LDA_IGA-NTX

Sven Morlock

28/03/2022

§ 1

Load required packages:

```
#if (Sys.info()["sysname"] != "Darwin"){
#  default_theme <- theme_minimal() +
#  theme(text = element_text(family = "Decima WE", size = 15)) +
#  theme(panel.grid.major = element_line(color = "grey", size = 0.3)) +
#  theme(axis.line = element_line(color = "black", size = 0.4))
#
#} else {
#  default_theme <- theme_minimal()
#}
default_theme <- theme_minimal()
two_scale_fill <- scale_fill_manual(values=c("#69b3a2", "#404080"))
```

Read data using read_data.R scrip:

```
source("read_data.R")
```

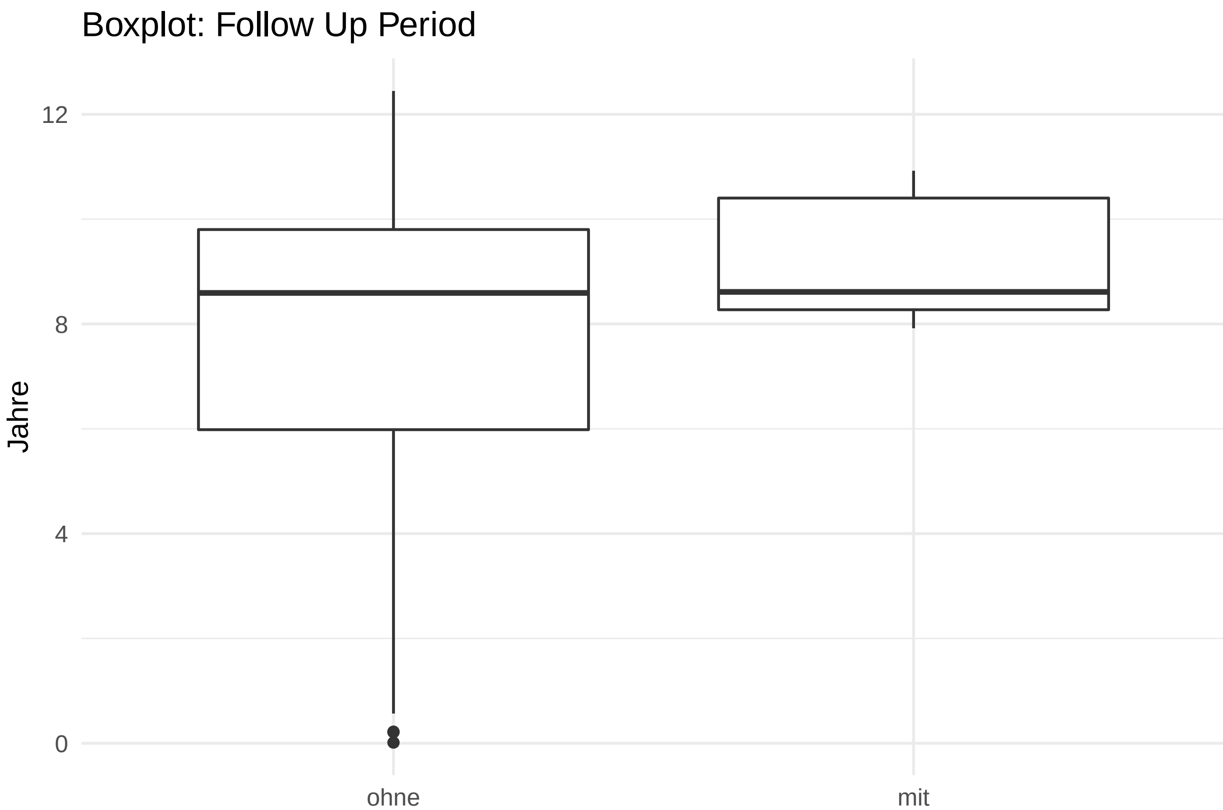
§ 2 EDA

```
follow_up <- years(10)
data_iga$follow_up_truncated <-pmin(data_iga$`T-dls`, data_iga$`T-date` + follow_up)

data_iga_pos <- data_iga[`biopsy proven recurrence (0=no, 1=yes)` == 1]
data_iga_neg <- data_iga[`biopsy proven recurrence (0=no, 1=yes)` == 0]
```

§ 2.1 IGA

```
# follow-up mean / median
tbl_iga_follow_up_mean <- data.frame(
  iga_all = mean((interval(data_iga$`T-date`, data_iga$`T-dls`) / years(1))),
  iga_pos = mean((interval(data_iga_pos$`T-date`, data_iga_pos$`T-dls`) / years(1))),
  iga_neg = mean((interval(data_iga_neg$`T-date`, data_iga_neg$`T-dls`) / years(1)))
)
tbl_iga_follow_up_median <- data.frame(
  iga_all = median((interval(data_iga$`T-date`, data_iga$`T-dls`) / years(1))),
  iga_pos = median((interval(data_iga_pos$`T-date`, data_iga_pos$`T-dls`) / years(1))),
  iga_neg = median((interval(data_iga_neg$`T-date`, data_iga_neg$`T-dls`) / years(1)))
)
```

```
ggplot() +
  geom_boxplot(aes(x = `biopsy proven recurrence (0=no, 1=yes)`, y = interval(data_iga$`T-date`, data_i
  default_theme +
  ggtitle("Boxplot: Follow Up Period") +
  ylab("Jahre") +
  xlab("") +
  scale_x_discrete(labels = c("ohne", "mit"))
```




```r
# follow_up mean/median (truncated)
tbl_iga_follow_up_mean_truncated <- data.frame(
  iga_all = mean((interval(data_iga$`T-date`, data_iga$follow_up_truncated) / years(1))),
  iga_pos = mean((interval(data_iga_pos$`T-date`, data_iga_pos$follow_up_truncated) / years(1))),
  iga_neg = mean((interval(data_iga_neg$`T-date`, data_iga_neg$follow_up_truncated) / years(1)))
)
tbl_iga_follow_up_median_truncated <- data.frame(
  iga_all = median((interval(data_iga$`T-date`, data_iga$follow_up_truncated) / years(1))),
  iga_pos = median((interval(data_iga_pos$`T-date`, data_iga_pos$follow_up_truncated) / years(1))),
  iga_neg = median((interval(data_iga_neg$`T-date`, data_iga_neg$follow_up_truncated) / years(1)))
)
```

```r
# patient death within follow up
tbl_iga_pat_death <- data.frame(
  iga_all = nrow(data_iga[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 1]),
  iga_pos = nrow(data_iga_pos[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 1
  iga_neg = nrow(data_iga_neg[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 1
)
```

```r
# patient drop out within follow up
tbl_iga_pat_drop <- data.frame(
  iga_all = nrow(data_iga[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 0]),
  iga_pos = nrow(data_iga_pos[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 0
  iga_neg = nrow(data_iga_neg[(`T-dls` <= (`T-date` + follow_up)) & `Pat death (0=alive, 1= dead)` == 0
)
```

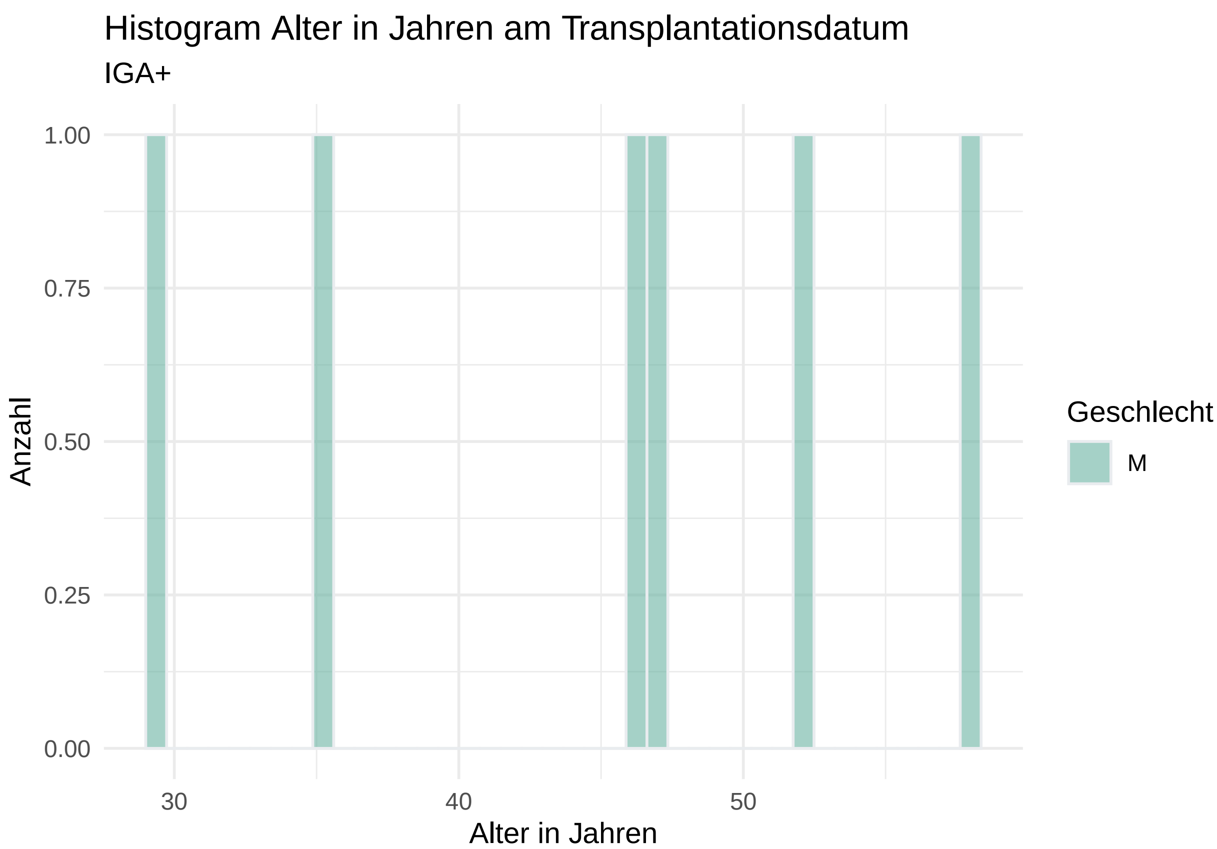
```r
# patients with graft loss
tbl_iga_graft_loss <- data.frame(
  iga_all = nrow(data_iga[`graft loss (0=functial, 1=loss)` == 1]),
```

```r
  iga_pos = nrow(data_iga_pos[`graft loss (0=functial, 1=loss)` == 1]),
  iga_neg = nrow(data_iga_neg[`graft loss (0=functial, 1=loss)` == 1])
)
```

```r
# patients with graft loss within follow up period
tbl_iga_graft_loss_follow_up <- data.frame(
  iga_all = nrow(data_iga[`graft loss date` < `T-date` + follow_up]),
  iga_pos = nrow(data_iga_pos[`graft loss date` < `T-date` + follow_up]),
  iga_neg = nrow(data_iga_neg[`graft loss date` < `T-date` + follow_up])
)
```

