Hassoun et al_ctDNA in Testicular Cancer Clinical Analysis

Code

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(rms) library(DT)

#ctDNA Detection rate by Stage and Window

Hide

```
#Pre-Orchiectomy
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$preSx.ctDNA.anytime!="",]
circ_data$preSx.ctDNA.anytime <- factor(circ_data$preSx.ctDNA.anytime, levels=c("NEGATIV
E","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, preSx.ctDNA.anytime %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$preSx.ctDNA.anytime == "POSITIVE", by=li
st(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$preSx.ctDNA.anytime, by=list(circ_data$Stag
e), 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$x) * 100  # Convert to perc
entage
)
combined_data$Rate <- sprintf("%.2f%%", combined_data$Rate)
overall_total_count <- nrow(circ_data)
overall_positive_count <- nrow(circ_data[circ_data$preSx.ctDNA.anytime == "POSITIVE",])
overall_positivity_rate <- (overall_positive_count / overall_total_count) * 100  # Conve
rt 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	11	91.67%
II	1	1	100.00%
III	2	2	100.00%
Overall	15	14	93.33%

4 rows

Hide

```
#MRD Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD.Window!="",]
circ_data$ctDNA.MRD.Window <- factor(circ_data$ctDNA.MRD.Window, levels=c("NEGATIVE","PO
SITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, ctDNA.MRD.Window %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$ctDNA.MRD.Window == "POSITIVE", by=list
(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.MRD.Window, by=list(circ_data$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$x) * 100  # Convert to perc
entage
)
combined_data$Rate <- sprintf("%.2f%%", combined_data$Rate)
overall_total_count <- nrow(circ_data)
overall_positive_count <- nrow(circ_data[circ_data$ctDNA.MRD.Window == "POSITIVE",])
overall_positivity_rate <- (overall_positive_count / overall_total_count) * 100  # Conve
rt 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	20	3	15.00%
II	9	7	77.78%
III	13	11	84.62%
Overall	42	21	50.00%

4 rows

Hide

```
#Surveillance Window
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.surveillance!="",]
circ_data$ctDNA.MRD.Window <- factor(circ_data$ctDNA.surveillance, levels=c("NEGATIV
E","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$ctDNA.surveillance == "POSITIVE", by=lis
t(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$ctDNA.surveillance, by=list(circ_data$Stag
e), 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$x) * 100  # Convert to perc
entage
)
combined_data$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  # Conve
rt 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	27	4	14.81%
II	12	4	33.33%
III	12	5	41.67%
Overall	51	13	25.49%

4 rows

Hide

```
#Anytime post-surgery
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$postSx.ctDNA.anytime!="",]
circ_data$postSx.ctDNA.anytime <- factor(circ_data$postSx.ctDNA.anytime, levels=c("NEGAT
IVE","POSITIVE"))
circ_data$Stage <- factor(circ_data$Stage, levels=c("I","II","III"))
circ_data <- subset(circ_data, postSx.ctDNA.anytime %in% c("NEGATIVE", "POSITIVE"))
positive_counts_by_stage <- aggregate(circ_data$postSx.ctDNA.anytime == "POSITIVE", by=l
ist(circ_data$Stage), FUN=sum)
total_counts_by_stage <- aggregate(circ_data$postSx.ctDNA.anytime, by=list(circ_data$Sta
ge), 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$x) * 100  # Convert to perc
entage
)
combined_data$Rate <- sprintf("%.2f%%", combined_data$Rate)
overall_total_count <- nrow(circ_data)
overall_positive_count <- nrow(circ_data[circ_data$postSx.ctDNA.anytime == "POSITIVE",])
overall_positivity_rate <- (overall_positive_count / overall_total_count) * 100  # Conve
rt to percentage
overall_row <- data.frame(
  Stage = "Overall",
  Total_Count = overall_total_count,
  Positive_Count = overall_positive_count,
  Rate = sprintf("%.2f%%", overall_positivity_rate)
)
combined_data <- rbind(combined_data, overall_row)
print(combined_data)
```

Stage <fctr>	Total_Count <int>	Positive_Count <int>	Rate <chr>
I	29	5	17.24%
II	17	12	70.59%
III	26	21	80.77%
Overall	72	38	52.78%

4 rows

#Demographics Table

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]

circ_data_subset <- circ_data %>%
  select(
    Age,
    pT,
    Stage,
    Histology,
    Treatment,
    RPLND.Pathology,
    preSx.markers.anytime,
    Markers.postSx.anytime,
    PFS.Event,
    OS.Event,
    OS.months) %>%
  mutate(
    Age = as.numeric(Age),
    pT = factor(pT, levels = c("0","1", "2", "3", "X"), labels = c("pT0","pT1", "pT2",
"pT3","pTX")),
    Stage = factor(Stage, levels = c("I", "II", "III")),
    Histology = factor(Histology, levels = c("Non seminoma", "Seminoma")),
    Treatment = factor(Treatment, levels = c("Surveillance", "Chemotherapy", "RPLND", "R
PLND, Chemotherapy")),
    RPLND.Pathology = factor(RPLND.Pathology, levels = c("Negative", "Germ Cell/Teratom
a")),
    preSx.markers.anytime = factor(preSx.markers.anytime, levels = c("NEGATIVE", "POSITI
VE"), labels = c("Normal", "Elevated")),
    Markers.postSx.anytime = factor(Markers.postSx.anytime, levels = c("NEGATIVE", "POSI
TIVE"), labels = c("Normal", "Elevated")),
    PFS.Event = factor(PFS.Event, levels = c("FALSE", "TRUE"), labels = c("No Event", "E
vent")),
    OS.Event = factor(OS.Event, levels = c("FALSE", "TRUE"), labels = c("Alive", "Deceas
ed")),
    OS.months = as.numeric(OS.months))
table1 <- circ_data_subset %>%
  tbl_summary(
    statistic = list(
      all_continuous() ~ "{median} ({min} - {max})",
      all_categorical() ~ "{n} ({p}%)"),
    label = list(
      Age ~ "Age",
      pT ~ "Pathological T Stage",
      Stage ~ "Overall Stage",
      Histology ~ "Histology",
      Treatment ~ "Post-surgery treatment",
      RPLND.Pathology ~ "RPLND Pathology",
      preSx.markers.anytime ~ "Serum Markers pre-op",
      Markers.postSx.anytime ~ "Serum Markers anytime post-op",
      PFS.Event ~ "Event During Follow-up",
```

```
      OS.Event ~ "Vital Status",
      OS.months ~ "Total Follow up (months)"),) %>%
  bold_labels()
table1
```

Characteristic	N = 74[1]
Age	34 (16 - 67)
Pathological T Stage	
pT0	1 (1.4%)
pT1	39 (53%)
pT2	28 (38%)
pT3	3 (4.1%)
pTX	3 (4.1%)
Overall Stage	
I	30 (41%)
II	17 (23%)
III	27 (36%)
Histology	
Non seminoma	51 (69%)
Seminoma	23 (31%)
Post-surgery treatment	
Surveillance	17 (23%)
Chemotherapy	30 (41%)
RPLND	5 (6.8%)
RPLND, Chemotherapy	22 (30%)
RPLND Pathology	
Negative	10 (37%)
Germ Cell/Teratoma	17 (63%)
Unknown	47

[1] Median (Range); n (%)

Characteristic	N = 74[1]
Serum Markers pre-op	
Normal	22 (30%)
Elevated	52 (70%)
Serum Markers anytime post-op	
Normal	30 (41%)
Elevated	44 (59%)
Event During Follow-up	
No Event	39 (53%)
Event	35 (47%)
Vital Status	
Alive	69 (93%)
Deceased	5 (6.8%)
Total Follow up (months)	17 (2 - 84)

[1] Median (Range); n (%)

Hide

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

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

Hide

```
fit1
```

Characteristic	N = 74[1]
Age	34 (16 - 67)
Pathological T Stage	

[1]Median (Range); n (%)

Characteristic	N = 74[1]
pT0	1 (1.4%)
pT1	39 (53%)
pT2	28 (38%)
pT3	3 (4.1%)
pTX	3 (4.1%)
Overall Stage	
I	30 (41%)
II	17 (23%)
III	27 (36%)
Histology	
Non seminoma	51 (69%)
Seminoma	23 (31%)
Post-surgery treatment	
Surveillance	17 (23%)
Chemotherapy	30 (41%)
RPLND	5 (6.8%)
RPLND, Chemotherapy	22 (30%)
RPLND Pathology	
Negative	10 (37%)
Germ Cell/Teratoma	17 (63%)
Unknown	47
Serum Markers pre-op	
Normal	22 (30%)
Elevated	52 (70%)
Serum Markers anytime post-op	
Normal	30 (41%)
Elevated	44 (59%)
Event During Follow-up	
No Event	39 (53%)
Event	35 (47%)

[1]Median (Range); n (%)

Characteristic	N = 74[1]
Vital Status	
Alive	69 (93%)
Deceased	5 (6.8%)
Total Follow up (months)	17 (2 - 84)

[1]Median (Range); n (%)

Hide

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

#Heatmap with Clinical & Genomics Factors

Hide

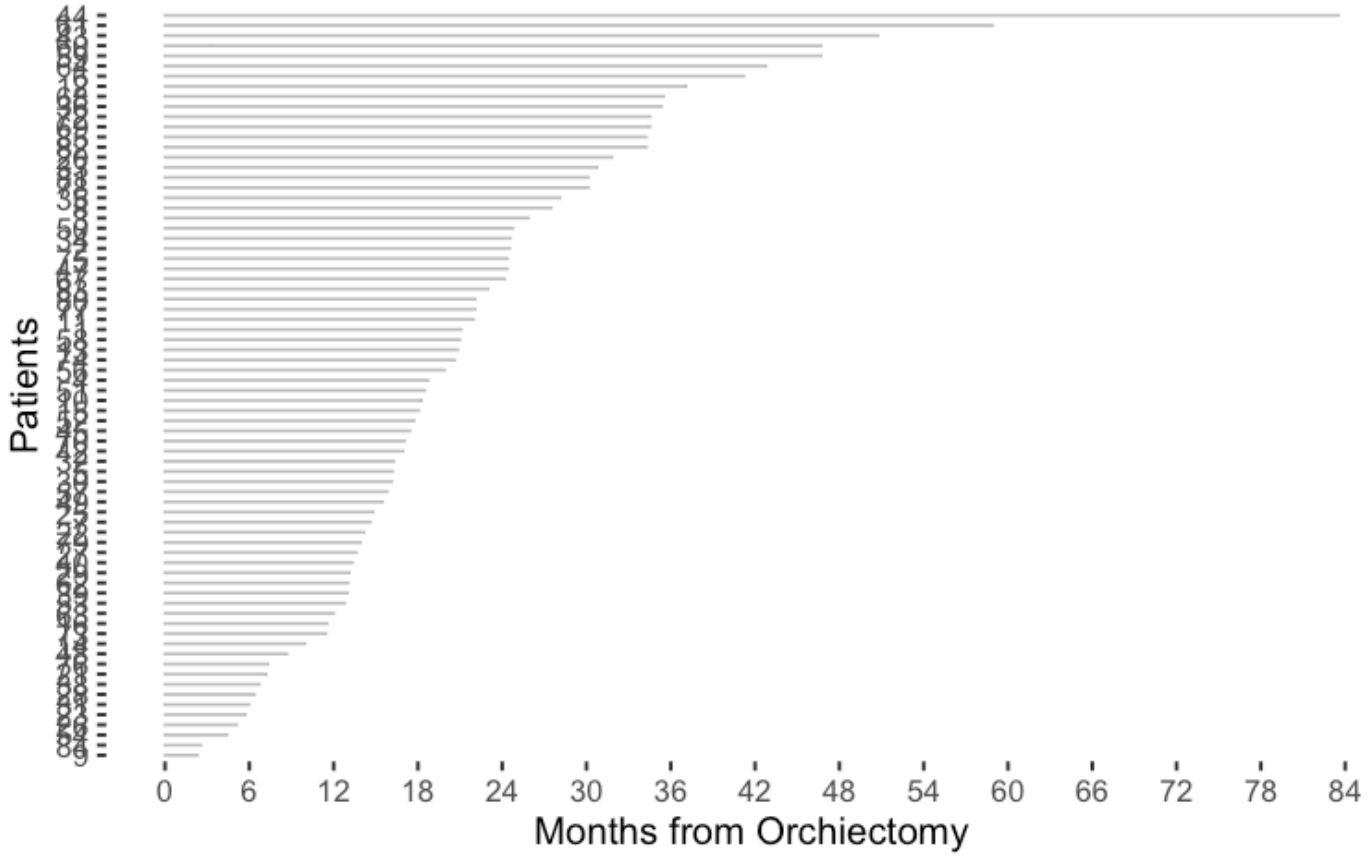
```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data %>% arrange(Stage)
circ_datadf <- as.data.frame(circ_data)

ha <- HeatmapAnnotation(
  Stage = circ_data$Stage,
  Histology = circ_data$Histology,
  pT = circ_data$pT,
  Treatment = circ_data$Treatment,
  Platinum.Response = circ_data$Platinum.Response,
  PFS.Event = circ_data$PFS.Event,
  preSx.markers.anytime = circ_data$preSx.markers.anytime,
  Markers.MRD.Window = circ_data$Markers.MRD.Window,
  Markers.Surveillance = circ_data$Markers.Surveillance,
  Markers.post.ACT = circ_data$Markers.post.ACT,
  Markers.post.RPLND = circ_data$Markers.post.RPLND,
  Markers.postSx.anytime = circ_data$Markers.postSx.anytime,
  Markers.anytime = circ_data$Markers.anytime,
  preSx.ctDNA.anytime = circ_data$preSx.ctDNA.anytime,
  ctDNA.MRD.Window = circ_data$ctDNA.MRD.Window,
  ctDNA.surveillance = circ_data$ctDNA.surveillance,
  ctDNA.post.ACT = circ_data$ctDNA.post.ACT,
  ctDNA.post.RPLND = circ_data$ctDNA.post.RPLND,
  postSx.ctDNA.anytime = circ_data$postSx.ctDNA.anytime,
  ctDNA.anytime = circ_data$ctDNA.anytime,

  col = list(Stage = c("I" = "seagreen1", "II" = "orange", "III" = "purple"),
    Histology = c("Non seminoma" = "goldenrod" , "Seminoma" = "blue4"),
    pT = c("0" = "coral","1" = "coral", "2" ="darkgreen", "3" = "yellow3", "X" = "gra
y"),
    Treatment = c("Surveillance" = "lightblue", "Chemotherapy" = "orange2", "RPLND" ="kh
aki", "RPLND, Chemotherapy" = "brown2"),
    Platinum.Response = c("SD/PD" = "red3", "PR/CR" ="blue"),
    PFS.Event = c("TRUE" = "red3", "FALSE" ="blue"),
    preSx.markers.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.MRD.Window = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.Surveillance = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.post.ACT = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.post.RPLND = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.postSx.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    Markers.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    preSx.ctDNA.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    ctDNA.MRD.Window = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    ctDNA.surveillance = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    ctDNA.post.ACT = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    ctDNA.post.RPLND = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    postSx.ctDNA.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue"),
    ctDNA.anytime = c("POSITIVE" = "red3", "NEGATIVE" ="blue")
)
```

```
)
ht <- Heatmap(matrix(nrow = 0, ncol = length(circ_data$Stage)),show_row_names = FALSE,cl
uster_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

Hide

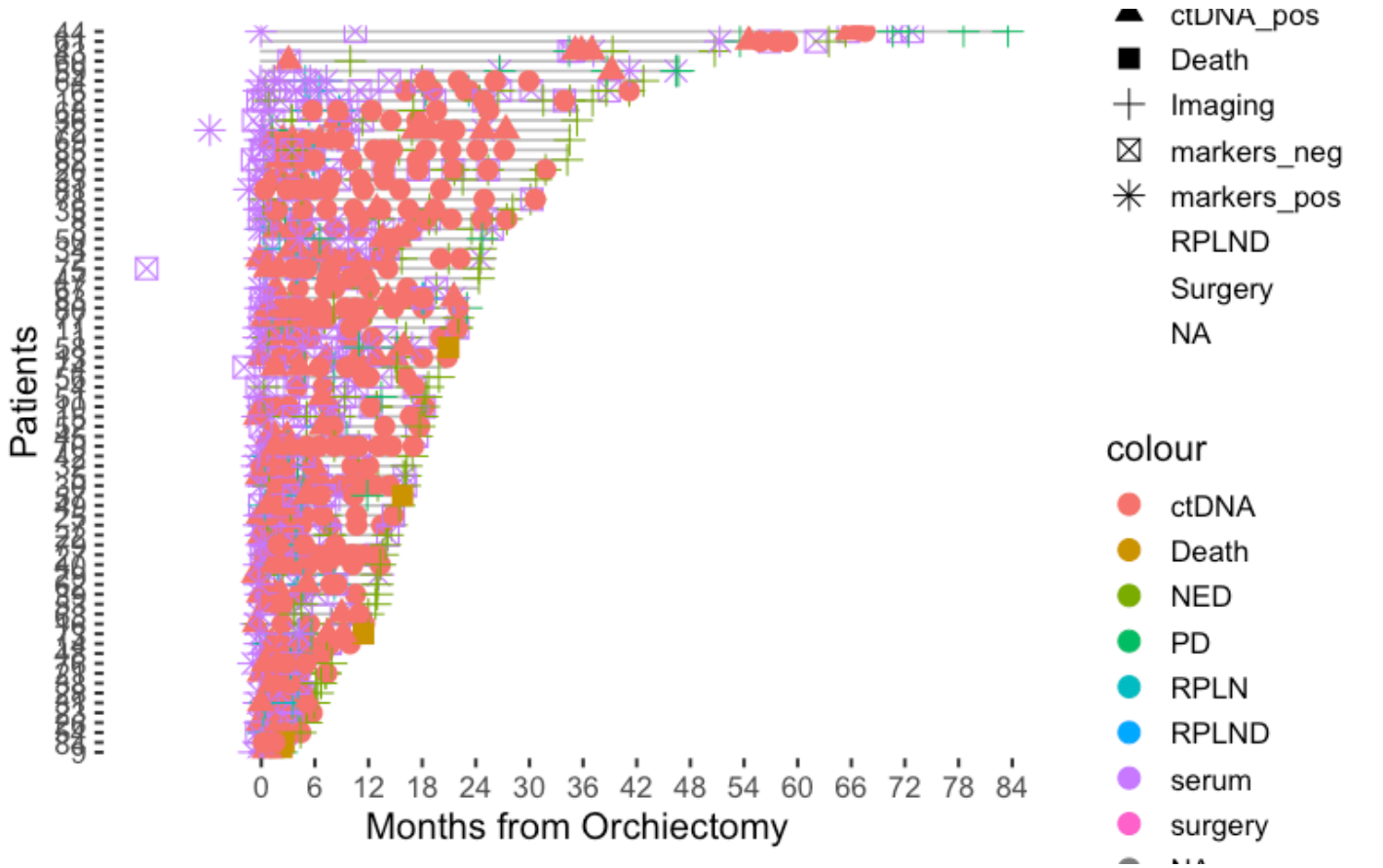
```
library(swimplot)
library(ggplot2)
library(grid)
library(gtable)

