Code merge

Emmanuelle Comets

22/01/2021

Summary

Objective

Integrate new code to package !

Notes

- to investigate
  - mclust library not loaded automatically (check when running the package)
  - transformation using a reference profile **Eco TODO**
    - check if applies to pd (not sure it makes sense), and if there is a non-parametric version (using quantiles instead of E/SD)
    - check if need to compute the transformation after exponential transformation (probably not, the current plots seem to work also in log-scale)
- improvements
  - need an option to facet the plots split by covariate **Romain TODO +++**
- bugs:
  - plot.opt option **main** doesn't work in *aux.npdeplot.plot()*
  - Romain **TODO +++** : reprendre la partie avec avoid_code() dabs NpdeData qui empêche de lire toutes les covariables

Code changes

- **new functions in kompareCode**
  - **plotNpde-scatterplot.R**: main mid-level function called using a NpdeObject
  - **plotNpde-auxScatter.R**: auxiliary functions to compute and transform obsmat, pimat, ref profiles.
    - **aux.npdeplot.computepi** : Compute prediction intervals, the size of which depends on the number of observations in each bin
    - **aux.npdeplot.meanprof** : Compute a reference profile based on simulations from the model
    - **aux.npdeplot.transformPI** : transform pimat with a reference profile **TODO** check if applies to pd (not sure it makes sense), and if there is a non-parametric version (using quantiles instead of E/SD)
    - **aux.npdeplot.transformObs** : transform obsmat with a reference profile **TODO** same
    - **aux.npdeplot.pimat** : create pimat for plot function
  - **plotNpde-auxScatterPlot.R**: actual plot function => rename this file (plotNpde-auxPlotScatter.R ?) or include it in plotNpde-auxScatter.R
- functions in npde/R
  - plotNpde-unitFunctionsPI.R: bins the X data, creating plot.opt$bin.number bins, computing the mean value of x for each bin, and associating groups to the observed data
- **functions renamed**
  - for consistency, **aux.plot.hist** and **aux.plot.dist** renamed to **aux.npdeplot.hist** and **aux.npdeplot.dist**

  - **aux.npdeplot.plot** renamed to **aux.npdeplot.scatter**
  - **npde.plot.meanprofile** renamed to **npde.plot.scatterplot**
- **functions removed***
  - **compute.vpc.pi** : old code to compute PI for VPC, now computed in the same way as the other PI using *aux.npdeplot.computepi*

Functions to create

- plot functions
  - **aux.npdeplot.scatter.facet** : not done yet, but should be a modified version of the plot with a facet layout (same y scales ? or add a graphical option ?) instead of a list of plots OR make this the default option for covsplit and simply add an option in NpdeControl for free scale on x and y axes
- testthat files for unitary tests
  - **Romain TODO +++** : test replacement options (are we capturing the ..., are we capturing errors, are we superseding the right things ?)

