1.Qc_data_Checking

NikolayNikolaev

2023-10-26

Set Enviorment

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.1
```

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.3.1
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.3.1
```

```r
library(xgxr)
```

```
## Warning: package 'xgxr' was built under R version 4.3.1
```

```r
#flag for labeling figures as draft
status = "DRAFT"

# ggplot settings
xgx_theme_set()

#directories for saving individual graphs
```

```
dirs = list(
  parent_dir = "D:/REPOSITORIES/PkPd-visul-exploration-",
  rscript_dir  = "./",
  rscript_name = "Example.R",
  results_dir  = "./",
  filename_prefix   = "",
  filename      = "Example.png")
```

Load Dataset

```
sd_pkpd_data  <- read.csv("data/Single_Ascending_Dose_Dataset2.csv")
#mul_pkpd_data <- read.csv("data/Multiple_Ascending_Dose_Dataset2.csv")

# set columns
sd_pkpd_data = sd_pkpd_data %>%
    mutate(ID      = ID,     #ID   column
           TIME    = TIME,   #TIME column name, time relative to first dose
           NOMTIME = NOMTIME,#NOMINAL TIME column name
           EVID    = EVID   ,#EVENT ID, >=1 is dose, otherwise measurement
           LIDV    = LIDV,   #DEPENDENT VARIABLE column name
           CENS    = CENS,   #CENSORING column name
           CMT     = CMT,    #COMPARTMENT column
           DOSE    = DOSE,   #DOSE column here (numeric value)
           TRTACT  = TRTACT, #DOSE REGIMEN column here (character, with units),
           LIDV_NORM = LIDV/DOSE,
           LIDV_UNIT    = ifelse(CMT==2, "ng/ml", NA )
    )

# create a factor for the treatment vaiable for plotting
sd_pkpd_data = sd_pkpd_data %>%
  arrange(DOSE) %>%
  mutate(TRTACT_low2high = factor(TRTACT, levels = unique(TRTACT)),
         TRTACT_high2low = factor(TRTACT, levels = rev(unique(TRTACT)))) %>%
  select(-TRTACT)

#create pk dataset
pk_data <- sd_pkpd_data %>% filter(CMT==2)

#perform NCA, for additional plots
NCA = pk_data %>%
  group_by(ID, DOSE) %>%
  filter(!is.na(LIDV)) %>%
  summarize(AUC_last = caTools::trapz(TIME,LIDV),
            Cmax     = max(LIDV),
            SEX      = SEX[1], #this part just keeps the SEX and WEIGHTB covariates
            WEIGHTB  = WEIGHTB[1]) %>%
  gather(PARAM, VALUE,-c(ID, DOSE, SEX, WEIGHTB)) %>%
  ungroup() %>%
  mutate(VALUE_NORM = VALUE/DOSE)
```

```
## 'summarise()' has grouped output by 'ID'. You can override using the '.groups'
```

```
## argument.
```
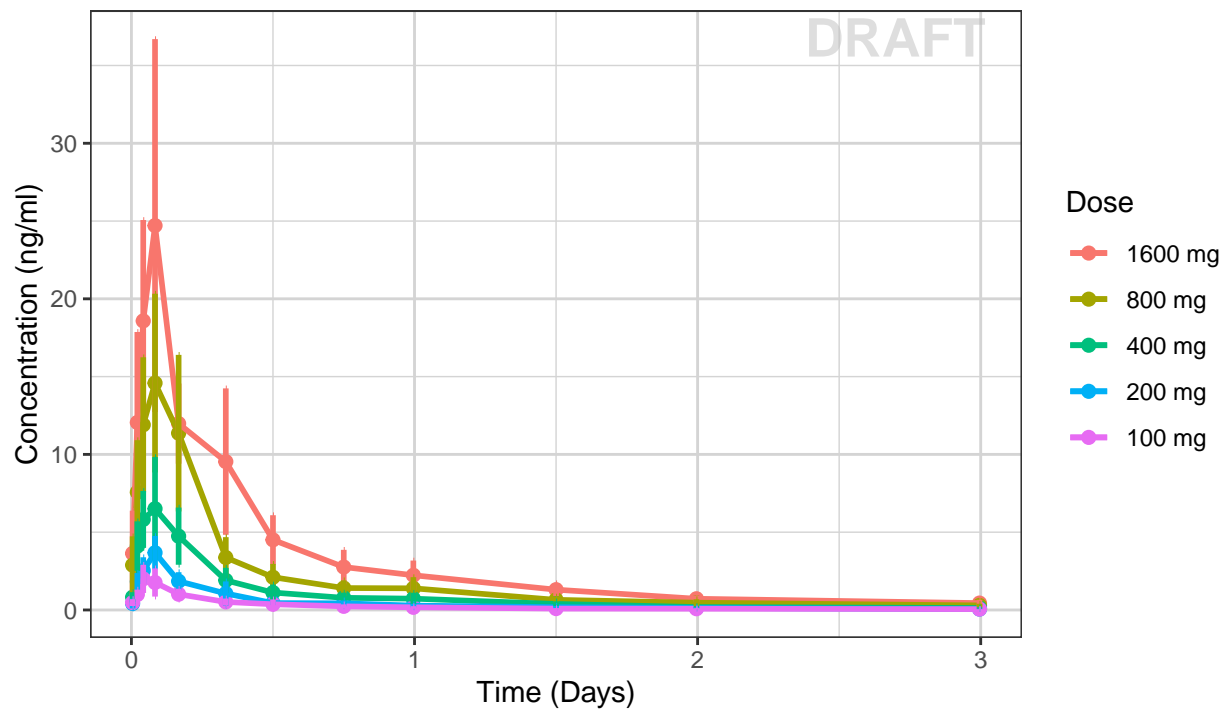
```
#units and labels
  time_units_dataset = "hours"
  time_units_plot    = "days"
  trtact_label       = "Dose"
  dose_label         = "Dose (mg)"
  conc_units         = "ng/ml"
  AUC_units          = paste0("h.", conc_units)
  conc_label         = paste0("Concentration (", conc_units, ")")
  concnorm_label     = paste0("Normalized Concentration (", conc_units, ")/mg")
```