#Overview plot
rm(list=ls())
setwd("~/Downloads")
clinstage<- read.csv("CLIA Testicular_OP.csv")
clinstage_df<- as.data.frame(clinstage)

oplot_stratify <-swimmer_plot(df=clinstage_df,
                              id='PatientName',
                              end='fu.diff.months',
                              #name_fill='Arm',
                              col="gray",
                              alpha=0.75,
                              width=.01,
                              base_size = 14,
                              )
oplot_stratify <- oplot_stratify + theme(panel.border = element_blank())
oplot_stratify <- oplot_stratify + scale_y_continuous(breaks = seq(0, 108, by = 6))
oplot_stratify <- oplot_stratify + labs(x ="Patients" , y="Months from Orchiectomy")
oplot_stratify
```



Hide

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

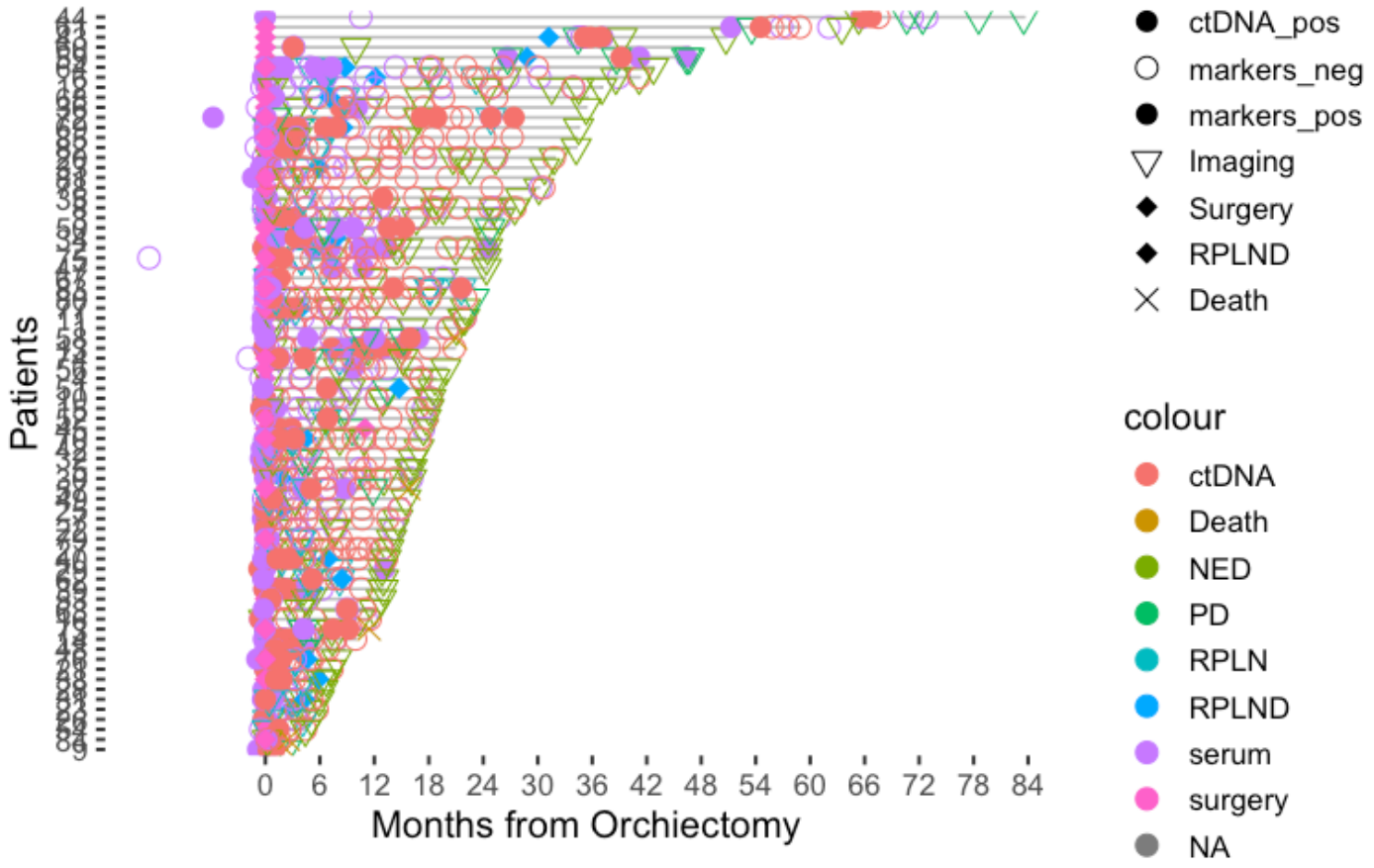
```
Warning: The shape palette can deal with a maximum of 6 discrete values because more tha
n 6 becomes difficult to discriminate
ℹ you have requested 8 values. Consider specifying shapes manually if you need that many
have them.
Warning: Removed 178 rows containing missing values or values outside the scale range (`
geom_point()`).
```



Hide

```
#Shape customization to Event_type

oplot_ev3.1 <- oplot_ev3 + ggplot2::scale_shape_manual(name="Event_type",values=c(1,16,
1,16,6,18,18,4),breaks=c('ctDNA_neg','ctDNA_pos','markers_neg','markers_pos','Imagin
g','Surgery','RPLND', 'Death'))

oplot_ev3.1
```

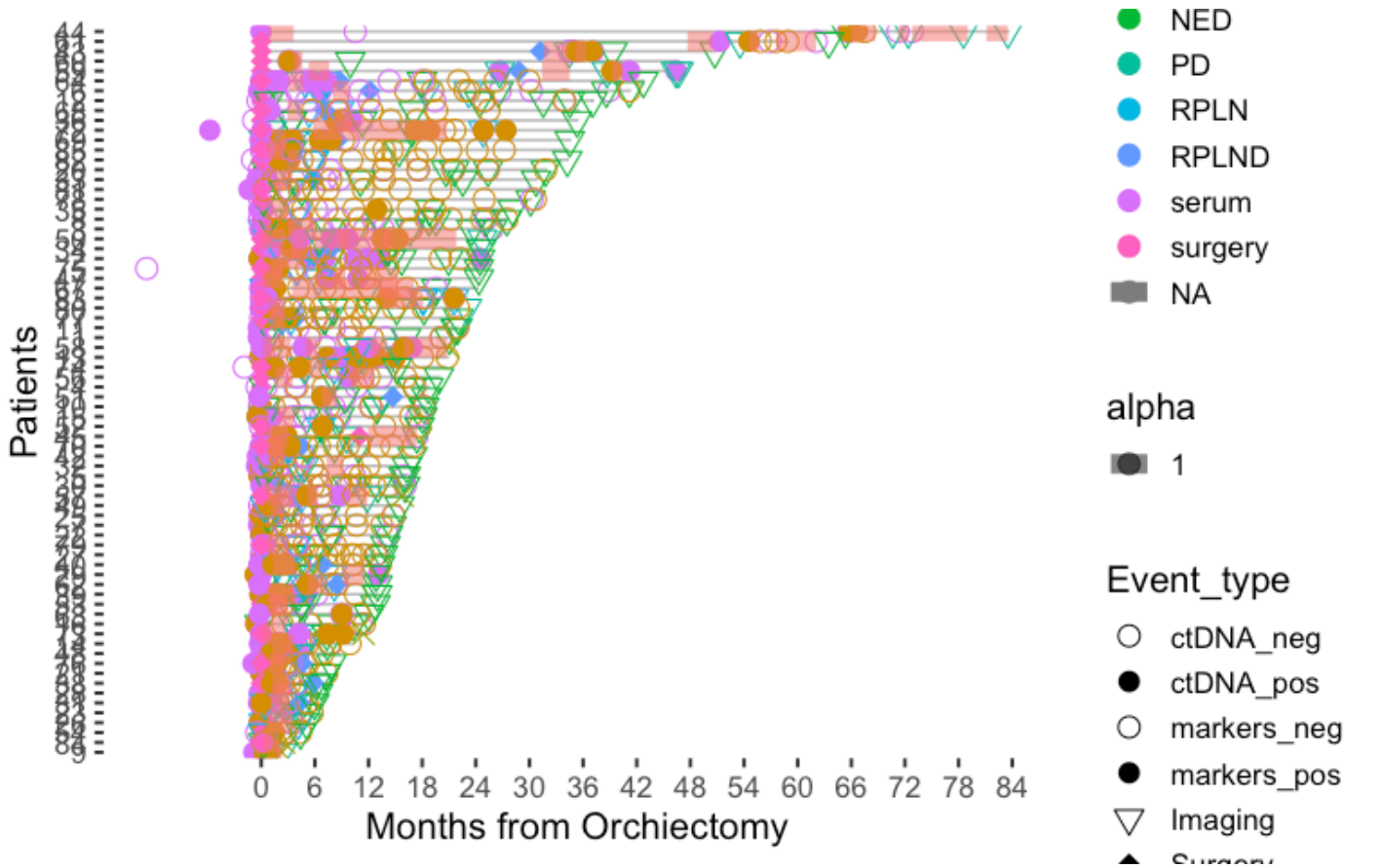
```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
```



Hide

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

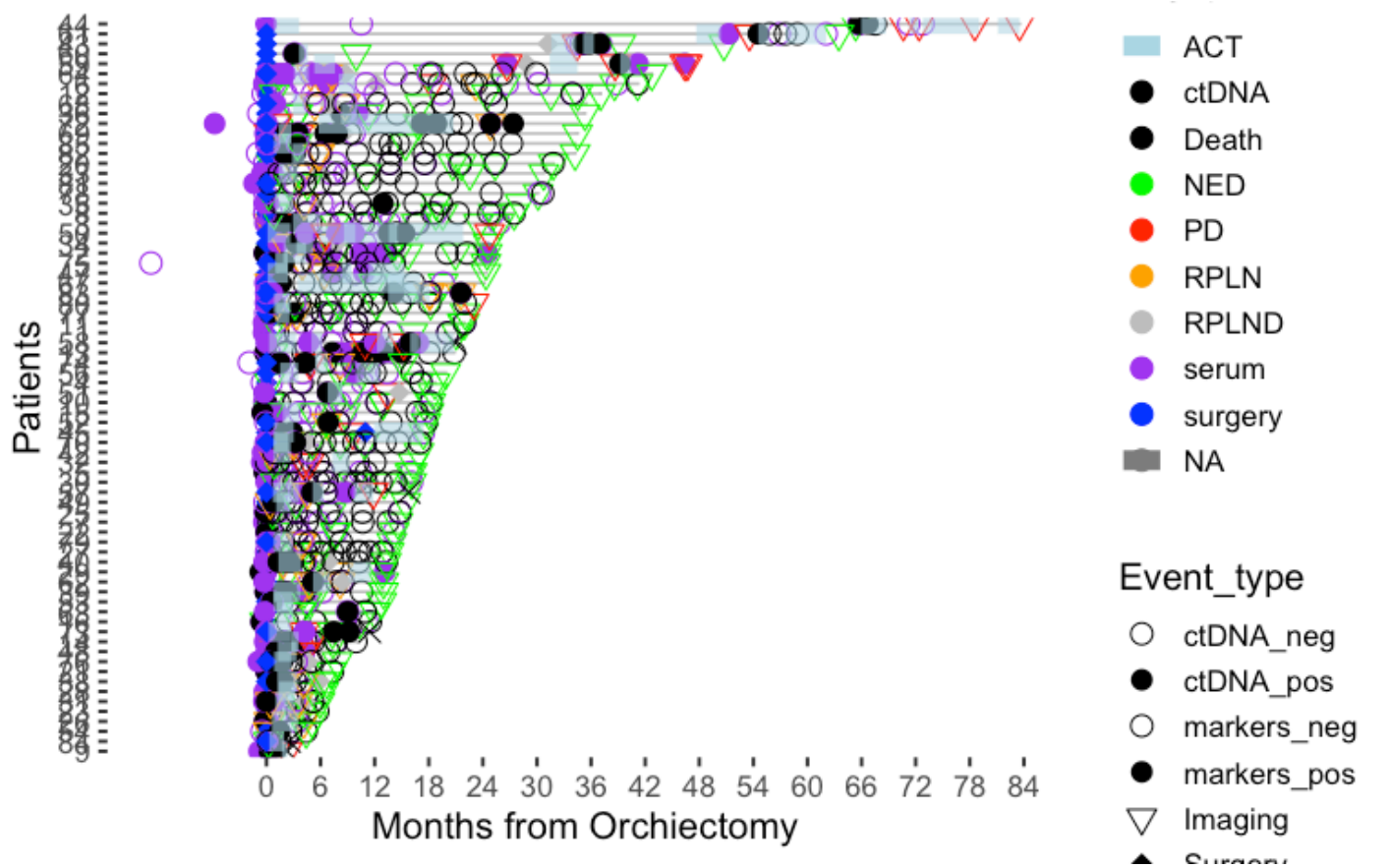
```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
Warning: Removed 997 rows containing missing values or values outside the scale range (`
geom_segment()`).
```



Hide