Running new code

Defining folders, loading libraries

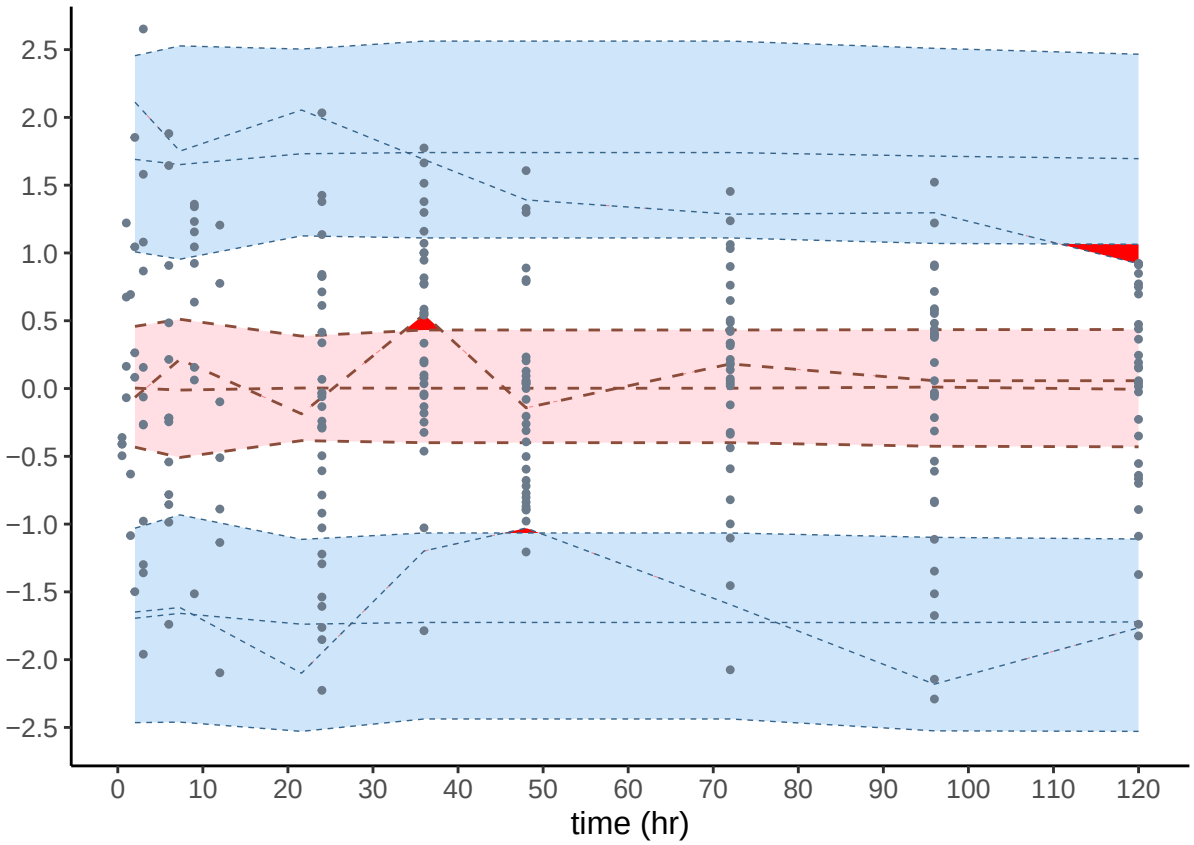
Loading functions

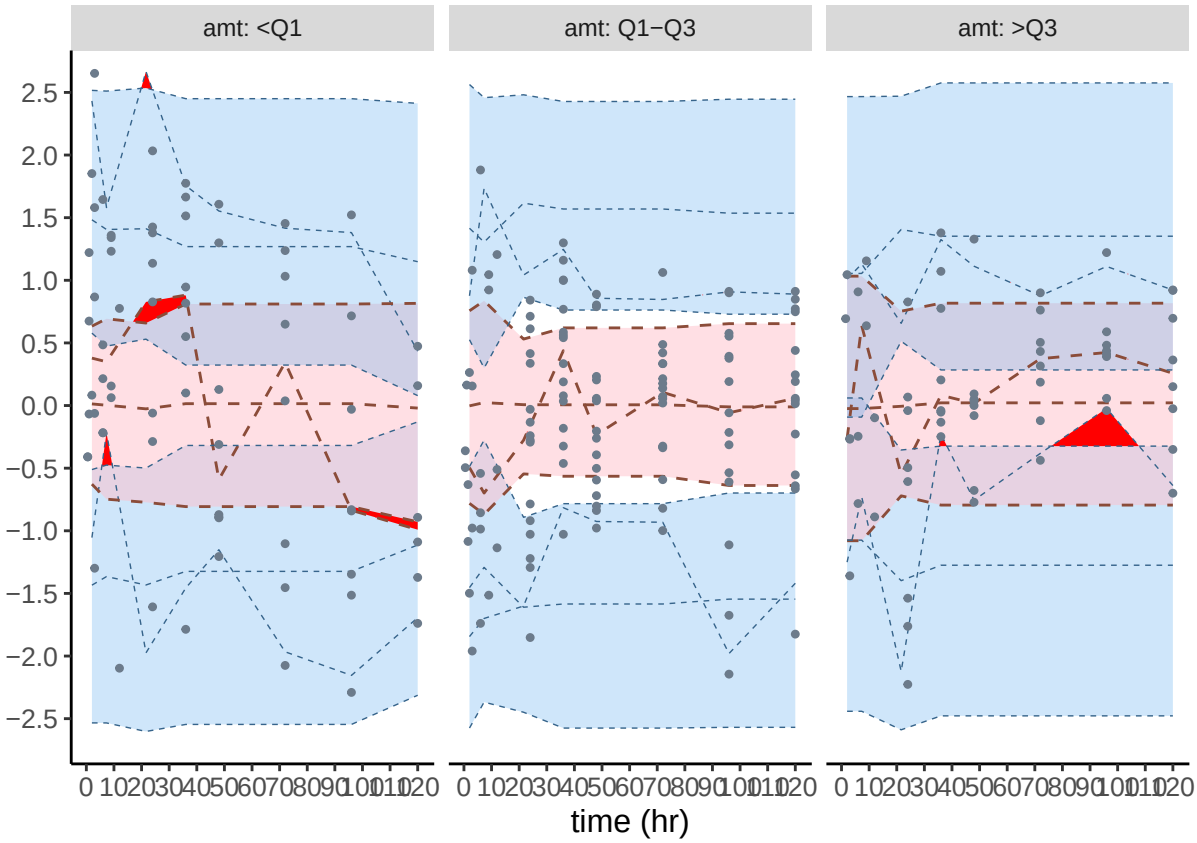
Run npde

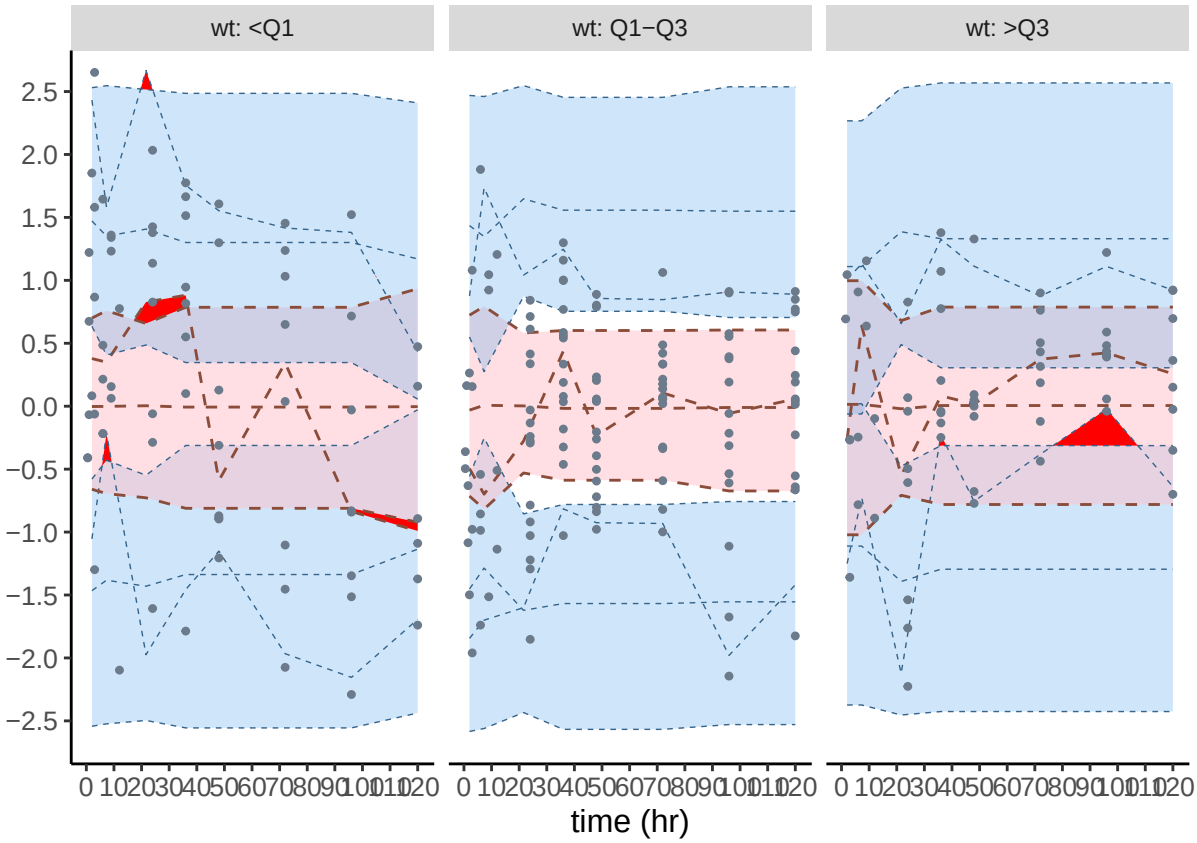
- tests
  - what happens if which=c("npd","npde") => for the moment, bugs :-) but create loop around typmet in plotNpdeObject => test loop
- questions
  - why the argument "new=force.new" (and why is it only for eg pred.scatter but not x.scatter ?)

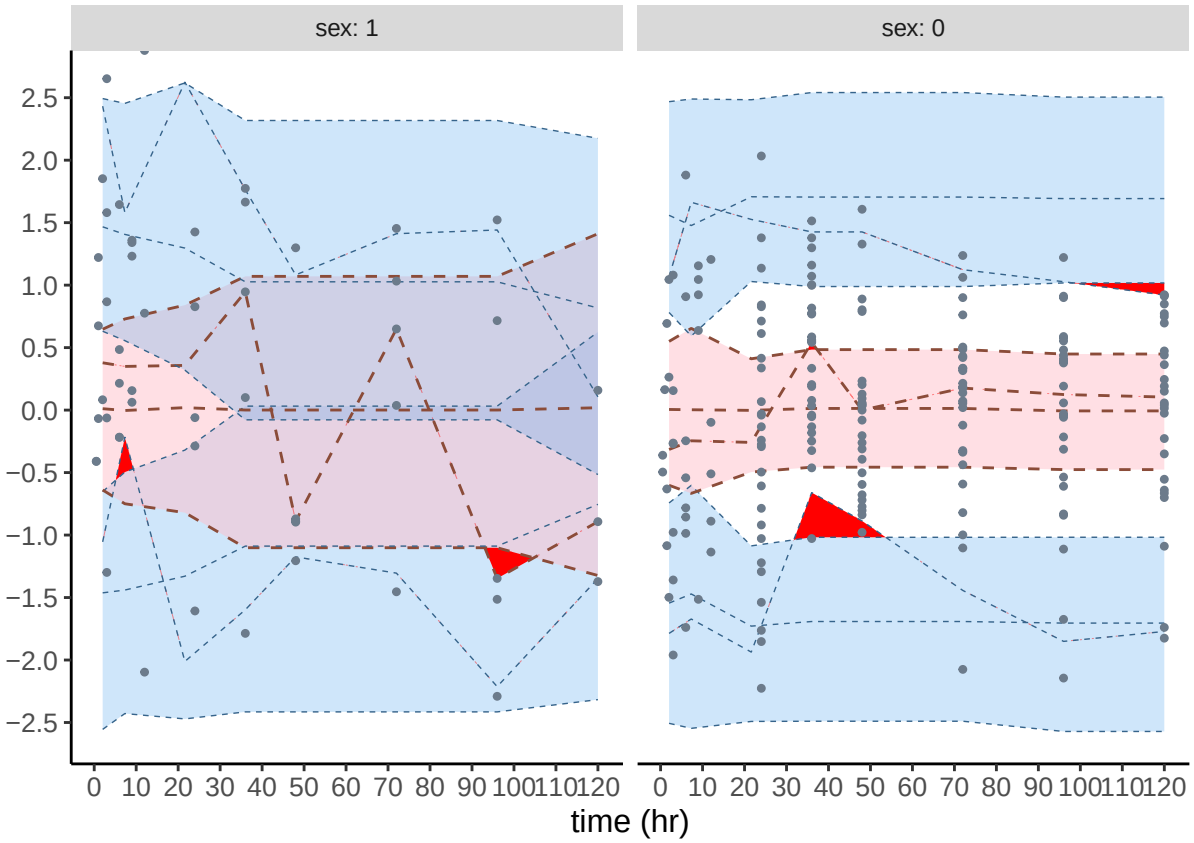
Graphs with the new functions

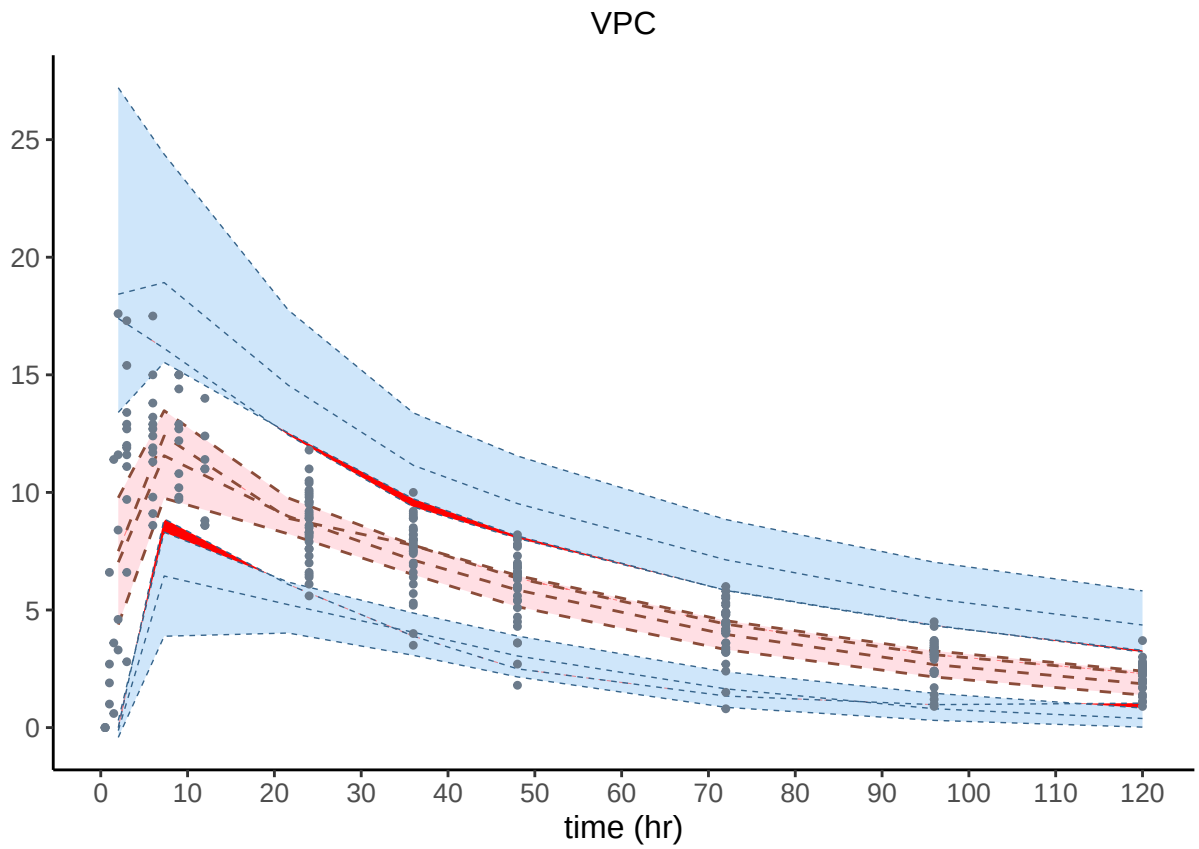
- Works !!!
- issues
  - sorting problem in weight when using covsplit: need to sort level of categories properly => done
- plot.box
  - not sure box size is set correctly
  - warnings from ggplot2 to use variable name instead of table$varname
- labels: xlab, ylab, main don't work
  - should be captured by the intersect at the beginning of function *npde.plot.scatterplot()* ?
  - note: maybe title wouldn't work unless we use a facet plot with covsplit (then categories would be facet titles and a main title could come above)

