DECK with R

Umberto Maggiore

2025-10-04

ADJUSTED ANALYSIS OF eGFR TRAJECTORIES (INCLUDING SPLINES)

```r
# Load necessary libraries
library(haven)
library(tidyverse)
library(lme4)
library(splines)
library(ggplot2)
library(dplyr)
library(ggeffects)
library(car)
library(kableExtra)
library(jtools)
library(Hmisc)

# transform data
deck_long <- read_dta("C:/Documenti/Furian/DECK/deck_long.dta")
deck_long <- deck_long %>%
  filter(pz_categoria == 2 | pz_categoria == 4) %>%
  filter(!is.na(eGFR_EPI2021))

deck_long$pz_categoria <- factor(deck_long$pz_categoria, levels = c(2, 4), labels = c("CIK", "LD"))
deck_long$volume_attivitàDEC_centrotx <- factor(deck_long$volume_attivitàDEC_centrotx,
                                        levels = c(1, 2), labels = c("<70", "70+"))
deck_long$ric_sesso <- factor(deck_long$ric_sesso, levels = c(1, 2), labels = c("Female", "Male"))

deck_long <- deck_long %>%
  mutate(across(c(don_quality_score, don_eta, ric_CIT, ric_eta, ric_PRA), scale))

# Perform eGFR using splines
deck_long <- deck_long %>%
  mutate(mese_scaled = (mese - 1) / (72 - 1))  # Assuming mese ranges from 1 to 60

# Define knots at mese = 6, 24, 48 and rescale them
knots_mese <- c(6, 24, 48)
knots_scaled <- (knots_mese - 1) / (72 - 1)

# Generate RCS basis using these knots
rcs_terms <- rcspline.eval(deck_long$mese_scaled, knots = knots_scaled, inclx = TRUE)

# Add spline terms to the dataset
```

```
deck_long <- deck_long %>%
  mutate(rcs_mese_1 = rcs_terms[, 2])

# Fit the model and test interaction term including spline
model.a <- lmer(eGFR_EPI2021 ~ mese_scaled * pz_categoria + rcs_mese_1 * pz_categoria +
                  don_quality_score + don_eta + volume_attivitàDEC_centrotx +
                  ric_CIT + ric_eta + ric_sesso + ric_PRA +
                  (mese_scaled | id), data = deck_long)

linearHypothesis(model.a,
                 c("mese_scaled:pz_categoriaLD",
                   "pz_categoriaLD:rcs_mese_1"))
```

```
##
## Linear hypothesis test:
## mese_scaled:pz_categoriaLD = 0
## pz_categoriaLD:rcs_mese_1 = 0
##
## Model 1: restricted model
## Model 2: eGFR_EPI2021 ~ mese_scaled * pz_categoria + rcs_mese_1 * pz_categoria +
##     don_quality_score + don_eta + volume_attivitàDEC_centrotx +
##     ric_CIT + ric_eta + ric_sesso + ric_PRA + (mese_scaled |
##     id)
##
##   Df  Chisq Pr(>Chisq)
## 1
## 2   2 1.5131     0.4693
```

```
test.int <- linearHypothesis(model.a,
                             c("mese_scaled:pz_categoriaLD",
                               "pz_categoriaLD:rcs_mese_1"))

round(test.int$"Pr(>Chisq)"[2], 2)
```

```
## [1] 0.47
```

