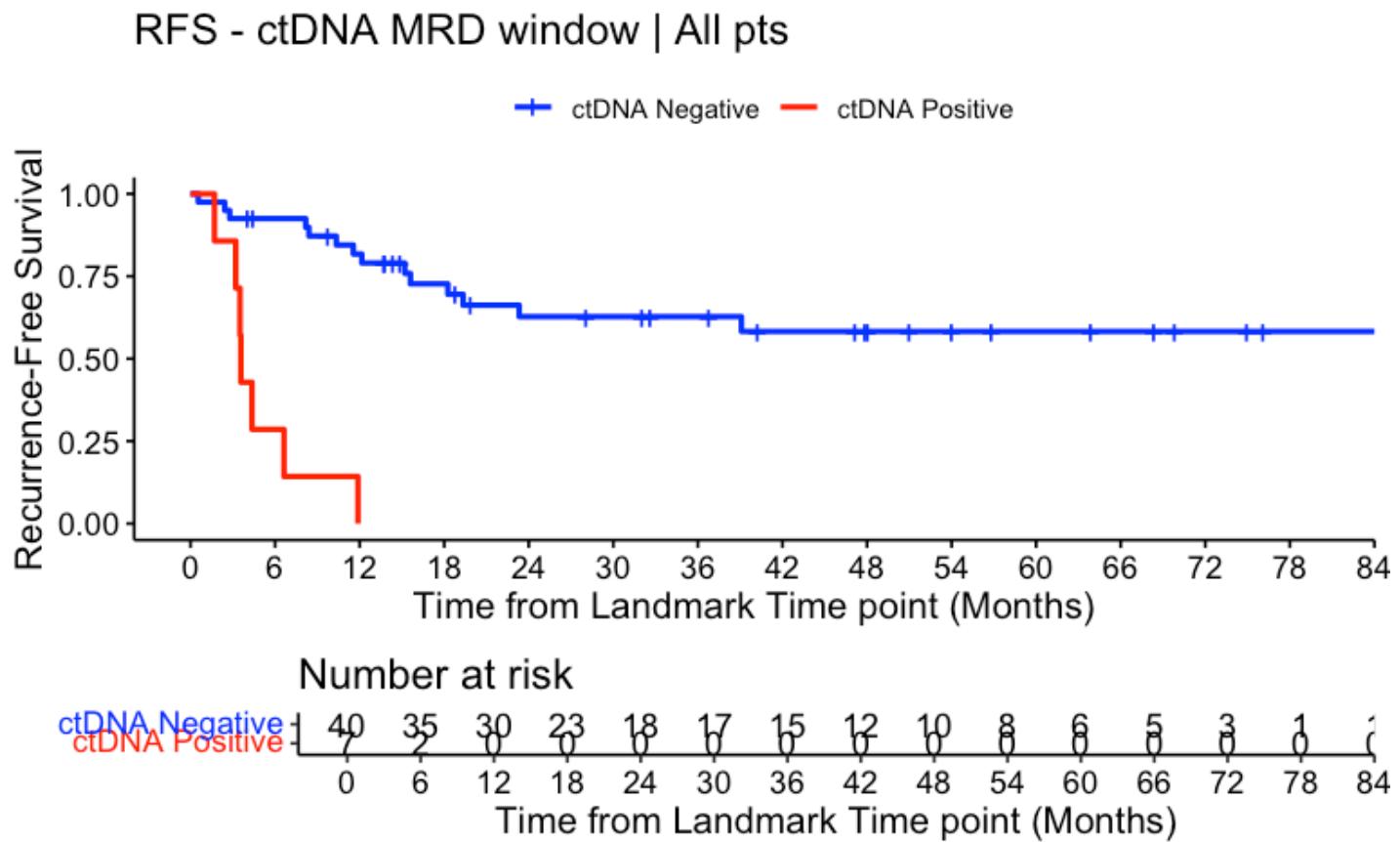
	n	events	median	0.95LCL	0.95UCL
ctDNA.MRD=NEGATIVE	40	14	NA	23.32	NA
ctDNA.MRD=POSITIVE	7	7	3.57	3.21	NA

[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	

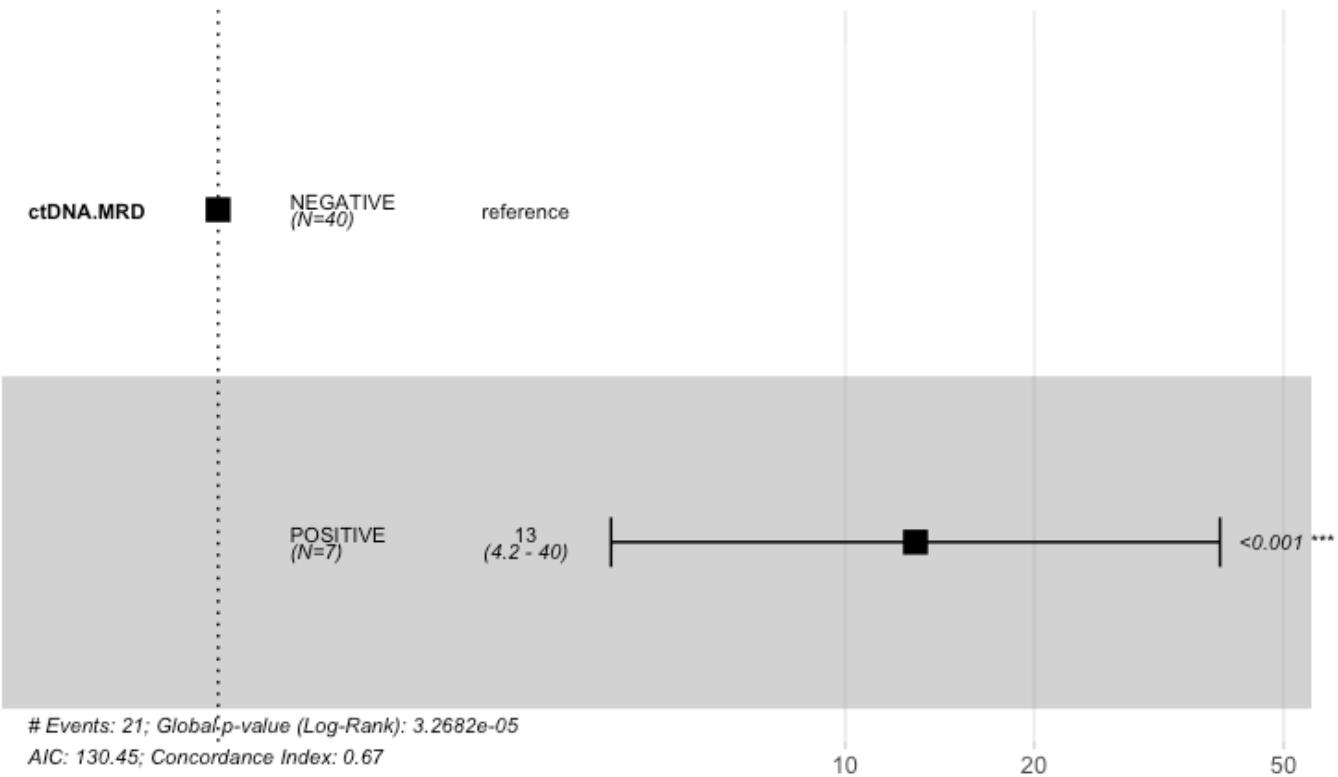
ctDNA.MRD=POSITIVE

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

[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
```

```
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

```
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)

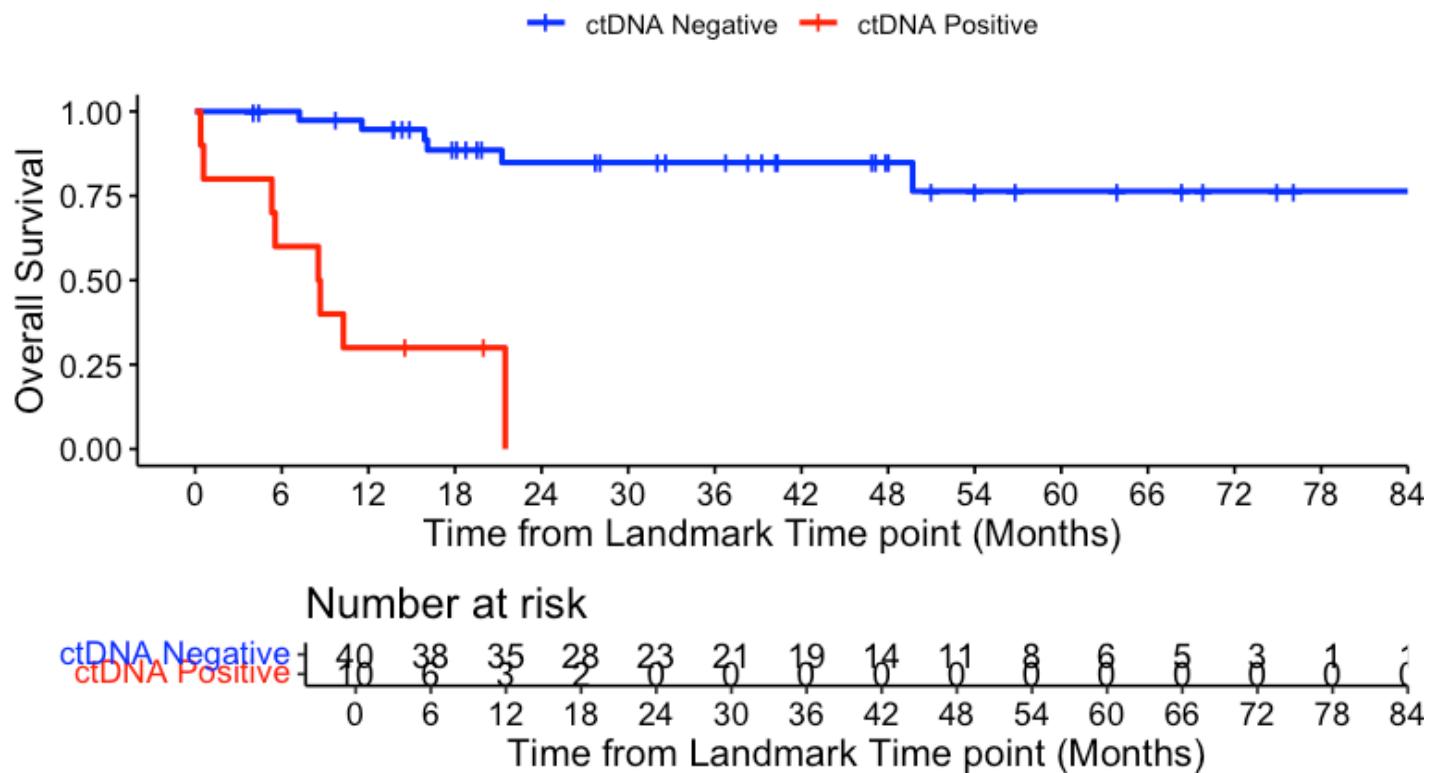
survfit(Surv(time = circ_data$FU.months, event = circ_data$OS.Event)~ctDNA.MRD, data = circ_data)
```

```
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	NA
ctDNA.MRD=POSITIVE	10	8	8.59	5.31	NA