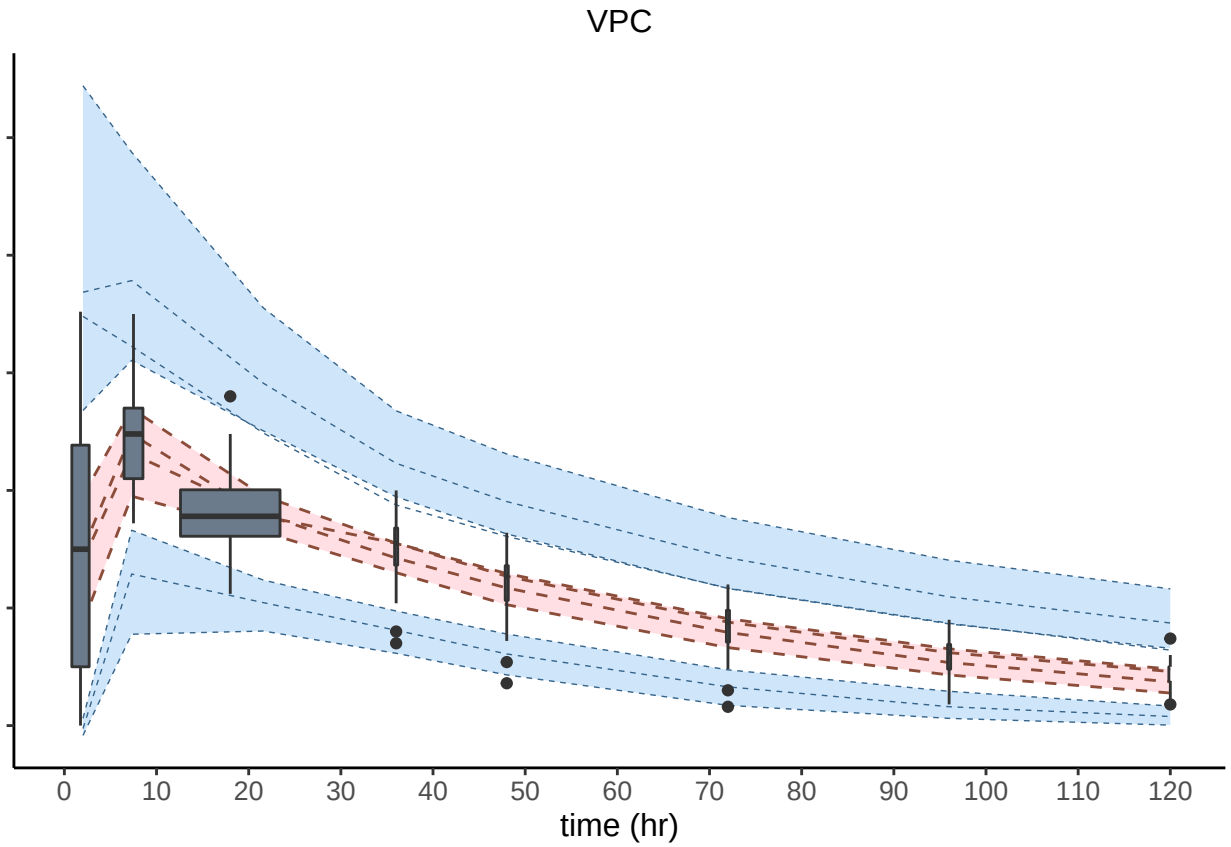

2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0
10
20
30
40
50
60
70
80
90
100
110
120
time (hr)