Make the plot

```
# Generate the prediction grid
mese_seq <- seq(0, 1, length.out = 100)
pz_levels <- levels(deck_long$pz_categoria)
new_data <- expand.grid(
  mese_scaled = mese_seq,
  pz_categoria = pz_levels
)
# Calculate the splines for the grid
new_data$rcs_mese_1 <- rcspline.eval(new_data$mese_scaled, knots = knots_scaled, inclx = TRUE)[, 2]
# Add average values for covariates
covariates <- c("don_quality_score", "don_eta", "ric_CIT", "ric_eta", "ric_PRA")
for (var in covariates) {
```

```r
  new_data[[var]] <- mean(deck_long[[var]], na.rm = TRUE)
}
# Expand grid to include all combinations of factor levels
factor_grid <- expand.grid(
  volume_attivitàDEC_centrotx = levels(deck_long$volume_attivitàDEC_centrotx),
  ric_sesso = levels(deck_long$ric_sesso)
)
# Repeat new_data for each combination of factor levels
new_data <- new_data[rep(1:nrow(new_data), each = nrow(factor_grid)), ]
new_data <- cbind(new_data, factor_grid[rep(1:nrow(factor_grid), times = nrow(new_data) / nrow(factor_g
# Predictions and associated confidence intervals
pred <- predict(model.a, newdata = new_data, re.form = NA, se.fit = TRUE)
new_data$predicted <- pred$fit
new_data$conf.low <- pred$fit - 1.96 * pred$se.fit
new_data$conf.high <- pred$fit + 1.96 * pred$se.fit
# Average over factor levels
new_data_avg <- new_data %>%
  group_by(mese_scaled, pz_categoria) %>%
  summarise(
    predicted = mean(predicted),
    conf.low = mean(conf.low),
    conf.high = mean(conf.high),
    .groups = "drop"
  )
# cut at month > 60
x.up.lim <- (60-1) / (72-1)
# Plot with confidence bands
ggplot(new_data_avg, aes(x = mese_scaled, y = predicted, color = pz_categoria)) +
  geom_line(linewidth = 1.2) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high, fill = pz_categoria), alpha = 0.2, color = NA) +