```
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



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

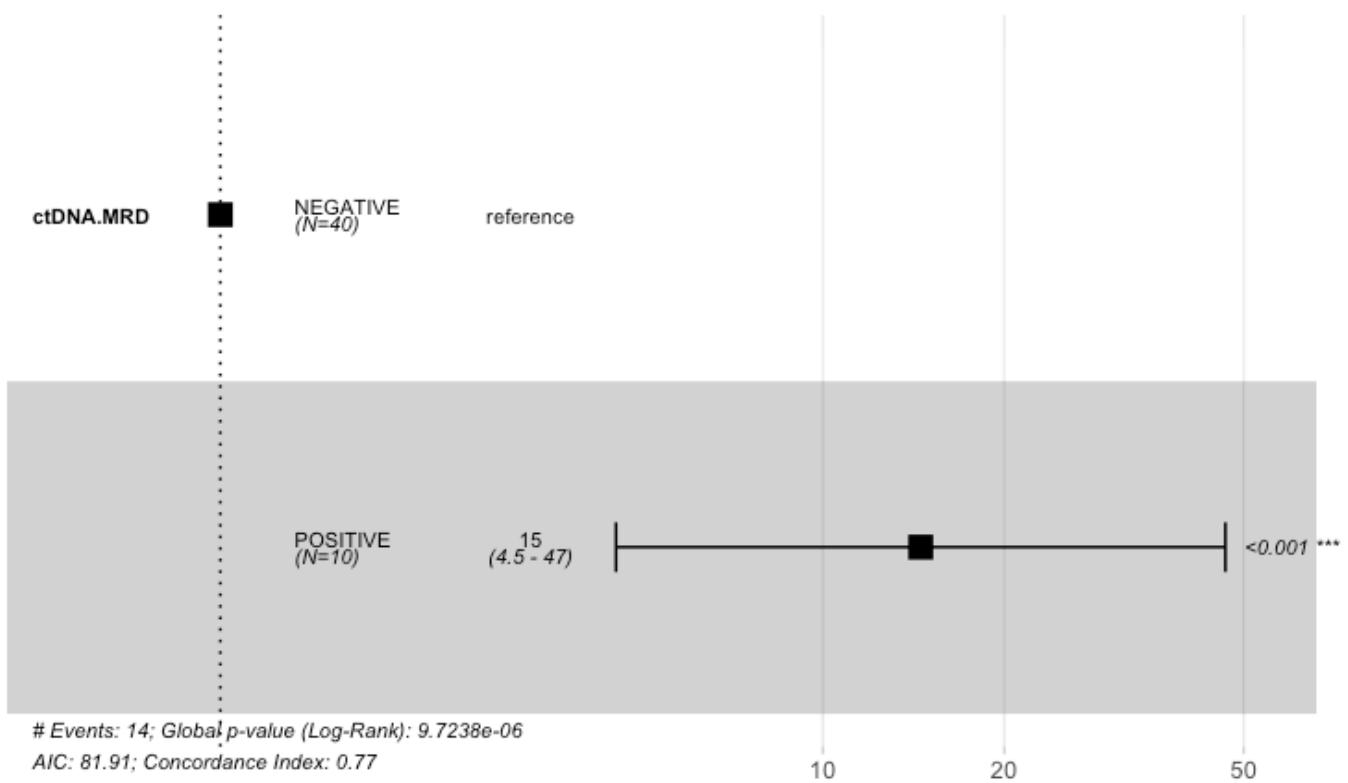
```
Call: survfit(formula = surv_object ~ ctDNA.MRD, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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

```
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



```
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

```
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 - ACT pts only

[Hide](#)

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

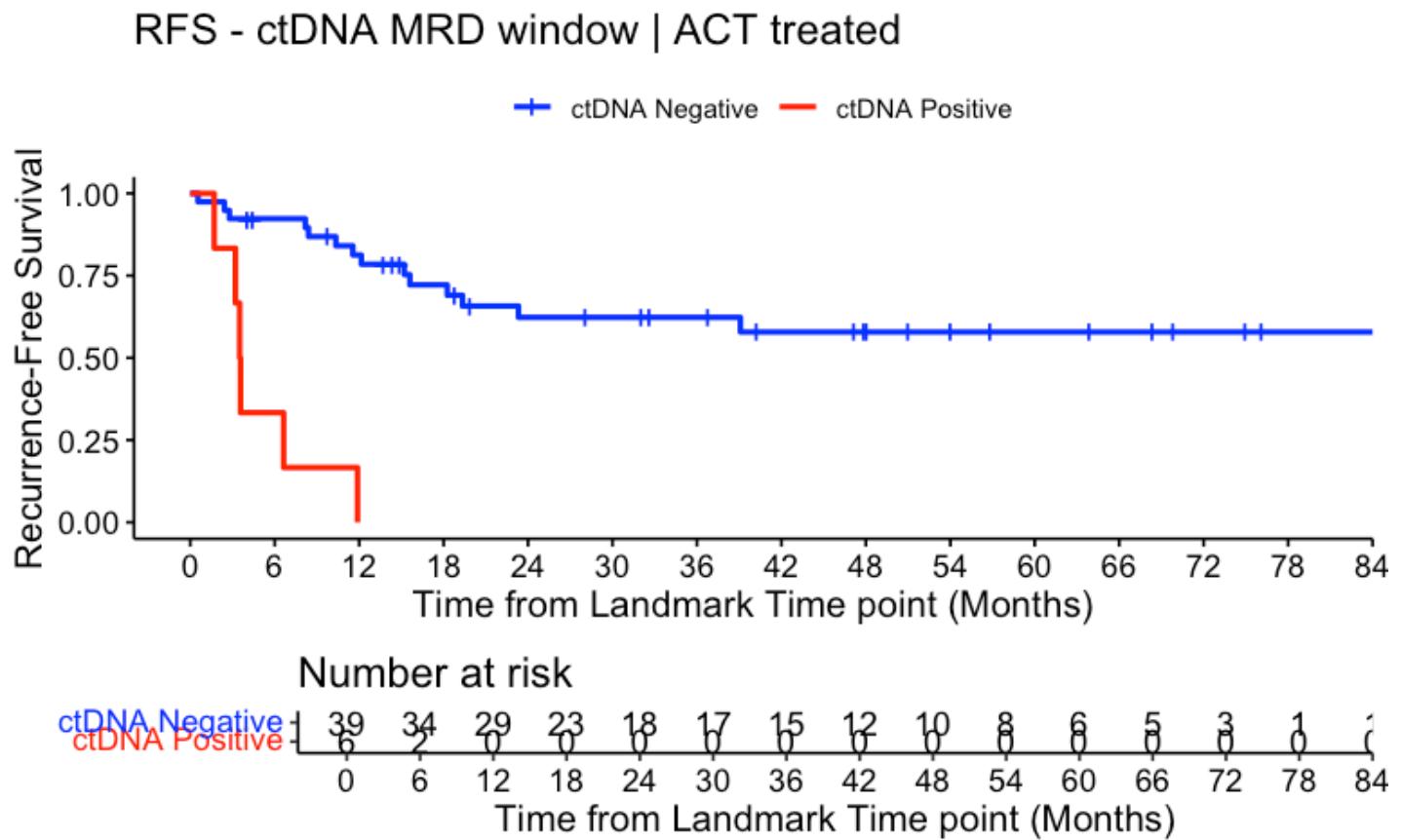
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.MRD, data = circ_data)
```

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	39	14	NA	23.32	NA
ctDNA.MRD=POSITIVE	6	6	3.54	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 | ACT treated",
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	39	0	1.000	0.000	1.000		1.000	
24	18	13	0.623	0.084	0.437		0.763	

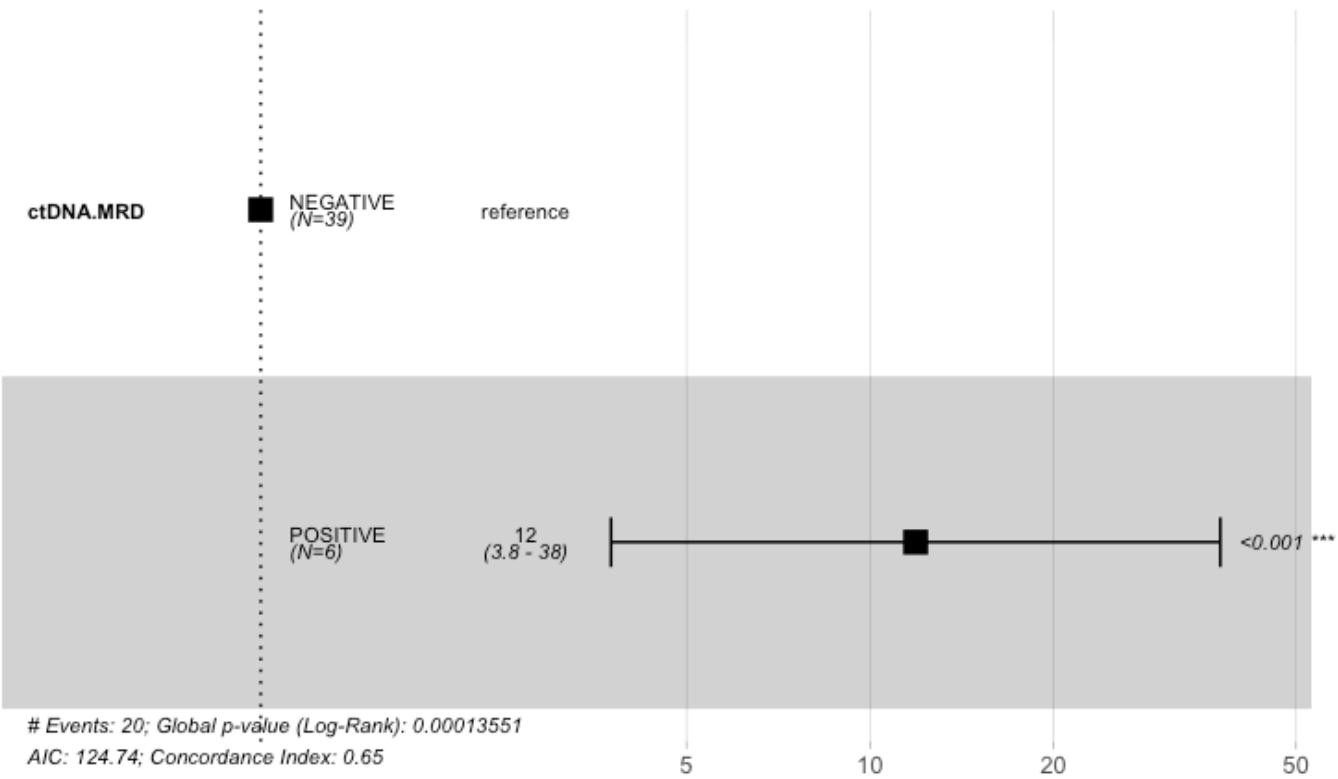
ctDNA.MRD=POSITIVE

time	n.risk	n.event	survival	std.err	lower	95% CI	upper	95% CI
0	6	0	1	0	0		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= 45, number of events= 20

            coef exp(coef) se(coef)      z Pr(>|z|)    
ctDNA.MRDPOSITIVE 2.4748    11.8797   0.5874 4.213 2.52e-05 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

            exp(coef) exp(-coef) lower .95 upper .95    
ctDNA.MRDPOSITIVE     11.88     0.08418     3.757     37.57   

Concordance= 0.651  (se = 0.048 ) 
Likelihood ratio test= 14.56  on 1 df,  p=1e-04 
Wald test             = 17.75  on 1 df,  p=3e-05 
Score (logrank) test = 27.27  on 1 df,  p=2e-07
```