```
#colour customization
oplot_ev4.2 <- oplot_ev4 + ggplot2::scale_color_manual(name="Event",values=c( "lightblu
e", "black", "black", "green", "red", "orange", "grey", "purple", "blue"))
oplot_ev4.2
```

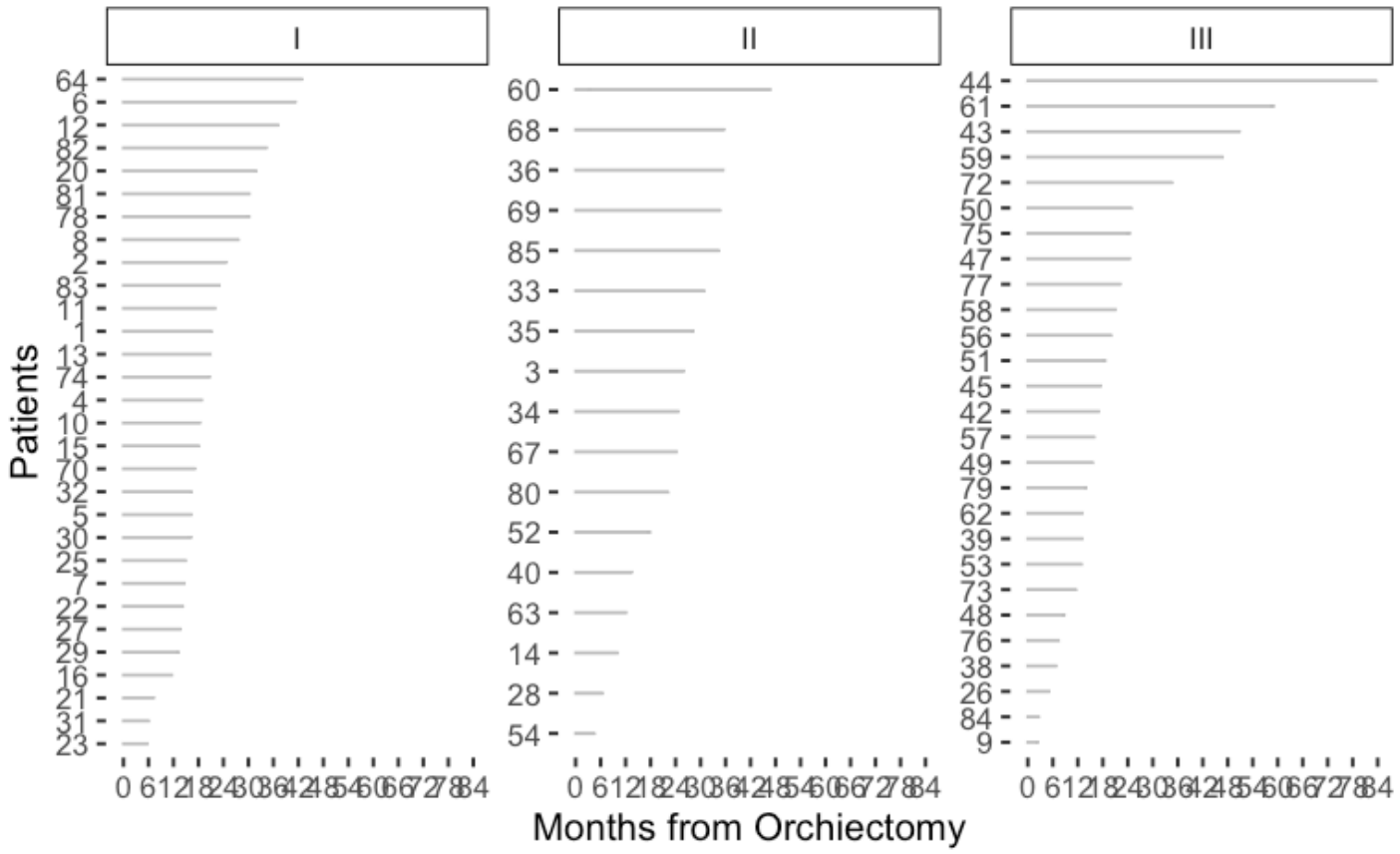
```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
Warning: Removed 997 rows containing missing values or values outside the scale range (`
geom_segment()`).
```



#Overview plot Stratified by Stage

Hide

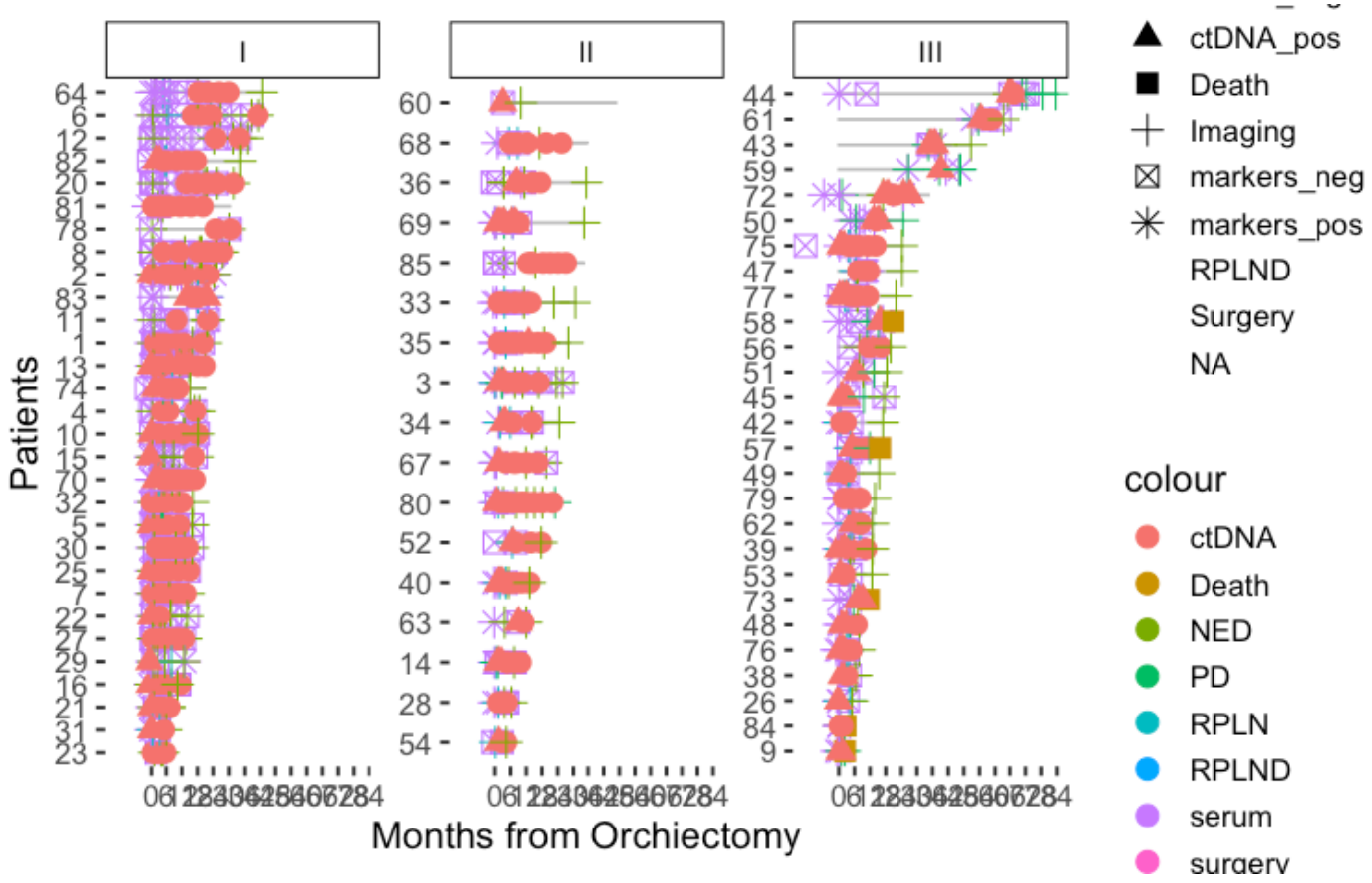
```
library(swimplot)
library(ggplot2)
library(grid)
library(gtable)

#Overview plot
rm(list=ls())
setwd("~/Downloads")
clinstage<- read.csv("CLIA Testicular_OP.csv")
clinstage_df<- as.data.frame(clinstage)

oplot_stratify <-swimmer_plot(df=clinstage_df,
                              id='PatientName',
                              end='fu.diff.months',
                              #name_fill='Arm',
                              col="gray",
                              alpha=0.75,
                              width=.01,
                              base_size = 14,
                              stratify= c('Stage'))
oplot_stratify <- oplot_stratify + theme(panel.border = element_blank())
oplot_stratify <- oplot_stratify + scale_y_continuous(breaks = seq(0, 108, by = 6))
oplot_stratify <- oplot_stratify + labs(x ="Patients" , y="Months from Orchiectomy")
oplot_stratify
```



Hide

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

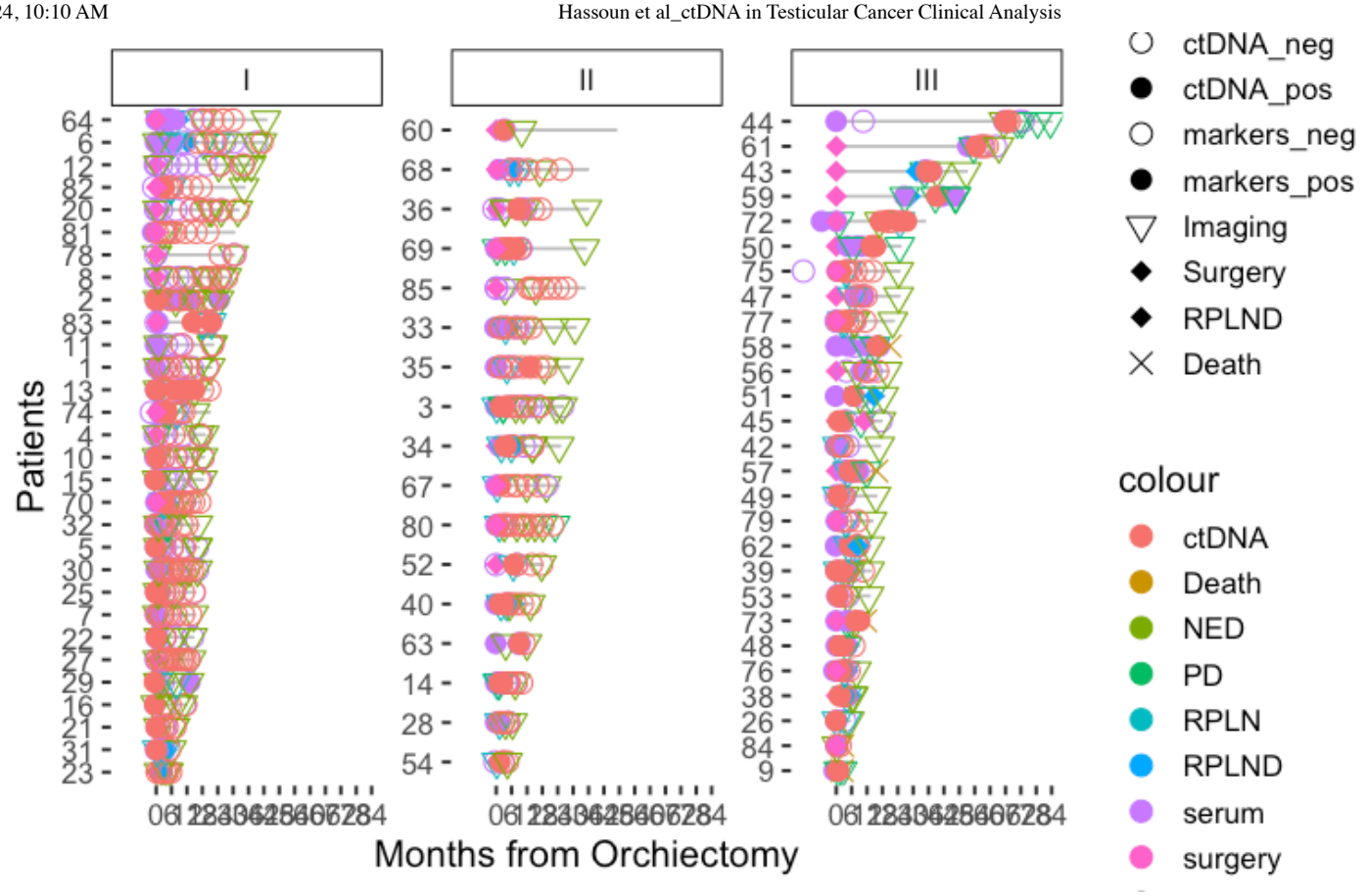
```
Warning: The shape palette can deal with a maximum of 6 discrete values because more tha
n 6 becomes difficult to discriminate
ℹ you have requested 8 values. Consider specifying shapes manually if you need that many
have them.
Warning: Removed 178 rows containing missing values or values outside the scale range (`
geom_point()`).
```



Hide