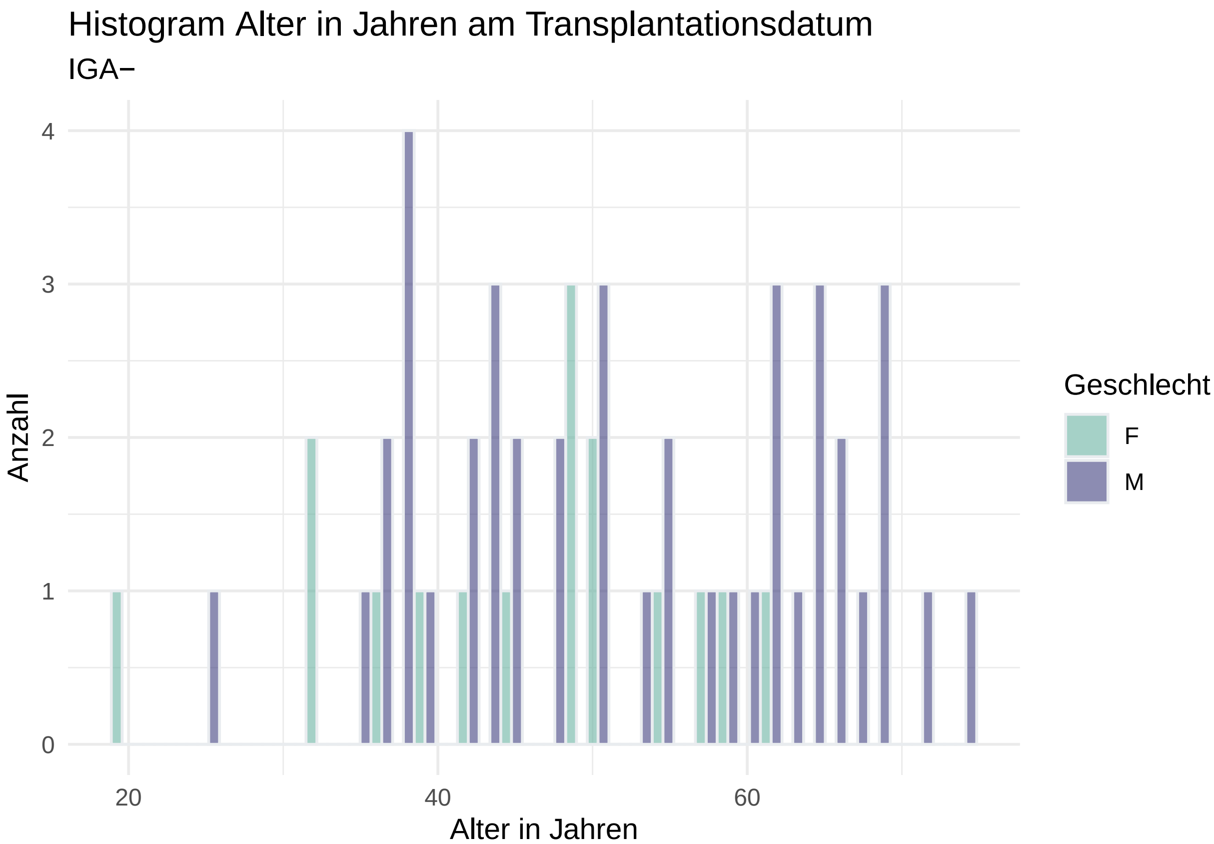
```r
# age patients (yrs.)
ggplot(data = data_iga) +
  geom_histogram(mapping = aes(x = interval(`Date of birth`,`T-date`) / years(1),
                               fill = `R-sex`),
                 color="#e9ecef",
                 alpha=0.6,
                 bins = 40,
                 position = "dodge") +
  two_scale_fill +
  ylab("Anzahl") +
  xlab("Alter in Jahren") +
  ggtitle("Histogram Alter in Jahren am Transplantationsdatum",
          subtitle =  "IGA all") +
  labs(fill = "Geschlecht") +
  default_theme
```

```r
ggplot(data = data_iga_pos) +
 geom_histogram(mapping = aes(x = interval(`Date of birth`,`T-date`) / years(1),
                              fill = `R-sex`),
               color="#e9ecef",
               alpha=0.6,
               bins = 40,
               position = "dodge") +
 two_scale_fill +
 ylab("Anzahl") +
 xlab("Alter in Jahren") +
 ggtitle("Histogram Alter in Jahren am Transplantationsdatum",
        subtitle =  "IGA+") +
 labs(fill = "Geschlecht") +
 default_theme
```




```r
ggplot(data = data_iga_neg) +
 geom_histogram(mapping = aes(x = interval(`Date of birth`,`T-date`) / years(1),
                              fill = `R-sex`),
               color="#e9ecef",
               alpha=0.6,
               bins = 40,
               position = "dodge") +
 two_scale_fill +
 ylab("Anzahl") +
 xlab("Alter in Jahren") +
 ggtitle("Histogram Alter in Jahren am Transplantationsdatum",
        subtitle =  "IGA-") +
```

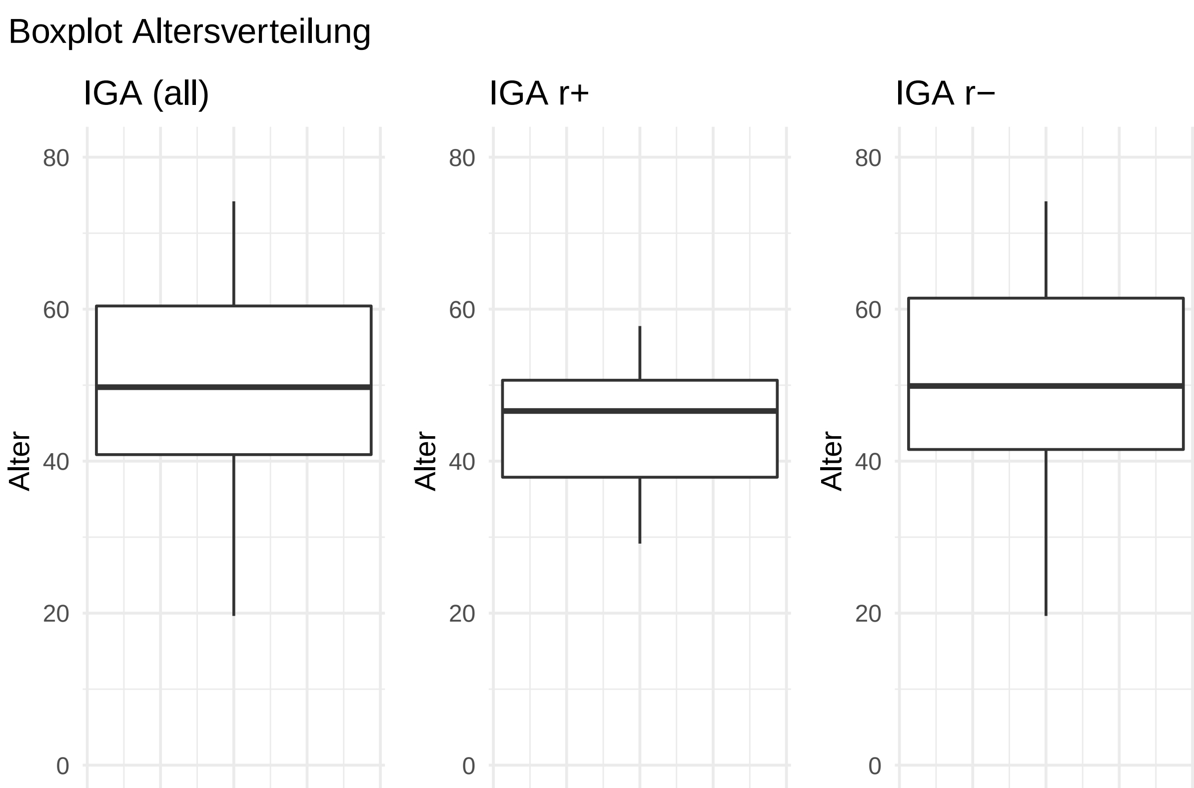
```
labs(fill = "Geschlecht") +
default_theme
```




```
## IQR
tbl_iga_iqr_age <- data.frame(
  iga_all = IQR(data_iga$R_age_Tdate),
  iga_po = IQR(data_iga_pos$R_age_Tdate),
  iga_neg = IQR(data_iga_neg$R_age_Tdate)
)
```

```
## boxpot

p1 <- ggplot(data = data_iga) +
  geom_boxplot(aes(x = R_age_Tdate)) +
  coord_flip() +
  ggtitle("IGA (all)") +
  default_theme

p2 <- ggplot(data = data_iga_pos) +
  geom_boxplot(aes(x = R_age_Tdate)) +
  coord_flip() +
  ggtitle("IGA r+") +
  default_theme

p3 <- ggplot(data = data_iga_neg) +
  geom_boxplot(aes(x = R_age_Tdate)) +
  coord_flip()+
```