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 11.88 (3.76-37.57); p = 0"
```

#OS by ctDNA at the MRD Window - ACT pts only

Hide

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

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

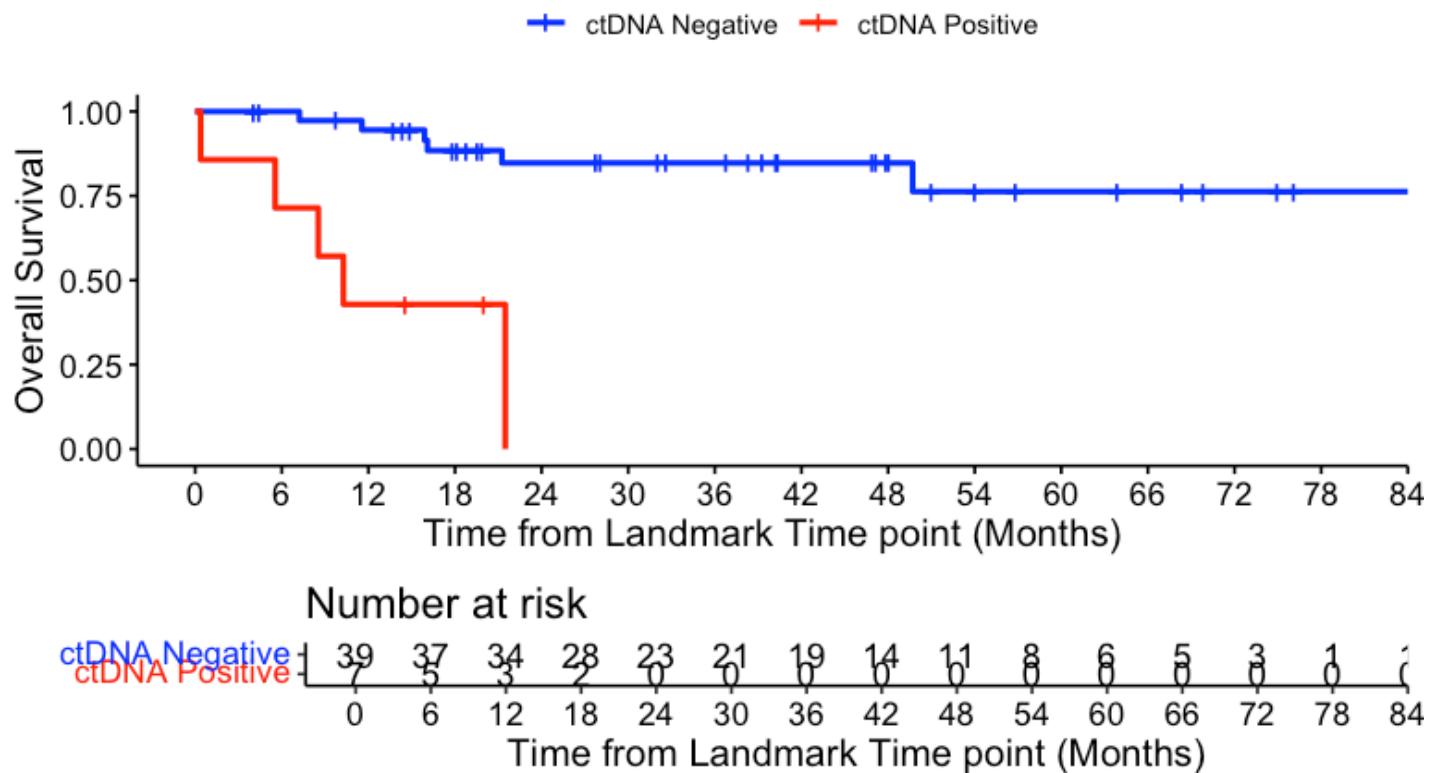
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.MRD, data = circ_data)

      n  events  median  0.95LCL  0.95UCL
ctDNA.MRD=NEGATIVE 39      6      NA      NA
ctDNA.MRD=POSITIVE  7      5  10.3    5.54      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 | ACT treated",
ylab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.labs
=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

OS - ctDNA MRD window | ACT treated



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

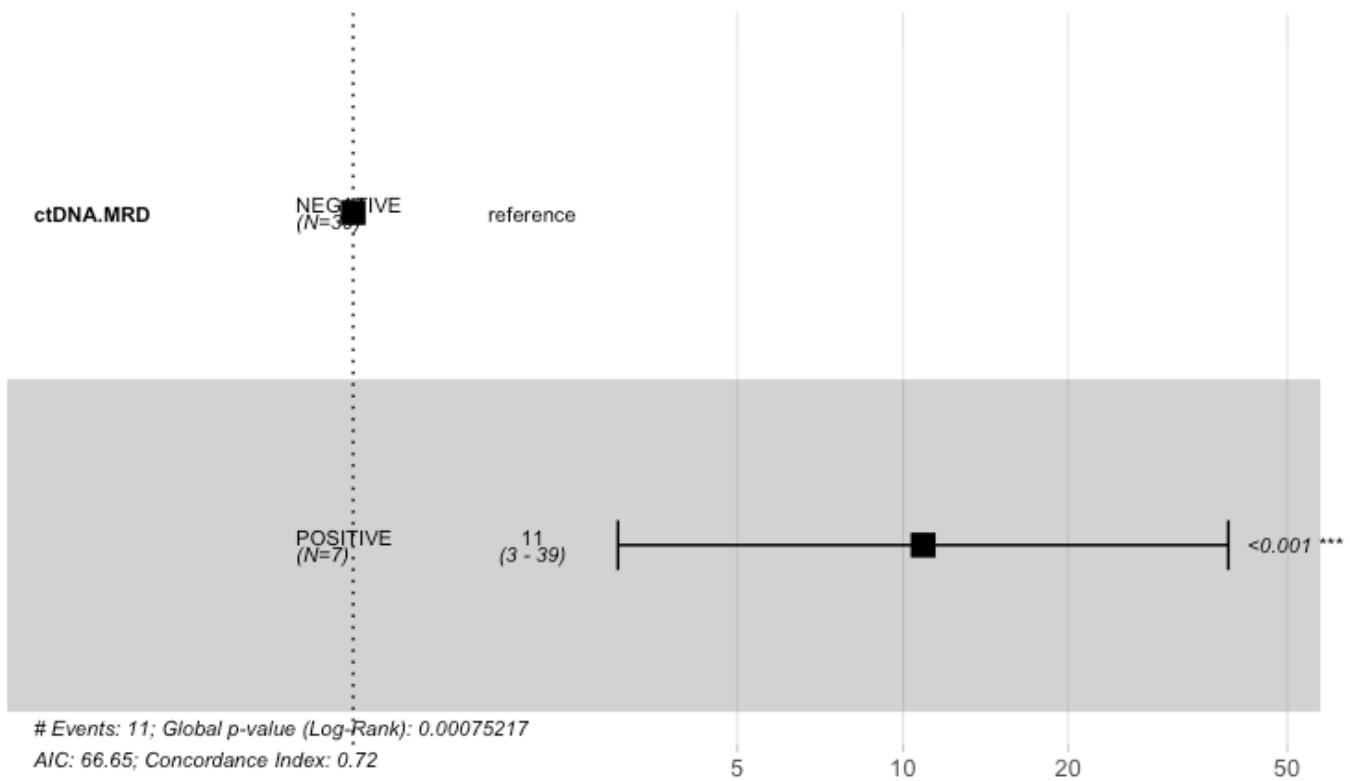
```
Call: survfit(formula = surv_object ~ ctDNA.MRD, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.MRD=NEGATIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	39	0	1.000	0.0000	1.000	1.000
24	23	5	0.847	0.0636	0.669	0.934

ctDNA.MRD=POSITIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	7	0	1	0	0	1

```
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= 46, number of events= 11

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.MRDPOSITIVE	2.3872	10.8835	0.6522	3.66	0.000252 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.MRDPOSITIVE	10.88	0.09188	3.031	39.08

Concordance= 0.717 (se = 0.071)
 Likelihood ratio test= 11.36 on 1 df, p=8e-04
 Wald test = 13.4 on 1 df, p=3e-04
 Score (logrank) test = 20.32 on 1 df, p=7e-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 = 10.88 (3.03-39.08); p = 0"
```

#RFS by ctDNA at the MRD Window - NAC pts only

[Hide](#)

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

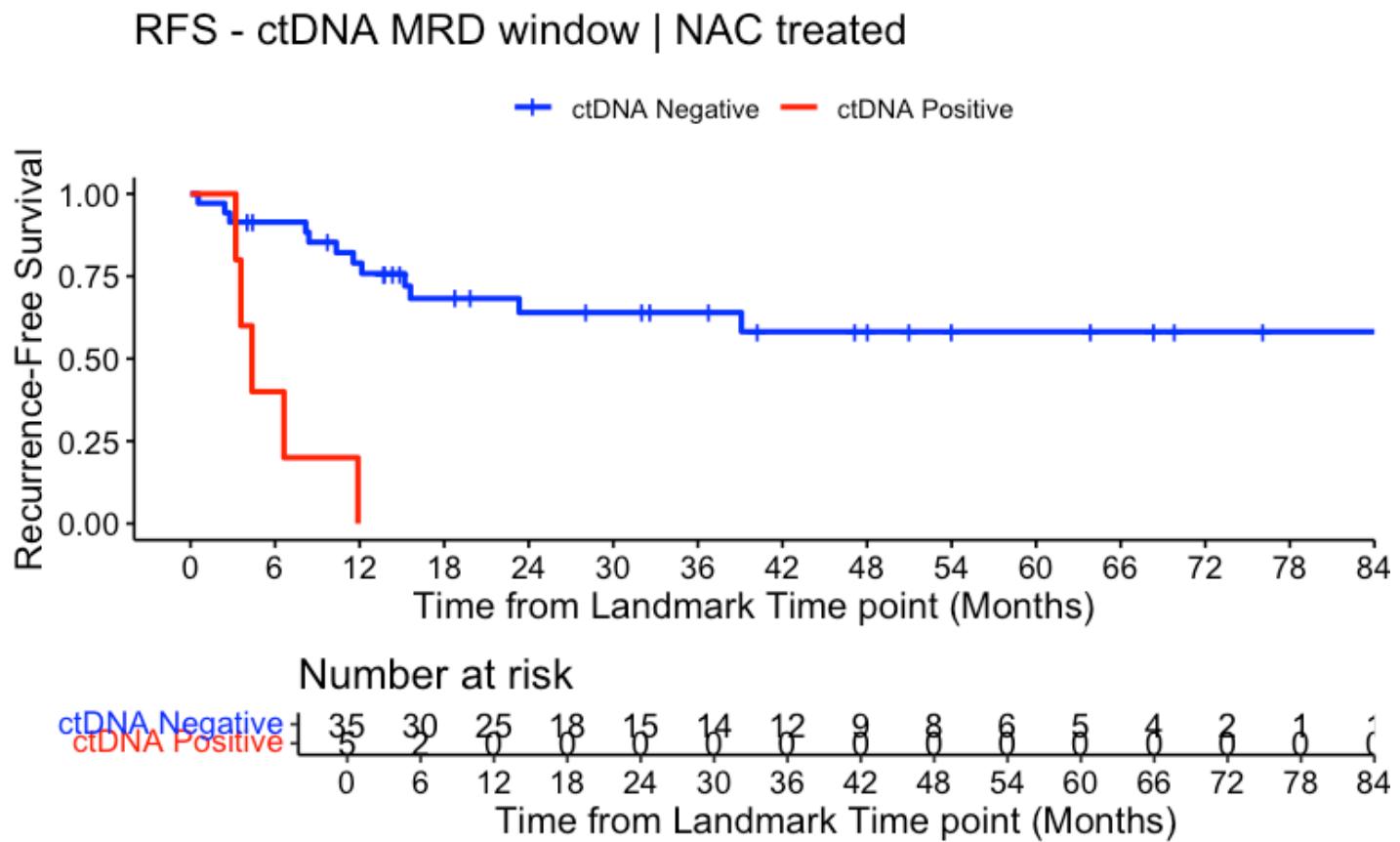
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.MRD, data = circ_data)
```

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	35	12	NA	23.32	NA
ctDNA.MRD=POSITIVE	5	5	4.36	3.57	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 | NAC treated",
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	35	0	1.00	0.0000	1.000		1.000	
24	15	11	0.64	0.0892	0.439		0.785	

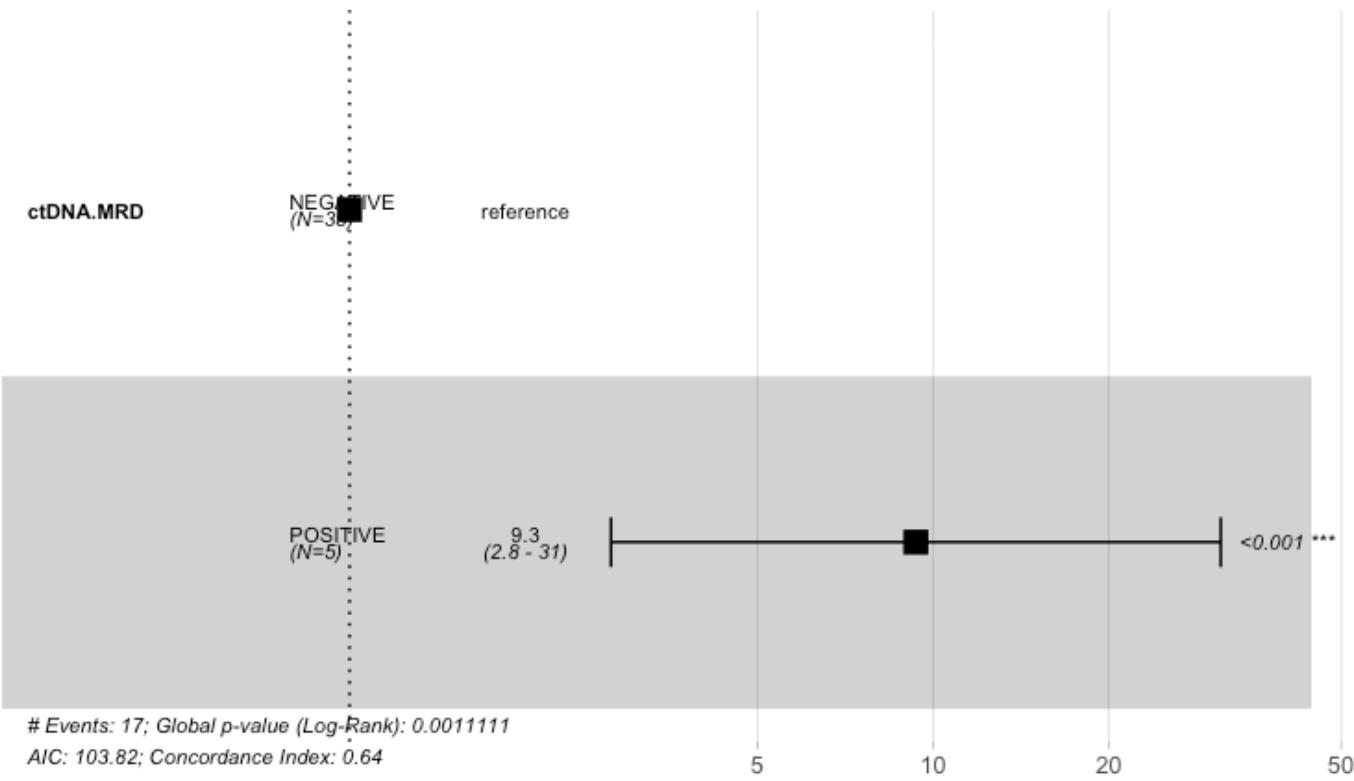
ctDNA.MRD=POSITIVE

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$ctDNA.MRD <- factor(circ_data$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= 40, number of events= 17

            coef exp(coef)  se(coef)      z Pr(>|z|)    
ctDNA.MRDPOSITIVE 2.2339     9.3362   0.6137  3.64 0.000272 *** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

            exp(coef) exp(-coef) lower .95 upper .95    
ctDNA.MRDPOSITIVE     9.336      0.1071     2.804     31.08   

Concordance= 0.639  (se = 0.051 )
Likelihood ratio test= 10.63  on 1 df,  p=0.001
Wald test             = 13.25  on 1 df,  p=3e-04
Score (logrank) test = 18.98  on 1 df,  p=1e-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 = 9.34 (2.8-31.08); p = 0"
```