COncentration over time, colored by Dose, mean+/- 95% CI

```
gg <- ggplot(data = pk_data,
             aes(x = NOMTIME,
                 y = LIDV,
                 group= TRTACT_high2low,
                 color = TRTACT_high2low))
gg <- gg + xgx_stat_ci(conf_level = .95)
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y=conc_label,
                color = trtact_label)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```

```
## Warning: Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
```
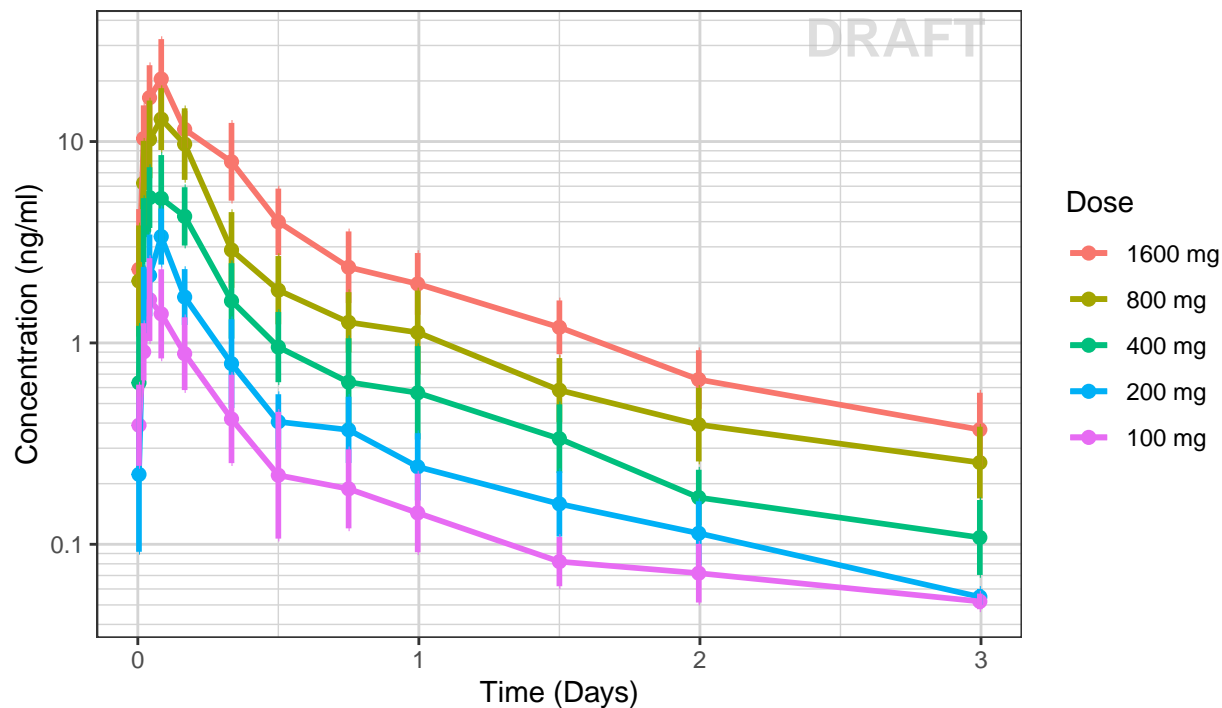
D:/REPOSITORIES/PkPd−visul−exploration−
.//Example.R
.//Example.png
Created: 2023−10−26 22:08:10.64162

```
print(gg + xgx_scale_y_log10() )
```

```
## Warning: Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
```

D:/REPOSITORIES/PkPd–visul–exploration–
.//Example.R
.//Example.png
Created: 2023–10–26 22:08:10.64162

Concentration over time, faceted by Dose, mean +/- 95% CI, overlaid on gray spaghetti plots

```r
gg <- ggplot(data = pk_data, aes(x = TIME, y = LIDV))
gg <- gg + geom_line(aes(group = ID), color = rgb(0.5,0.5,0.5), size = 1, alpha = 0.3)
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```r
gg <- gg + geom_point(aes(color = factor(CENS), shape = factor(CENS), alpha = 0.3), size = 2, alpha = 0
gg <- gg + scale_shape_manual(values=c(1,8))
gg <- gg + scale_color_manual(values=c("grey50","red"))
gg <- gg + xgx_stat_ci(aes(x = NOMTIME, color=NULL, group=NULL), conf_level = 0.95)
gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y=conc_label,color = trtact_label)
gg <- gg + theme(legend.position="none") + facet_grid(.~TRTACT_low2high)
gg <- gg + xgx_annotate_status(status)
```
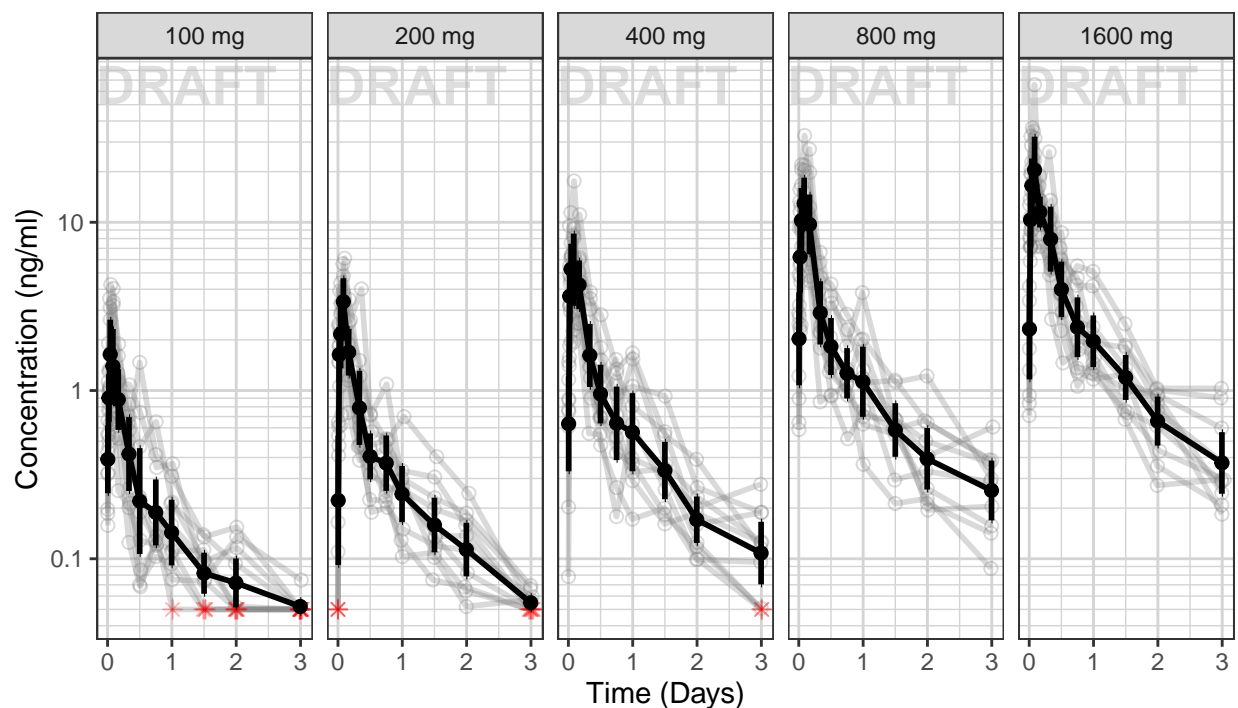
```
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```