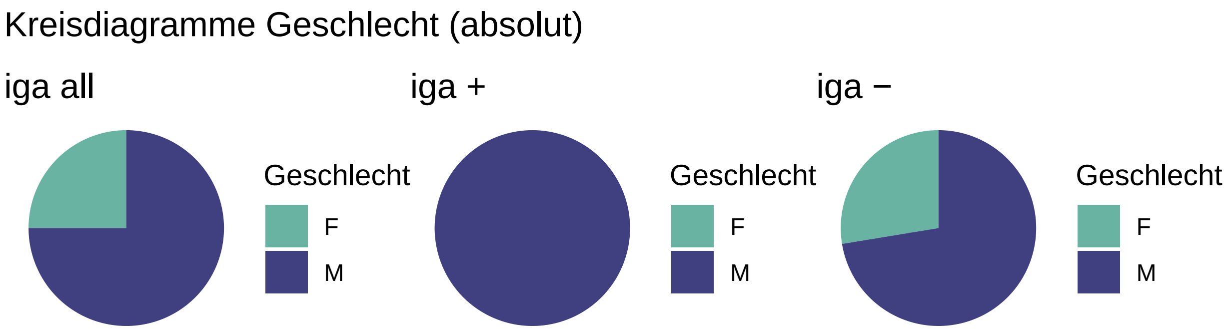
```
  ggtitle("IGA r-") +
  default_theme

p1$labels$x <- p2$labels$x <- p3$labels$x <- ""
p1$labels$y<- p2$labels$y <- p3$labels$y <- ""

patch <- p1 | p2 | p3
patch <- patch + plot_annotation(
  title = "Boxplot Altersverteilung"
)
patch & xlim(0, 80) & theme(axis.ticks.x = element_blank(),
                            axis.text.x = element_blank()) & xlab("Alter")
```




```
# sex
tbl_iga_sex_abs <- data.frame(
  iga_all = summary(data_iga$`R-sex`),
  iga_pos = summary(data_iga_pos$`R-sex`),
  iga_neg = summary(data_iga_neg$`R-sex`)
)
```

```
p1 <- data.frame(group = rownames(tbl_iga_sex_abs),
           value = tbl_iga_sex_abs$iga_all) %>%
  ggplot(aes(x = "", y = value, fill = group)) +
  geom_bar(width = 1, stat = "identity") +
  coord_polar("y", start = 0) +
  guides(fill=guide_legend(title = "Geschlecht")) +
  theme_void() +
  two_scale_fill +
  ggtitle("iga all")
```

```r
p2 <- data.frame(group = rownames(tbl_iga_sex_abs),
           value = tbl_iga_sex_abs$iga_pos) %>%
  ggplot(aes(x = "", y = value, fill = group)) +
  geom_bar(width = 1, stat = "identity") +
  coord_polar("y", start = 0) +
  guides(fill=guide_legend(title = "Geschlecht")) +
  theme_void() +
  two_scale_fill +
  ggtitle("iga +")

p3 <- data.frame(group = rownames(tbl_iga_sex_abs),
           value = tbl_iga_sex_abs$iga_neg) %>%
  ggplot(aes(x = "", y = value, fill = group)) +
  geom_bar(width = 1, stat = "identity") +
  coord_polar("y", start = 0) +
  guides(fill=guide_legend(title = "Geschlecht")) +
  theme_void() +
  two_scale_fill +
  ggtitle("iga -")

patch <- p1 | p2 | p3
patch + plot_annotation(title = "Kreisdiagramme Geschlecht (absolut)")
```

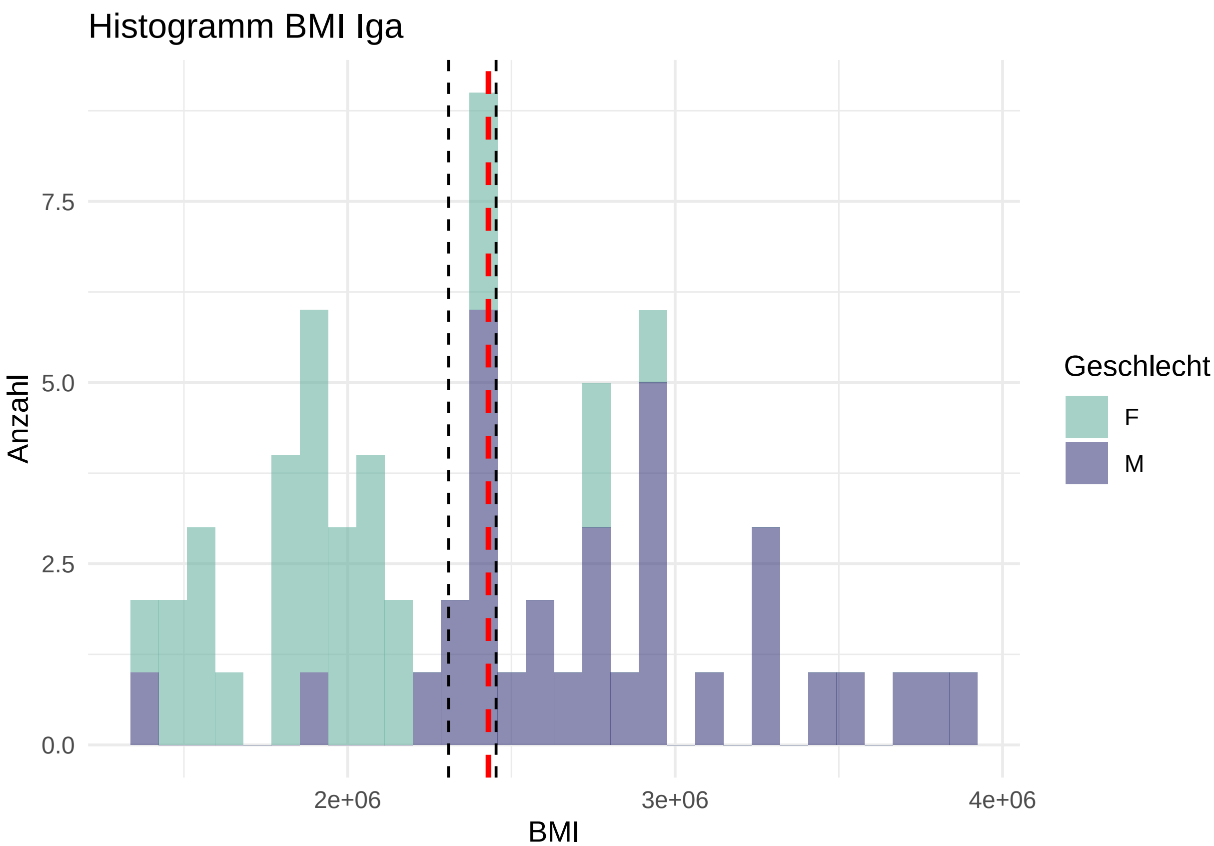



```r
# sex percent
tbl_iga_sex_percent <- data.frame(
  iga_all = round(summary(data_iga$`R-sex`) / nrow(data_iga), 2),
  iga_pos = round(summary(data_iga_pos$`R-sex`) / nrow(data_iga_pos), 2),
  iga_neg = round(summary(data_iga_neg$`R-sex`) / nrow(data_iga_neg), 2)
)
```

```r
# BMI (mean.)
ggplot(data = data_iga) +
  geom_histogram(mapping = aes(`D-weight` * (`D-height`)^2,
                               fill = `D-sex`),
                 alpha = 0.6) +
  two_scale_fill +
  # median BMI total
  geom_vline(aes(xintercept = median(`D-weight` * (`D-height`)^2)),
             size = 1.0, color = "red", linetype = "dashed") +
  # median BMI F
  geom_vline(aes(xintercept = median(`D-weight` * (`D-height`)^2)),
             data = data_iga[`R-sex` == "F"], linetype = "dashed") +
  # median BMI M
```

```
  geom_vline(aes(xintercept = median(`D-weight` * (`D-height`)^2)),
             data = data_iga[`R-sex` == "M"], linetype = "dashed") +
  ylab("Anzahl") +
  xlab("BMI") + # Einheit???
  ggtitle("Histogramm BMI Iga") +
  labs(fill = "Geschlecht") +
  default_theme
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# deceased D.
# living D.
tbl_iga_1 <- summary(data_iga[`graft loss (0=functial, 1=loss)` == 1]$`D-type`)
tbl_iga_1 <- round(tbl_iga_1 / sum(tbl_iga_1), 3)
tbl_iga_2 <- summary(data_iga[`graft loss (0=functial, 1=loss)` == 0]$`D-type`)
tbl_iga_2 <- round(tbl_iga_2 / sum(tbl_iga_2),3)
rbind("loss" = tbl_iga_1, "functional" = tbl_iga_2) %>%
  as.data.frame()
```

```
##            Cadaver Living
## loss         0.867  0.133
## functional   0.735  0.265
```

```
# dead/alive
tbl_iga_1 <- summary(data_iga[`Pat death (0=alive, 1= dead)` == 1]$`D-type`)
tbl_iga_1 <- round(tbl_iga_1 / sum(tbl_iga_1), 3)
tbl_iga_2<- summary(data_iga[`Pat death (0=alive, 1= dead)` == 0]$`D-type`)
tbl_iga_2 <-  round(tbl_iga_2 / sum(tbl_iga_2),3)  # in %
```

```
rbind("dead" = tbl_iga_1, "alive" = tbl_iga_2) %>%
  as.data.frame()
```

```
##       Cadaver Living
## dead    0.917  0.083
## alive   0.675  0.325
```

```
# BMI (mean.)
tbl_iga_bmi <- data.table(
  iga_all = mean(data_iga$`D-weight` * (data_iga$`D-height`)^2),
  iga_pos = mean(data_iga_pos$`D-weight` * (data_iga_pos$`D-height`)^2),
  iga_neg = mean(data_iga_neg$`D-weight` * (data_iga_neg$`D-height`)^2)
)
```

```
# HLA-mm (0-6)
## data.table of mm-A, mm-B and mm-DR
tbl_iga_1 <- as.data.table(lapply(data_iga[, c("mm-A", "mm-B", "mm-DR")],
                                  as.numeric))
## calculate mean for each column
tbl_iga_2 <- apply(X = tbl_iga_1,
                   MARGIN = 2,
                   FUN = mean, na.rm = TRUE)
round(tbl_iga_2, 3)
```

```
##  mm-A  mm-B mm-DR
## 1.792 2.146 1.917
```

```
# mean of means
round(mean(tbl_iga_2),3)
```

```
## [1] 1.951
```

```
## sum of mm-A, mm-B and mm-DR
tbl_iga_3 <- apply(X = tbl_iga_1,
                   MARGIN = 1,
                   FUN = sum, na.rm = TRUE)
# mean value of col sum
mean(tbl_iga_3)
```

```
## [1] 4.390625
```

```
# median value of col sum
sd(tbl_iga_3)
```

```
## [1] 2.711628
```

```
# PRA current (mean)
tbl_iga_pra_curr_mean <- data.frame(
  iga_all = mean(data_iga$`Current PRA%`, na.rm = TRUE),
  iga_pos = mean(data_iga_pos$`Current PRA%`, na.rm = TRUE),
  iga_neg = mean(data_iga_neg$`Current PRA%`, na.rm = TRUE)
)
# PRA highest (mean)
tbl_iga_pra_high <- data.frame(
  iga_all = mean(data_iga$`Highest PRA%`, na.rm = TRUE),
  iga_pos = mean(data_iga_pos$`Highest PRA%`, na.rm = TRUE),
  iga_neg = mean(data_iga_neg$`Highest PRA%`, na.rm = TRUE)
)
```