#OS by ctDNA at the MRD Window - NAC pts only

Hide

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

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

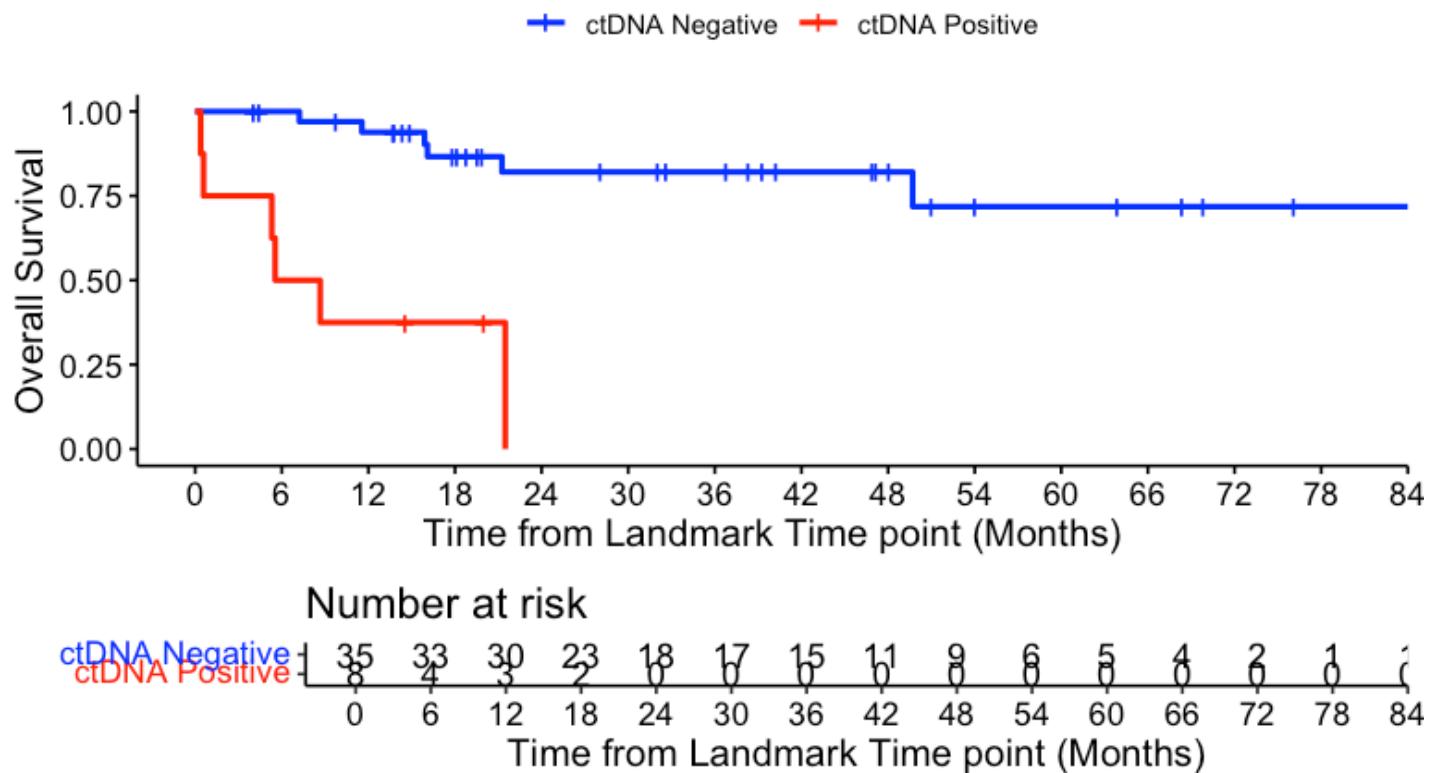
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.MRD, data = circ_data)
```

	n	events	median	0.95LCL	0.95UCL
ctDNA.MRD=NEGATIVE	35	6	NA	NA	NA
ctDNA.MRD=POSITIVE	8	6	7.1	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 | NAC treated",
ylab= "Overall Survival", xlab="Time from Landmark Time point (Months)", legend.labs
=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```

OS - ctDNA MRD window | NAC treated



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

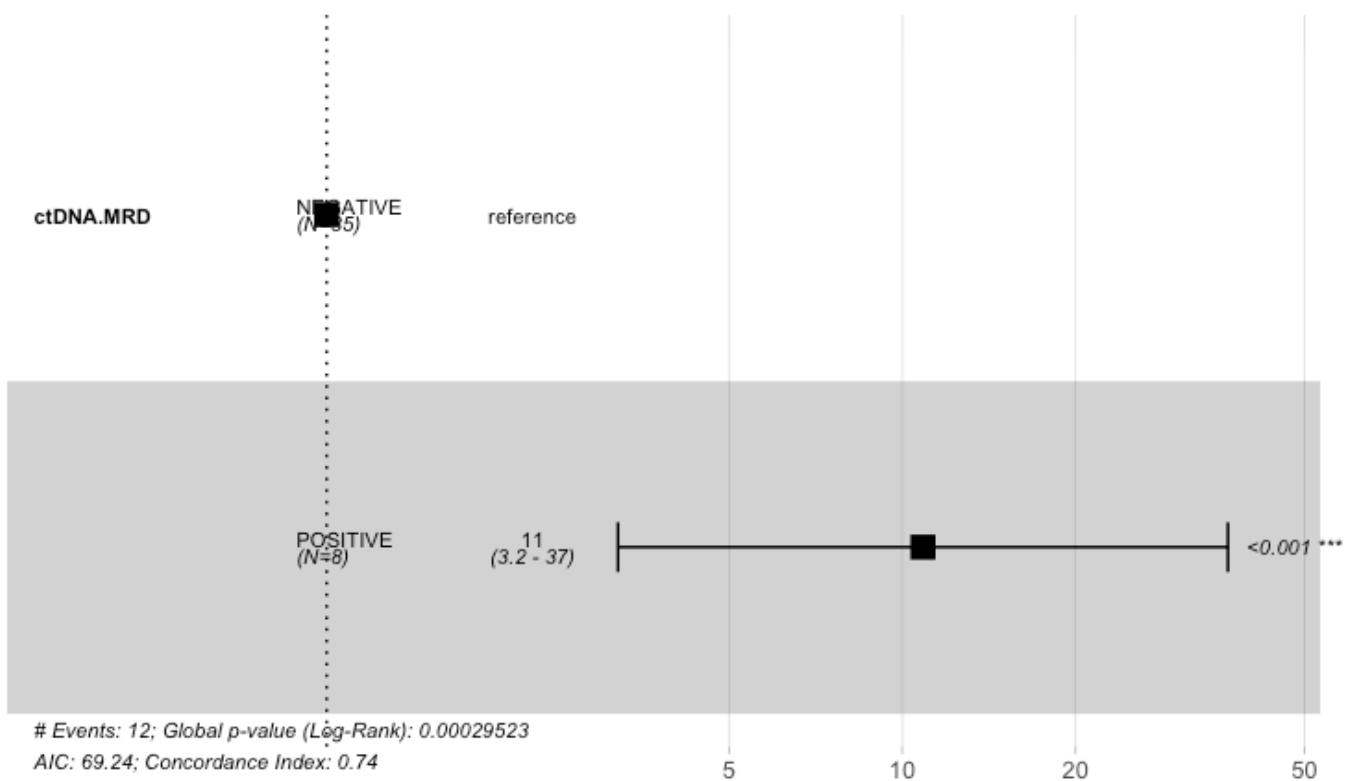
```
Call: survfit(formula = surv_object ~ ctDNA.MRD, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

ctDNA.MRD=NEGATIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	35	0	1.000	0.0000	1.000	1.000
24	18	5	0.821	0.0741	0.616	0.922

ctDNA.MRD=POSITIVE						
time	n.risk	n.event	survival	std.err	lower	95% CI
0	8	0	1	0	0	1
1						

```
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= 43, number of events= 12

	coef	exp(coef)	se(coef)	z	Pr(> z)
ctDNA.MRDPOSITIVE	2.3851	10.8601	0.6224	3.832	0.000127 ***

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

	exp(coef)	exp(-coef)	lower .95	upper .95
ctDNA.MRDPOSITIVE	10.86	0.09208	3.206	36.78

Concordance= 0.742 (se = 0.065)
 Likelihood ratio test= 13.1 on 1 df, p=3e-04
 Wald test = 14.68 on 1 df, p=1e-04
 Score (logrank) test = 22.15 on 1 df, p=3e-06

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 10.86 (3.21-36.78); 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)
```

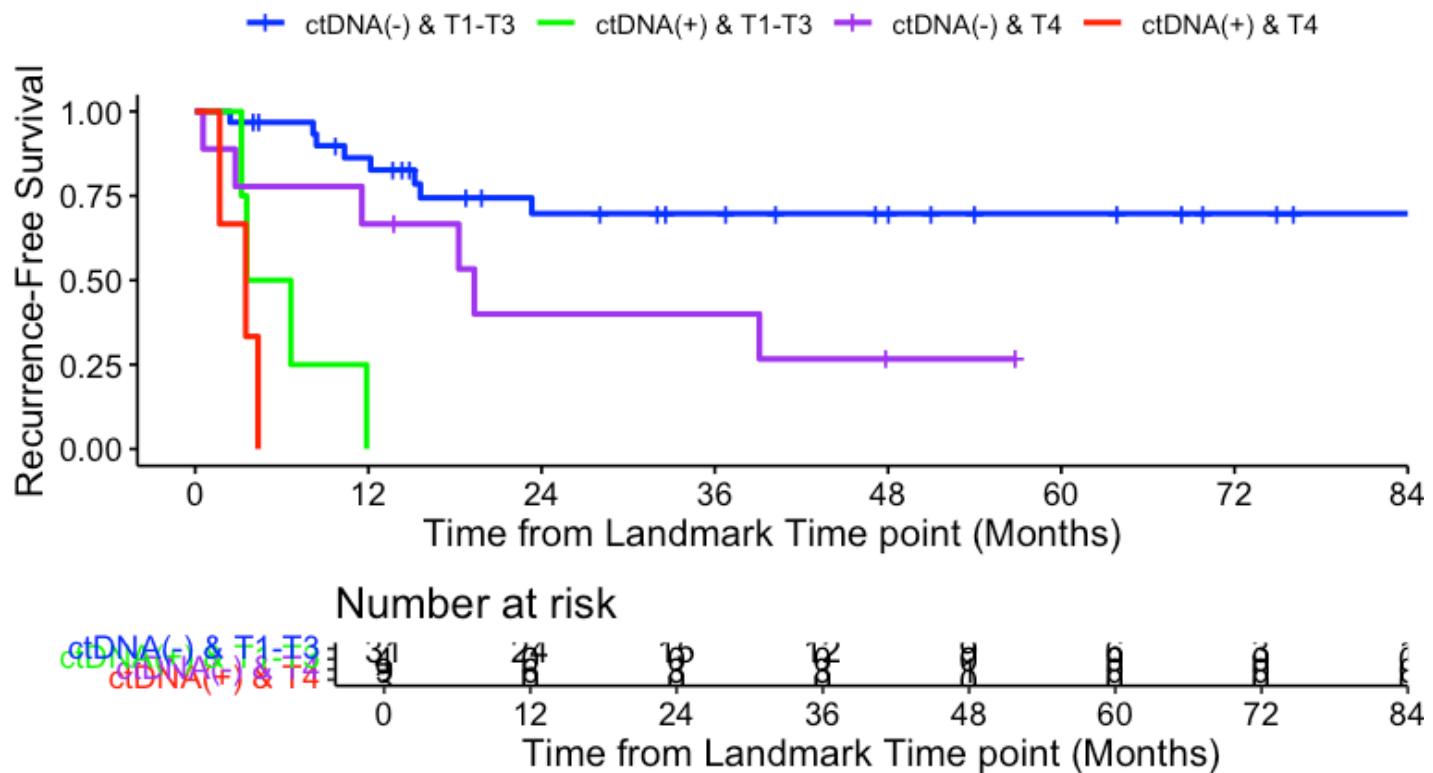
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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="RFS - ctDNA MRD & ypTN",
ylab= "Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)",
legend.labs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"),
legend.title="")
```

RFS - ctDNA MRD & ypTN



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)
```