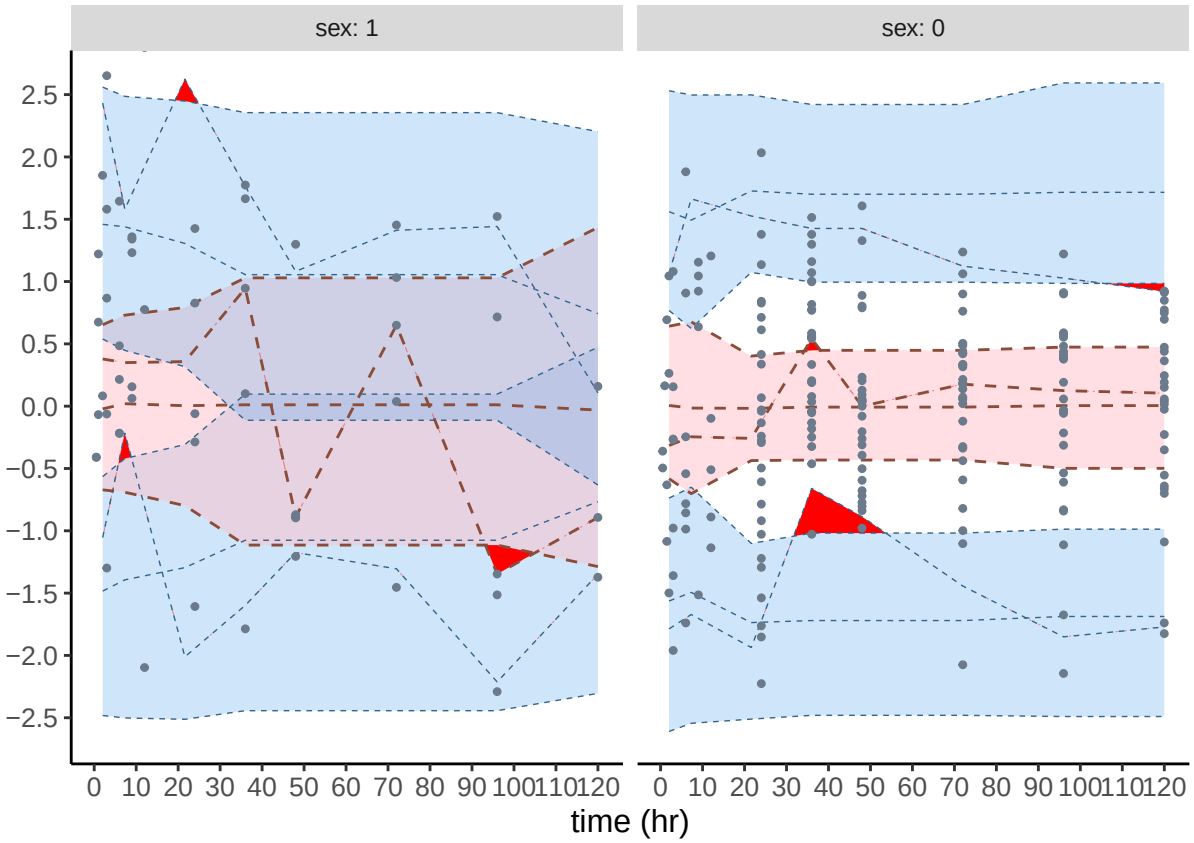

amt: <Q1
amt: Q1−Q3
amt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

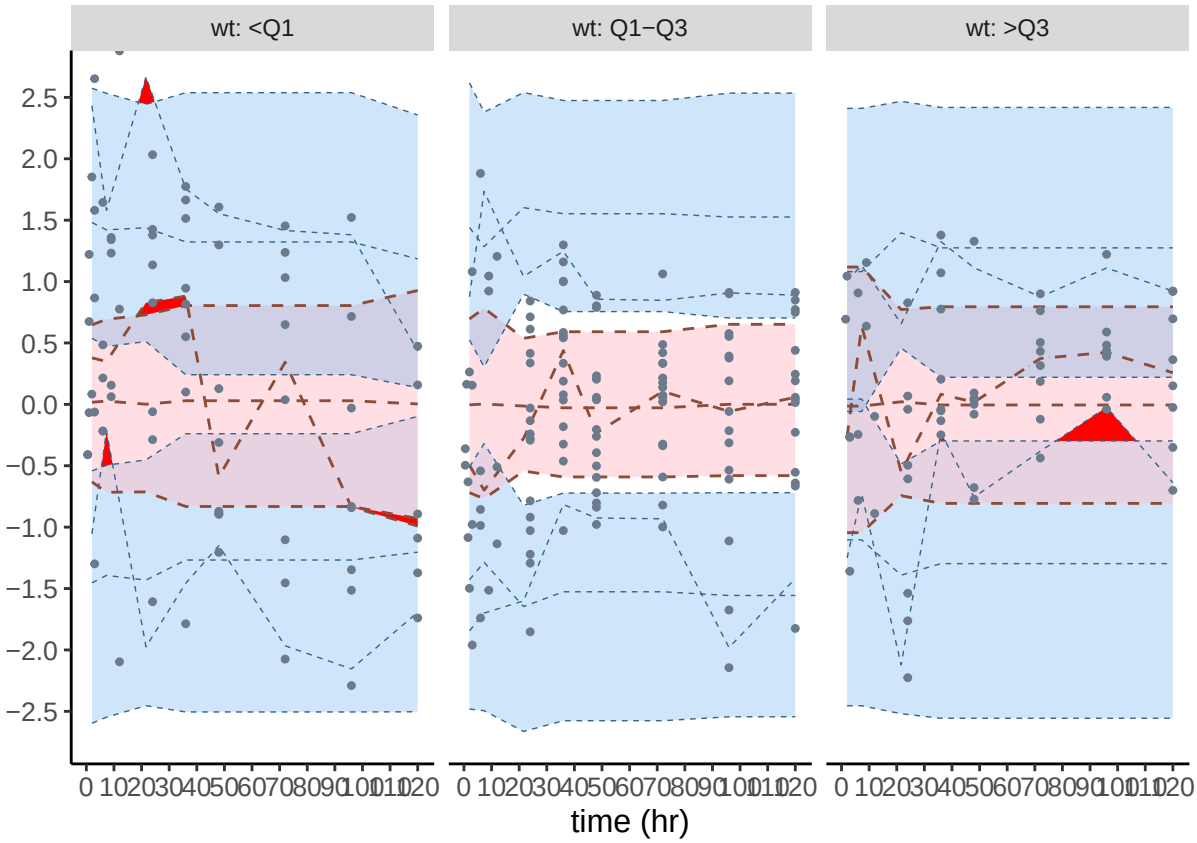

wt: <Q1
wt: Q1−Q3
wt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

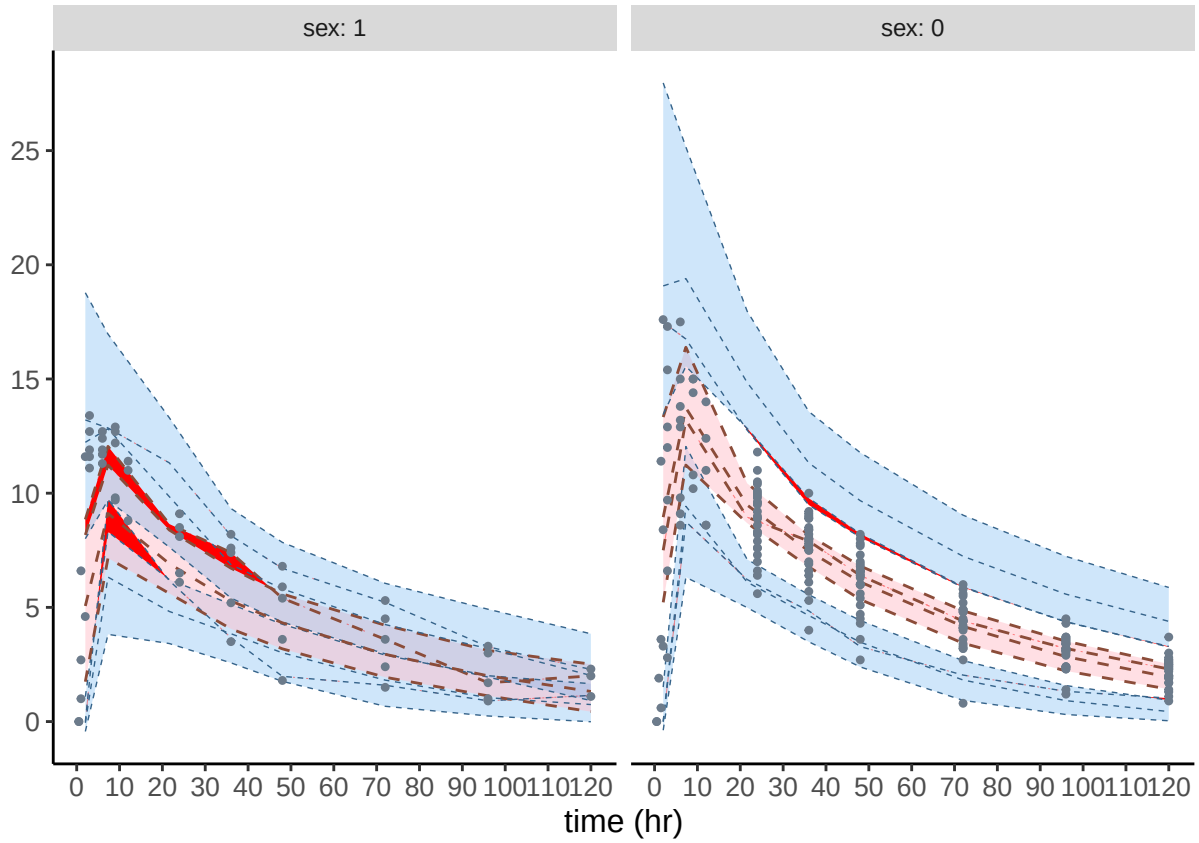

sex: 1
sex: 0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

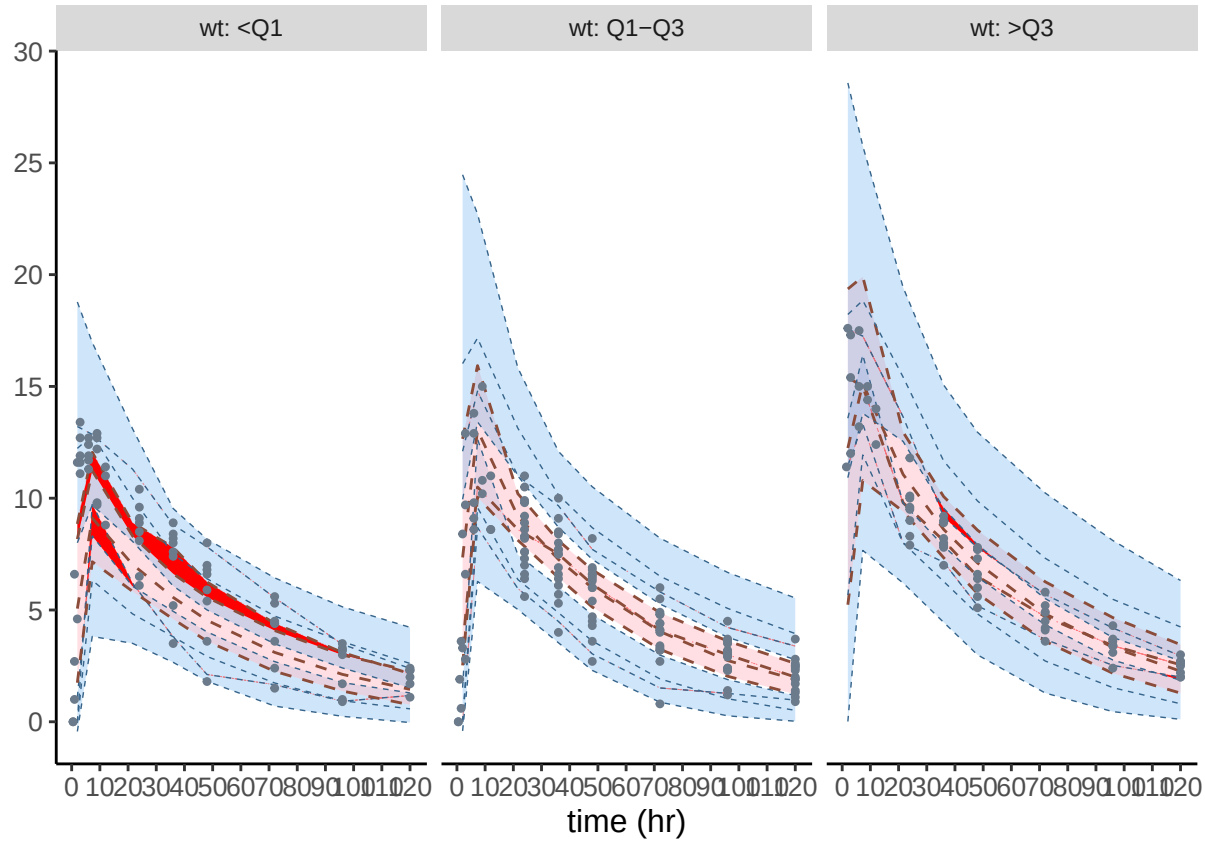

VPC
25
20
15
10
5
0
0
10
20
30
40
50
60
70
80
90
100
110
120
time (hr)