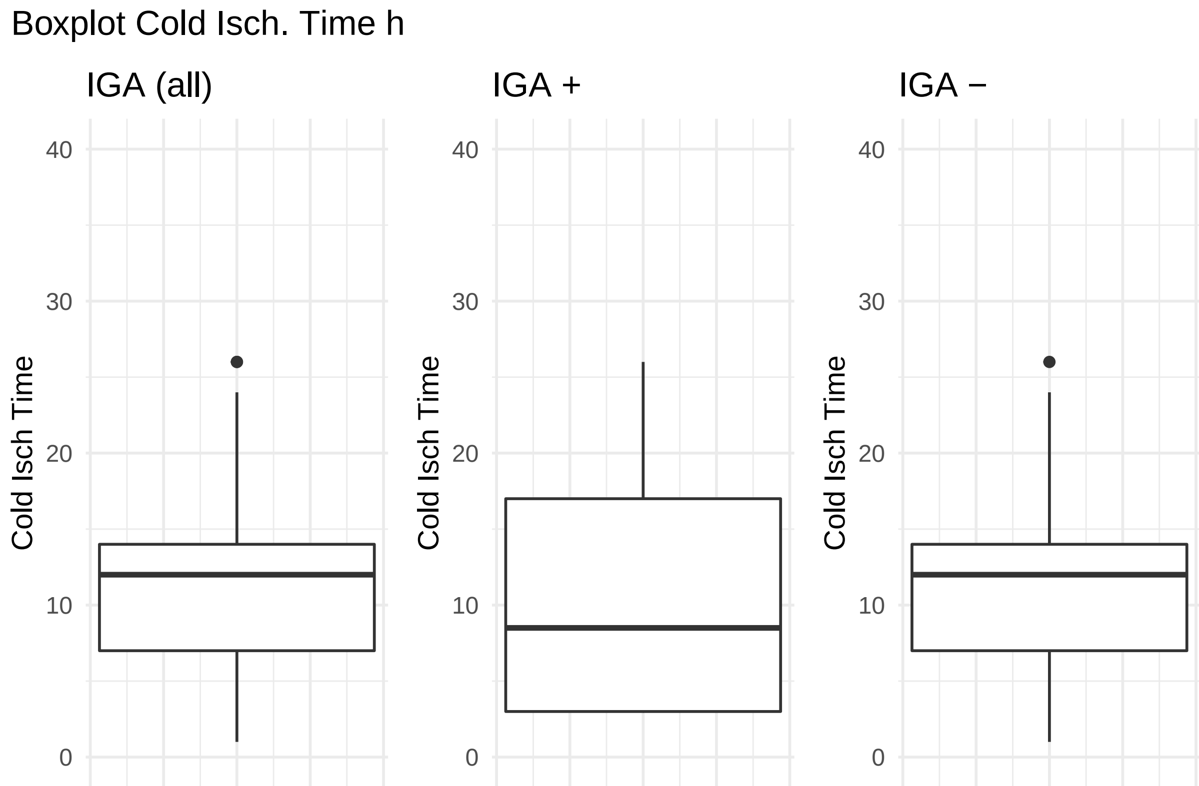
```r
# age donor (mean.)
tbl_iga_age_donor <- data.frame(
  iga_all = mean(data_iga$`D-age`),
  iga_pos =  mean(data_iga_pos$`D-age`),
  iga_neg =  mean(data_iga_neg$`D-age`)
)
```

```r
# cold-ischemia time (hours)
# mean
tbl_iga_cis_mean <- data.frame(
  iga_all = mean(data_iga$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_pos = mean(data_iga_pos$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_neg = mean(data_iga_neg$`Cold ischaemic period hours`, na.rm = TRUE)
)
# median
tbl_iga_cis_median <- data.frame(
  iga_all = median(data_iga$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_pos = median(data_iga_pos$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_neg = median(data_iga_neg$`Cold ischaemic period hours`, na.rm = TRUE)
)
# standard error
tbl_iga_cis_sd <- data.frame(
  iga_all = sd(data_iga$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_pos = sd(data_iga_pos$`Cold ischaemic period hours`, na.rm = TRUE),
  iga_neg = sd(data_iga_neg$`Cold ischaemic period hours`, na.rm = TRUE)
)
```

```r
p1 <- data_iga %>%
ggplot() +
geom_boxplot(aes(x = `Cold ischaemic period hours`)) +
  coord_flip() +
  ggtitle("IGA (all)") +
  default_theme

p2 <- data_iga_pos %>%
  ggplot() +
  geom_boxplot(aes(x = `Cold ischaemic period hours`)) +
  coord_flip() +
  ggtitle("IGA +") +
  default_theme

p3 <- data_iga_neg %>%
  ggplot() +
  geom_boxplot(aes(x = `Cold ischaemic period hours`)) +
  coord_flip() +
  ggtitle("IGA -") +
  default_theme

p1$labels$x <- p2$labels$x <- p3$labels$x <- ""
p1$labels$y<- p2$labels$y <- p3$labels$y <- ""

patch <- p1 | p2 | p3
patch <- patch + plot_annotation(
  title = "Boxplot Cold Isch. Time h"
```

```
)
patch & xlim(0, 40) & theme(axis.ticks.x = element_blank(),
                            axis.text.x = element_blank()) & xlab("Cold Isch Time")
```

```
## Warning: Removed 15 rows containing non-finite values (stat_boxplot).

## Warning: Removed 2 rows containing non-finite values (stat_boxplot).

## Warning: Removed 13 rows containing non-finite values (stat_boxplot).
```




```
# living vs dead donator
tbl_iga_don_living_abs <- data.frame(
  iga_all = nrow(data_iga[`D-type` == "Living"]),
  iga_pos = nrow(data_iga_pos[`D-type` == "Living"]),
  iga_neg = nrow(data_iga_neg[`D-type` == "Living"])
)

tbl_iga_don_dead_abs <- data.frame(
  iga_all = nrow(data_iga[`D-type` == "Cadaver"]),
  iga_pos = nrow(data_iga_pos[`D-type` == "Cadaver"]),
  iga_neg = nrow(data_iga_neg[`D-type` == "Cadaver"])
)


tbl_iga_don_living_rel <- data.frame(
  iga_all = nrow(data_iga[`D-type` == "Living"]) / nrow(data_iga),
  iga_pos = nrow(data_iga_pos[`D-type` == "Living"]) / nrow(data_iga_pos),
  iga_neg = nrow(data_iga_neg[`D-type` == "Living"]) / nrow(data_iga_neg)
)
```

```r
tbl_iga_don_dead_rel <- data.frame(
  iga_all = nrow(data_iga[`D-type` == "Cadaver"]) / nrow(data_iga),
  iga_pos = nrow(data_iga_pos[`D-type` == "Cadaver"]) / nrow(data_iga_pos),
  iga_neg = nrow(data_iga_neg[`D-type` == "Cadaver"]) / nrow(data_iga_neg)
)
```

```r
iga_table <- rbindlist(
  list(
    iga_dropout = data.frame(
      cbind(name = "dropout",tbl_iga_pat_drop)),
    iga_graftloss = data.frame(
      cbind(name = "graftloss", tbl_iga_graft_loss)),
    iga_graftloss_follow_up = data.frame(
      cbind(name = "graftloss_followup", tbl_iga_graft_loss_follow_up)),
    iga_pat_death = data.frame(
      cbind(name = "pat_death", tbl_iga_pat_death)),
    iga_don_dead_abs = data.frame(
      cbind(name = "don_dead_abs", tbl_iga_don_dead_abs)),
    iga_don_dead_rel = data.frame(
      cbind(name = "don_dead_rel", tbl_iga_don_dead_rel)),
    iga_don_living_abs = data.frame(
      cbind(name = "don_living_abs", tbl_iga_don_living_abs)),
    iga_don_living_rel = data.frame(
      cbind(name = "don_living_rel", tbl_iga_don_living_rel)),
    iga_cis_h_mean = data.frame(
      cbind(name = "cold_h_mean", tbl_iga_cis_mean)),
    iga_cis_h_sd = data.frame(
      cbind(name = "cold_h_sd", tbl_iga_cis_sd))
    ))

kable(iga_table, col.names = c("Merkmal", "IgA (all)", "IgA r+", "IgA r-")) %>%
  kable_styling(latex_options = "hold_position")
```

Merkmal	IgA (all)	IgA r+	IgA r-
dropout	27.000000	3.0000000	24.0000000
graftloss	15.000000	2.0000000	13.0000000
graftloss_followup	14.000000	2.0000000	12.0000000
pat_death	23.000000	1.0000000	22.0000000
don_dead_abs	49.000000	4.0000000	45.0000000
don_dead_rel	0.765625	0.6666667	0.7758621
don_living_abs	15.000000	2.0000000	13.0000000
don_living_rel	0.234375	0.3333333	0.2241379
cold_h_mean	11.612245	11.5000000	11.6222222
cold_h_sd	6.330535	10.9696551	5.9592895