```
#Shape customization to Event_type

oplot_ev3.1 <- oplot_ev3 + ggplot2::scale_shape_manual(name="Event_type",values=c(1,16,
1,16,6,18,18,4),breaks=c('ctDNA_neg','ctDNA_pos','markers_neg','markers_pos','Imagin
g','Surgery','RPLND', 'Death'))

oplot_ev3.1
```

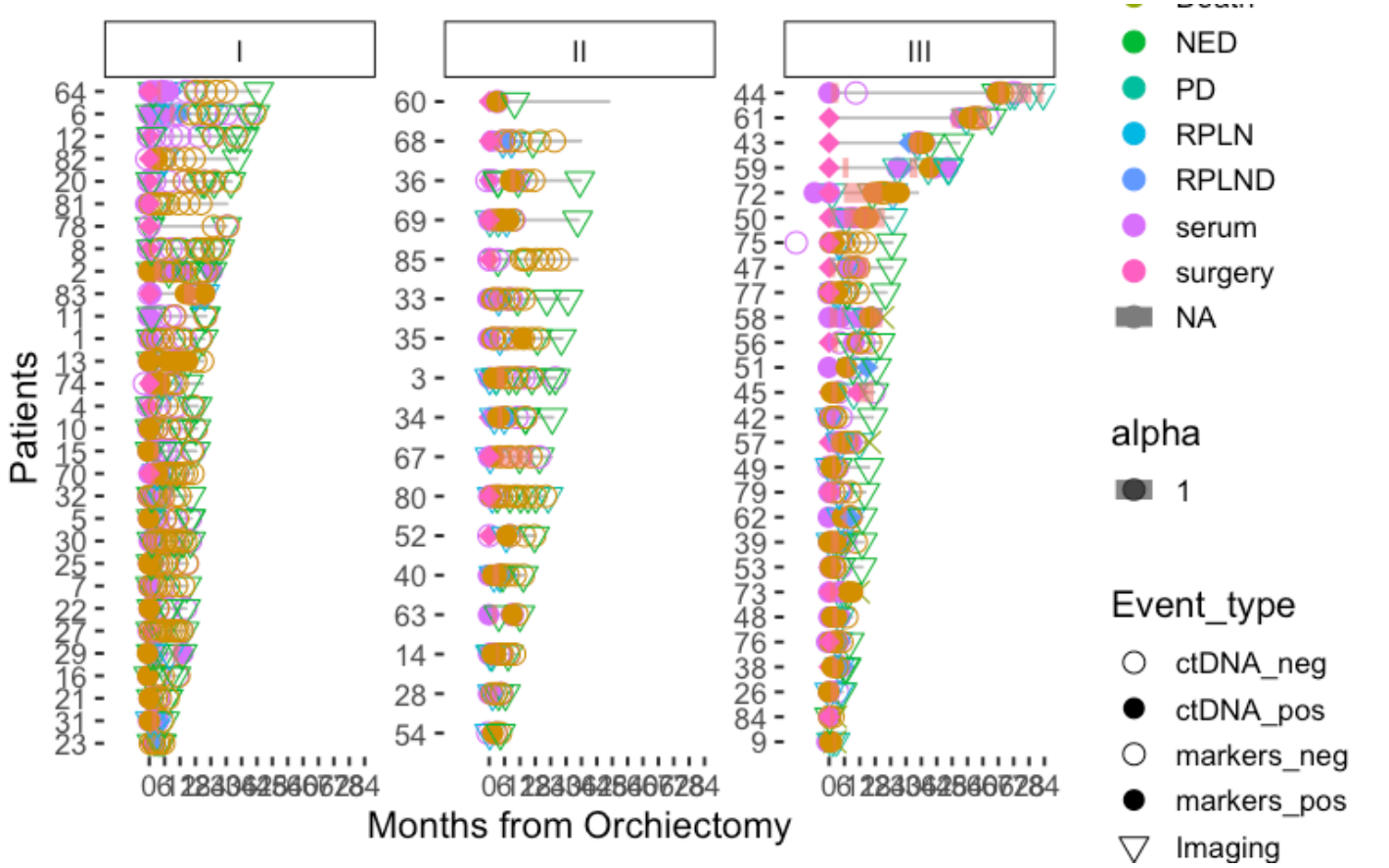
```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
```



Hide

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

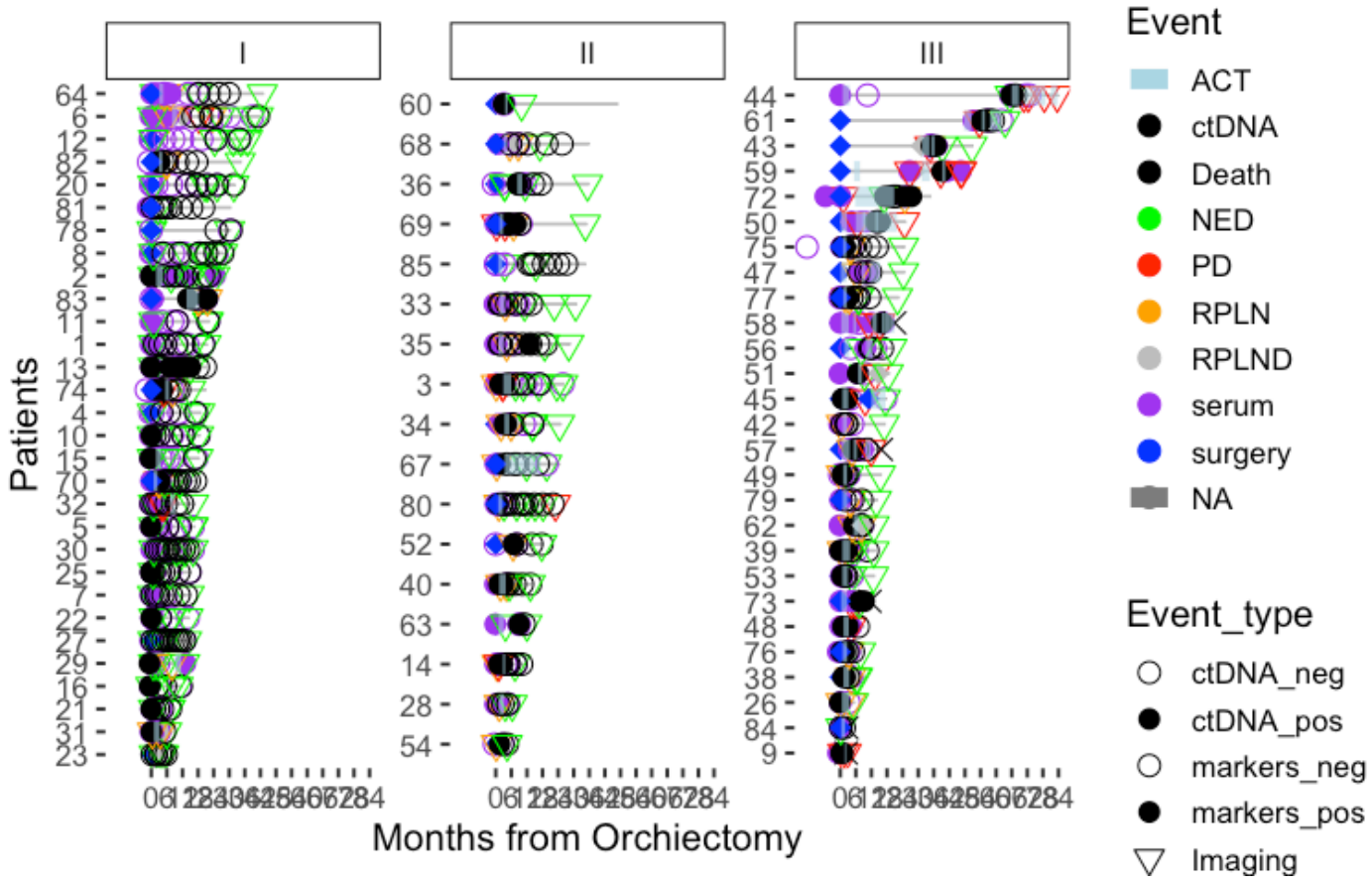
```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
Warning: Removed 997 rows containing missing values or values outside the scale range (`
geom_segment()`).
```



Hide

```
#colour customization
oplot_ev4.2 <- oplot_ev4 + ggplot2::scale_color_manual(name="Event",values=c( "lightblu
e", "black", "black", "green", "red", "orange", "grey", "purple", "blue"))
oplot_ev4.2
```

```
Warning: Removed 76 rows containing missing values or values outside the scale range (`g
eom_point()`).
Warning: Removed 997 rows containing missing values or values outside the scale range (`
geom_segment()`).
```



#Sankey Plots for all pts with post-orchiectomy timepoint available

Hide

```
##To run this commands, please visit: https://sankeymatic.com/build/
#No pre-orch [20] MRD positive #E67272
#No pre-orch [12] MRD negative #87EA86
#No pre-orch [27] no MRD timepoint #808080
#Pre-orch Positive [1] MRD positive #E67272
#Pre-orch Positive [8] MRD negative #87EA86
#Pre-orch Positive [5] no MRD timepoint #808080
#Pre-orch Negative [1] MRD negative #87EA86

#MRD positive [7] Surveillance positive #E67272
#MRD positive [7] Surveillance negative #87EA86
#MRD positive [7] no Surveillance timepoint #80808
#MRD negative [1] Surveillance positive #E67272
#MRD negative [18] Surveillance negative #87EA86
#MRD negative [2] no Surveillance timepoint #80808
#no MRD timepoint [5] Surveillance positive #E67272
#no MRD timepoint [13] Surveillance negative #87EA86
#no MRD timepoint [14] no Surveillance timepoint #808080

#Surveillance positive [13] Event #E67272
#Surveillance positive [0] No Event #87EA86
#Surveillance negative [5] Event #E67272
#Surveillance negative [33] No Event #87EA86
#no Surveillance timepoint [17] Event #E67272
#no Surveillance timepoint [6] No Event #87EA86
```

#EFS by ctDNA at the MRD Window

Hide

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

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

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

                           n events median 0.95LCL 0.95UCL
ctDNA.MRD.Window=NEGATIVE 21      3     NA      NA      NA
ctDNA.MRD.Window=POSITIVE 21     13   5.36    4.07      NA
```

Hide

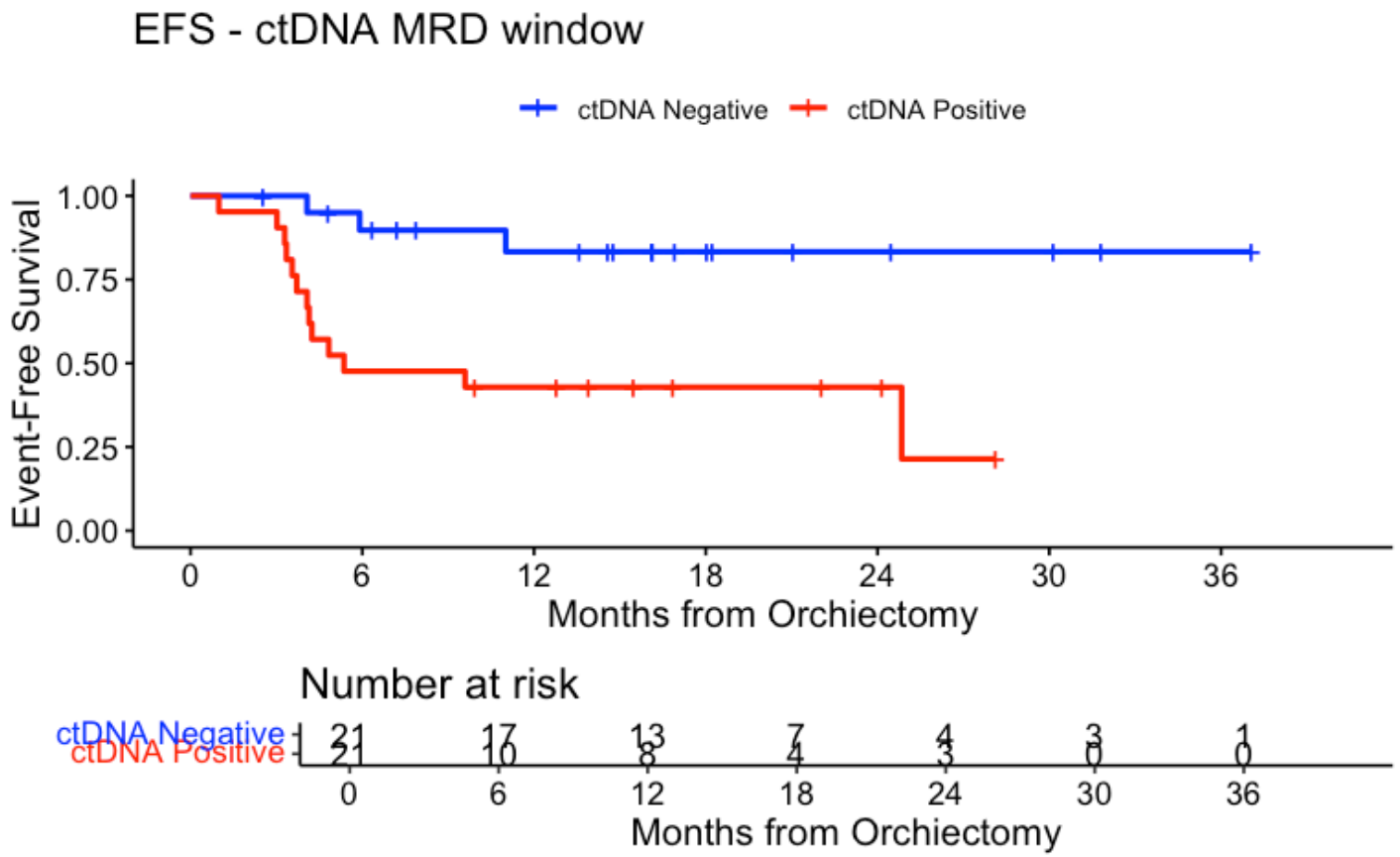
```
event_summary <- circ_data %>%
  group_by(ctDNA.MRD.Window) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.MRD.Window <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	21	3	0.1428571	14.28571
POSITIVE	21	13	0.6190476	61.90476

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.MRD.Window, 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="EFS - ctDNA MRD window", ylab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



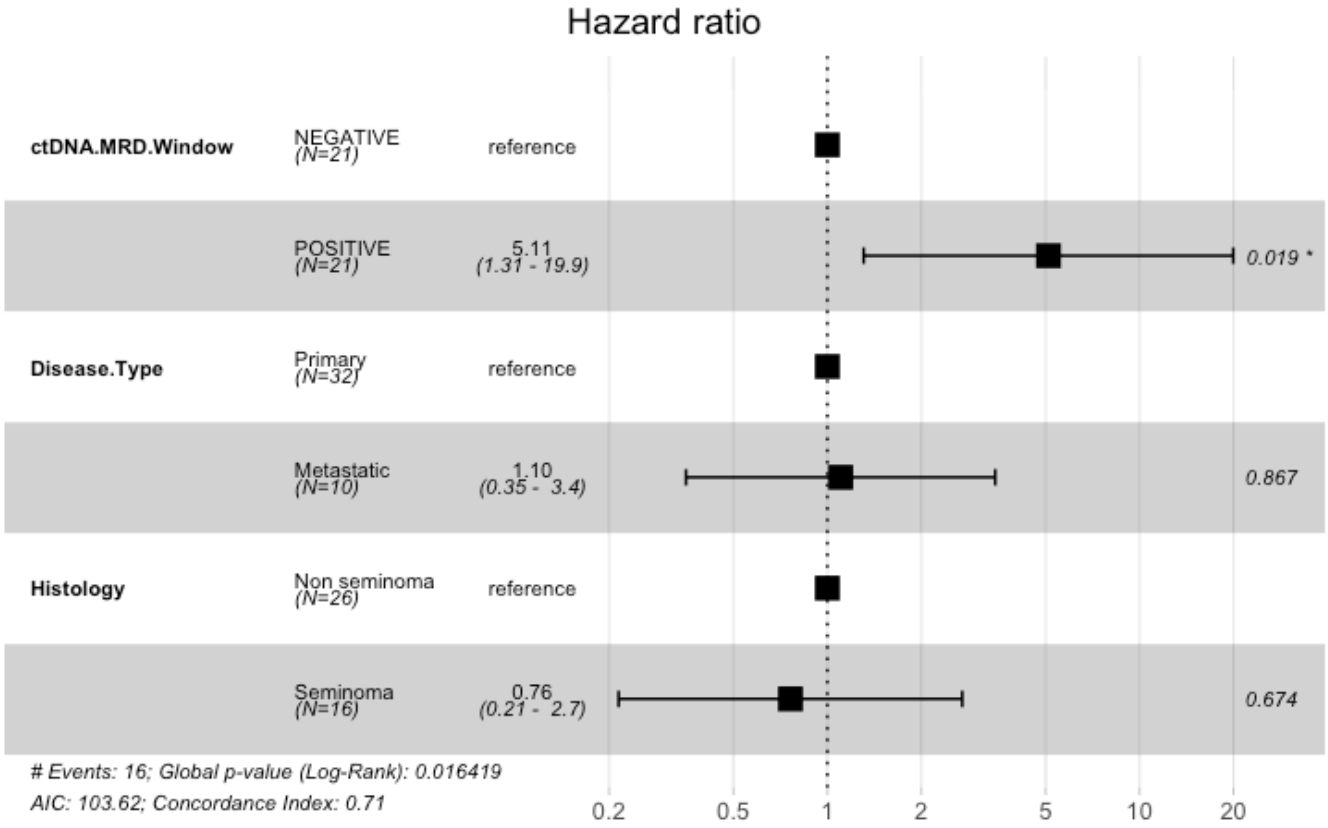
Hide

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

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

                ctDNA.MRD.Window=NEGATIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   12     13       3    0.833  0.0889        0.563        0.944
   24      4       0    0.833  0.0889        0.563        0.944

                ctDNA.MRD.Window=POSITIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   12      8      12    0.429   0.108        0.219        0.623
   24      3       0    0.429   0.108        0.219        0.623
```