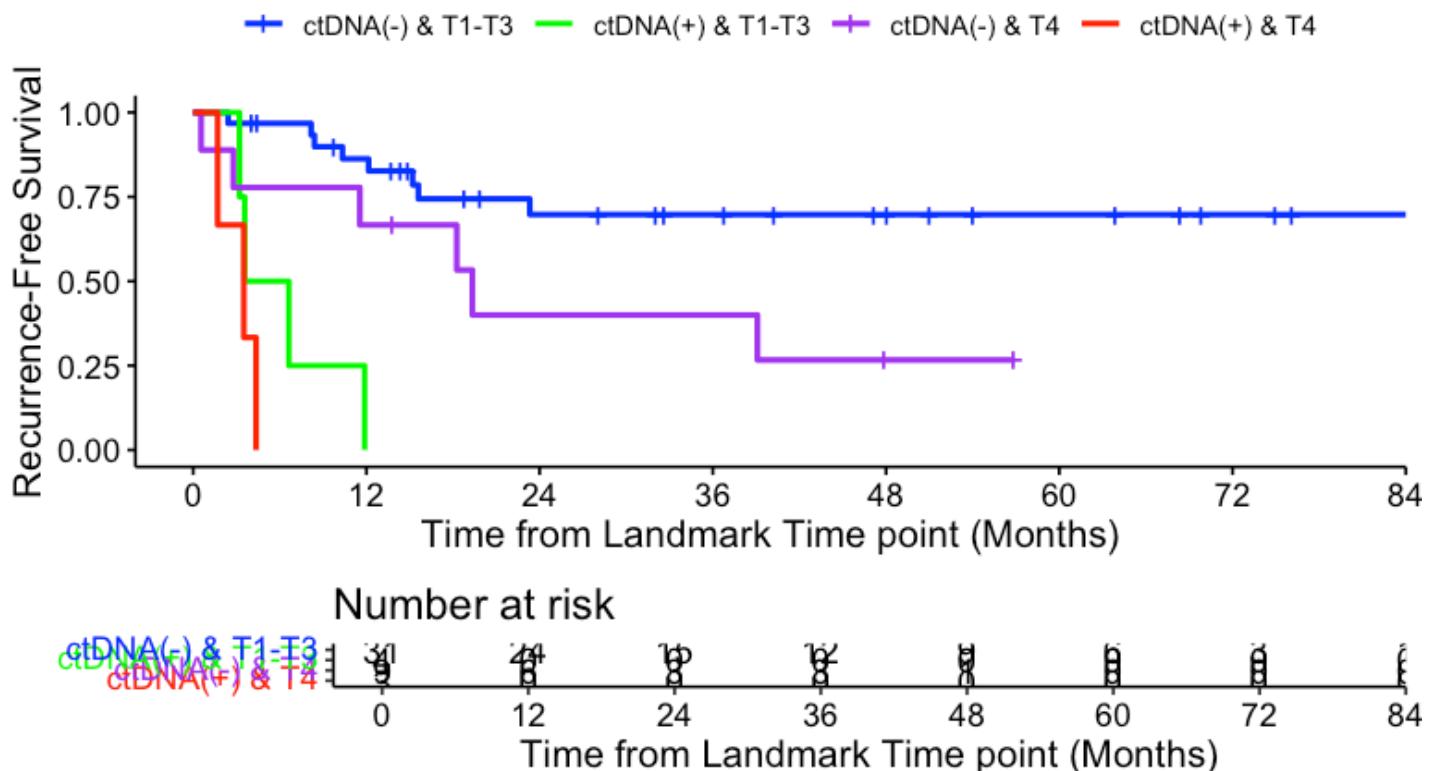
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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="RFS - ctDNA MRD &
ypTN", ylab= "Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)", legend.labs=c("ctDNA(-) & T1-T3", "ctDNA(+) & T1-T3", "ctDNA(-) & T4", "ctDNA(+) & T4"), legend.title="")
```

RFS - ctDNA MRD & ypTN



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)

```

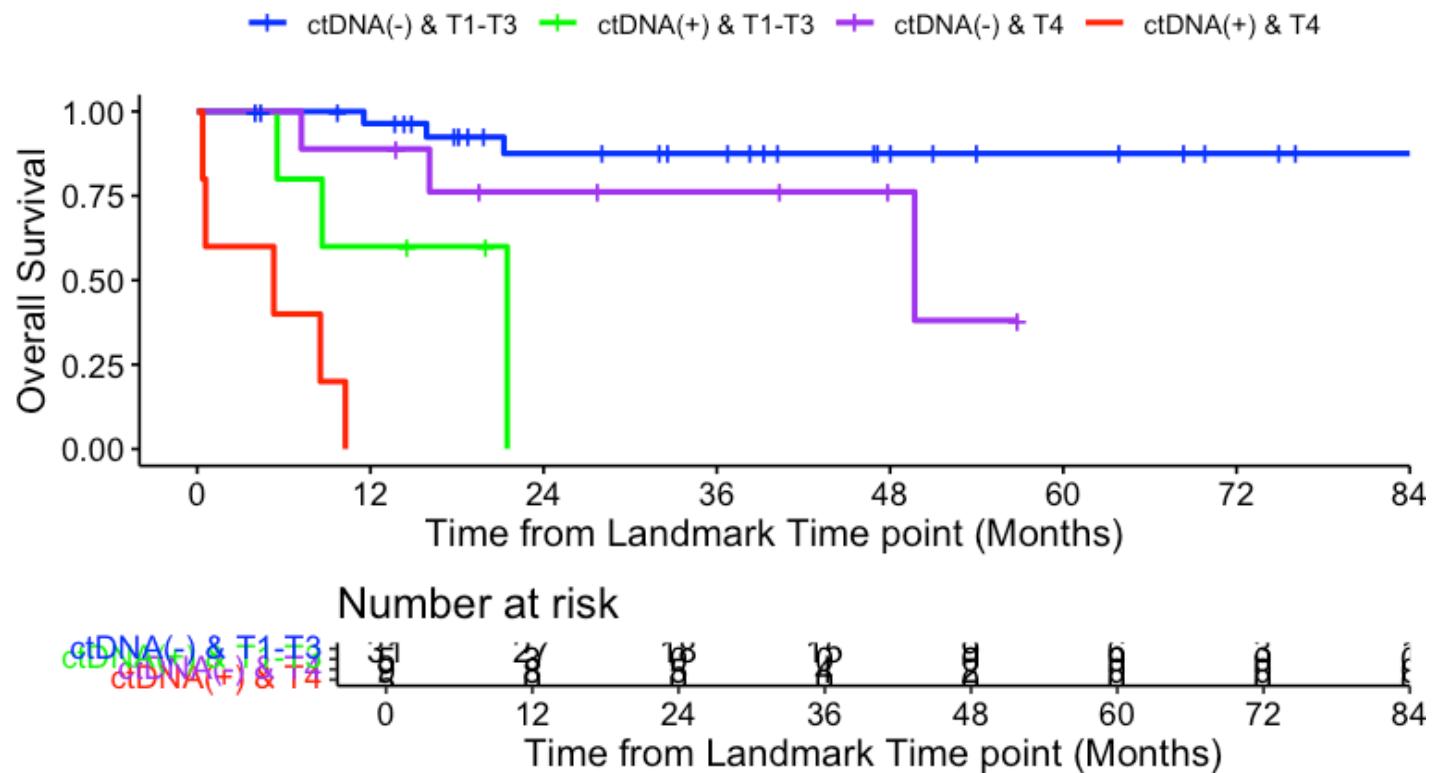
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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)
```

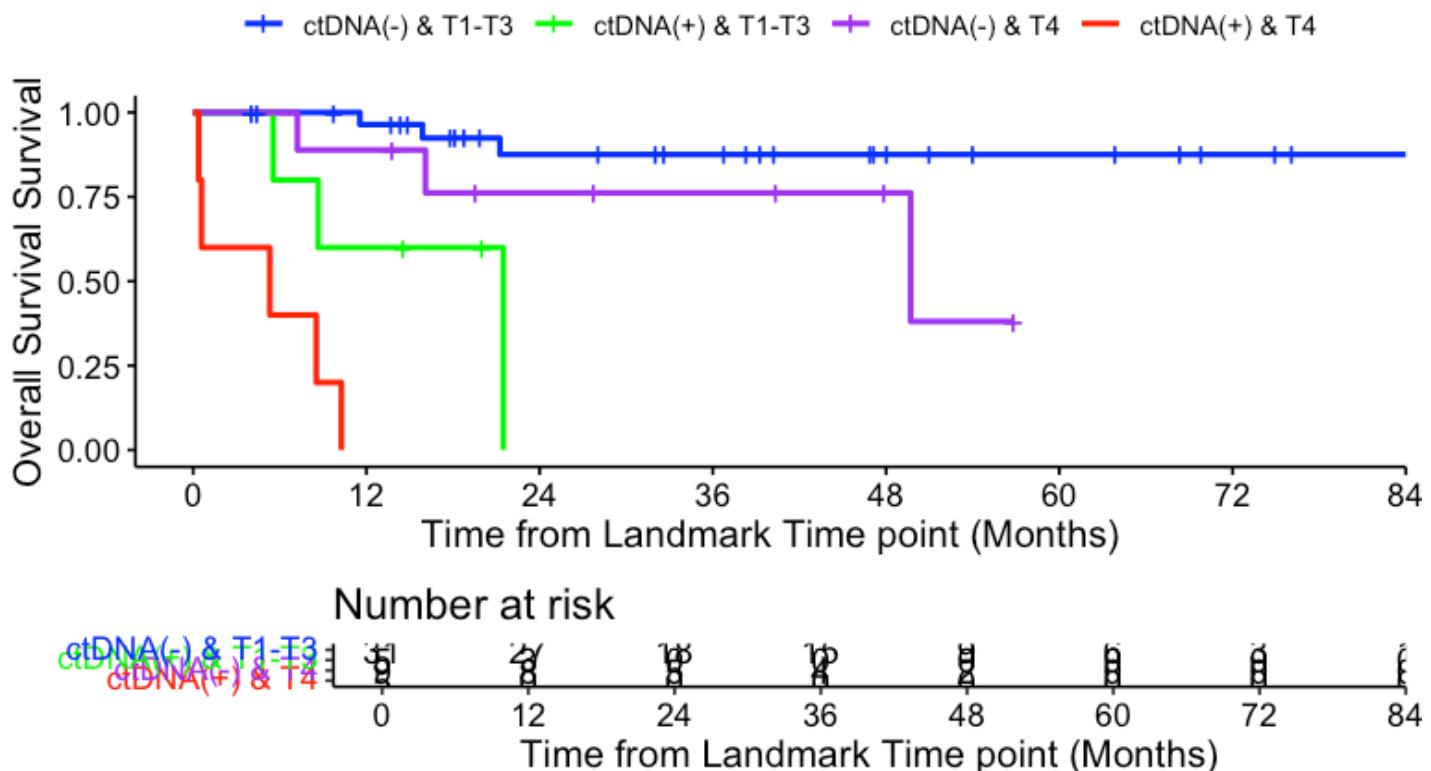
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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



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

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=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

CI	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)