§ 2.2 NTX

```r
# follow_up mean
pmin(
  # follow up
  interval(data_ntx$Datum, (data_ntx$Datum + follow_up)) / years(1),
  # last seen
  interval(data_ntx$Datum, data_ntx$tdls) / years(1)
```

```
) %>%
  mean()
```

```
## [1] 7.41754
```

```
# patient death within follow up
nrow(data_ntx[`Todesdatum[NTX PatientenInformation]` < (Datum + follow_up)])
```

```
## [1] 317
```

```
# patient drop out
nrow(data_ntx[`Date last seen[NTX PatientenInformation]` <= (Datum + follow_up) & `Patienten Status[NTX
```

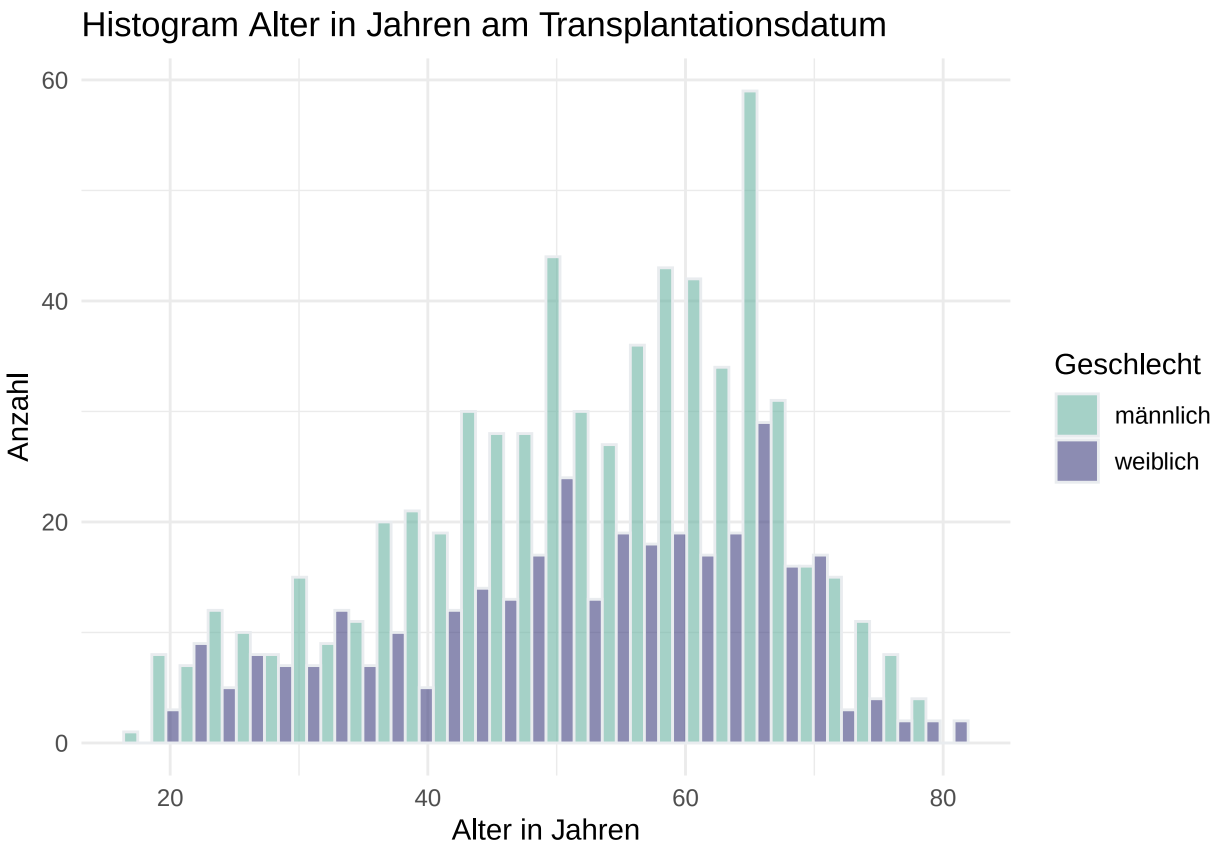
```
## [1] 370
```

```
# patients with graft loss
nrow(data_ntx[data_ntx$`TX Status[NTX PatientenInformation]` == "2 - ohne Transplantatfunktion"])
```

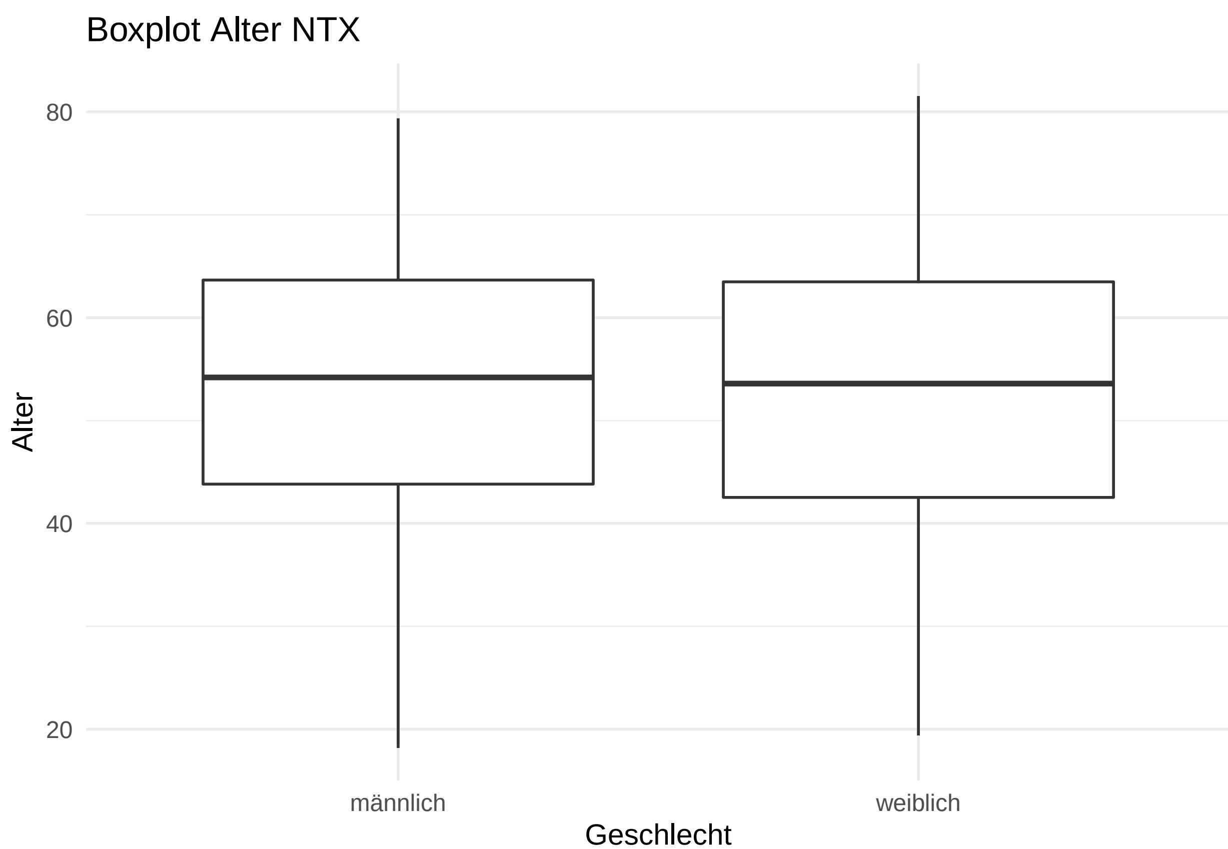
```
## [1] 179
```

```
# age patients (yrs.)
ggplot(data = data_ntx) +
  geom_histogram(mapping = aes(x = R_age_Datum, fill = Geschlecht),
                 color = "#e9ecef", alpha = 0.6, position = "dodge") +
  two_scale_fill +
  ylab("Anzahl") +
  xlab("Alter in Jahren") +
  ggtitle("Histogram Alter in Jahren am Transplantationsdatum") +
  labs(fill = "Geschlecht") +
  default_theme
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(data = data_ntx) +
  geom_boxplot(aes(x = data_ntx$Geschlecht, y = data_ntx$R_age_Datum))  +
  default_theme +
  ggtitle("Boxplot Alter NTX") +
  xlab("Geschlecht") +
  ylab("Alter")
```

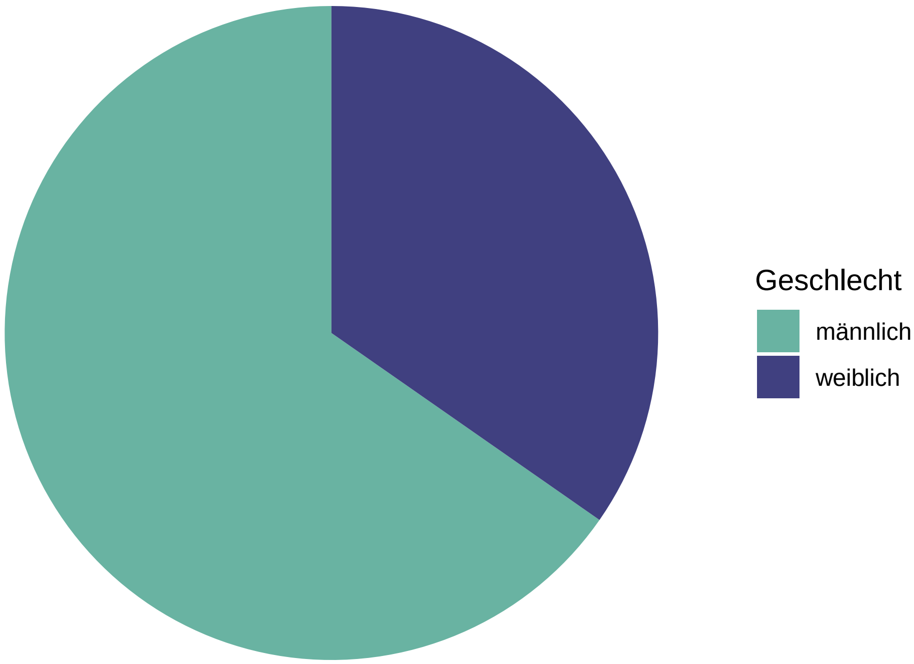



```
# male sex
summary(data_ntx$Geschlecht)
```

```
## männlich weiblich
##      627      333
```

```
data.frame(group = names(summary(data_ntx$Geschlecht)),
              value = c(summary(data_ntx$Geschlecht))) %>%
  ggplot(aes(x = "", y = value, fill = group)) +
  geom_bar(width = 1, stat = "identity") +
  coord_polar("y", start = 0) +
  guides(fill=guide_legend(title = "Geschlecht")) +
  theme_void() +
  two_scale_fill +
  ggtitle("NTX")
```