Hide

```
circ_data$ctDNA.MRD.Window <- factor(circ_data$ctDNA.MRD.Window, levels=c("NEGATIVE","POSITIVE"))
circ_data$Disease.Type <- factor(circ_data$Disease.Type, levels = c("Primary", "Metastatic"))
circ_data$Histology <- factor(circ_data$Histology, levels = c("Non seminoma", "Seminoma"), labels = c("Non seminoma", "Seminoma"))
cox_fit <- coxph(surv_object ~ ctDNA.MRD.Window + Disease.Type + Histology, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ ctDNA.MRD.Window + Disease.Type +
    Histology, data = circ_data)

  n= 42, number of events= 16

                             coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.MRD.WindowPOSITIVE  1.63048   5.10634  0.69515  2.346    0.019 *
Disease.TypeMetastatic    0.09773   1.10267  0.58170  0.168    0.867
HistologySeminoma        -0.27236   0.76158  0.64651 -0.421    0.674
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                         exp(coef) exp(-coef) lower .95 upper .95
ctDNA.MRD.WindowPOSITIVE    5.1063     0.1958    1.3074    19.945
Disease.TypeMetastatic      1.1027     0.9069    0.3526     3.448
HistologySeminoma           0.7616     1.3131    0.2145     2.704

Concordance= 0.707  (se = 0.053 )
Likelihood ratio test= 10.27  on 3 df,   p=0.02
Wald test            = 7.8  on 3 df,   p=0.05
Score (logrank) test = 9.96  on 3 df,   p=0.02
```

#EFS by STMs at the MRD Window

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.MRD.Window!="",]
circ_data <- circ_data[circ_data$Markers.MRD.Window!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~Markers.MRD.Windo
w, data = circ_data)
```

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

                              n events median 0.95LCL 0.95UCL
Markers.MRD.Window=NEGATIVE 17      3     NA      NA      NA
Markers.MRD.Window=POSITIVE 21     11   24.8    4.14      NA
```

Hide

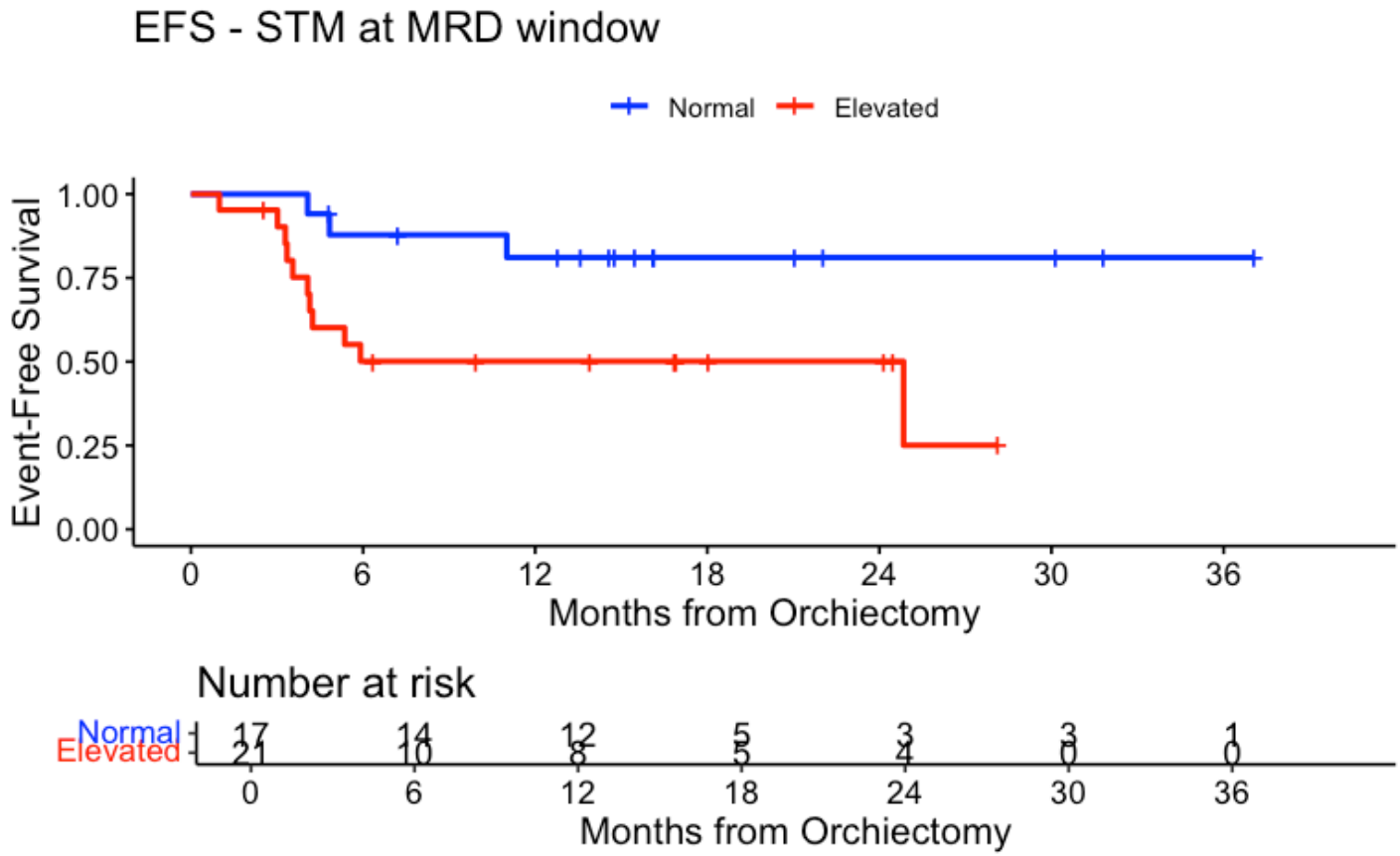
```
event_summary <- circ_data %>%
  group_by(Markers.MRD.Window) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

Markers.MRD.Window <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	17	3	0.1764706	17.64706
POSITIVE	21	11	0.5238095	52.38095

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ Markers.MRD.Window, 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=6, palette=c("blue","red"), title="EFS - STM at MRD window", ylab= "Eve
nt-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("Normal", "Elevated"),
legend.title="")
```



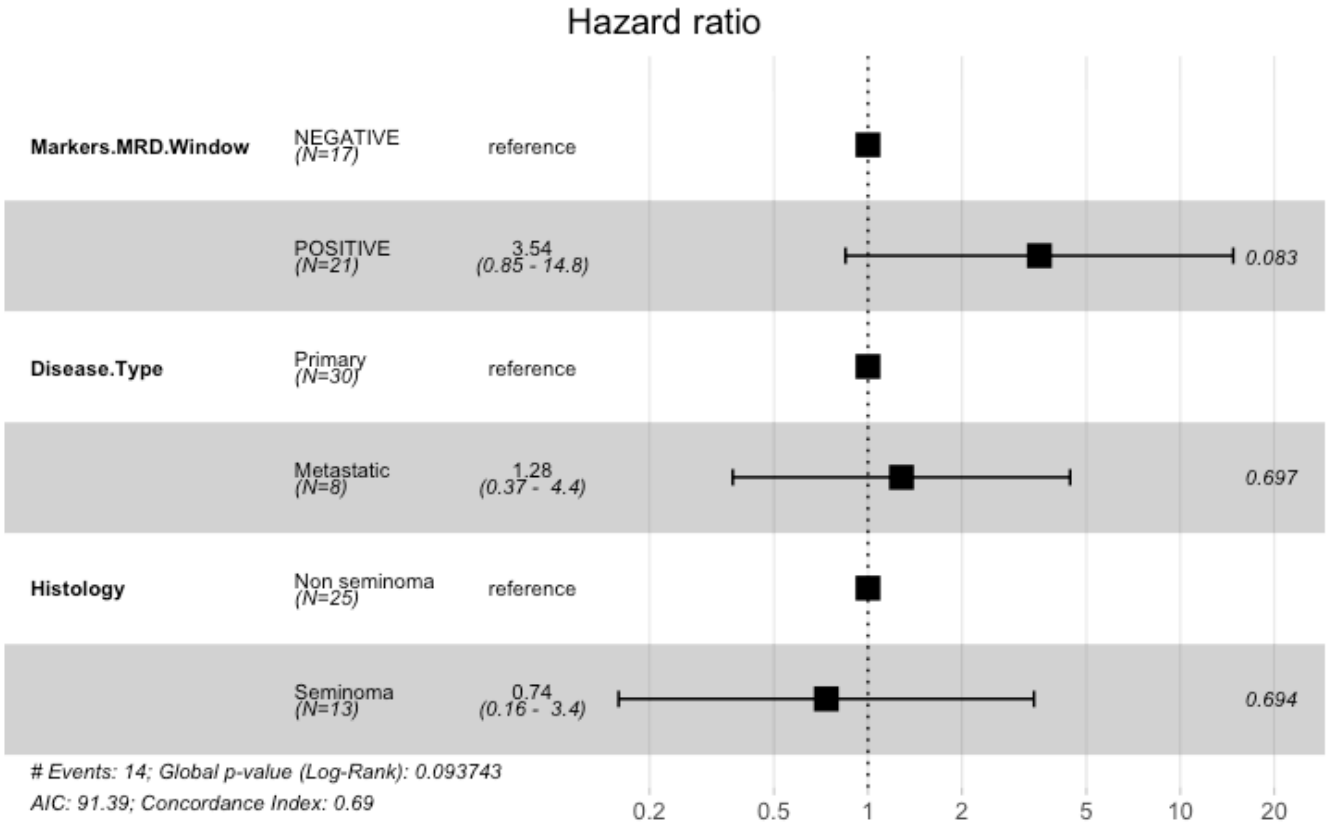
Hide

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

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

                Markers.MRD.Window=NEGATIVE
        time        n.risk       n.event      survival       std.err lower 95% CI upper 95%
CI
     12.0000      12.0000       3.0000       0.8109       0.0988       0.5194       0.93
51

                Markers.MRD.Window=POSITIVE
        time        n.risk       n.event      survival       std.err lower 95% CI upper 95%
CI
      12.000        8.000       10.000        0.501        0.112        0.272        0.6
93
```

Hide

```
circ_data$Markers.MRD.Window <- factor(circ_data$Markers.MRD.Window, levels=c("NEGATIV
E","POSITIVE"))
circ_data$Disease.Type <- factor(circ_data$Disease.Type, levels = c("Primary", "Metastat
ic"))
circ_data$Histology <- factor(circ_data$Histology, levels = c("Non seminoma", "Seminom
a"), labels = c("Non seminoma", "Seminoma"))
cox_fit <- coxph(surv_object ~ Markers.MRD.Window + Disease.Type + Histology, data=circ_
data)
ggforest(cox_fit,data = circ_data)
```



Hide

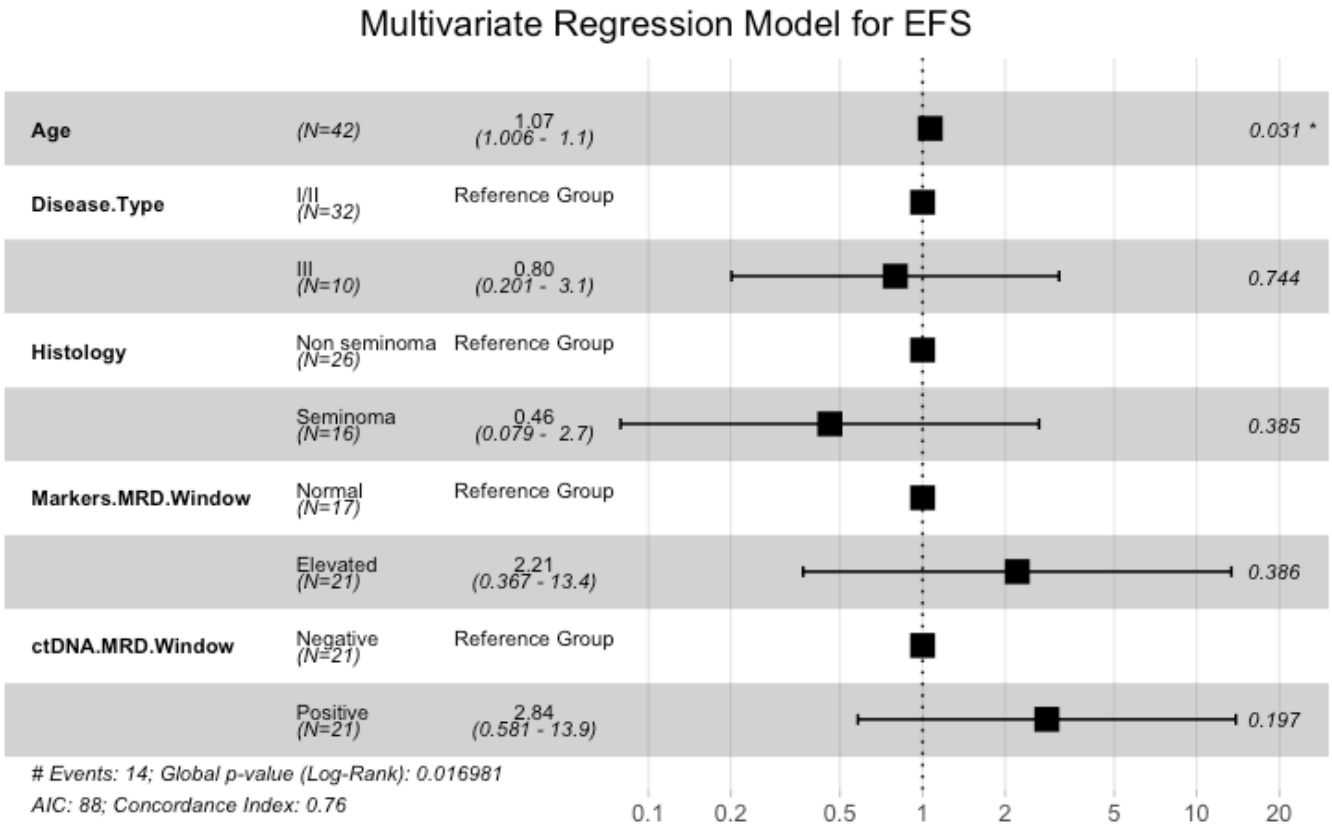
```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ Markers.MRD.Window + Disease.Type +
    Histology, data = circ_data)

  n= 38, number of events= 14

                             coef exp(coef) se(coef)      z Pr(>|z|)
Markers.MRD.WindowPOSITIVE  1.2635    3.5379   0.7290  1.733   0.0831 .
Disease.TypeMetastatic      0.2467    1.2798   0.6342  0.389   0.6973
HistologySeminoma          -0.3072    0.7355   0.7812 -0.393   0.6942
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                           exp(coef) exp(-coef) lower .95 upper .95
Markers.MRD.WindowPOSITIVE    3.5379     0.2827    0.8476    14.767
Disease.TypeMetastatic        1.2798     0.7814    0.3692     4.436
HistologySeminoma             0.7355     1.3596    0.1591     3.401