```

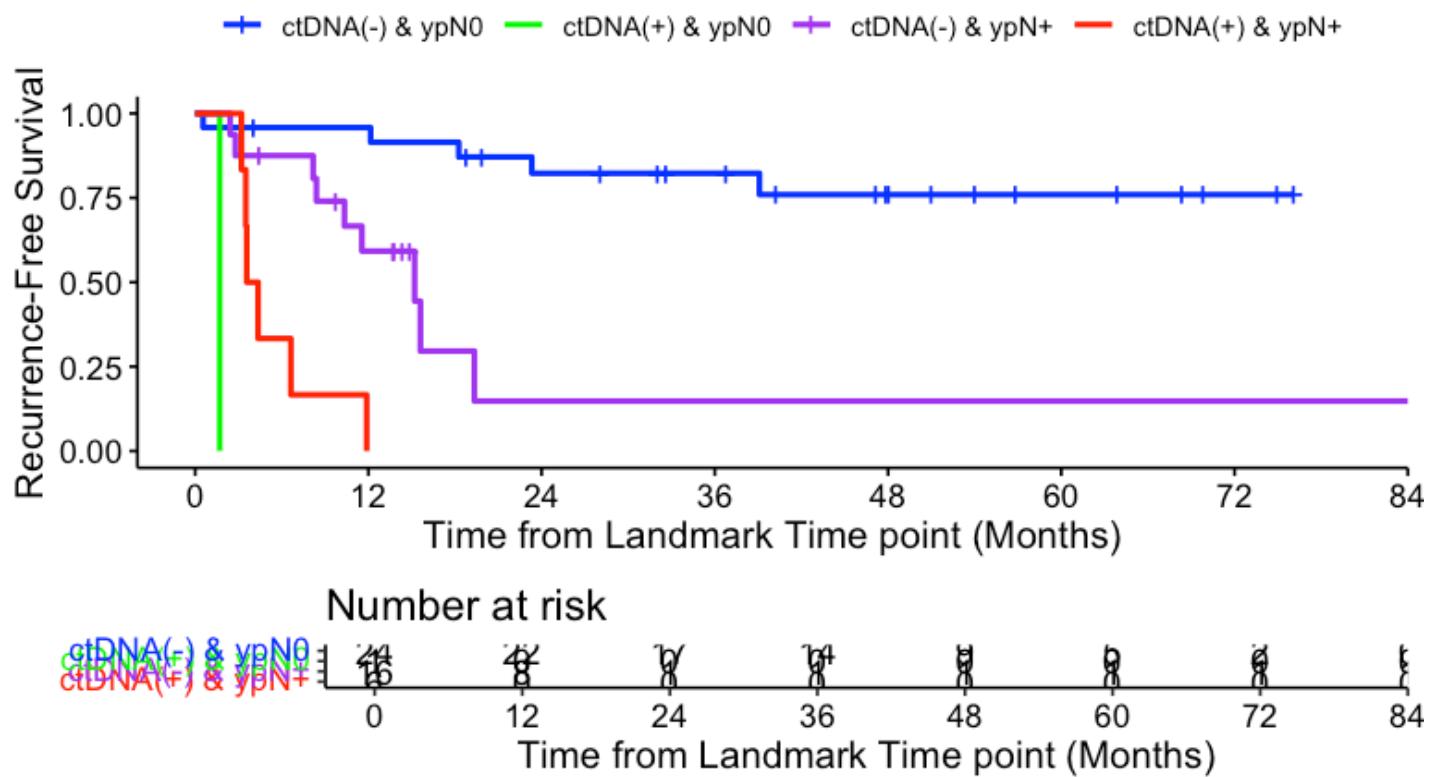
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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="RFS - ctDNA MRD &
ypN", ylab= "Recurrence-Free Survival", xlab="Time from Landmark Time point (Months)", l-
egend.labs=c("ctDNA(-) & ypN0", "ctDNA(+) & ypN0", "ctDNA(-) & ypN+", "ctDNA(+) & ypN-
+"), legend.title="")
```

RFS - ctDNA MRD & 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      1.00
   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
  0           1           0           1           0           1
  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
  0           6           0           1           0           1
  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)

```

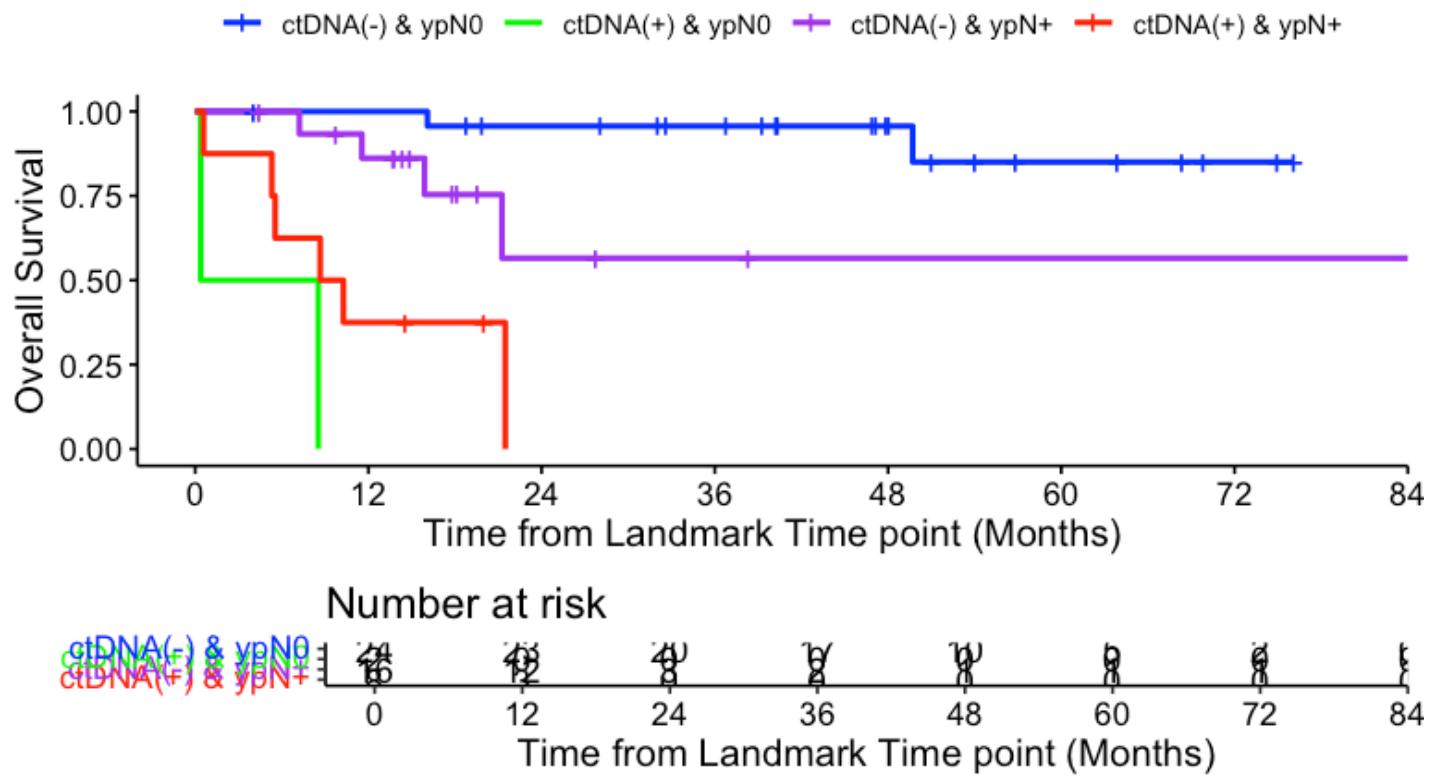
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.Stage.II.TNM, data = circ_data)
```

	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



```
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
  0           2           0           1           0           1
  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
  0           8           0           1           0           1
  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

#RFS by ctDNA at the MRD Window & TRG combination (TRG1/2/3 vs TRG3-5)

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$TRG <- factor(circ_data$TRG, levels = c("TRG1/2/3", "TRG4/5"))

circ_data$ctDNA.pCR <- NA
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)

```

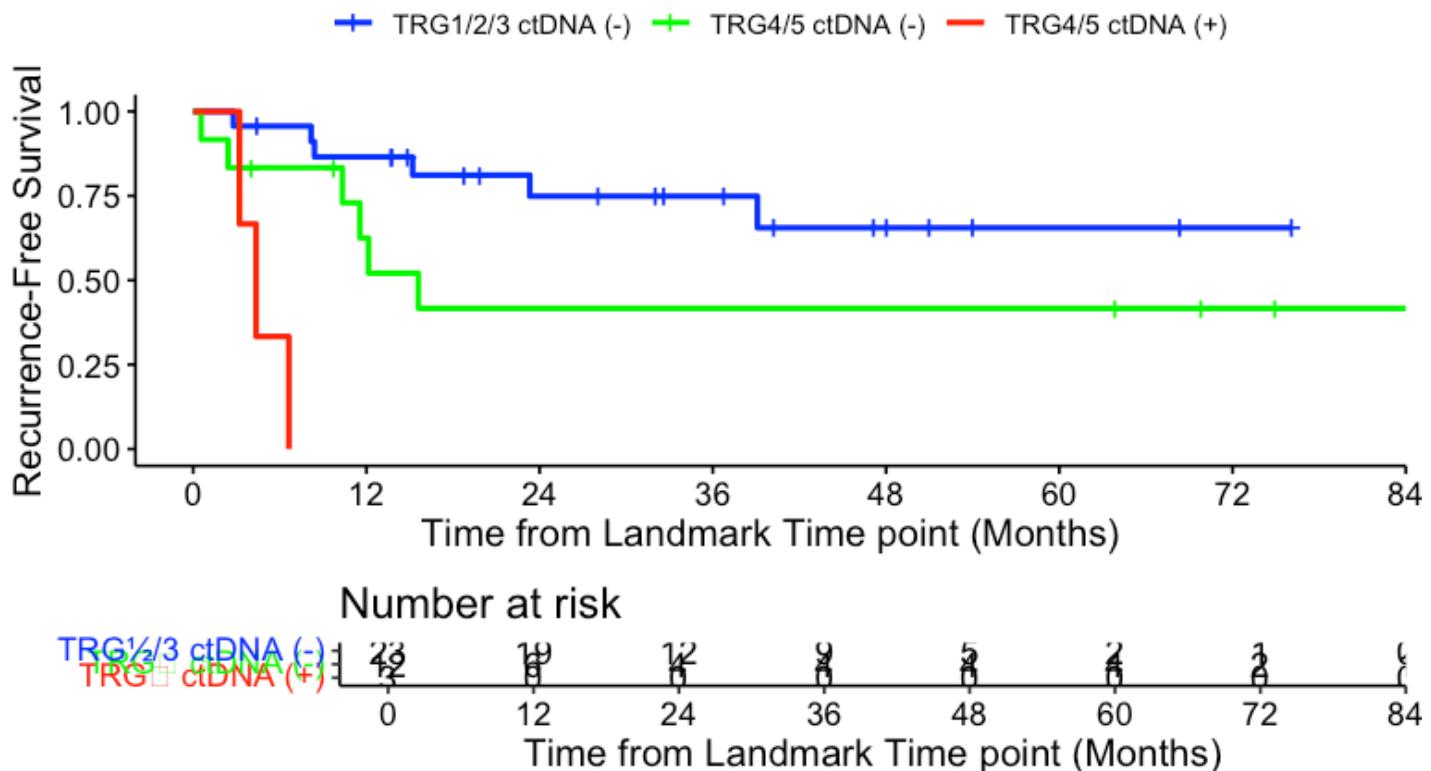
```
Call: survfit(formula = Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event) ~
  ctDNA.pCR, data = circ_data)
```

	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



```
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
 24     12      5     0.749  0.0993     0.492     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
 24      4      6     0.417  0.157     0.131     0.686
```

```
ctDNA.pCR=3
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     3      0     1.000  0.000     1.000     1.000
 1     1      0     1.000  0.000     1.000     1.000
```

```
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 (TRG1/2/3 vs TRG3-5)