NTX




```
# BMI (mean.)
## NO DATA AVAILABLE
```

```
# deceased D.
# living D.
## NO DATA AVAILABLE
```

```
# HLA-mm
## NO DATA AVAILABLE
```

```
# PPR
## NO DATA AVAILABLE
```

```
# living vs dead donator
## NO DATA AVAILABLE
```

§ 3 Kaplan-Meier

§ 3.1 IGA

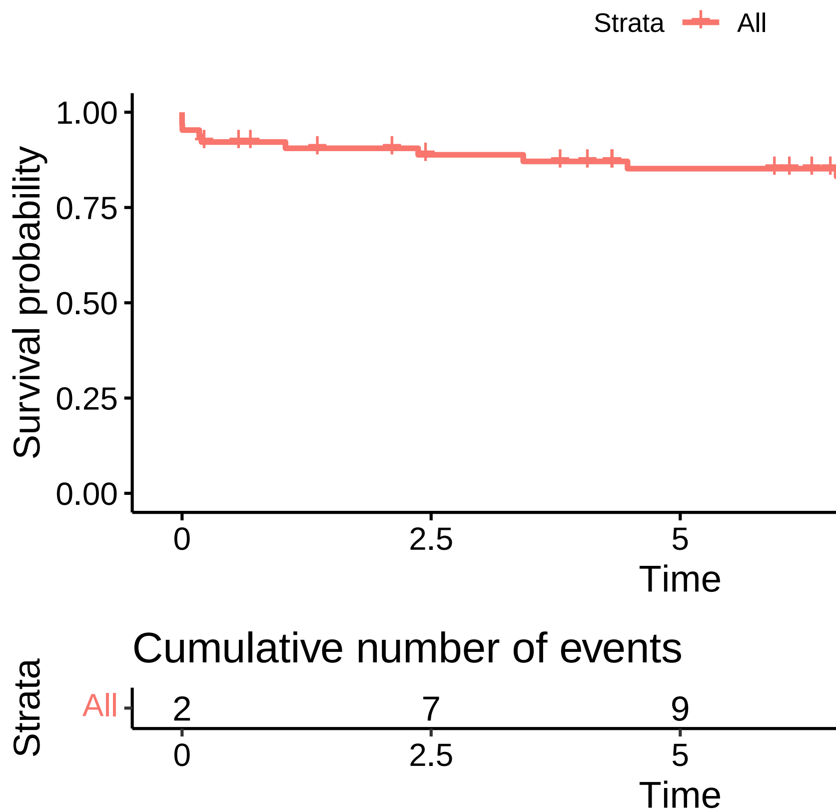
```
#  functions in survival and surviminer package need numeric-type input
data_iga[, time_date_biopsy := interval(`T-date`, `date of biopsy`) / years(1)]
data_iga[, time_t_dls := interval(`T-date`, `T-dls`) / years(1)]
data_iga[, time_date_birth := interval(`T-date`, `Date of birth`) / years(1)]
data_iga[, time_graft_loss := interval(`T-date`, `graft loss date`) / years(1)]
data_iga[, time_date_follow_up := interval(`T-date`,`T-date` + follow_up) / years(1)]
```

§ 3.1.1

- Event:
  - graft-loss within the follow up period.
- Censoring scheme:
  - if `graft loss date` after follow up period, censored by end of follow up.
  - if `T-dls` (date last seen) within follow up period, censored by `T-dls`.
- Time period:
  - 10 years after `T-date` (kidney transplantaion).

```
data_iga <- data_iga %>%
  mutate(status_date = case_when(
    ## graft-loss within follow up period
    !is.na(`graft loss date`) & `graft loss date` <  `T-date` + follow_up ~ time_graft_loss,
    ## graft-loss after follow up period
    !is.na(`graft loss date`) & `graft loss date` >  `T-date` + follow_up ~ time_date_follow_up,
    ## no graft-loss and last seen within follow up
    is.na(`graft loss date`) & !is.na(`T-dls`) & `T-dls` < `T-date` + follow_up ~ time_t_dls,
    ## no graft-loss and last seen after follow up
    is.na(`graft loss date`) & !is.na(`T-dls`) & `T-dls` > `T-date` + follow_up ~ time_date_follow_up,
    ## no graft loss and no last seen
    is.na(`graft loss date`) & is.na(`T-dls`) ~ time_date_follow_up
                                )
  )
data_iga <- data_iga %>%
  mutate(status = case_when(
    ## graft-loss within follow up period
    !is.na(`graft loss date`) & `graft loss date` <  `T-date` + follow_up ~ 1,
    ## else censored
    TRUE ~ 0,
                                )
  )
```

```
model_iga_1 <- survfit(formula = Surv(time = status_date,
                                   event = status, type = "right")~ 1,
                    data = data_iga)
ggsurvplot(model_iga_1,
           conf.int = FALSE,
           cumevents = TRUE)
```



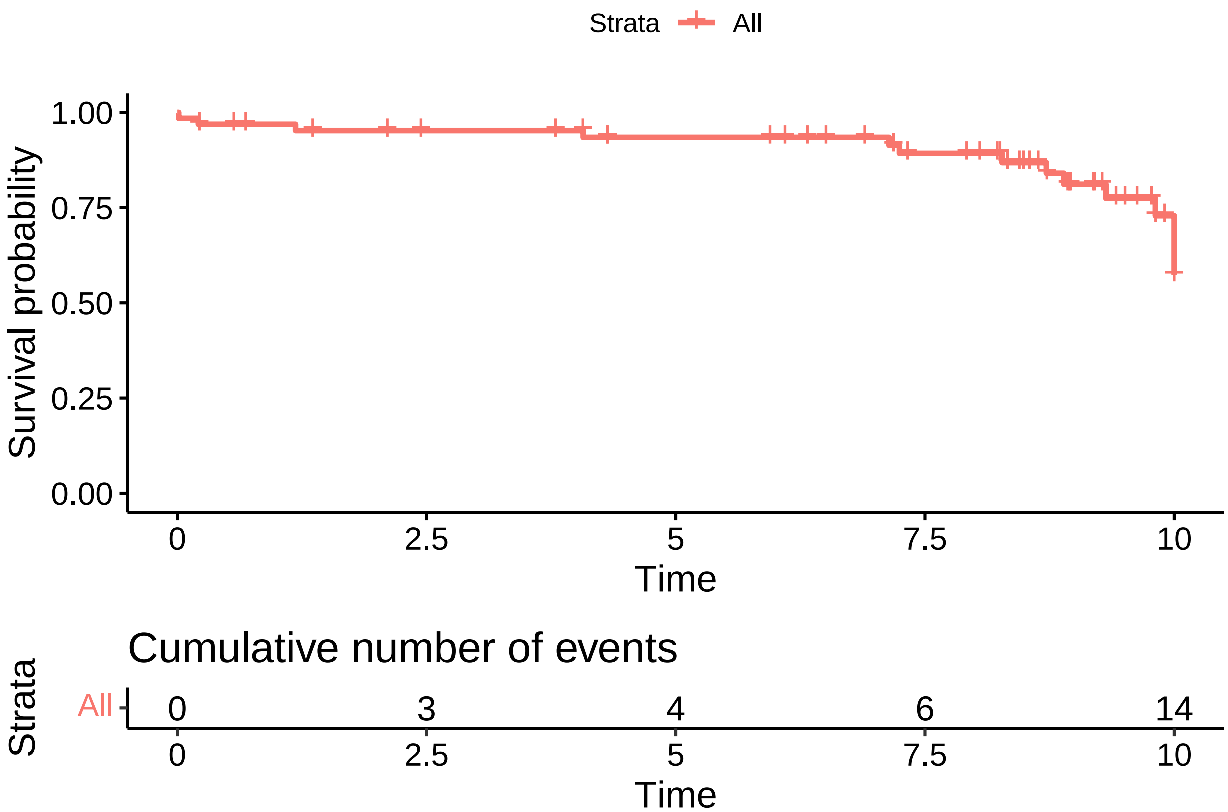
Overall kaplan-Meier curve (no stratification)

§ 3.1.2

- Event:
  - `T-dls` & ‘`Pat death (0=alive, 1= dead)` (patient death).
- Censoring scheme:
  - `T-dls` (date last seen) within follow up period.
- Time period:
  - 10 years after `T-date` (kidney transplantation).

```
data_iga <- data_iga %>%
  mutate(status_date = case_when(
    ## patient death and death date within follow up
    (`Pat death (0=alive, 1= dead)` == 1) & `T-dls` < `T-date` + follow_up ~ time_t_dls,
    ## patient dead but after follow up
    (`Pat death (0=alive, 1= dead)` == 1) & `T-dls` > `T-date` + follow_up ~ time_date_follow_up,
    ## patient not death but dropped within follow up
    (`Pat death (0=alive, 1= dead)` == 0) & `T-dls` < `T-date` + follow_up ~ time_t_dls,
    ## patient not death but dropped after follow up
    (`Pat death (0=alive, 1= dead)` == 0) & `T-dls` > `T-date` + follow_up ~ time_date_follow_up,
    ## NOTE: T-dls never NA
  )
)
```

```
model_iga_2 <- survfit(formula = Surv(time = status_date,
                                      event = status, type = "right")~ 1,
                       data = data_iga)
ggsurvplot(model_iga_2,
           conf.int = FALSE,
           cumevents = TRUE)
```