Concordance= 0.69  (se = 0.06 )
Likelihood ratio test= 6.4  on 3 df,   p=0.09
Wald test            = 5.2  on 3 df,   p=0.2
Score (logrank) test = 6.11  on 3 df,   p=0.1
```

#Multivariate regression model for EFS at the MRD Window

Hide

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

circ_datadf$Disease.Type <- factor(circ_datadf$Disease.Type, levels = c("Primary", "Meta
static"), labels = c("I/II", "III"))
circ_datadf$Histology <- factor(circ_datadf$Histology, levels = c("Non seminoma", "Semin
oma"), labels = c("Non seminoma", "Seminoma"))
circ_datadf$Markers.MRD.Window <- factor(circ_datadf$Markers.MRD.Window, levels=c("NEGAT
IVE","POSITIVE"), labels = c("Normal", "Elevated"))
circ_datadf$ctDNA.MRD.Window <- factor(circ_datadf$ctDNA.MRD.Window, levels=c("NEGATIV
E","POSITIVE"), labels = c("Negative", "Positive"))
surv_object<-Surv(time = circ_datadf$PFS.months, event = circ_datadf$PFS.Event)
cox_fit <- coxph(surv_object ~ Age + Disease.Type + Histology + Markers.MRD.Window + ctD
NA.MRD.Window, data=circ_datadf)
ggforest(cox_fit, data = circ_datadf, main = "Multivariate Regression Model for EFS", re
fLabel = "Reference Group")
```



Hide

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

#EFS by ctDNA at the Surveillance Window

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ctDNA.surveillance!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~ctDNA.surveillanc
e, data = circ_data)
```

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

                              n events median 0.95LCL 0.95UCL
ctDNA.surveillance=NEGATIVE 38      5     NA      NA      NA
ctDNA.surveillance=POSITIVE 13     13   4.83    3.35      NA
```

Hide

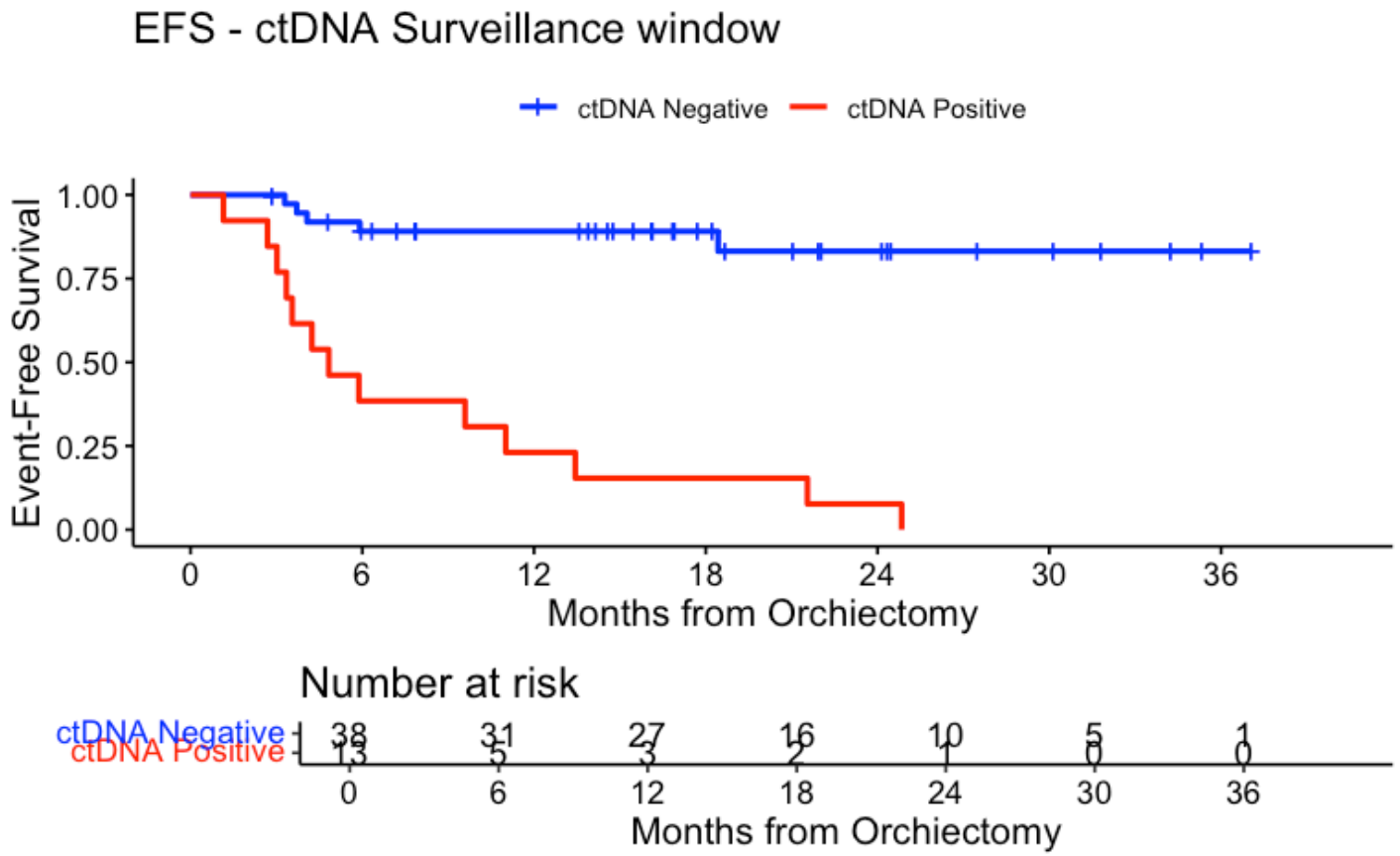
```
event_summary <- circ_data %>%
  group_by(ctDNA.surveillance) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.surveillance	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	38	5	0.1315789	13.15789
POSITIVE	13	13	1.0000000	100.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.surveillance, 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=6, palette=c("blue","red"), title="EFS - ctDNA Surveillance window", yl
ab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("ctDNA Negativ
e", "ctDNA Positive"), legend.title="")
```



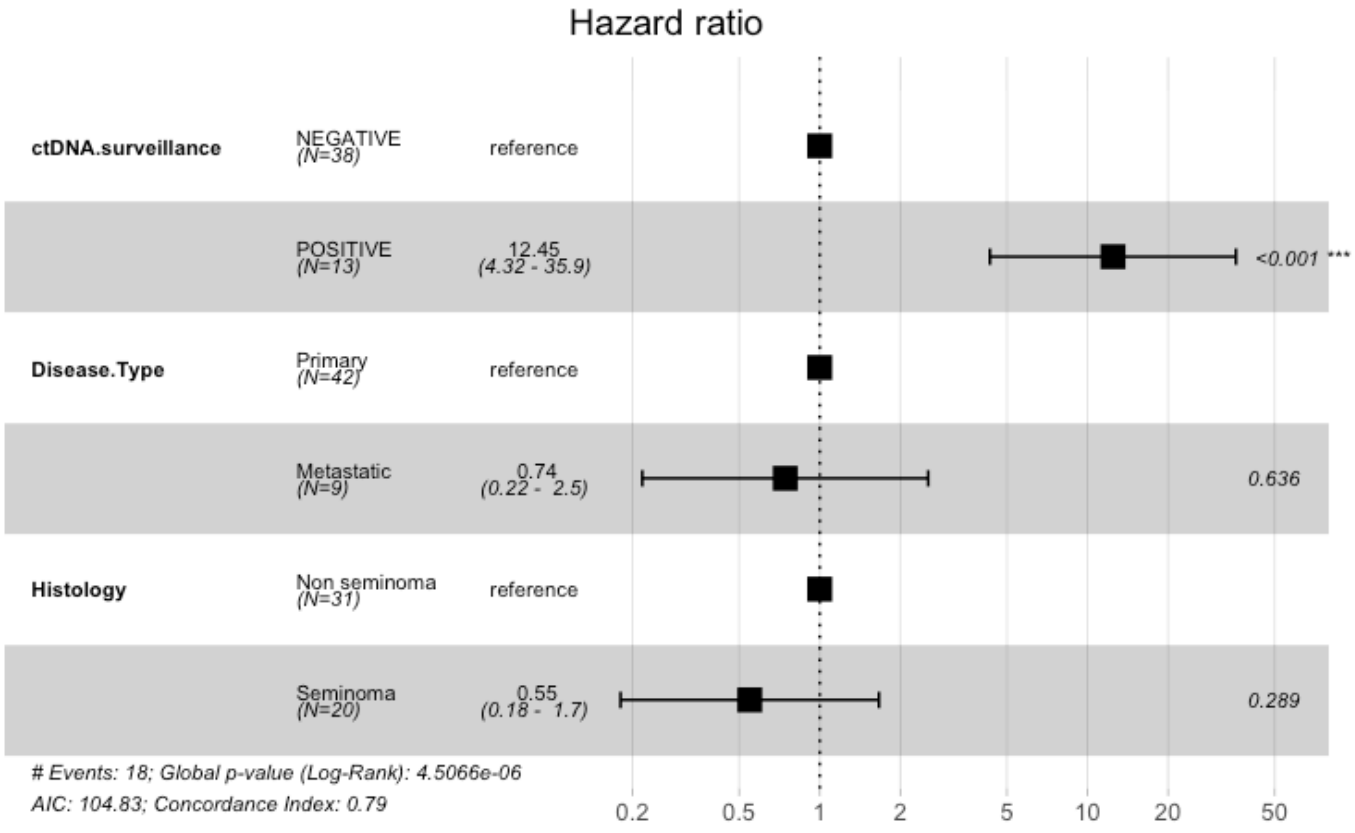
Hide

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

```
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
     12.0000      27.0000       4.0000       0.8911       0.0514       0.7352       0.95
77

                ctDNA.surveillance=POSITIVE
        time        n.risk      n.event     survival      std.err lower 95% CI upper 95%
CI
     12.0000       3.0000      10.0000       0.2308       0.1169       0.0558       0.47
46
```

Hide

```
circ_data$ctDNA.surveillance <- factor(circ_data$ctDNA.surveillance, levels=c("NEGATIV
E","POSITIVE"))
circ_data$Disease.Type <- factor(circ_data$Disease.Type, levels = c("Primary", "Metastat
ic"))
circ_data$Histology <- factor(circ_data$Histology, levels = c("Non seminoma", "Seminom
a"), labels = c("Non seminoma", "Seminoma"))
cox_fit <- coxph(surv_object ~ ctDNA.surveillance + Disease.Type + Histology, data=circ_
data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ ctDNA.surveillance + Disease.Type +
    Histology, data = circ_data)

  n= 51, number of events= 18

                             coef exp(coef) se(coef)      z Pr(>|z|)
ctDNA.surveillancePOSITIVE  2.5216   12.4485   0.5397  4.672 2.98e-06 ***
Disease.TypeMetastatic     -0.2967    0.7433   0.6270 -0.473    0.636
HistologySeminoma          -0.6017    0.5479   0.5670 -1.061    0.289
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                           exp(coef) exp(-coef) lower .95 upper .95
ctDNA.surveillancePOSITIVE   12.4485    0.08033    4.3222    35.854
Disease.TypeMetastatic        0.7433    1.34535    0.2175     2.540
HistologySeminoma             0.5479    1.82529    0.1803     1.665

Concordance= 0.793  (se = 0.069 )
Likelihood ratio test= 27.55  on 3 df,   p=5e-06
Wald test            = 23.95  on 3 df,   p=3e-05
Score (logrank) test = 38.6  on 3 df,   p=2e-08
```

#EFS by Serum Markers at the Surveillance Window

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$Markers.Surveillance!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~Markers.Surveilla
nce, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event) ~
    Markers.Surveillance, data = circ_data)

                               n events median 0.95LCL 0.95UCL
Markers.Surveillance=NEGATIVE 48     15   65.3      NA      NA
Markers.Surveillance=POSITIVE 14     11   18.4    7.82      NA
```

Hide

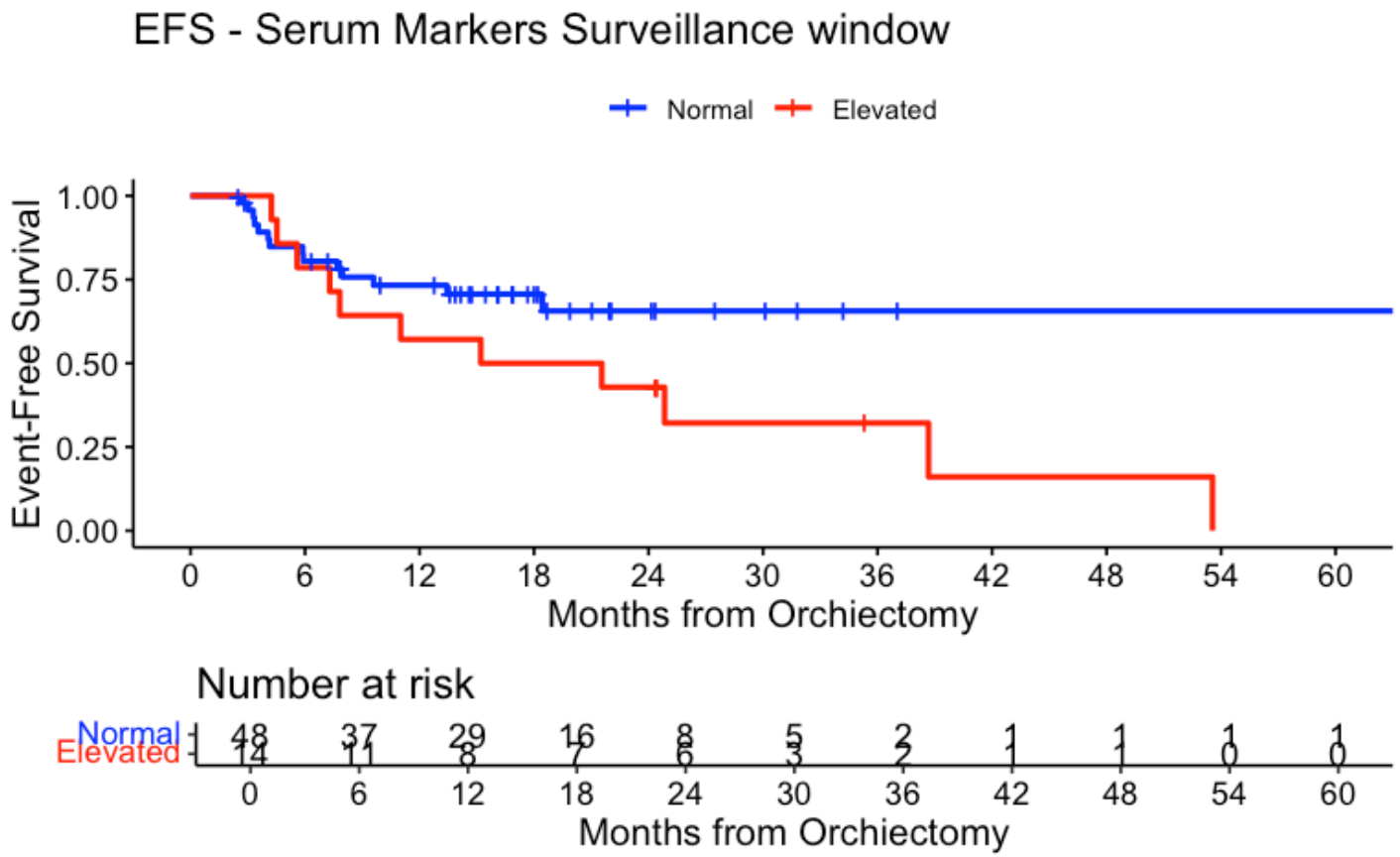
```
event_summary <- circ_data %>%
  group_by(Markers.Surveillance) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

Markers.Surveillance <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	48	15	0.3125000	31.25000
POSITIVE	14	11	0.7857143	78.57143