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$TRG <- factor(circ_data$TRG, levels = c("TRG1/2/3", "TRG4/5"))

circ_data$ctDNA.pCR <- NA
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)
```

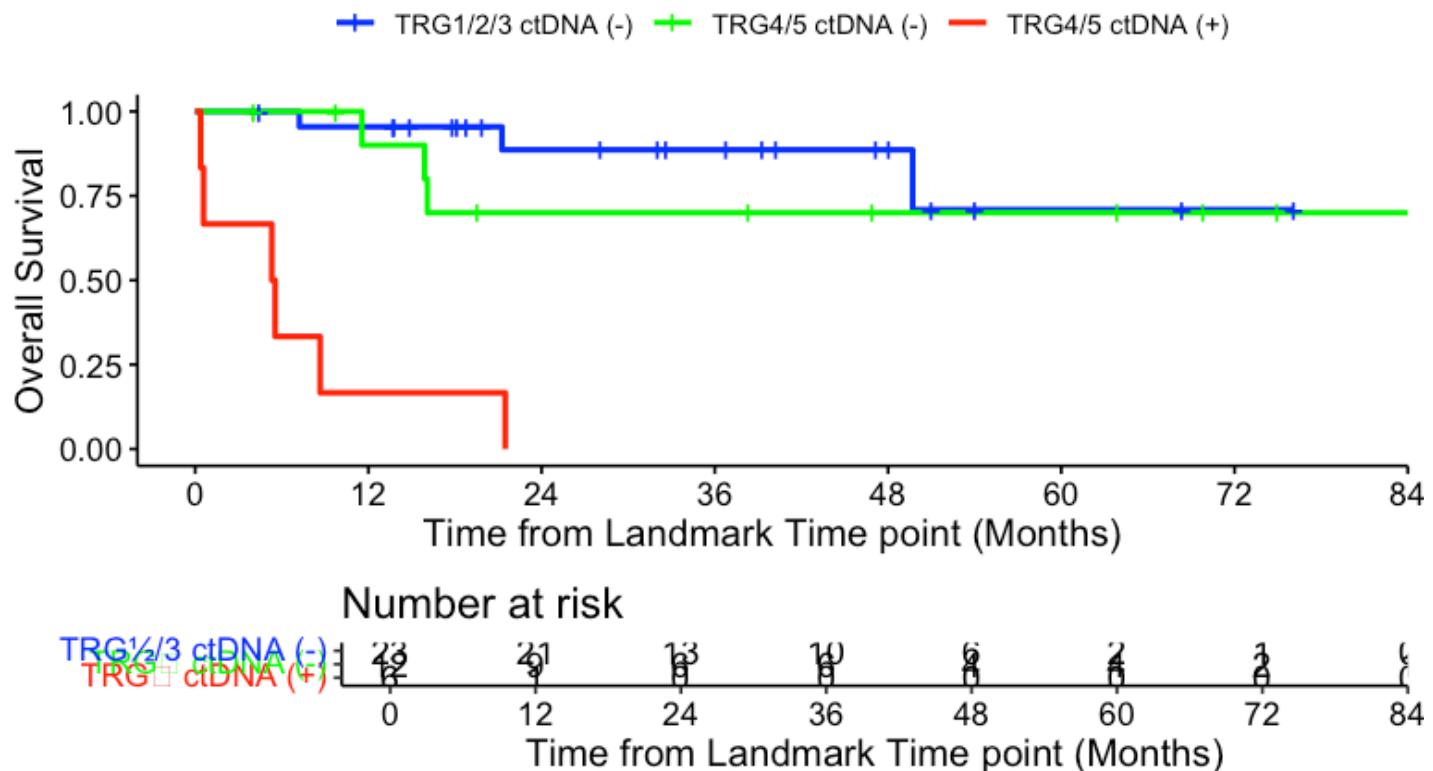
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
  ctDNA.pCR, data = circ_data)
```

	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



```
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
 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     1.000  0.0000    1.000    1.000
 24      6      3     0.700  0.1450    0.329    0.892
```

```
ctDNA.pCR=3
time n.risk n.event survival std.err lower 95% CI upper 95% CI
  0     6      0     1.000  0.0000    1.000    1.000
 1     1      0     1.000  0.0000    1.000    1.000
```

```
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

#RFS 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$DFS.Event)~ctDNA.pCR, data = circ_data)
surv_object <- Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.pCR, data = circ_data, conf.int = 0.95, conf.type = "log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRUE,
            break.time.by = 12, palette = c("blue", "red"),
            title = "RFS - ctDNA MRD/Clinical Characteristics",
            ylab = "Recurrence-Free Survival", xlab = "Time from Landmark Time point (Months)",
            legend.labs = c("MRD neg, T1-T3, N0", "MRD pos or T4 or N1-N3"),
            legend.title = "")
summary(KM_curve, times = c(0, 24))
circ_data$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)
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)

```

#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$OS.Event)~ctDNA.pCR, data = circ_data)
surv_object <- Surv(time = circ_data$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))
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)

```

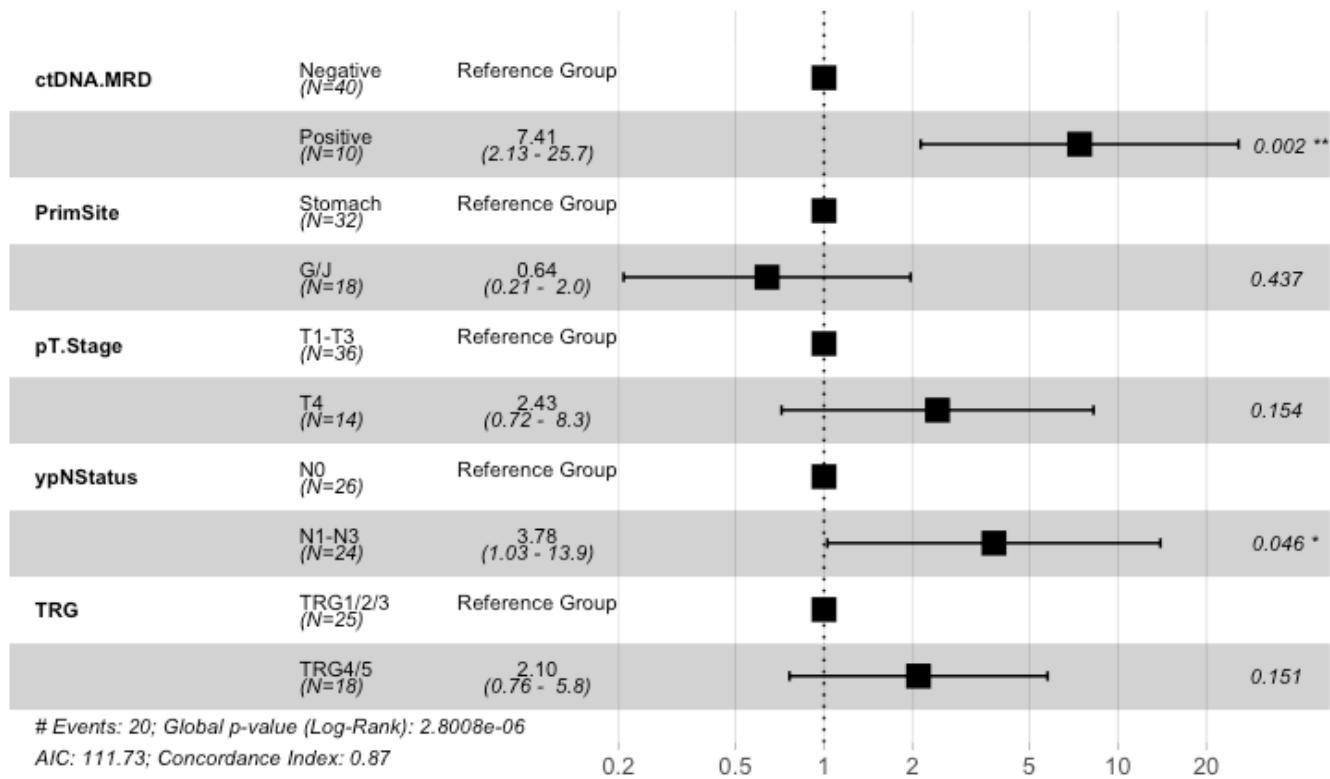
#Multivariate cox regression for RFS - ctDNA MRD Window

[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("Stomach","G/J"))
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 + 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)
```

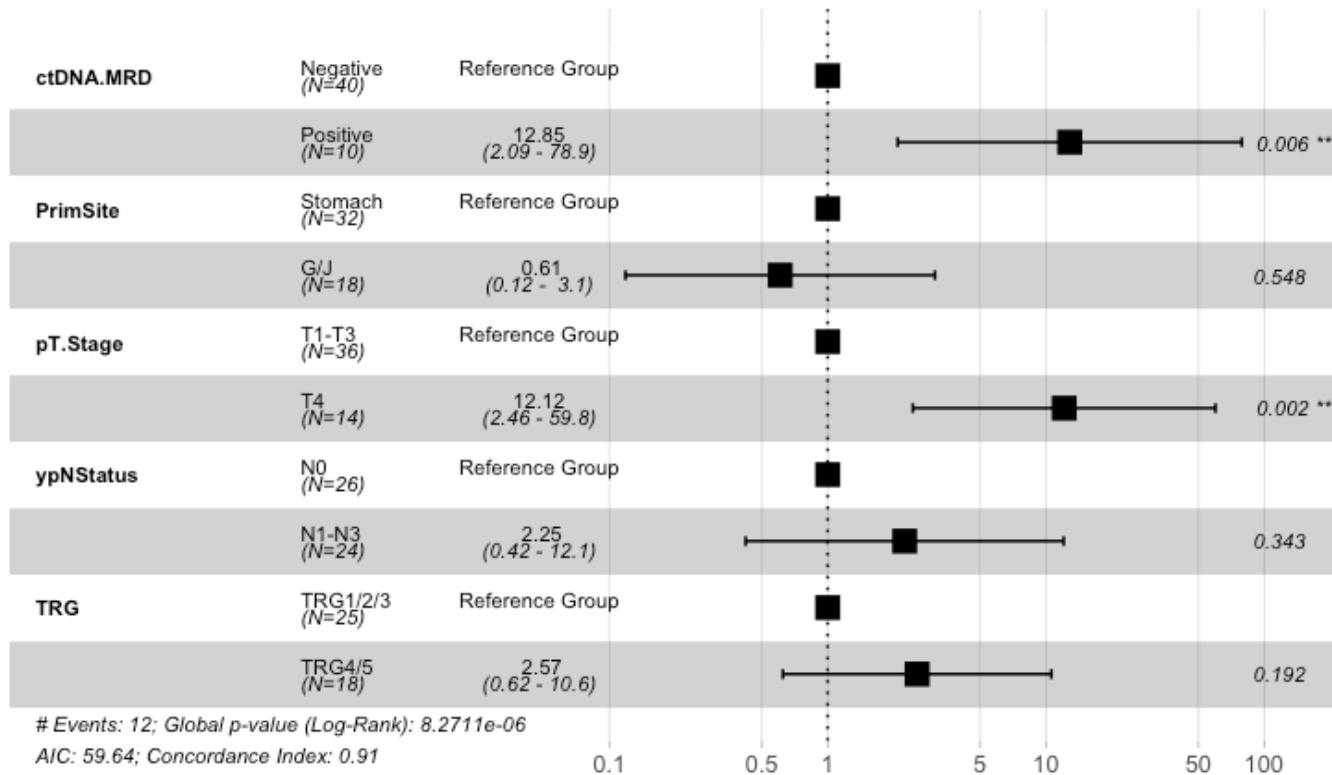
```
#Multivariate cox regression for OS - ctDNA MRD Window
```