§ 3.1.3

- Event: graft loss and death
- Censoring scheme:
- Time period:

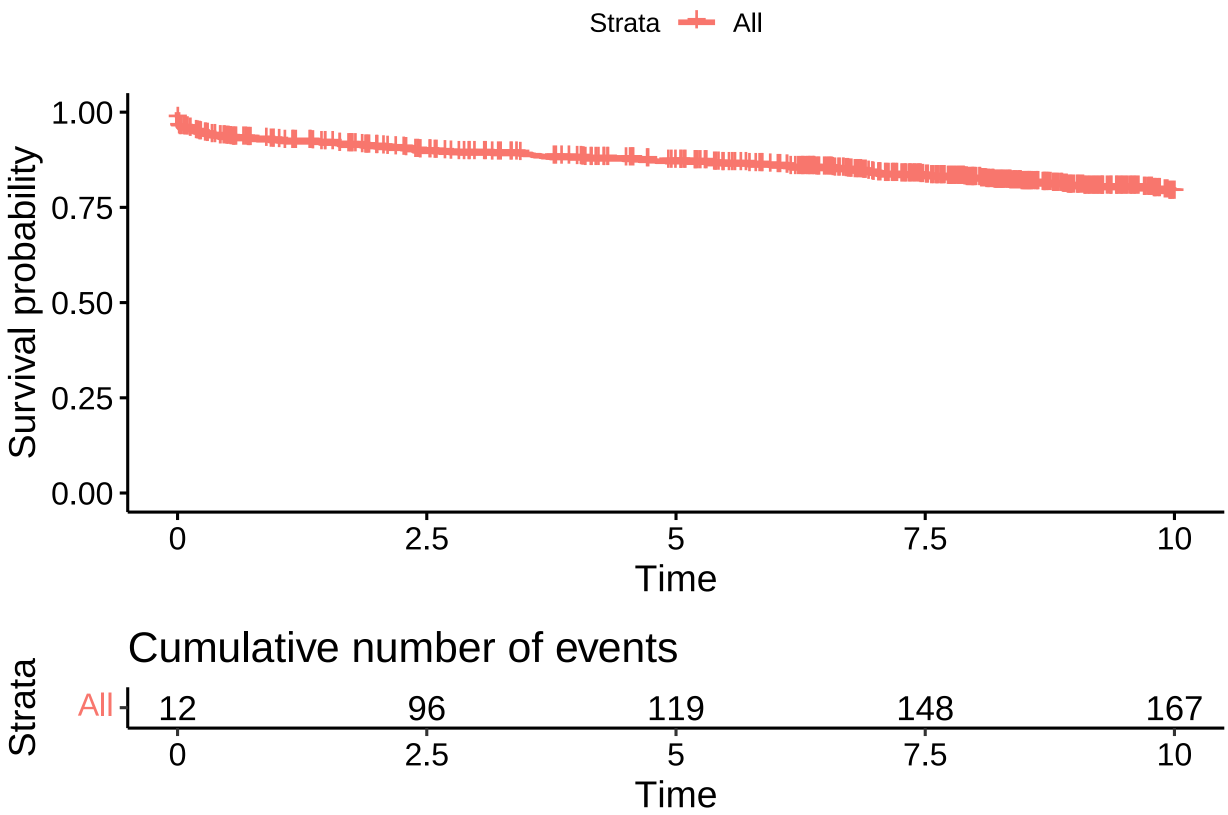
§ 3.2 IGA

§ 3.2.1

- Event: `TX Status[NTX PatientenInformation]`
- Censoring scheme: -if `TX Status[NTX PatientenInformation]` within follow up period then then event
  - if died within follow up and before no graft loss then censored
  - if last seen within follow up and before no graft loss then censored
  - if graft loss after follow up then censored
- Time period: 10 years after `Datum`

```
data_ntx <- data_ntx %>%
  mutate(status_date = case_when(
    ## patient experienced graft loss
    !is.na(Transplantatfunktionsende) & Transplantatfunktionsende <= (Datum + follow_up) ~ interval(Datu
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` < (Datum + follow_up) ~ interval(Datum, `Todesdatum[NTX Patie
    ## patiend last seen within follow up
    `Date last seen[NTX PatientenInformation]` < (Datum + follow_up) ~ interval(Datum, `Date last seen[N
    ## else follow up
    TRUE ~ interval(Datum, (Datum + follow_up)) / years(1)
    )
  )

data_ntx <- data_ntx %>%
  mutate(status = case_when(
    ## patient experienced graft loss
    !is.na(Transplantatfunktionsende) & Transplantatfunktionsende <= (Datum + follow_up) ~ 1,
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` < (Datum + follow_up) ~ 0,
    ## patiend last seen within follow up
    `Date last seen[NTX PatientenInformation]` < (Datum + follow_up) ~ 0,
    ## else follow up
    TRUE ~ 0
    )
  )
```

```
model_ntx_1 <- survfit(formula = Surv(time = status_date,
                                      event = status, type = "right") ~ 1,
                       data = data_ntx)

ggsurvplot(model_ntx_1,
           conf.int = FALSE,
           cumevents = TRUE)
```



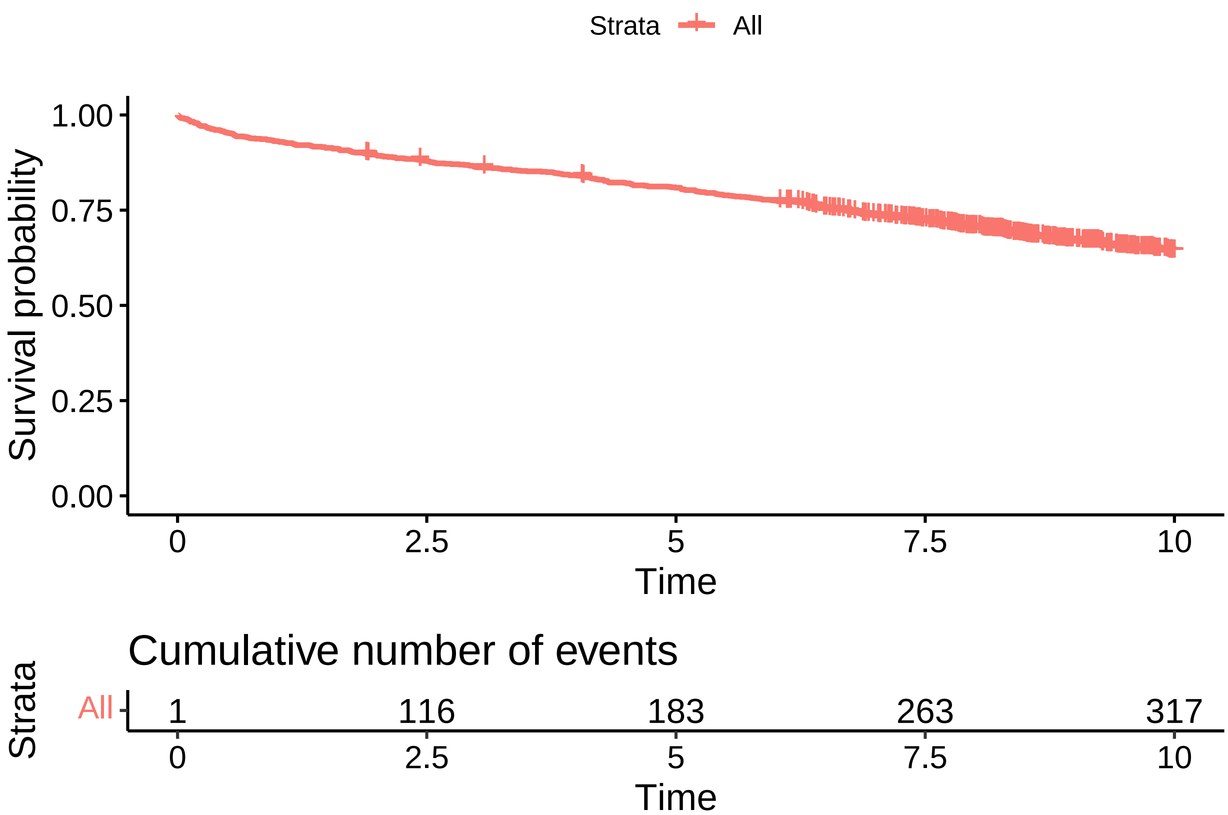
§ 3.2.2

- Event: patient died within follow up period
- Censoring scheme:
  - patient died within follow up then event
  - patient dropped from study within follow up then censored

```
data_ntx <- data_ntx %>%
  mutate(status_date = case_when(
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` <= (Datum + follow_up) ~ interval(Datum, `Todesdatum[NTX Pat
    ## patient died after follow up
    `Todesdatum[NTX PatientenInformation]` > (Datum + follow_up) ~ interval(Datum, (Datum + follow_up))
    ## patient dropped within follow up
    `Date last seen[NTX PatientenInformation]` <= (Datum + follow_up) ~ interval(Datum, `Date last seen
    ## patient dropped after follow up
    `Date last seen[NTX PatientenInformation]` > (Datum + follow_up) ~ interval(Datum, (Datum + follow_u
  )
)


data_ntx <- data_ntx %>%
  mutate(status = case_when(
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` <= (Datum + follow_up) ~ 1,
    ## patient died after follow up
    `Todesdatum[NTX PatientenInformation]` > (Datum + follow_up) ~ 0,
    ## patient dropped within follow up
```

```
    `Date last seen[NTX PatientenInformation]` <= (Datum + follow_up) ~ 0,
    ## patient dropped after follow up
    `Date last seen[NTX PatientenInformation]` > (Datum + follow_up) ~ 0
  )
)
```

```
model_ntx_2 <- survfit(Surv(time = status_date, event = status,
                            type = "right") ~1, data = data_ntx)

ggsurvplot(model_ntx_2,
           conf.int = FALSE,
           cumevents = TRUE)
```