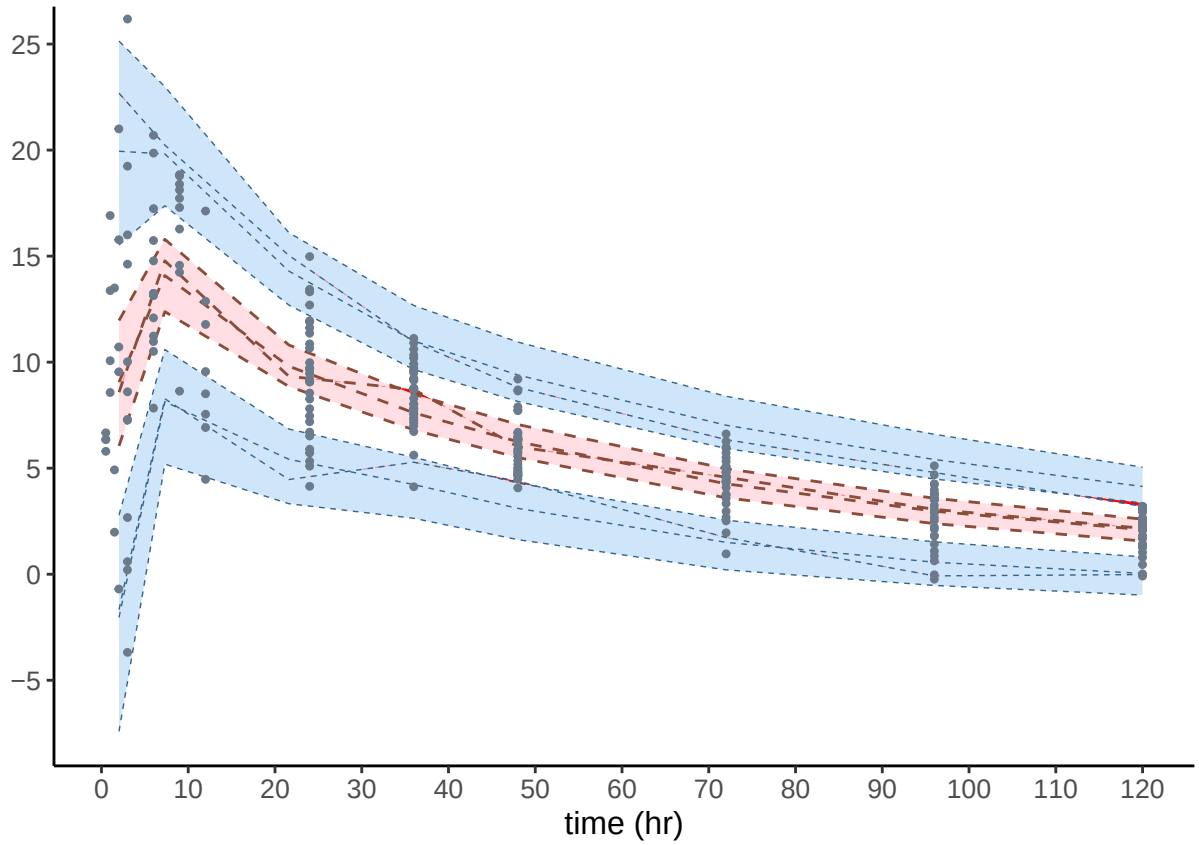

VPC
0
10
20
30
40
50
60
70
80
90
100
110
120
time (hr)