  scale_color_manual(values = c("CIK" = "blue", "LD" = "red")) +
  scale_fill_manual(values = c("CIK" = "blue", "LD" = "red")) +
  guides(color = guide_legend(title = NULL), fill = guide_legend(title = NULL)) +
  labs(x = "Month since Transplantation", y = "eGFR (ml/min/1.73m²)") +
  scale_x_continuous(breaks = c(unique(deck_long$mese_scaled)),
                     labels = c(1, 3, 6, 12, 24, 36, 48, 60), limits = c(0, x.up.lim)) +
  scale_y_continuous(breaks = c(0,15,30,45,60,90,120), limits = c(0, NA)) +
  theme_minimal() +
  theme(
    panel.grid.minor = element_blank(),
    axis.line = element_line(color = "black", linewidth = 0.5)
  ) +
  geom_line(data = deck_long, aes(x = mese_scaled, y = eGFR_EPI2021, group = id,
                                  color = pz_categoria), alpha = 0.3) +
  annotate("text", x = Inf, y = -Inf, label = paste("P value for different slopes=", round(test.int$"Pr
           hjust = 1.1, vjust = -1.5, size = 4, color = "black")
```
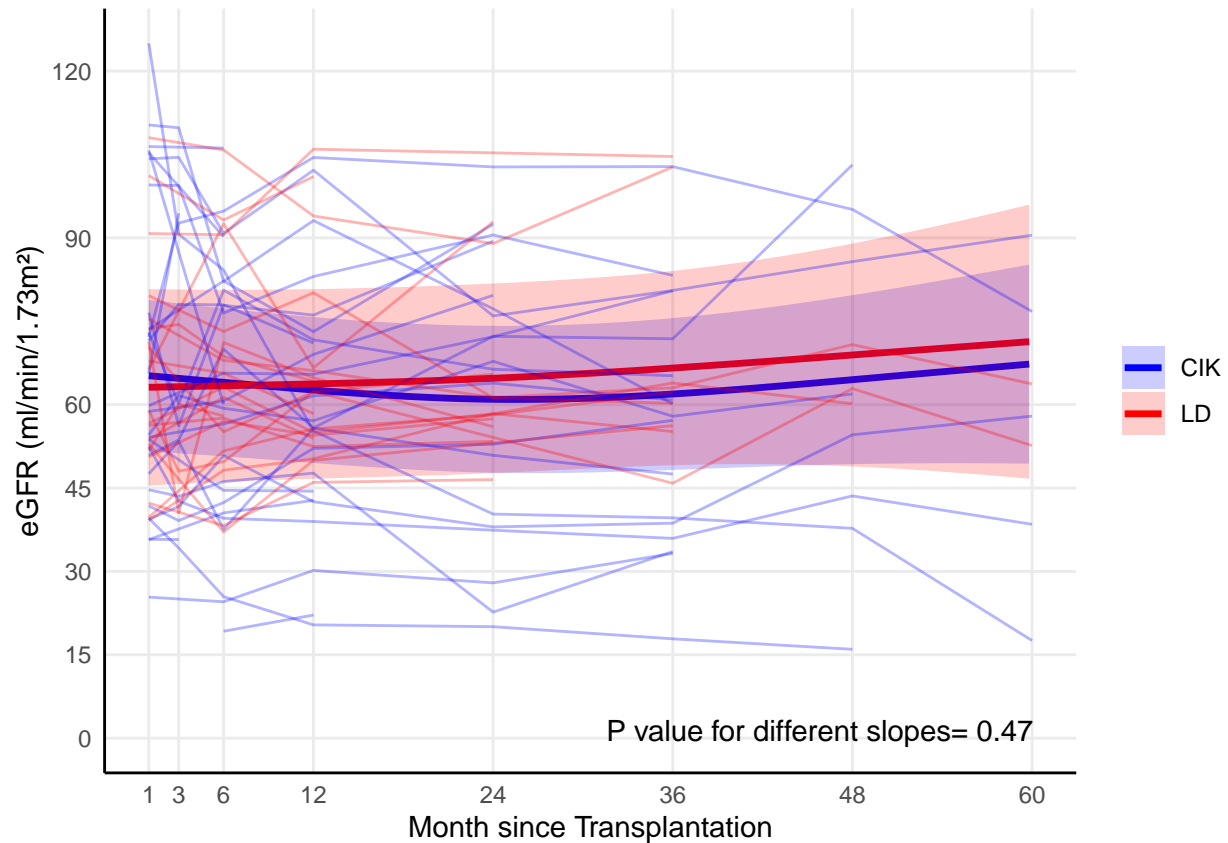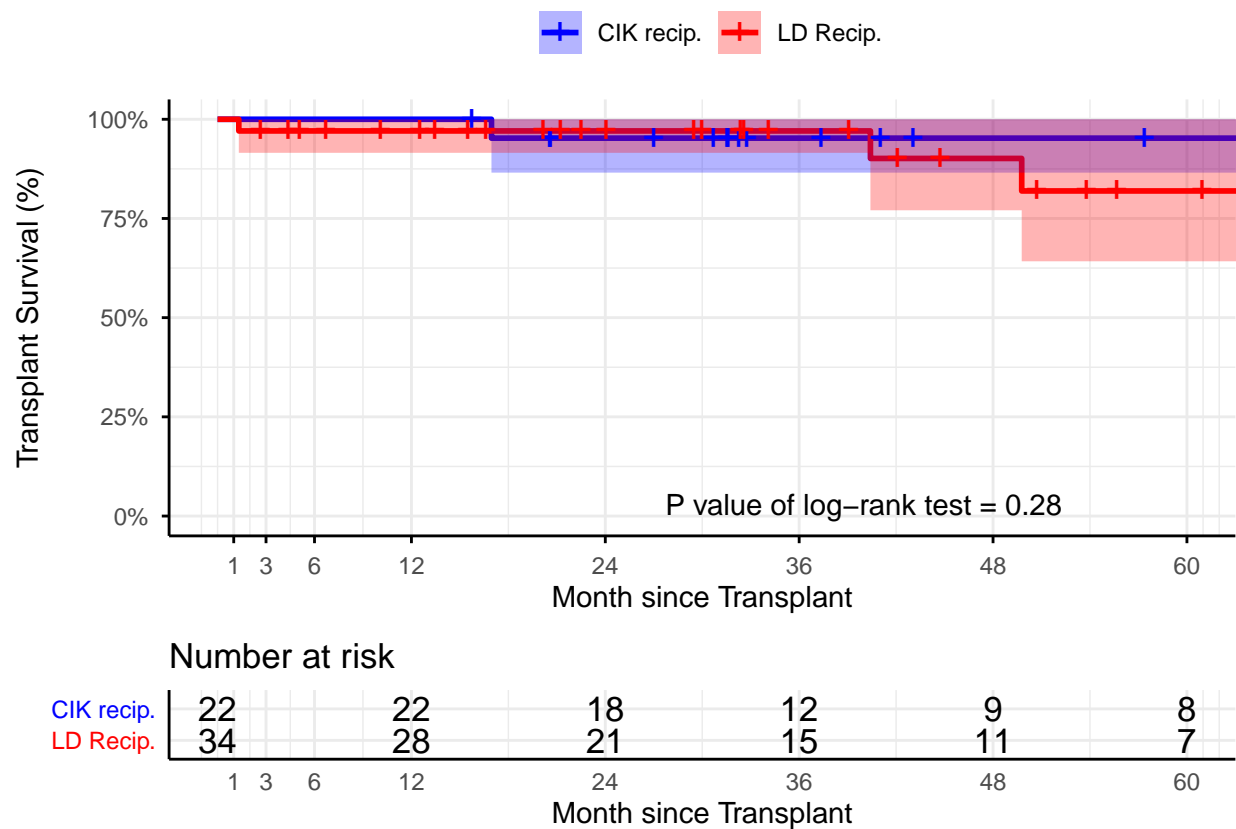
Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

ADJUSTED DIFFERENCE BETWEEN LINEAR eGFR SLOPES (NOT IN-CLUDING SPLINES)

```
## # A tibble: 1 x 5
##   term                        `Linear Diff` CI          `P value` Model
##   <chr>                               <dbl> <chr>           <dbl> <chr>
## 1 I(mese/12):pz_categoriaCIK             -2 -6.3 to 2.3      0.35 Adjusted
```

SURVIVAL ANALYSIS (KAPLAN-MEIER PLOT AND LOG-RANK TEST)

CRUDE AND ADJUSTED COX PH REGRESSION MODEL

```
## # A tibble: 2 x 5
##   term                    HR CI             `P value` Model
##   <chr>                <dbl> <chr>              <dbl> <chr>
## 1 factor(pz_categoria)CIK  3.13 0.35 - 28.09      0.31 Crude
## 2 factor(pz_categoria)CIK  2.25 0.12 - 43.21      0.59 Adjusted
```