Warning: Removed 50 rows containing non-finite values ('stat_summary()').

Warning: Removed 50 rows containing non-finite values ('stat_summary()').
Removed 50 rows containing non-finite values ('stat_summary()').

Warning: Removed 50 rows containing missing values ('geom_line()').

Warning: Removed 50 rows containing missing values ('geom_point()').

D:/REPOSITORIES/PkPd–visul–exploration–
.//Example.R
.//Example.png
Created: 2023–10–26 22:08:12.079077

Assess the dose linearity of exposure

Dose Normalized Concentration over time, colored by Dose, mean +/- 95% CI

```
gg <- ggplot(data = pk_data,
             aes(x = NOMTIME,
                 y = LIDV_NORM,
                 group = TRTACT_high2low,
```

```
                color = TRTACT_high2low))
gg <- gg + xgx_stat_ci(conf_level = 0.95,
                       alpha = 0.5,
                       position = position_dodge(1))
gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y=concnorm_label, color = trtact_label)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
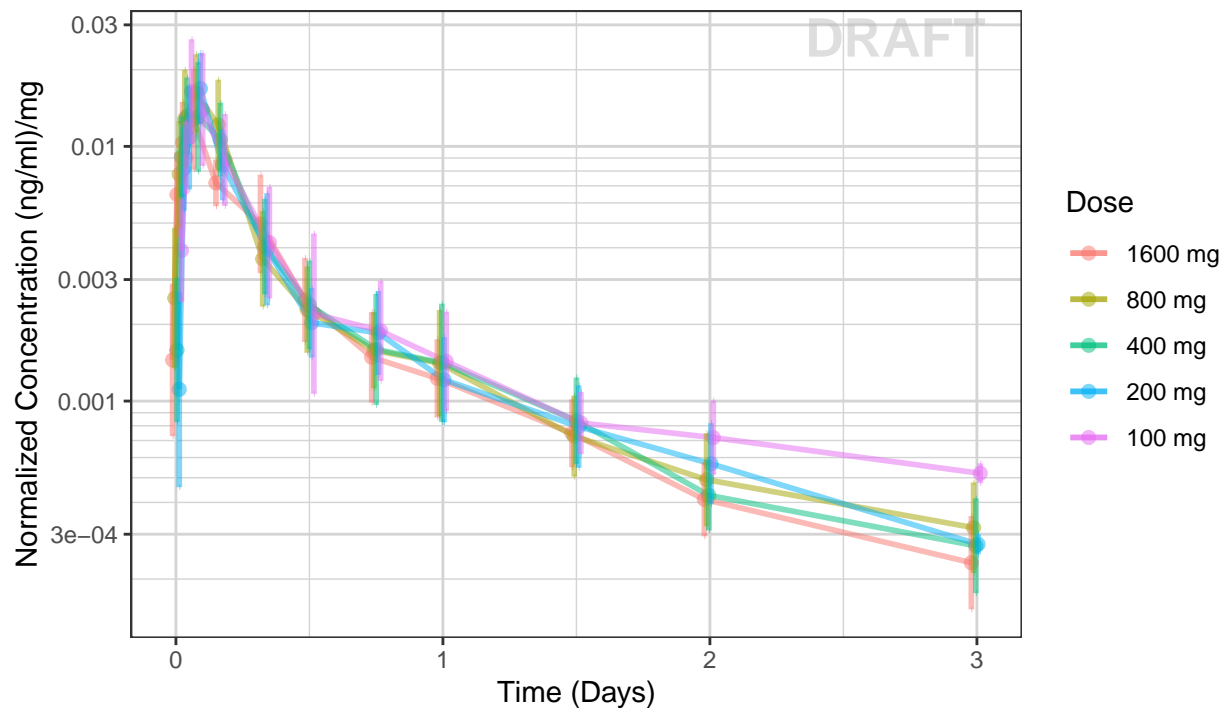
```
## Warning: Removed 50 rows containing non-finite values (`stat_summary()`).
## Removed 50 rows containing non-finite values (`stat_summary()`).
## Removed 50 rows containing non-finite values (`stat_summary()`).
```