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("Stomach","G/J"))
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 + 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)
```

#RFS 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)

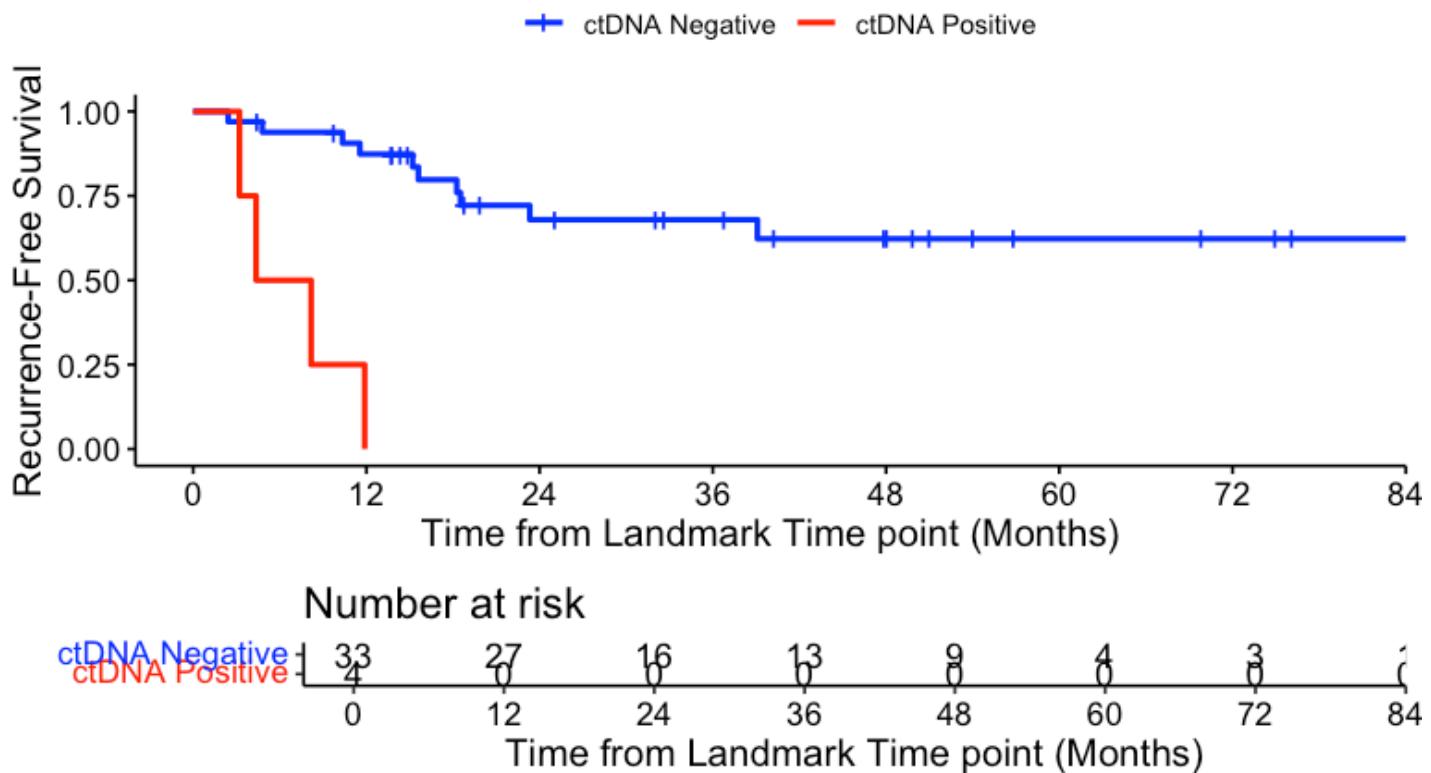
survfit(Surv(time = circ_data$DFS.months, event = circ_data$DFS.Event)~ctDNA.surveillance, data = circ_data)
```

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

```
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))
```

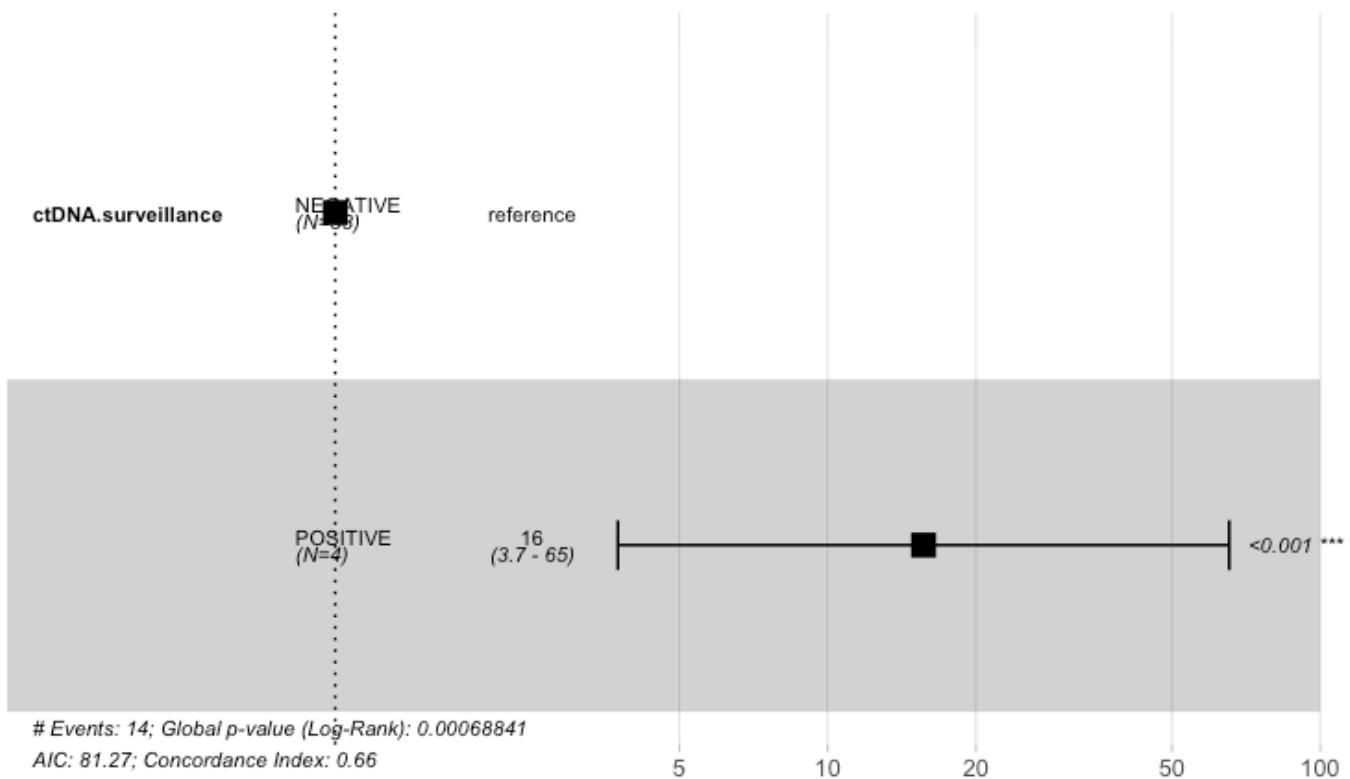
```
Call: survfit(formula = surv_object ~ ctDNA.surveillance, data = circ_data,
  conf.int = 0.95, conf.type = "log-log")
```

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)
```

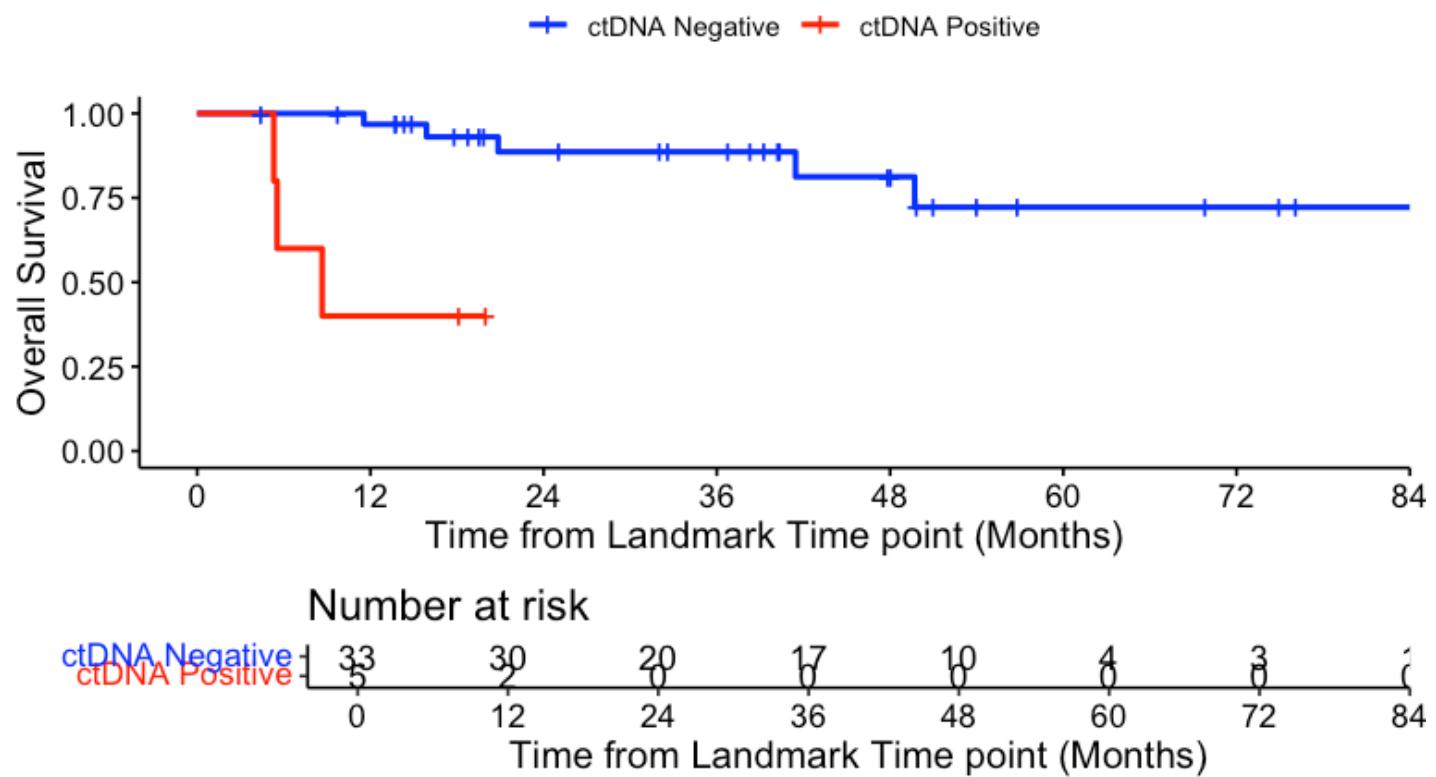
```
Call: survfit(formula = Surv(time = circ_data$FU.months, event = circ_data$OS.Event) ~
ctDNA.surveillance, data = circ_data)
```

	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="")
```

OS - 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	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	

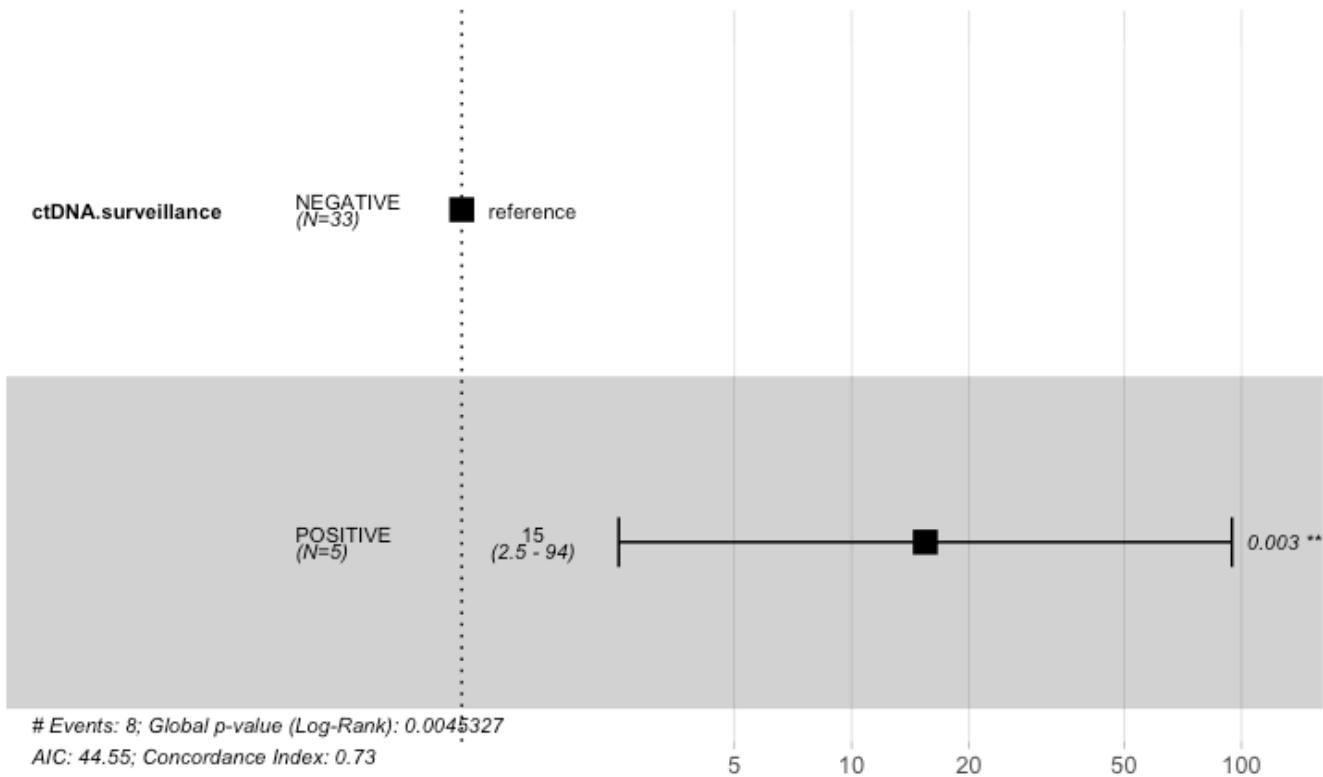
ctDNA.surveillance=POSITIVE

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"
```