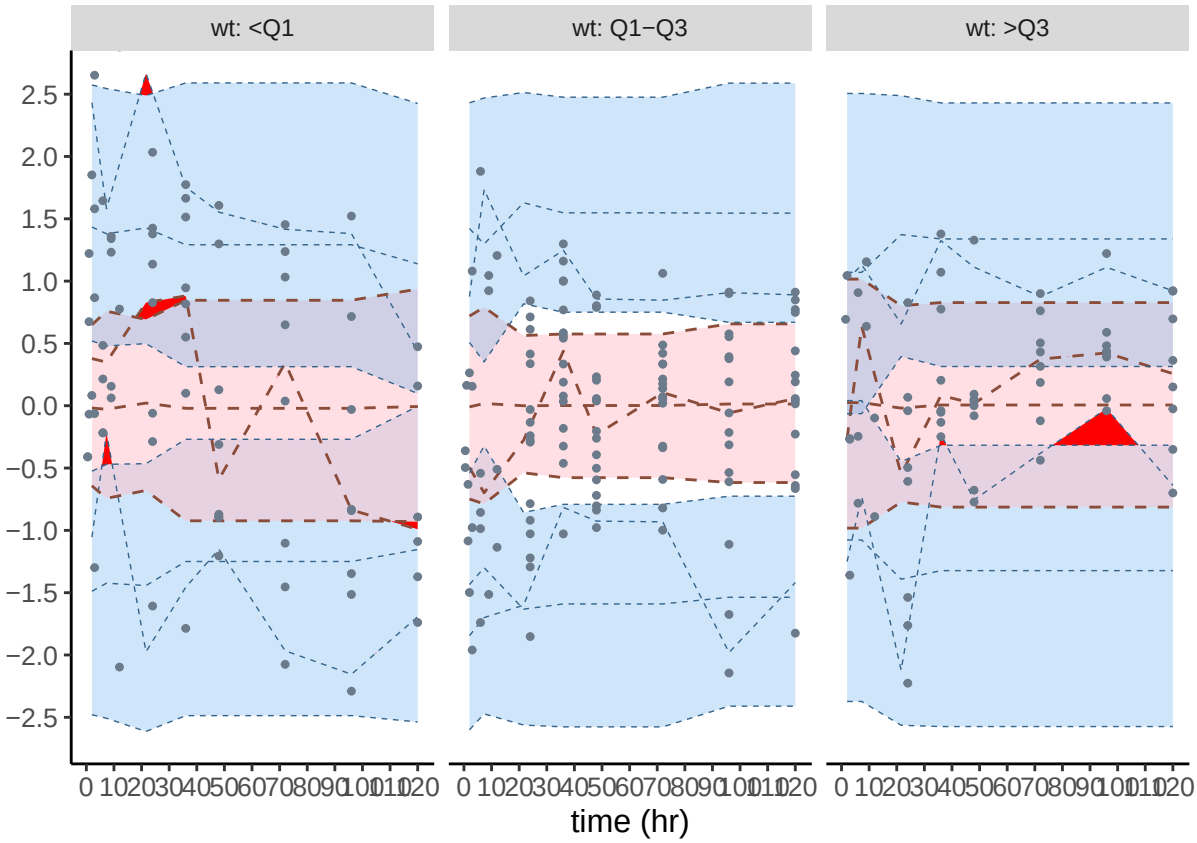

sex: 1
sex: 0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

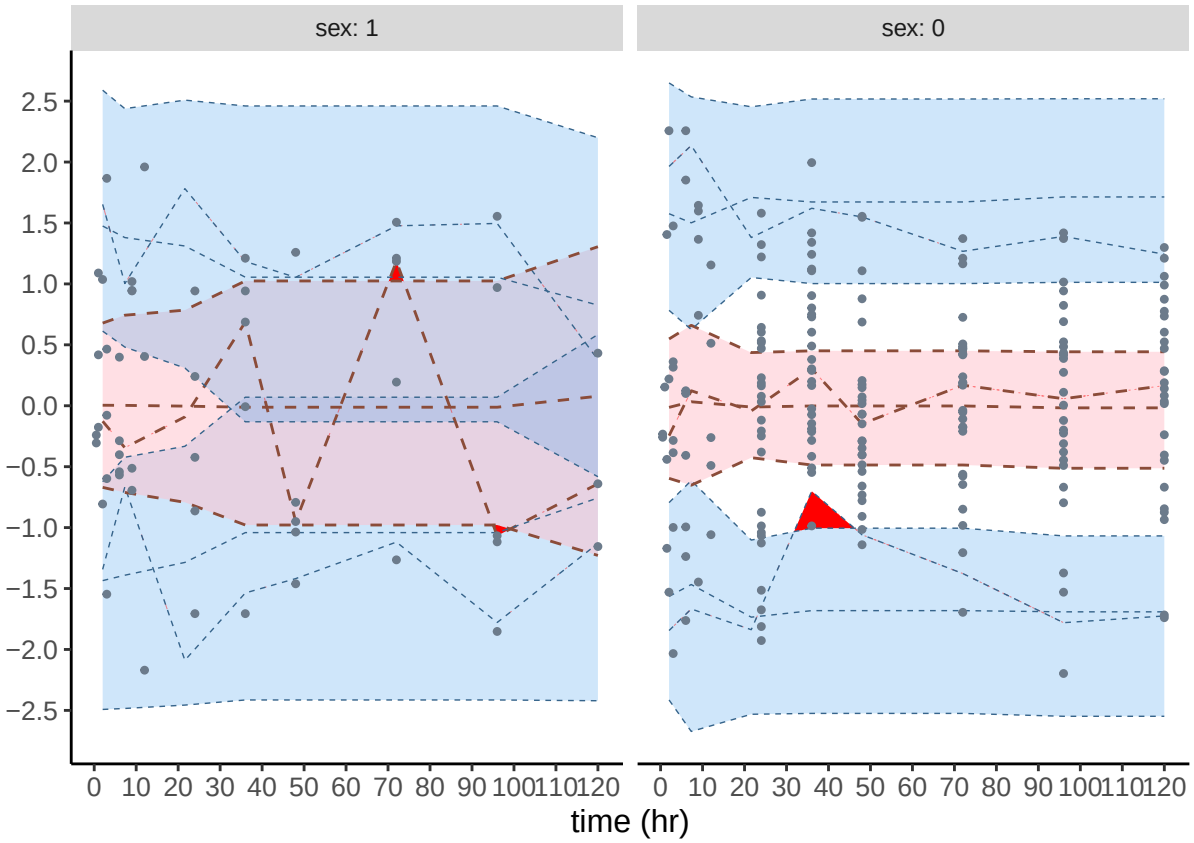

wt: <Q1
wt: Q1−Q3
wt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

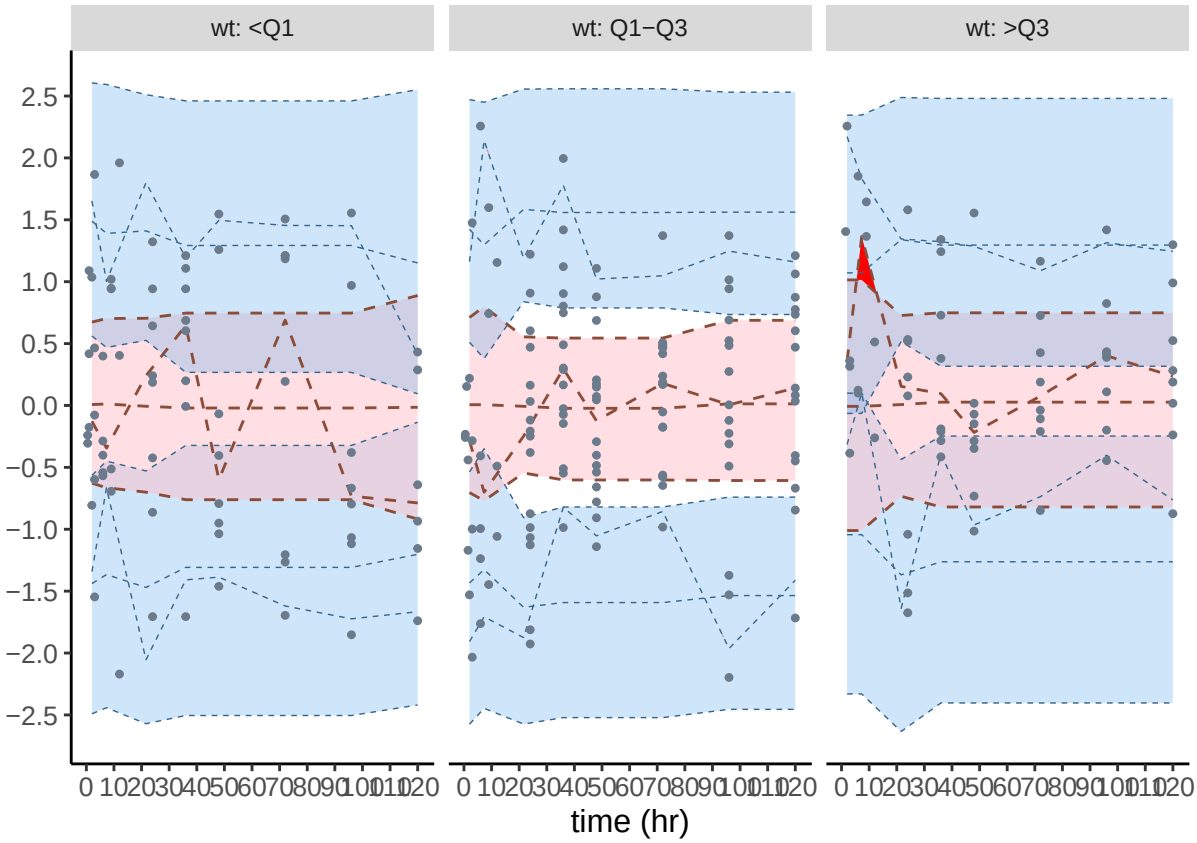

sex: 1
sex: 0
25
20
15
10
5
0
0 10 20 30 40 50 60 70 80 90 100 110 120
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

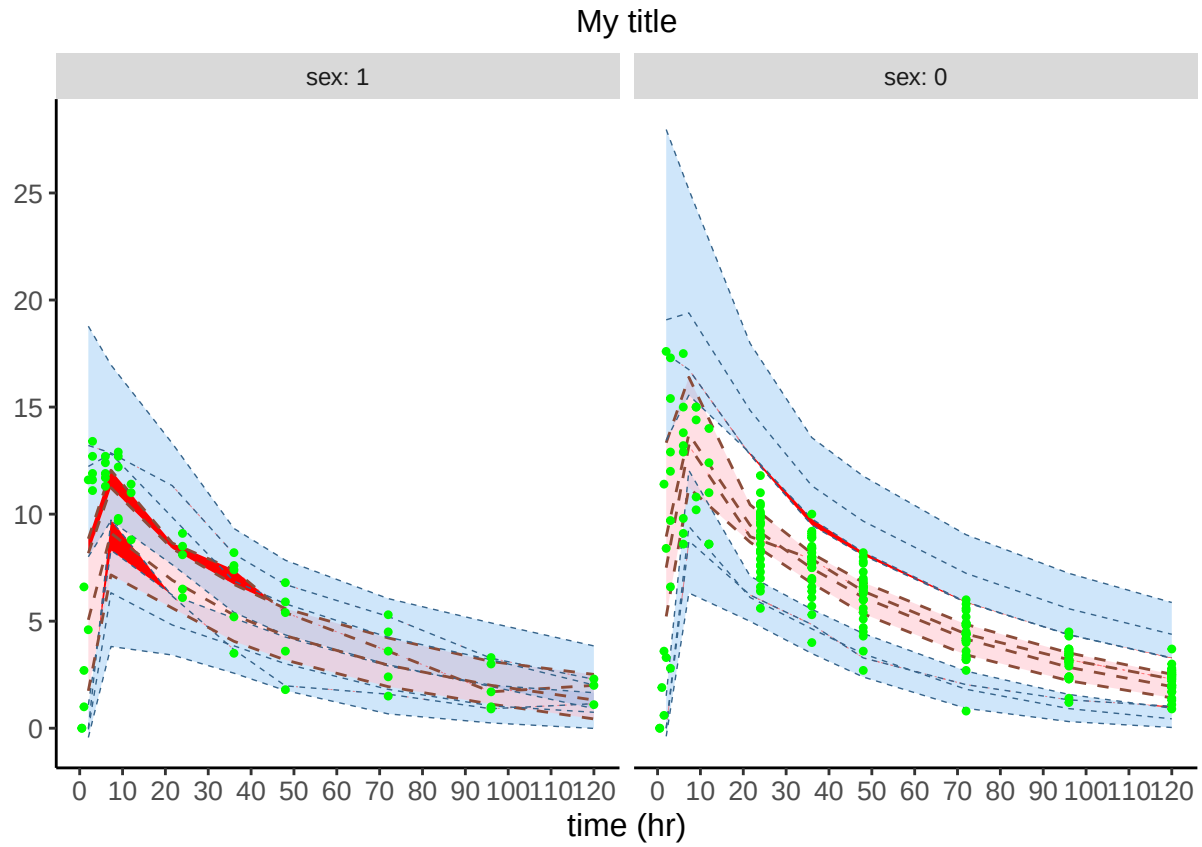

wt: <Q1
wt: Q1−Q3
wt: >Q3
30
25
20
15
10
5
0
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)