```
## Warning: `position_dodge()` requires non-overlapping x intervals
## `position_dodge()` requires non-overlapping x intervals
```


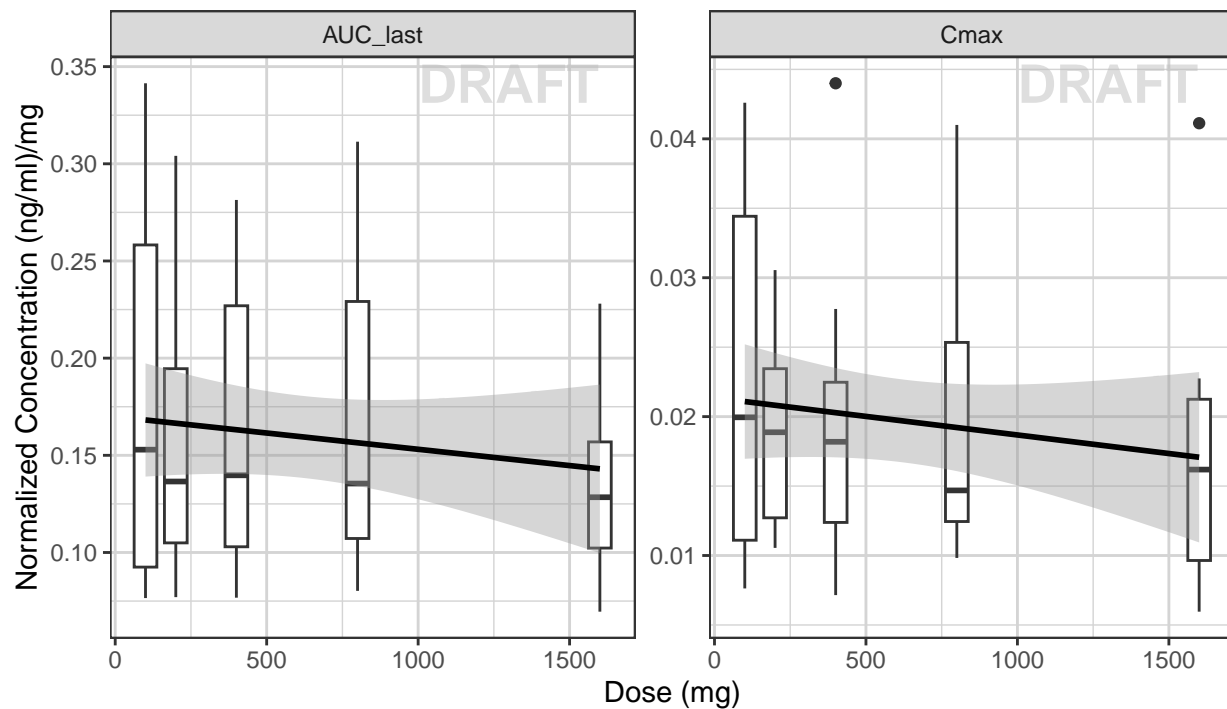NCA of dose normalized AUC and Cmax vs Dose

```
gg <- ggplot(data = NCA, aes(x = DOSE, y = VALUE_NORM))
gg <- gg + geom_boxplot(aes(group = DOSE))
gg <- gg + geom_smooth(method = "lm", color = "black")
gg <- gg + facet_wrap(~PARAM, scales = "free_y")
gg <- gg + labs(x = dose_label, y = concnorm_label)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```

`geom_smooth()` using formula = 'y ~ x'

D:/REPOSITORIES/PkPd–visul–exploration–
.//Example.R
.//Example.png
Created: 2023–10–26 22:08:13.602167

Explore variability

Concentration over time, colored by Dose, dots and lines grouped by individual

```
gg <- ggplot(data = pk_data,
             aes(x = TIME, y = LIDV))
gg <- gg + geom_line(aes(group = ID,
                         color = factor(TRTACT_high2low)),
                     size = 1,
```

```
                           alpha = 0.5)
gg <- gg + geom_point(data = pk_data %>%
                            filter(CENS==0),
                          aes(color = TRTACT_high2low),
                          size = 2,
                          alpha = 0.5)
gg <- gg + geom_point(data = pk_data %>%
                            filter(CENS==1),
                          color="red",
                          shape=8,
                          size = 2,
                          alpha = 0.5)
gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                     units_plot    = time_units_plot)
gg <- gg + labs(y = conc_label,
                 color = trtact_label)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
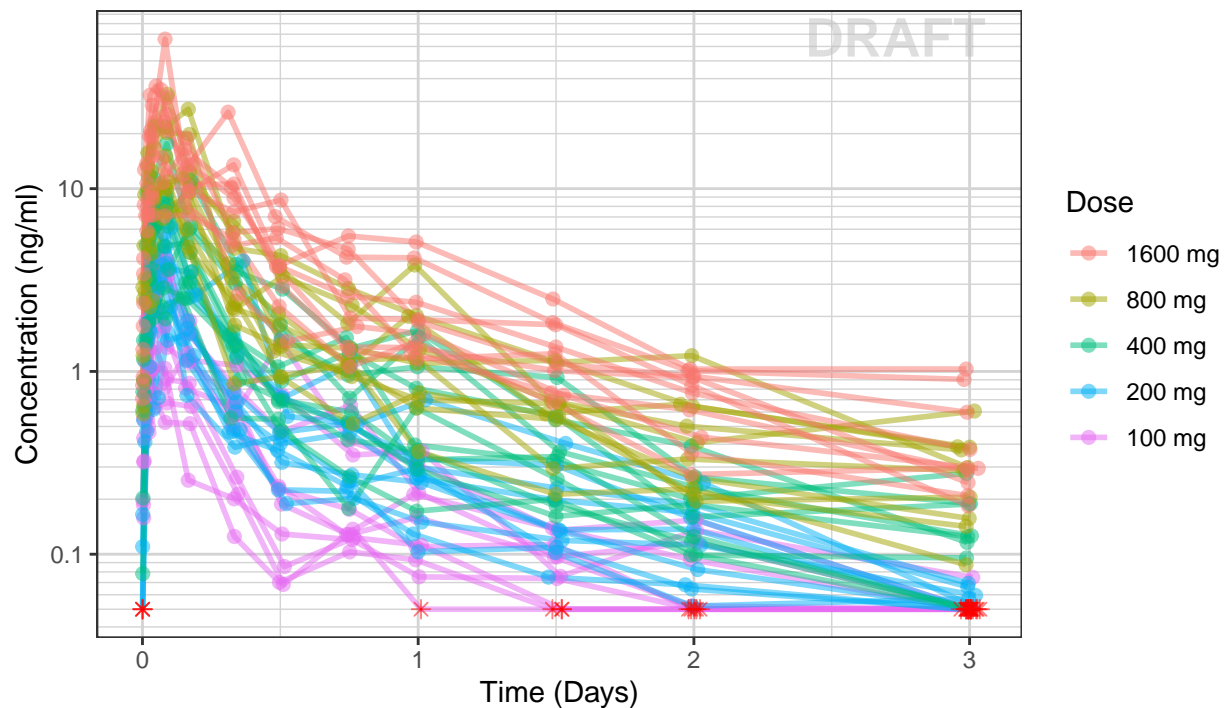
```
## Warning: Removed 50 rows containing missing values (‘geom_line()‘).
```

```
## Warning: Removed 50 rows containing missing values (‘geom_point()‘).
```

Concentration over time, faceted by Dose, dots and lines grouped by individual