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ Markers.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=6, palette=c("blue","red"), title="EFS - Serum Markers Surveillance window", ylab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("Normal", "Elevated"), legend.title="")
```



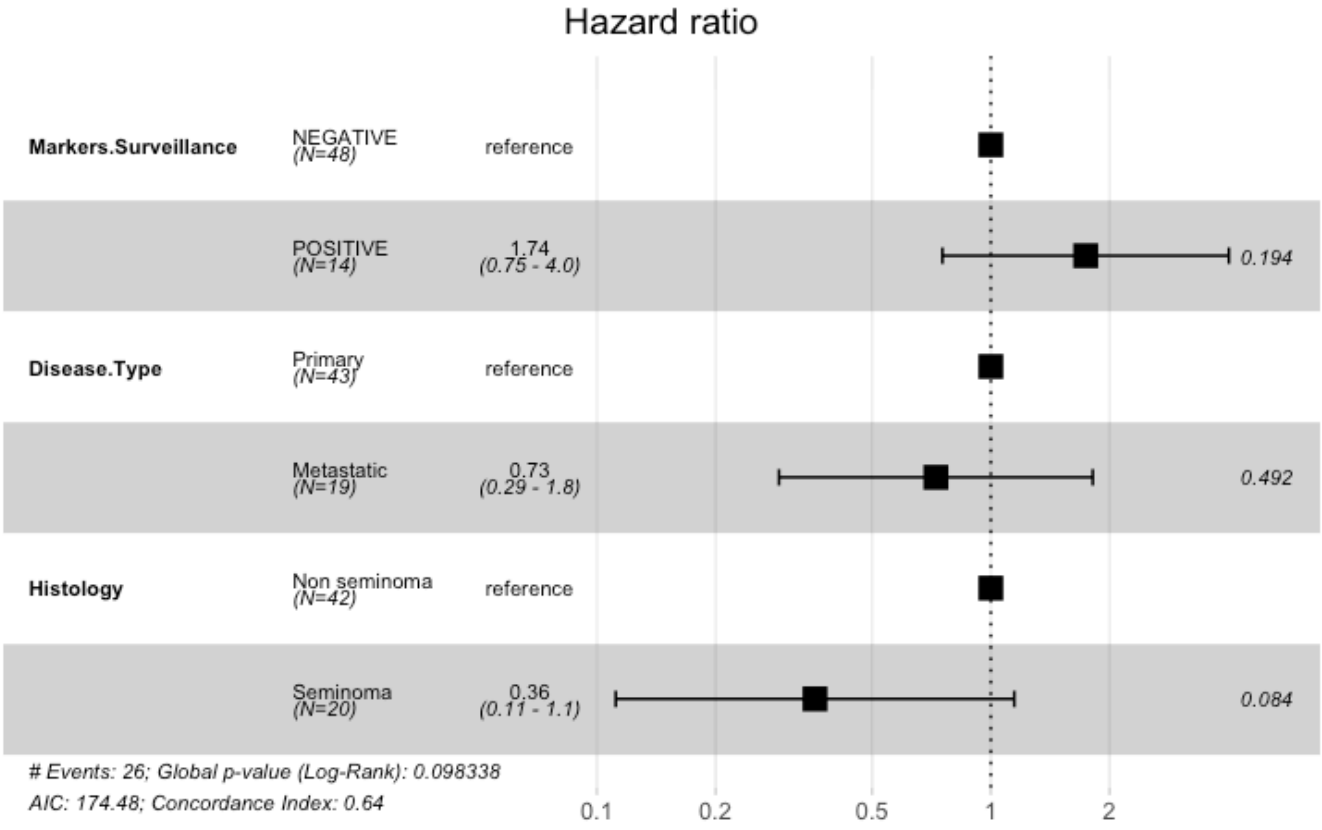
Hide

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

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

                Markers.Surveillance=NEGATIVE
        time        n.risk       n.event      survival       std.err lower 95% CI upper 95%
CI
     12.0000      29.0000      12.0000       0.7329       0.0663       0.5770       0.83
89

                Markers.Surveillance=POSITIVE
        time        n.risk       n.event      survival       std.err lower 95% CI upper 95%
CI
      12.000        8.000        6.000        0.571        0.132        0.284        0.7
80
```

Hide

```
circ_data$Markers.Surveillance <- factor(circ_data$Markers.Surveillance, levels=c("NEGAT
IVE","POSITIVE"))
circ_data$Disease.Type <- factor(circ_data$Disease.Type, levels = c("Primary", "Metastat
ic"))
circ_data$Histology <- factor(circ_data$Histology, levels = c("Non seminoma", "Seminom
a"), labels = c("Non seminoma", "Seminoma"))
cox_fit <- coxph(surv_object ~ Markers.Surveillance + Disease.Type + Histology, data=cir
c_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ Markers.Surveillance + Disease.Type +
    Histology, data = circ_data)

  n= 62, number of events= 26

                              coef exp(coef) se(coef)      z Pr(>|z|)
Markers.SurveillancePOSITIVE  0.5545    1.7410   0.4271  1.298   0.1942
Disease.TypeMetastatic       -0.3215    0.7251   0.4678 -0.687   0.4919
HistologySeminoma            -1.0279    0.3578   0.5944 -1.729   0.0838 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                            exp(coef) exp(-coef) lower .95 upper .95
Markers.SurveillancePOSITIVE    1.7410     0.5744    0.7538     4.021
Disease.TypeMetastatic          0.7251     1.3792    0.2899     1.814
HistologySeminoma               0.3578     2.7952    0.1116     1.147

Concordance= 0.638  (se = 0.05 )
Likelihood ratio test= 6.29  on 3 df,   p=0.1
Wald test            = 5.69  on 3 df,   p=0.1
Score (logrank) test = 6.19  on 3 df,   p=0.1
```

#EFS by ctDNA at post-RPLND - Stage II

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$RPLND==TRUE,]
circ_data <- circ_data[circ_data$Stage=="II",]
circ_data <- circ_data[circ_data$ctDNA.post.RPLND!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~ctDNA.post.RPLND,
data = circ_data)
```

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

                          n events median 0.95LCL 0.95UCL
ctDNA.post.RPLND=NEGATIVE 4      2   7.69    5.91      NA
ctDNA.post.RPLND=POSITIVE 2      2   1.92    1.15      NA
```

Hide

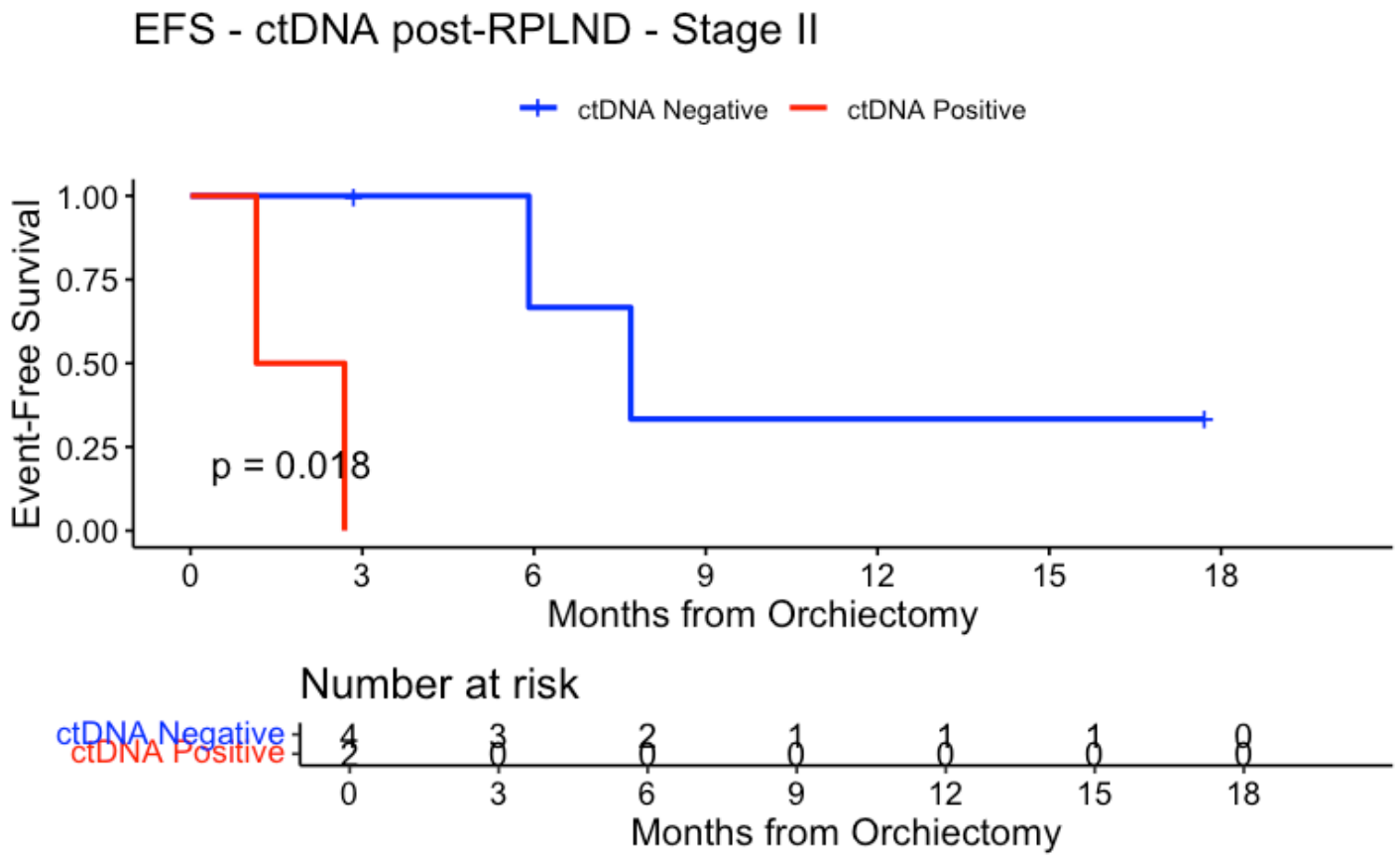
```
event_summary <- circ_data %>%
  group_by(ctDNA.post.RPLND) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.post.RPLND	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	4	2	0.5	50
POSITIVE	2	2	1.0	100

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.post.RPLND, data = circ_data,conf.int=0.95,conf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = TRUE, conf.int = FALSE, risk.table = TRUE, break.time.by=3, palette=c("blue","red"), title="EFS - ctDNA post-RPLND - Stage II", ylab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("ctDNA Negative", "ctDNA Positive"), legend.title="")
```



Hide

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

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

                ctDNA.post.RPLND=NEGATIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
    0      4       0    1.000   0.000      1.00000        1.000
   12      1       2    0.333   0.272      0.00896        0.774

                ctDNA.post.RPLND=POSITIVE
        time       n.risk      n.event     survival      std.err lower 95% CI upper 95%
CI
           0            2            0            1            0            1
1
```

Hide

```
circ_data$ctDNA.post.RPLND <- factor(circ_data$ctDNA.post.RPLND, levels=c("NEGATIVE","PO
SITIVE"))
cox_fit <- coxphf(surv_object ~ ctDNA.post.RPLND, data=circ_data)
summary(cox_fit)
```

```
coxphf(formula = surv_object ~ ctDNA.post.RPLND, 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.post.RPLNDPOSITIVE 2.715398 1.929072  15.11062   1.165602   2114.922 4.34548 0.037
10731

Likelihood ratio test=4.34548 on 1 df, p=0.03710731, n=6
Wald test = 1.981391 on 1 df, p = 0.159244

Covariance-Matrix:
                          ctDNA.post.RPLNDPOSITIVE
ctDNA.post.RPLNDPOSITIVE                 3.721317
```

#EFS by ctDNA at post-ACT - Stage II/III

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ACT==TRUE,]
circ_data <- circ_data[circ_data$Stage!="I",]
circ_data <- circ_data[circ_data$ctDNA.surveillance!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~ctDNA.surveillanc
e, data = circ_data)
```

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

                              n events median 0.95LCL 0.95UCL
ctDNA.surveillance=NEGATIVE 13      3     NA      NA      NA
ctDNA.surveillance=POSITIVE  9      9   4.24    3.35      NA
```

Hide

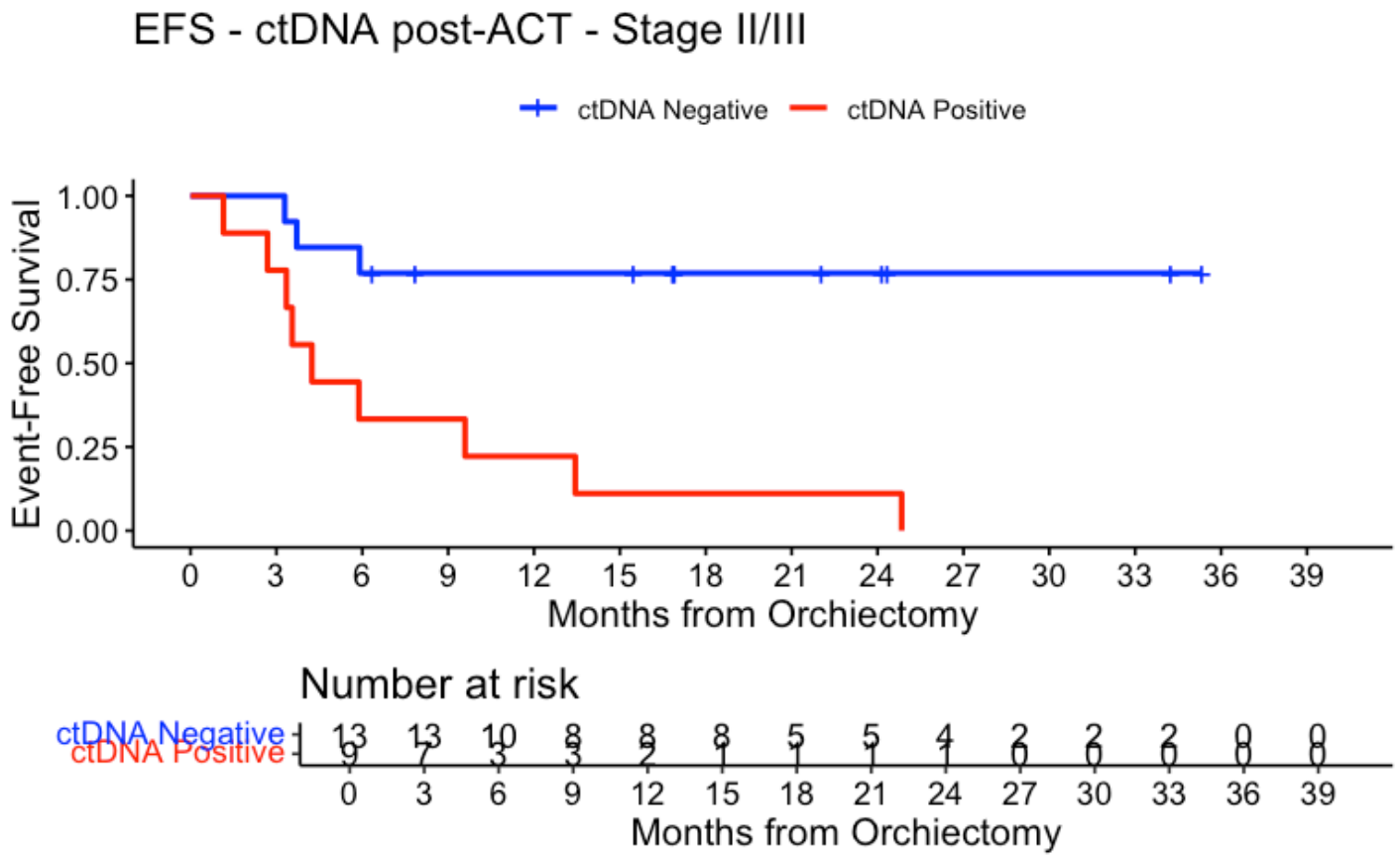
```
event_summary <- circ_data %>%
  group_by(ctDNA.surveillance) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