25
20
15
10
5
0
−5
0
10
20
30
40
50
60
70
80
90
100
110
120
time (hr)


wt: <Q1
wt: Q1−Q3
wt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

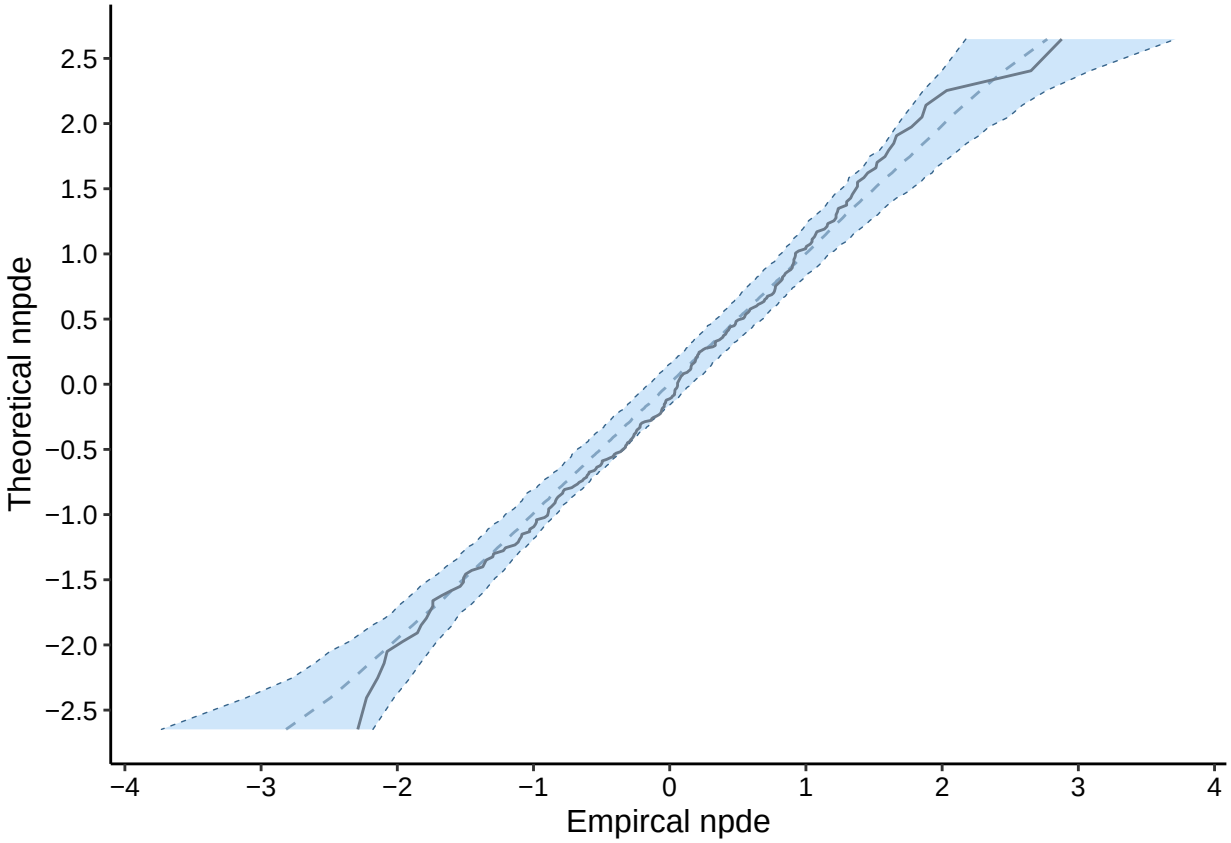

sex: 1
sex: 0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

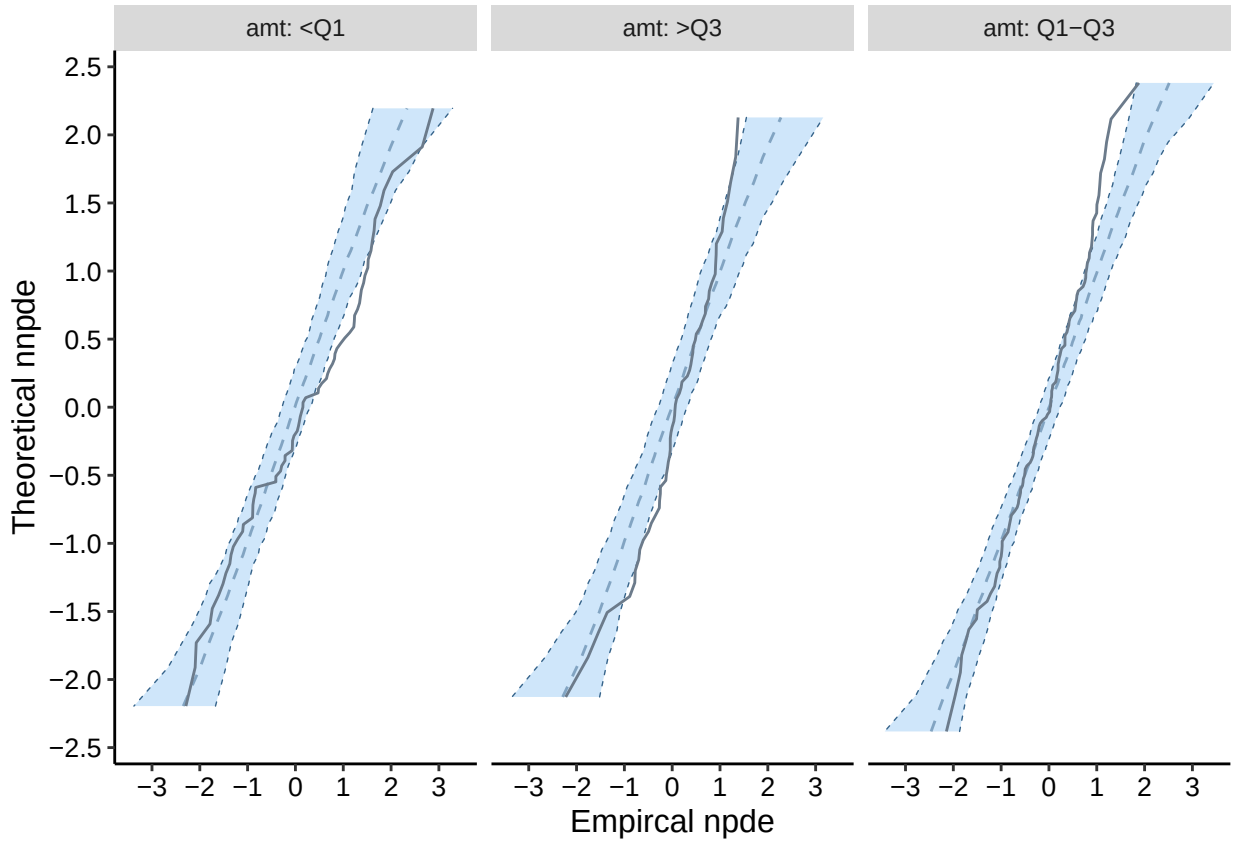

wt: <Q1
wt: Q1−Q3
wt: >Q3
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80 90 100 110 120
time (hr)