```r
gg <- ggplot(data = pk_data,
             aes(x = TIME,
                 y = LIDV))
gg <- gg + geom_line(aes(group = ID),
                     size = 1,
                     alpha = 0.2)

gg <- gg + geom_point(aes(color = factor(CENS),
                          shape = factor(CENS),
                          alpha = 0.3),
                      size = 2,
                      alpha = 0.2)
gg <- gg + scale_shape_manual(values=c(1,8))
gg <- gg + scale_color_manual(values=c("grey50","red"))

gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y = conc_label,
                shape = "BLQ",
                color = "BLQ")

gg <- gg + facet_grid(.~TRTACT_low2high)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
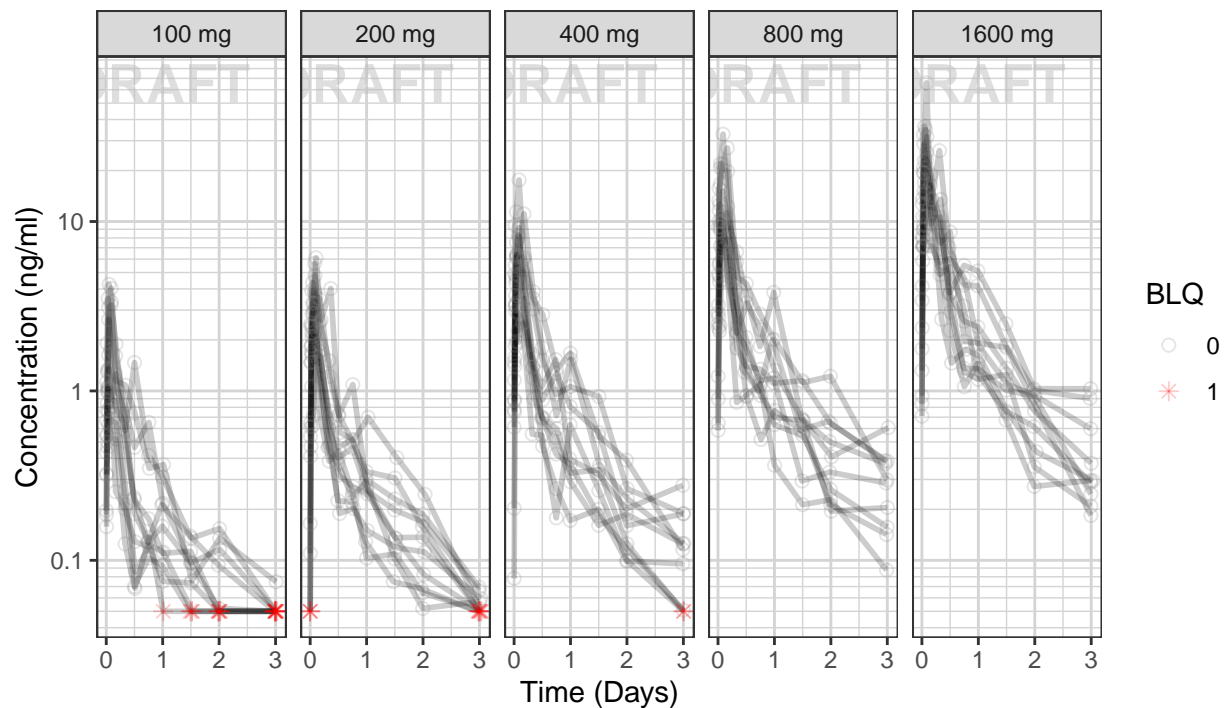
```
## Warning: Removed 50 rows containing missing values (`geom_line()`).
```

```
## Warning: Removed 50 rows containing missing values (`geom_point()`).
```

D:/REPOSITORIES/PkPd−visul−exploration−
.//Example.R
.//Example.png
Created: 2023−10−26 22:08:14.775804

Explore irregularities in profiles

1. Concentration over time, faceted by individual, individual line plots overlaid on gray spaghetti plots for that dose group

```r
pk_data_rep_by_trt <- list()
for(id in unique(pk_data$ID)){
  indiv_data <- pk_data %>% subset(ID == id)
  itrtact = unique(indiv_data$TRTACT_low2high)

  pk_data_rep_by_trt[[as.character(id)]] <- pk_data %>%
      subset(TRTACT_low2high == itrtact) %>%
      mutate(ID_rep_by_trt = ID, ID = id)
}
pk_data_rep_by_trt <- bind_rows(pk_data_rep_by_trt)

gg <- ggplot(mapping = aes(x = TIME, y = LIDV))
gg <- gg + geom_line(data = pk_data_rep_by_trt,
                aes(group = ID_rep_by_trt),
                size = 1, color = rgb(0.5,0.5,0.5), alpha = 0.3)
gg <- gg + geom_line(data = pk_data,
                aes(group = ID), size = 1)
gg <- gg + geom_point(data = pk_data %>% filter(CENS==1),
```

```
                    color="red",  shape=8, size = 2)

gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y = conc_label)
gg <- gg + theme(legend.position="none")
gg <- gg + facet_wrap(~ID + TRTACT_low2high,
            ncol = 10 )
gg <- gg + theme(panel.grid.minor.x = ggplot2::element_line(color = rgb(0.9,0.9,0.9)),
     panel.grid.minor.y = ggplot2::element_line(color = rgb(0.9,0.9,0.9)))
gg <- gg + xgx_annotate_status(status, fontsize = 4, color=rgb(0.5,0.5,1))
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
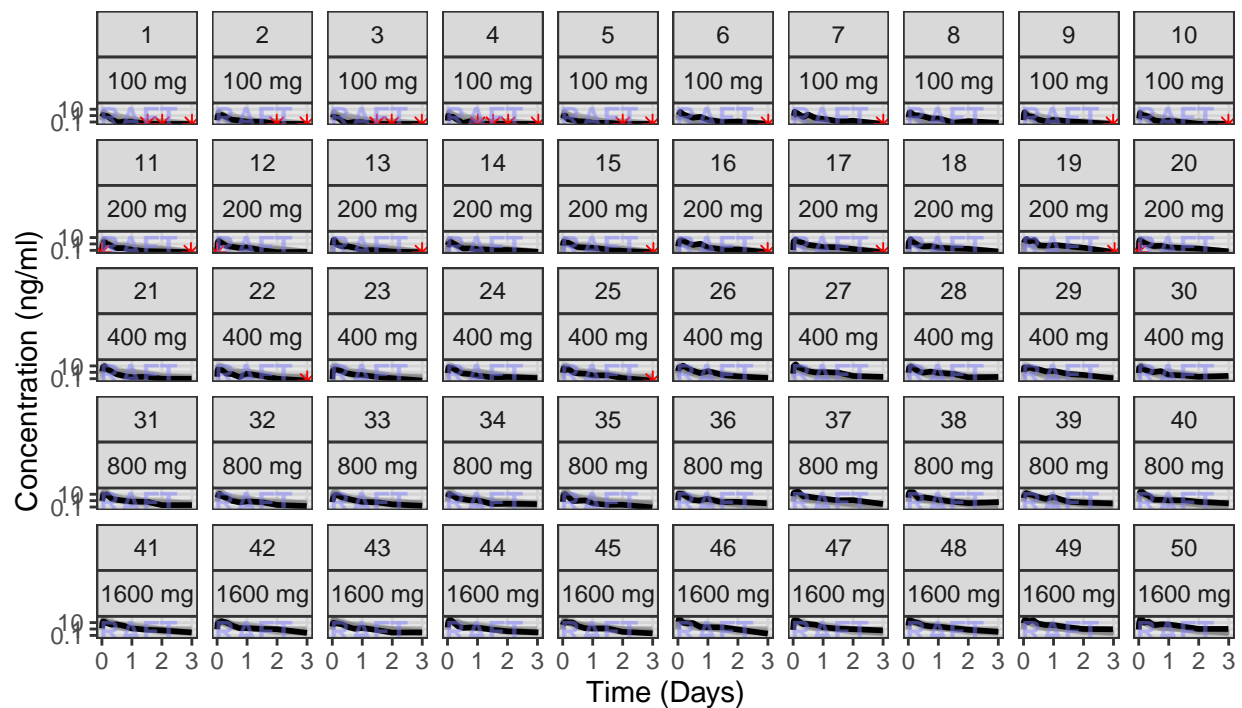
Warning: Removed 50 rows containing missing values (`geom_line()`).
Removed 50 rows containing missing values (`geom_line()`).

D:/REPOSITORIES/PkPd−visul−exploration−
.//Example.R
.//Example.png
Created: 2023−10−26 22:08:15.550014

Explore covariate effects on PK

1. Concentration over time, colored by categorical covariate, mean +/- 95% CI