ctDNA.surveillance	Total	Events	Fraction	Percentage
<chr>	<int>	<int>	<dbl>	<dbl>
NEGATIVE	13	3	0.2307692	23.07692
POSITIVE	9	9	1.0000000	100.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ ctDNA.surveillance, 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=3, palette=c("blue","red"), title="EFS - ctDNA post-ACT - Stage II/II
I", ylab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("ctDNA Ne
gative", "ctDNA Positive"), legend.title="")
```



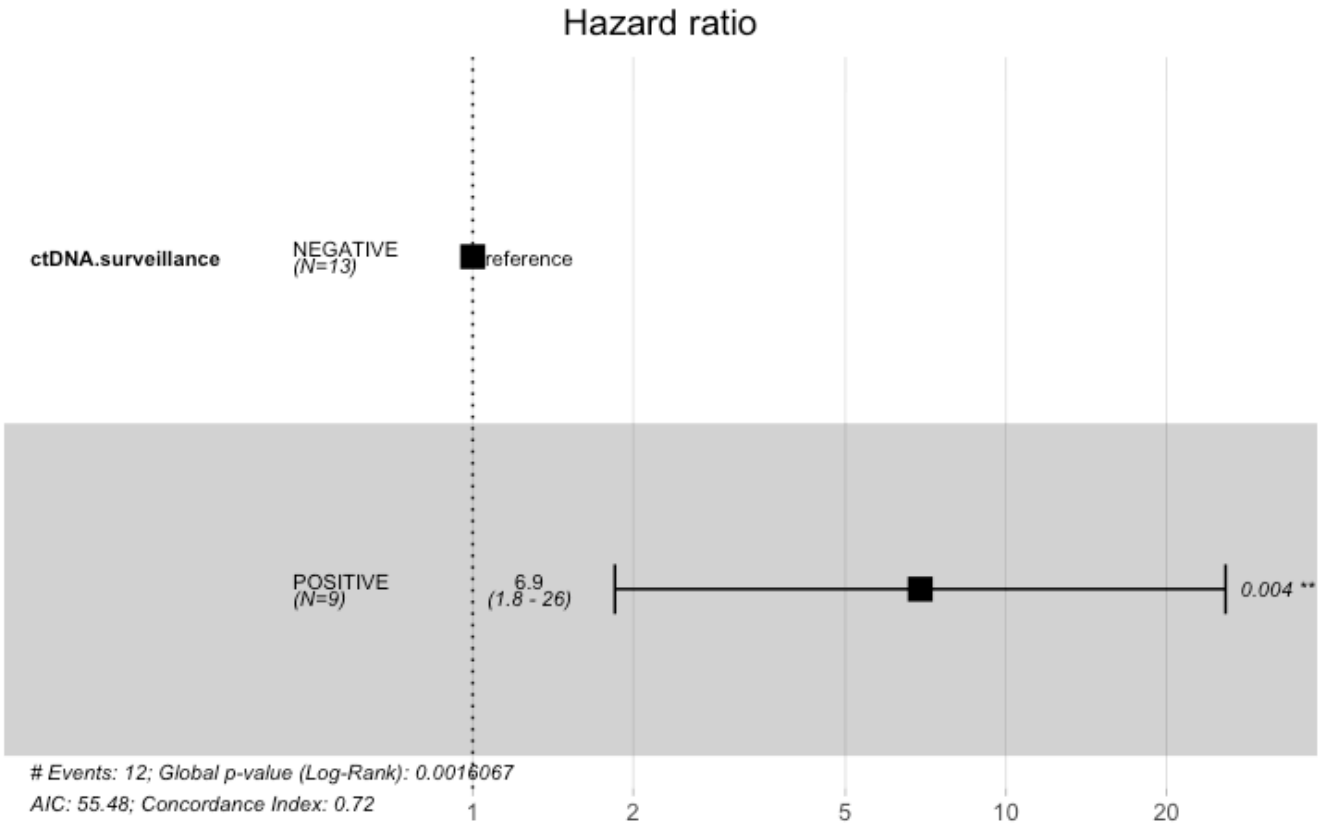
Hide

```
summary(KM_curve, times= c(12, 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
   12      8       3    0.769   0.117        0.442        0.919
   24      4       0    0.769   0.117        0.442        0.919

                ctDNA.surveillance=POSITIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   12      2       7    0.222   0.139      0.03371        0.513
   24      1       1    0.111   0.105      0.00613        0.388
```

Hide

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



Hide

```
summary(cox_fit)
```

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

  n= 22, number of events= 12

                             coef exp(coef) se(coef)     z Pr(>|z|)
ctDNA.surveillancePOSITIVE 1.9323    6.9055   0.6727 2.873  0.00407 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

                           exp(coef) exp(-coef) lower .95 upper .95
ctDNA.surveillancePOSITIVE     6.905     0.1448     1.848     25.81

Concordance= 0.716  (se = 0.067 )
Likelihood ratio test= 9.95  on 1 df,   p=0.002
Wald test            = 8.25  on 1 df,   p=0.004
Score (logrank) test = 10.97  on 1 df,   p=9e-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_C
I, 2), "); p = ", round(p_value, 3))
print(label_text)
```

```
[1] "HR = 6.91 (1.85–25.81); p = 0.004"
```

#EFS by STMs at post-ACT - Stage II/III

Hide

```
rm(list=ls())
setwd("~/Downloads")
circ_data <- read.csv("CLIA Testicular_GL.csv")
circ_data <- circ_data[circ_data$ctDNA.available==TRUE,]
circ_data <- circ_data[circ_data$ACT==TRUE,]
circ_data <- circ_data[circ_data$Stage!="I",]
circ_data <- circ_data[circ_data$Markers.Surveillance!="",]
circ_datadf <- as.data.frame(circ_data)

survfit(Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)~Markers.Surveilla
nce, data = circ_data)
```

```
Call: survfit(formula = Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event) ~
    Markers.Surveillance, data = circ_data)

                              n events median 0.95LCL 0.95UCL
Markers.Surveillance=NEGATIVE 23     11   65.3    7.69      NA
Markers.Surveillance=POSITIVE 10      8   20.0    5.59      NA
```

Hide

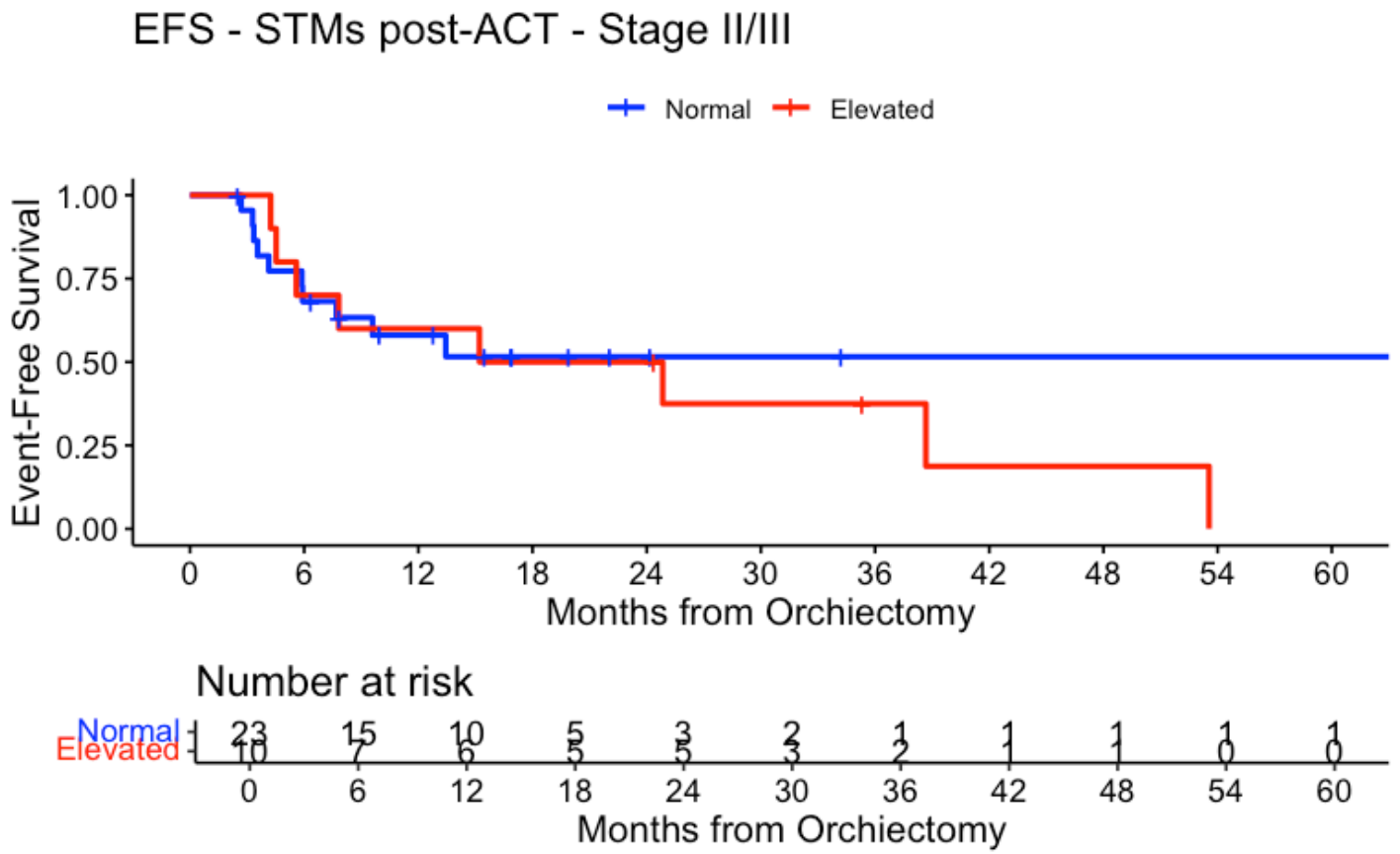
```
event_summary <- circ_data %>%
  group_by(Markers.Surveillance) %>%
  summarise(
    Total = n(),
    Events = sum(PFS.Event),
    Fraction = Events / n(),
    Percentage = (Events / n()) * 100
  )
print(event_summary)
```

Markers.Surveillance <chr>	Total <int>	Events <int>	Fraction <dbl>	Percentage <dbl>
NEGATIVE	23	11	0.4782609	47.82609
POSITIVE	10	8	0.8000000	80.00000

2 rows

Hide

```
surv_object <-Surv(time = circ_data$PFS.months, event = circ_data$PFS.Event)
KM_curve <- survfit(surv_object ~ Markers.Surveillance, data = circ_data,conf.int=0.95,c
onf.type="log-log")
ggsurvplot(KM_curve, data = circ_data, pval = FALSE, conf.int = FALSE, risk.table = TRU
E, break.time.by=6, palette=c("blue","red"), title="EFS - STMs post-ACT - Stage II/III",
ylab= "Event-Free Survival", xlab="Months from Orchiectomy", legend.labs=c("Normal", "El
evated"), legend.title="")
```



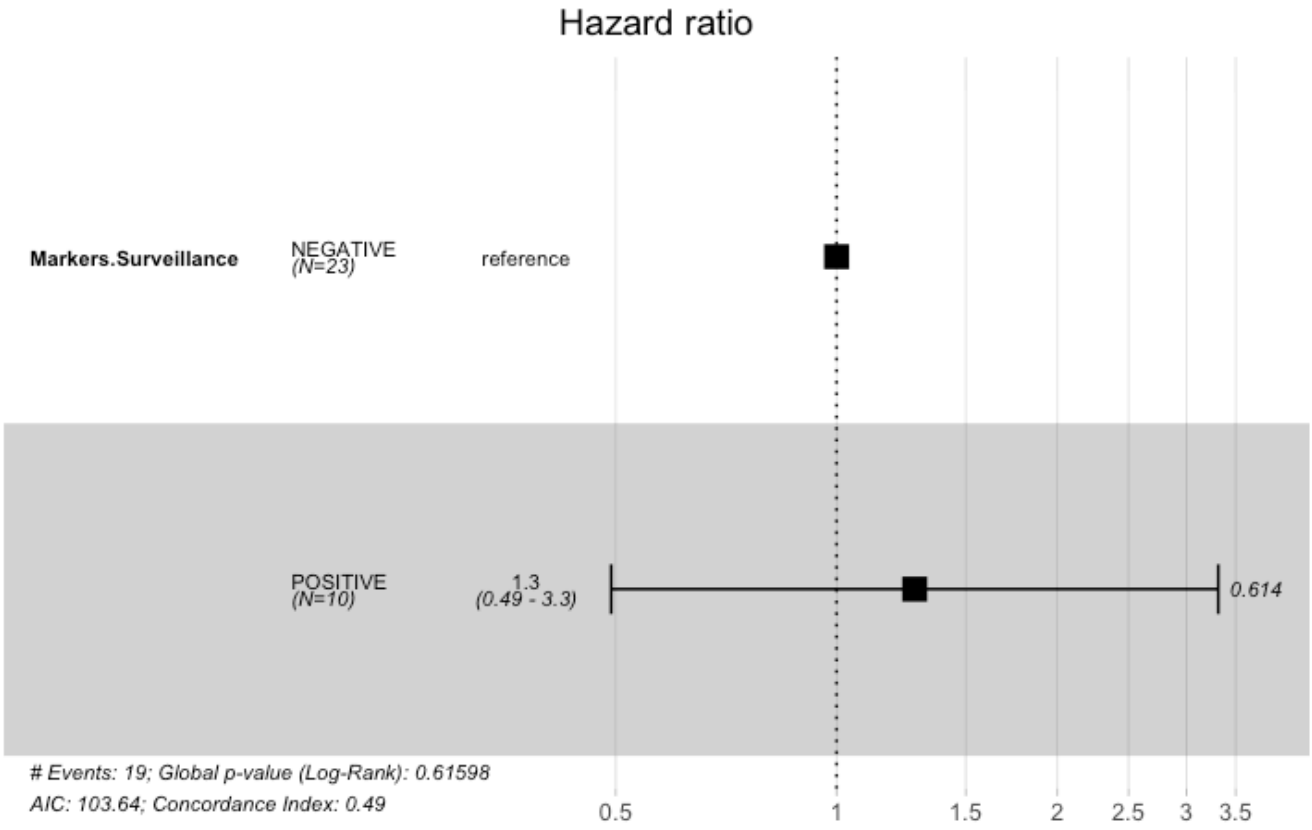
Hide

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

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

                Markers.Surveillance=NEGATIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   12     10       9    0.580   0.107        0.346        0.756
   24      3       1    0.516   0.113        0.281        0.708

                Markers.Surveillance=POSITIVE
 time n.risk n.event survival std.err lower 95% CI upper 95% CI
   12      6       4      0.6   0.155        0.253        0.827
   24      5       1      0.5   0.158        0.184        0.753
```

Hide

```
circ_data$Markers.Surveillance <- factor(circ_data$Markers.Surveillance, levels=c("NEGAT
IVE","POSITIVE"))
cox_fit <- coxph(surv_object ~ Markers.Surveillance, data=circ_data)
ggforest(cox_fit,data = circ_data)
```



Hide

```
summary(cox_fit)
```

```
Call:
coxph(formula = surv_object ~ Markers.Surveillance, data = circ_data)

  n= 33, number of events= 19

                              coef exp(coef) se(coef)     z Pr(>|z|)
Markers.SurveillancePOSITIVE 0.2453    1.2780   0.4862 0.505    0.614

                             exp(coef) exp(-coef) lower .95 upper .95
Markers.SurveillancePOSITIVE     1.278     0.7825    0.4928     3.314

Concordance= 0.49  (se = 0.058 )
Likelihood ratio test= 0.25  on 1 df,   p=0.6
Wald test            = 0.25  on 1 df,   p=0.6
Score (logrank) test = 0.26  on 1 df,   p=0.6
```

Hide

```
cox_fit_summary <- summary(cox_fit)

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

```
[1] "HR = 1.28 (0.49-3.31); p = 0.614"
```