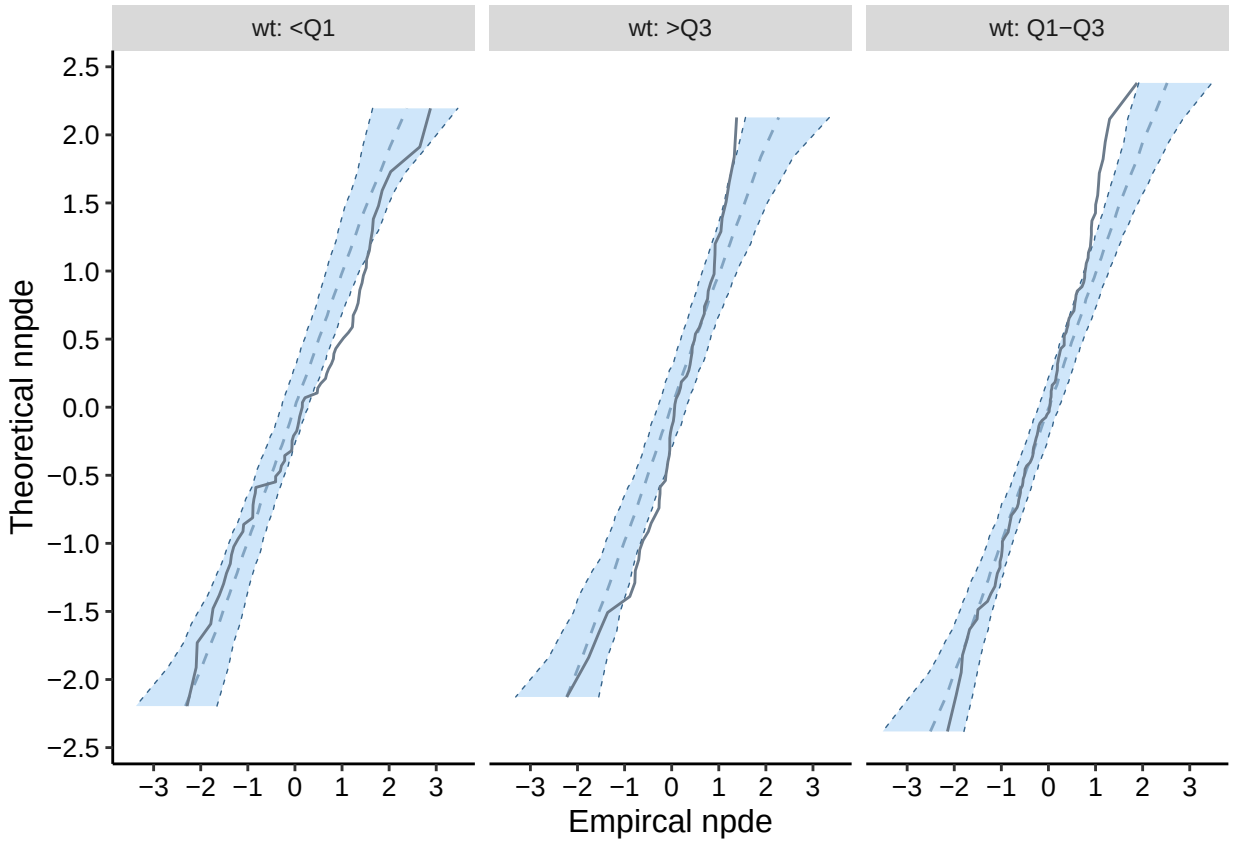


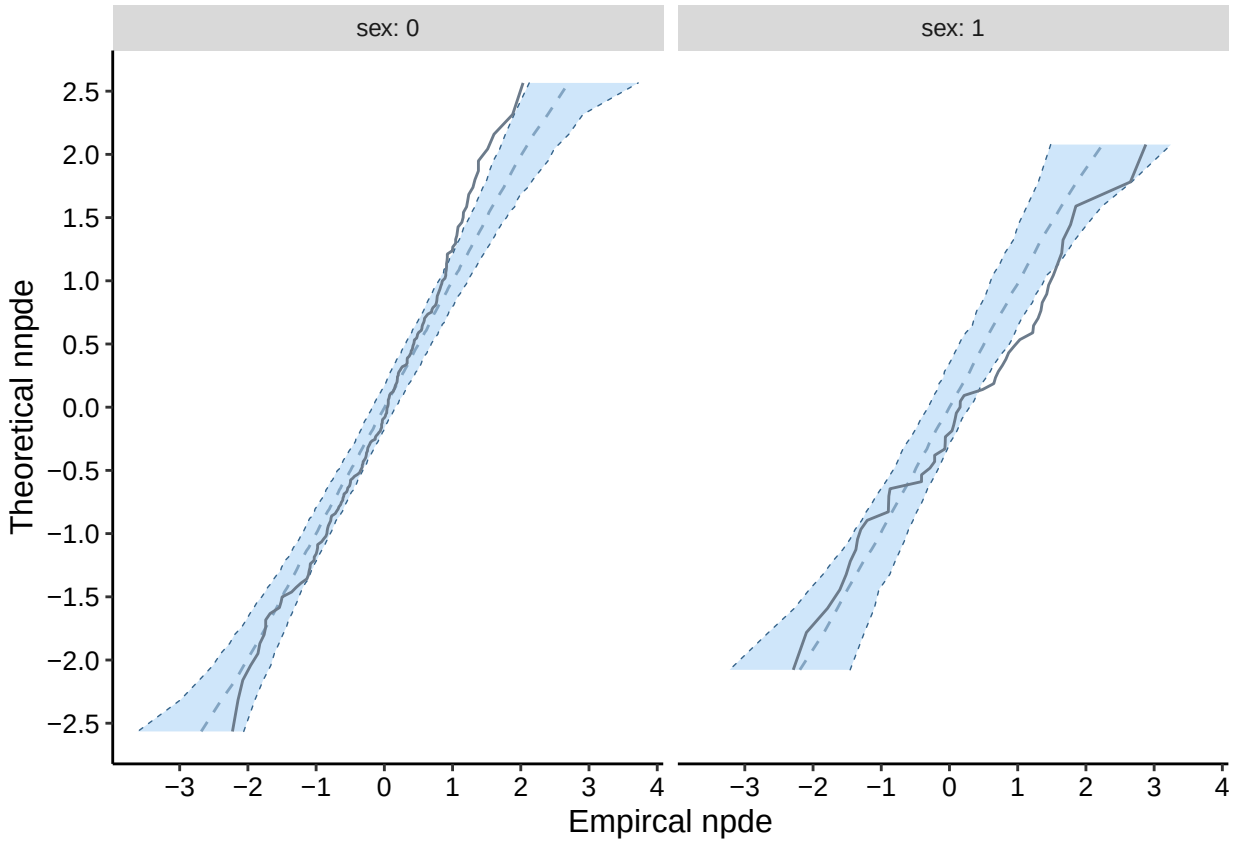
npde.plot.dist

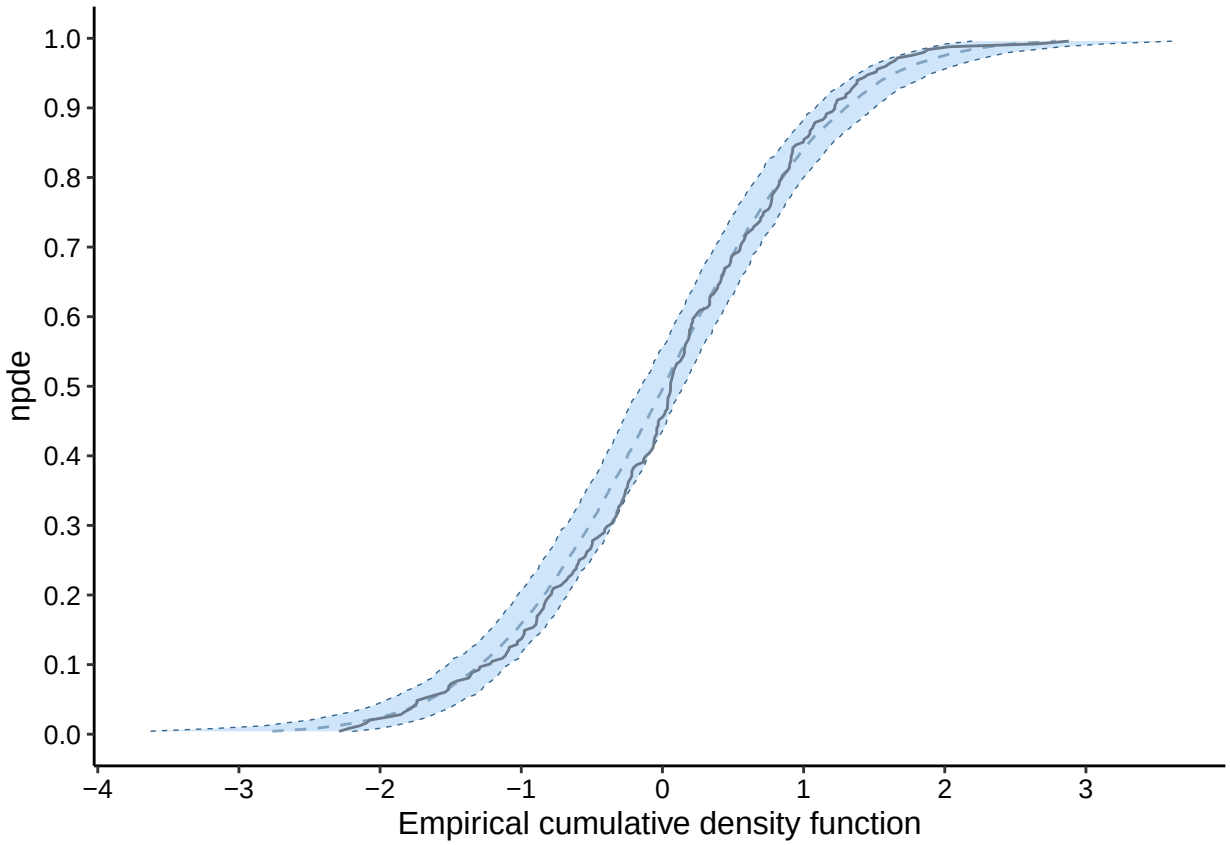
- fails => **unitary test to do +++***
  - debugging de la fonction: en fait elle n'a pas du tout la structure requise, la boucle sur les covariables ne devrait pas être dedans => ne correspond pas aux spécifications
  - absence de test unitaire pour voir à quoi devrait ressembler les entrées et les sorties, pas du tout optimal pour un travail collaboratif
- problem with argument list
  - added default to dist.type (qqplot)
  - **maybe (TBD)** covsplit should be passed on using prefs slot in object or in ... (see npde.plot.scatterplot())
- **Eco: re-coded this function and its dependencies from scratch, using only the coding of graphs**


Theoretical nnpde
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−4
−3
−2
−1
0
1
2
3
4
Empircal npde


amt: <Q1
amt: >Q3
amt: Q1−Q3
Theoretical nnpde
Empircal npde
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3
−2
−1
0
1
2
3


wt: <Q1
wt: >Q3
wt: Q1−Q3
Theoretical nnpde
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3 −2 −1 0 1 2 3
Empircal npde


sex: 0
sex: 1
Theoretical nnpde
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3
−2
−1
0
1
2
3
4
Empircal npde



Waffle plot

Fails, presumably because we don't return the proper object to use in grid.arrange... ? Tried to imitate distribution plot but doesn't seem to work => **TODO Romain**

- revoir ce que renvoie la (les) fonctions et refaire le waffle plot
- faire en sorte qu'on puisse extraire la même chose "à la main" pour pouvoir faire nous même des waffle plot

```
# Fails with "only 'grobs' allowed in "gList"" message
try(plot(wbase))
```

```
## Error in gList(structure(list(structure(list(data = structure(list(x = c(-0.361133033557212,  :
##   only 'grobs' allowed in "gList"
```

```
# What we want (user defined grid.arrange)
p1<-npde.plot.dist(wbase, dist.type="hist")
p2<-npde.plot.dist(wbase, dist.type="ecdf")
p3<-npde.plot.scatterplot(wbase)

try(grid.arrange(list(p1,p2,p3), nrow=2))
```

```
## Error in gList(structure(list(list(structure(list(data = structure(list( :
##   only 'grobs' allowed in "gList"
```

```
try(grid.arrange(p1, p2, nrow=2))
```

```
## Error in gList(structure(list(structure(list(data = structure(list(x = c(-0.361133033557212,  :
##   only 'grobs' allowed in "gList"
```