```r
gg <- ggplot(data = pk_data, aes(x = NOMTIME, y = LIDV, color = SEX))
gg <- gg + xgx_stat_ci(conf_level = 0.95)

gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + labs(y = conc_label)

gg <- gg + facet_grid(.~TRTACT_low2high)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
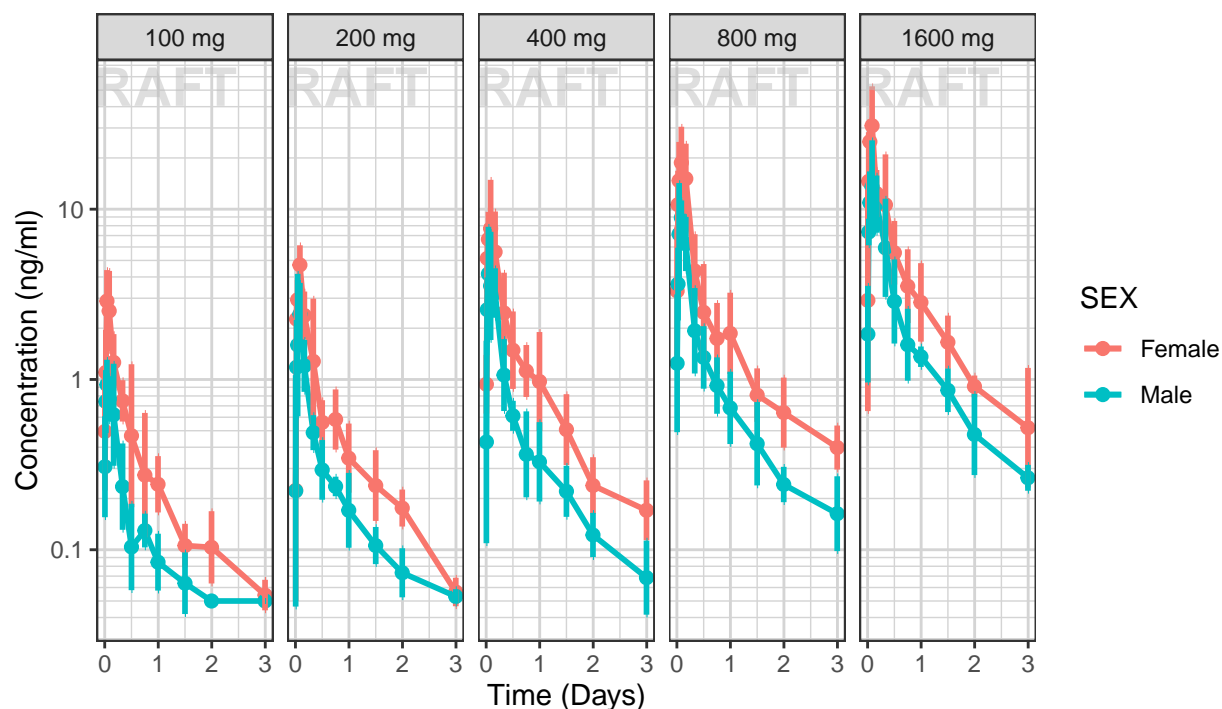
```
## Warning: Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
```
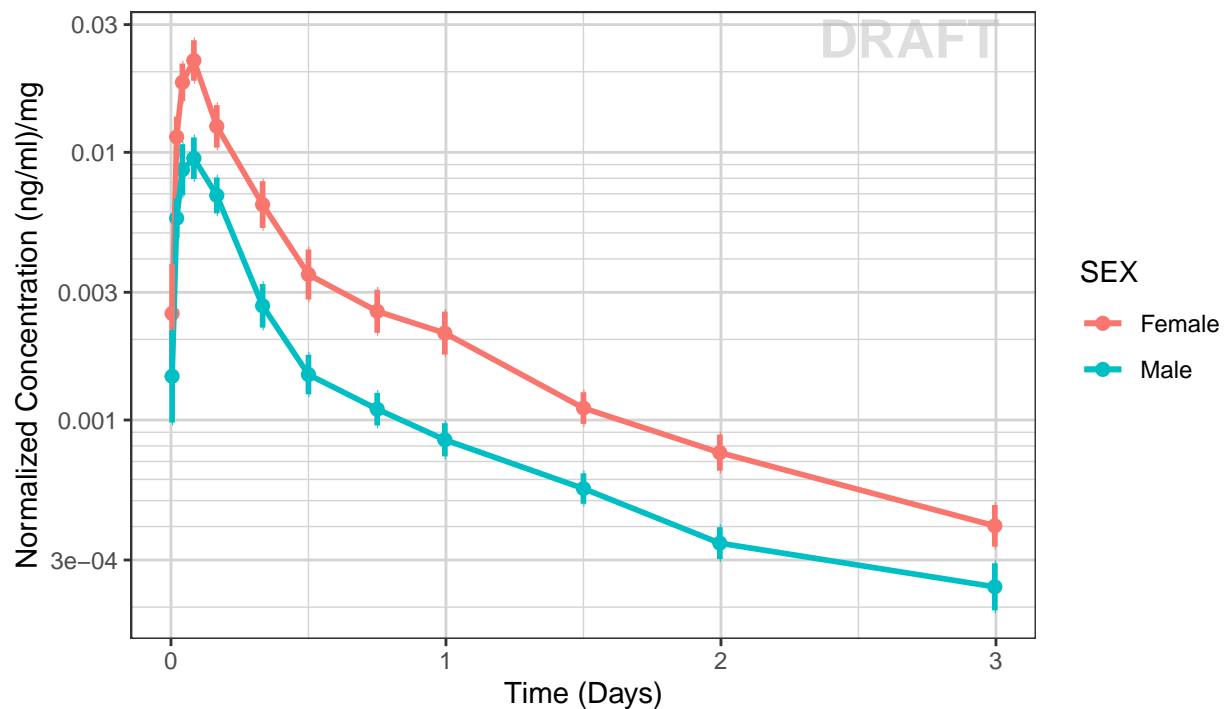