§ 4 Cox regression

§ 4.1 IGA

- Event:
  - graft loss within follow up
- Censoring scheme:
  - if patient dropped within follow up, then censored by time dropped
  - if patient patient experienced graft loss after follow up, then censored by follow up end
  - if patient experienced death within follow up, then censored by death date else by follow up end

```
data_iga <- data_iga %>%
  ## censor/event date
  mutate(status_date = case_when(
    ## patient dropped during follow up
    (`T-dls` <= `T-date` + follow_up) ~ `T-dls`,
    ## patient experienced graft loss but after follow up
    ((!is.na(`graft loss date`)) &  (`graft loss date` > `T-date` + follow_up)) ~ `T-date` + follow_up,
    ## patient experienced graft loss within follow up
    ((!is.na(`graft loss date`)) &  (`graft loss date` <= `T-date` + follow_up)) ~ `graft loss date`,
    ## patient experienced no graft loss within follow up, neither dropped
    (is.na(`graft loss date`) & (`T-dls` > `T-date` + follow_up)) ~ `T-date` + follow_up
  )) %>%
  ## status indicator
  mutate(status = case_when(
    ## patient dropped during follow up
    (`T-dls` <= `T-date` + follow_up) ~ 0,
    ## patient experienced graft loss but after follow up
    ((!is.na(`graft loss date`)) &  (`graft loss date` > `T-date` + follow_up)) ~ 0,
    ## patient experienced graft loss within follow up
    ((!is.na(`graft loss date`)) &  (`graft loss date` <= `T-date` + follow_up)) ~ 1,
    ## patient experienced no graft loss within follow up, neither dropped
    (is.na(`graft loss date`) & (`T-dls` > `T-date` + follow_up)) ~ 0
  ))
```

```
data_iga <- data_iga %>%
  ## censor/event date
  mutate(status_date = case_when(
    ## if graft loss within follow up
    `graft loss date` <= `T-date` + follow_up ~ `graft loss date`,
    ## else
    TRUE ~ `T-dls`
  )) %>%
  ## censor/event indicator
  mutate(status = case_when(
    ## if graft loss within follow up
    data_iga$`graft loss date` <= data_iga$`T-date` + follow_up ~ 1,
    ## else
    TRUE ~ 0
  ))
```

```
model_iga_cox <- coxph(data = data_iga, formula = Surv(time = as.numeric(status_date),
                                     event = status) ~ R_age_Tdate +
                         data_iga$`R-sex` +
                         data_iga$`D-type` +
```

```r
                    ## hla mismatch???
                    data_iga$`mm-A` +
                    data_iga$`Cold ischaemic period hours` +
                    data_iga$`D-age`)
```

```
## Warning in fitter(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 5 ; coefficient may be infinite.
```

```r
summary(model_iga_cox)
```

```
## Call:
## coxph(formula = Surv(time = as.numeric(status_date), event = status) ~
##     R_age_Tdate + data_iga$`R-sex` + data_iga$`D-type` + data_iga$`mm-A` +
##         data_iga$`Cold ischaemic period hours` + data_iga$`D-age`,
##     data = data_iga)
##
##   n= 38, number of events= 9
##    (26 observations deleted due to missingness)
##
##                                            coef  exp(coef)   se(coef)      z
## R_age_Tdate                          -1.101e-02  9.891e-01  3.584e-02 -0.307
## data_iga$`R-sex`M                     1.284e-01  1.137e+00  8.645e-01  0.149
## data_iga$`D-type`Living                      NA         NA  0.000e+00     NA
## data_iga$`mm-A`1                     -3.635e-02  9.643e-01  8.057e-01 -0.045
## data_iga$`mm-A`2                     -1.833e+01  1.100e-08  1.197e+04 -0.002
## data_iga$`Cold ischaemic period hours` -5.066e-03  9.949e-01  6.792e-02 -0.075
## data_iga$`D-age`                      4.052e-02  1.041e+00  2.901e-02  1.397
##                                      Pr(>|z|)
## R_age_Tdate                             0.759
## data_iga$`R-sex`M                       0.882
## data_iga$`D-type`Living                    NA
## data_iga$`mm-A`1                        0.964
## data_iga$`mm-A`2                        0.999
## data_iga$`Cold ischaemic period hours`  0.941
## data_iga$`D-age`                        0.162
##
##                                      exp(coef) exp(-coef) lower .95 upper .95
## R_age_Tdate                          9.891e-01  1.011e+00    0.9220     1.061
## data_iga$`R-sex`M                    1.137e+00  8.795e-01    0.2089     6.190
## data_iga$`D-type`Living                     NA         NA        NA        NA
## data_iga$`mm-A`1                     9.643e-01  1.037e+00    0.1988     4.677
## data_iga$`mm-A`2                     1.100e-08  9.091e+07    0.0000       Inf
## data_iga$`Cold ischaemic period hours` 9.949e-01  1.005e+00    0.8709     1.137
## data_iga$`D-age`                     1.041e+00  9.603e-01    0.9838     1.102
##
## Concordance= 0.701  (se = 0.095 )
## Likelihood ratio test= 3.73  on 6 df,   p=0.7
## Wald test            = 2.19  on 6 df,   p=0.9
## Score (logrank) test = 3.02  on 6 df,   p=0.8
```

```r
cox.zph(model_iga_cox)
```

```
##                                       chisq df    p
## R_age_Tdate                          2.2969  1 0.13
## data_iga$`R-sex`                     2.5696  1 0.11
```

```
## data_iga$`mm-A`                          0.0577  2 0.97
## data_iga$`Cold ischaemic period hours` 0.4896  1 0.48
## data_iga$`D-age`                         0.5799  1 0.45
## GLOBAL                                   3.6546  6 0.72
```

```
## all p-values are relatively large, therefore the Null hypothesis of proportional hazards can not be r
```

$ 4.2 NTX

```
data_ntx <- data_ntx %>%
  mutate(status_date = case_when(
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` <= (Datum + follow_up) ~ interval(Datum, `Todesdatum[NTX Pati
    ## patient died after follow up
    `Todesdatum[NTX PatientenInformation]` > (Datum + follow_up) ~ interval(Datum, Datum + follow_up) /
    ## patient dropped within follow up
    `Date last seen[NTX PatientenInformation]` <= (Datum + follow_up) ~ interval(Datum, `Date last seen
    ## patient dropped after follow up
    `Date last seen[NTX PatientenInformation]` > (Datum + follow_up) ~ interval(Datum, Datum + follow_up
  )
)

data_ntx <- data_ntx %>%
  mutate(status = case_when(
    ## patient died within follow up
    `Todesdatum[NTX PatientenInformation]` <= (Datum + follow_up) ~ 1,
    ## patient died after follow up
    `Todesdatum[NTX PatientenInformation]` > (Datum + follow_up) ~ 0,
    ## patient dropped within follow up
    `Date last seen[NTX PatientenInformation]` <= (Datum + follow_up) ~ 0,
    ## patient dropped after follow up
    `Date last seen[NTX PatientenInformation]` > (Datum + follow_up) ~ 0

  )
)
```

```
model_ntx_cox <- coxph(formula = Surv(time = status_date, event = status) ~ R_age_Datum +
                         Geschlecht + `TX Status[NTX PatientenInformation]`,
                       data = data_ntx)
```

```
## Warning in fitter(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 4 ; coefficient may be infinite.
```

```
summary(model_ntx_cox)
```

```
## Call:
## coxph(formula = Surv(time = status_date, event = status) ~ R_age_Datum +
##     Geschlecht + `TX Status[NTX PatientenInformation]`, data = data_ntx)
##
##   n= 960, number of events= 317
##
##                                                                    coef
## R_age_Datum                                                   6.693e-02
## Geschlechtweiblich                                           -2.412e-01
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion  7.966e-01
```

```
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion  -1.385e+01
##                                                                     exp(coef)
## R_age_Datum                                                         1.069e+00
## Geschlechtweiblich                                                  7.857e-01
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion  2.218e+00
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion   9.645e-07
##                                                                      se(coef)
## R_age_Datum                                                         5.395e-03
## Geschlechtweiblich                                                  1.224e-01
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion  1.308e-01
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion   1.086e+03
##                                                                          z
## R_age_Datum                                                        12.406
## Geschlechtweiblich                                                 -1.971
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion  6.089
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion  -0.013
##                                                                    Pr(>|z|)
## R_age_Datum                                                         < 2e-16 ***
## Geschlechtweiblich                                                   0.0487 *
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion 1.14e-09 ***
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion    0.9898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                                                                    exp(coef)
## R_age_Datum                                                        1.069e+00
## Geschlechtweiblich                                                 7.857e-01
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion 2.218e+00
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion  9.645e-07
##                                                                    exp(-coef)
## R_age_Datum                                                         9.353e-01
## Geschlechtweiblich                                                  1.273e+00
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion  4.509e-01
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion   1.037e+06
##                                                                    lower .95
## R_age_Datum                                                           1.0580
## Geschlechtweiblich                                                    0.6182
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion    1.7163
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion     0.0000
##                                                                    upper .95
## R_age_Datum                                                           1.0806
## Geschlechtweiblich                                                    0.9987
## `TX Status[NTX PatientenInformation]`2 - ohne Transplantatfunktion    2.8663
## `TX Status[NTX PatientenInformation]`2- ohne Transplantatfunktion        Inf
##
## Concordance= 0.716  (se = 0.015 )
## Likelihood ratio test= 211.1  on 4 df,   p=<2e-16
## Wald test            = 176.5  on 4 df,   p=<2e-16
## Score (logrank) test = 189.1  on 4 df,   p=<2e-16
```

```
cox.zph(model_ntx_cox)
```

```
##                                        chisq df     p
## R_age_Datum                            3.600  1 0.058
## Geschlecht                             0.336  1 0.562
```

```
## `TX Status[NTX PatientenInformation]` 0.641  2 0.726
## GLOBAL                                4.301  4 0.367
```

```
## all p-values are relatively large, therefore the Null hypothesis of proportional hazards can not be r
```