D:/REPOSITORIES/PkPd–visul–exploration–
.//Example.R
.//Example.png
Created: 2023–10–26 22:08:19.346567

13

2. Dose Normalized Concentration over time, colored by categorical covariate, mean +/- 95% CI

```r
gg <- ggplot(data = pk_data,
             aes(x = NOMTIME,
                 y = LIDV_NORM,
                 color = SEX))
gg <- gg + xgx_stat_ci(conf_level = 0.95)
gg <- gg + xgx_scale_y_log10()
gg <- gg + xgx_scale_x_time_units(units_dataset = time_units_dataset,
                                  units_plot    = time_units_plot)
gg <- gg + ylab(concnorm_label)
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```

```
## Warning: Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
## Removed 50 rows containing non-finite values ('stat_summary()').
```
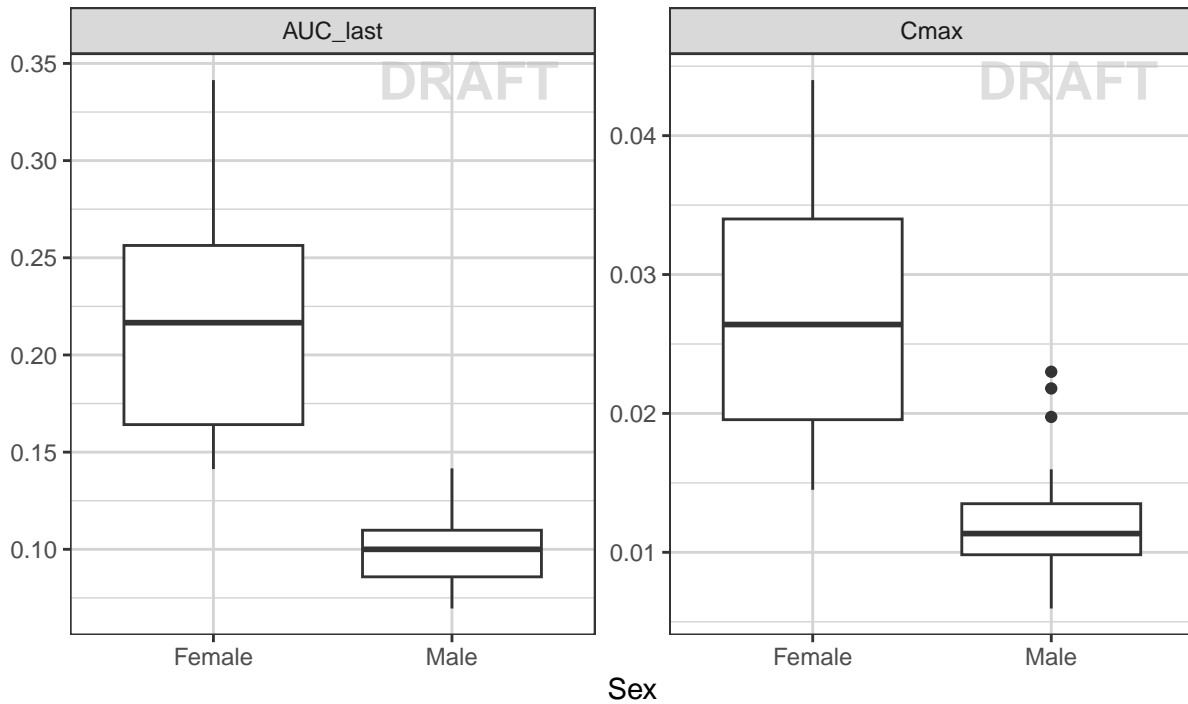


```r
gg <- ggplot(data = NCA, aes(x = SEX, y = VALUE_NORM))
gg <- gg + geom_boxplot(aes(group = SEX))
```

```
gg <- gg + ylab("") + xlab("Sex")
gg <- gg + facet_wrap(~PARAM, scales = "free_y")
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```


Sex

D:/REPOSITORIES/PkPd-visul-exploration-
.//Example.R
.//Example.png
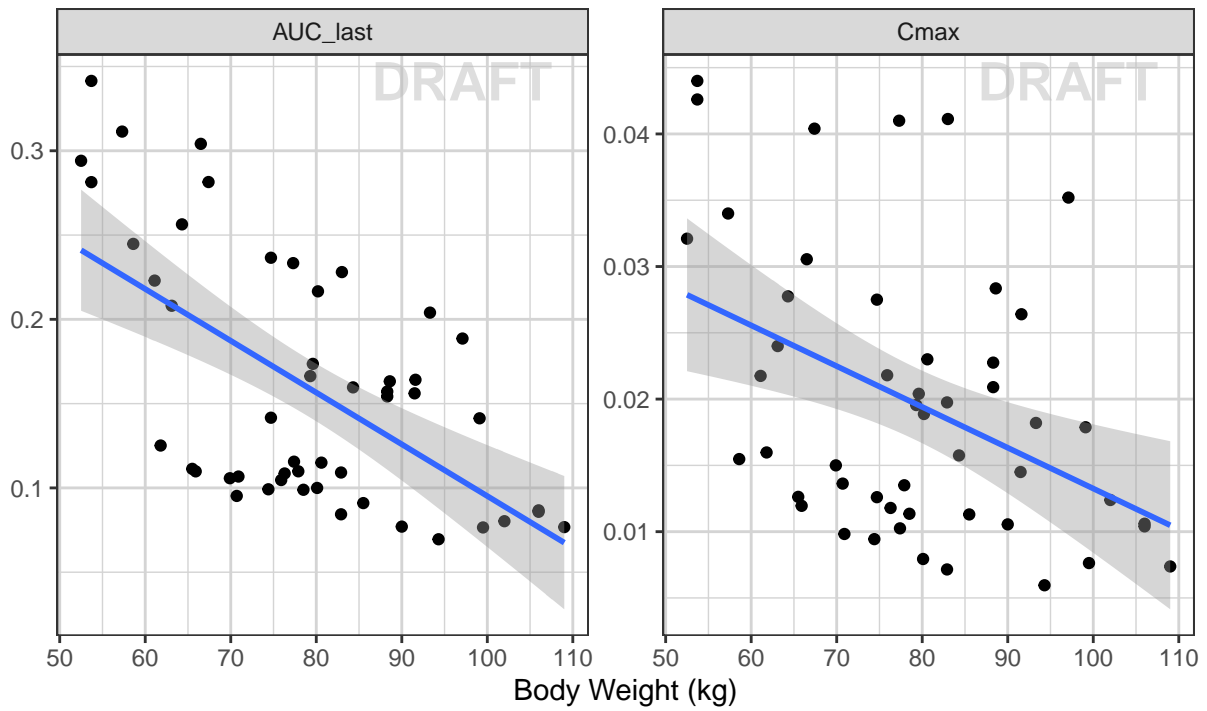Created: 2023-10-26 22:08:20.713162

```
gg <- ggplot(data = NCA, aes(x = WEIGHTB, y = VALUE_NORM))
gg <- gg + geom_point()
gg <- gg + ylab("") + xlab("Body Weight (kg)")
gg <- gg + facet_wrap(~PARAM, scales = "free_y")
gg <- gg + geom_smooth(method="lm")
gg <- gg + xgx_annotate_status(status)
gg <- gg + xgx_annotate_filenames(dirs)

print(gg)
```
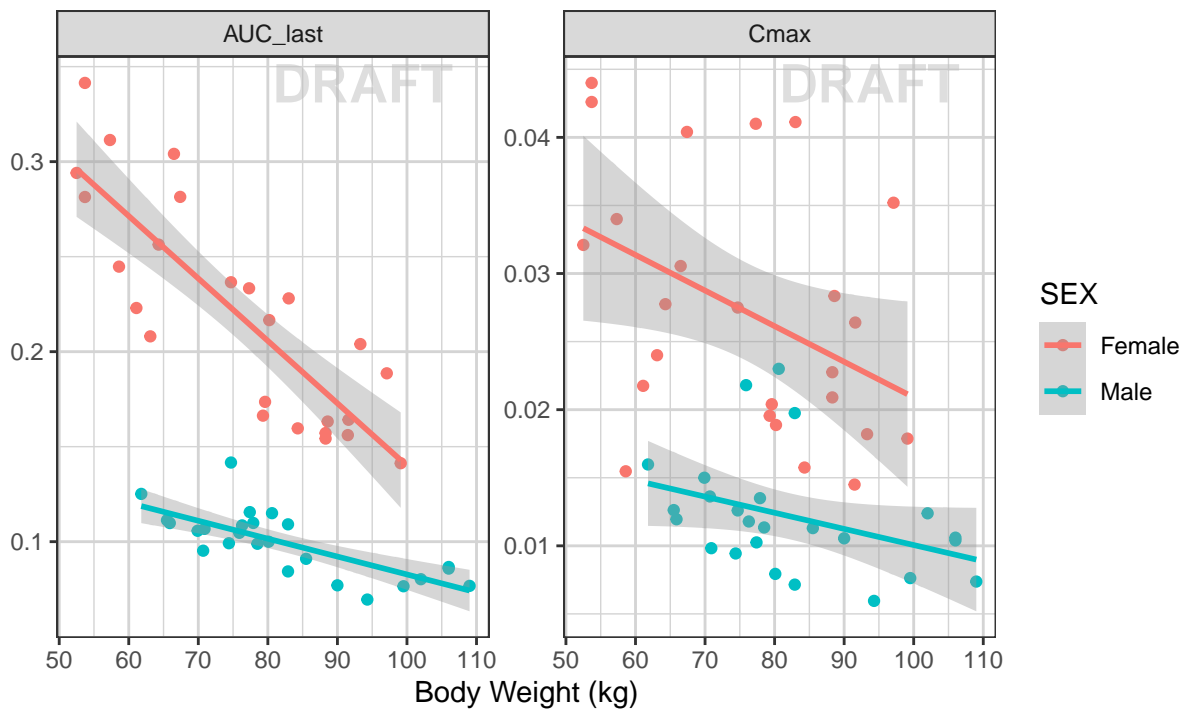
```
## 'geom_smooth()' using formula = 'y ~ x'
```

D:/REPOSITORIES/PkPd−visul−exploration−
.//Example.R
.//Example.png
Created: 2023−10−26 22:08:21.042409

```
gg + aes(color = SEX)
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

D:/REPOSITORIES/PkPd−visul−exploration−
.//Example.R
.//Example.png
Created: 2023−10−26 22:08:21.042409