Performance of npde for the evaluation of joint models with time to event data

M. Cerou[1,3], S. Peigné[3], M. Chenel[3], E. Comets[1,2]

[1] Inserm, IAME, University Paris Diderot, Paris, France
[2] Inserm, CIC 1414, University Rennes 1, Rennes, France
[3] Division of clinical pharmacokinetics and pharmacometrics, Servier, France

PAGE meeting – Stuart Beal methodology session

June, 13th 2019

Context

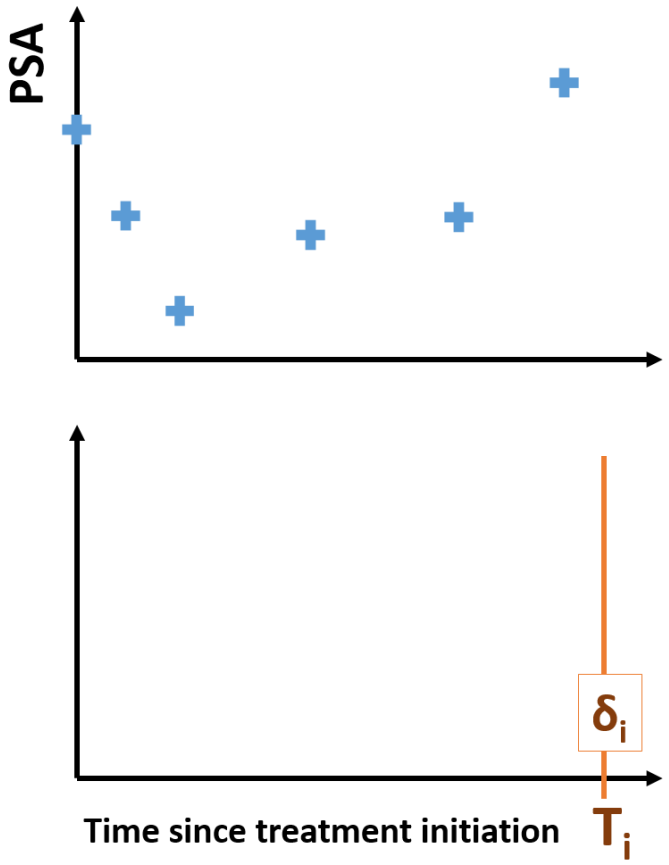
- Nonlinear Mixed Effect Models (NLMEM) increasingly more sophisticated
- Model evaluation
  - assessing the adequacy between the tested model and the data
  - important part in model development [1,2,3]
  - graphical and statistical methods available for continuous data
    - recommended methods include **visual predictive check (VPC)** and **npde** as a gold standard [4]

[1] Brendel K et al. Clin Pharmacokinet. 2007
[2] FDA 1999
[3] EMA 2006
[4] Nguyen THT et al. CPT Pharmacometrics Syst Pharmacol. 2017

Joint models

- Processes of interest are followed throughout clinical trials
- Typically in oncology with biomarkers (e.g. PSA, SLD) and time-to-event (e.g. death, relapse)
- Joint models provide a promising statistical framework to estimate this association
- Support clinical decisions and treatment choices
- Increased use of joint models [1,2] with NLMEM

- **How to extend npde for the evaluation of joint models ?**



[1] Sudell M et al. BMC Med Res Methodol. 2016
[2] Tardivon C et al. Clin Pharmacol Ther. 2019

Outline

- Development of npde for the evaluation of **joint model** with longitudinal and time-to-event (TTE) data
- Performance of the statistical test
- How can we visually diagnose model deficiencies ?

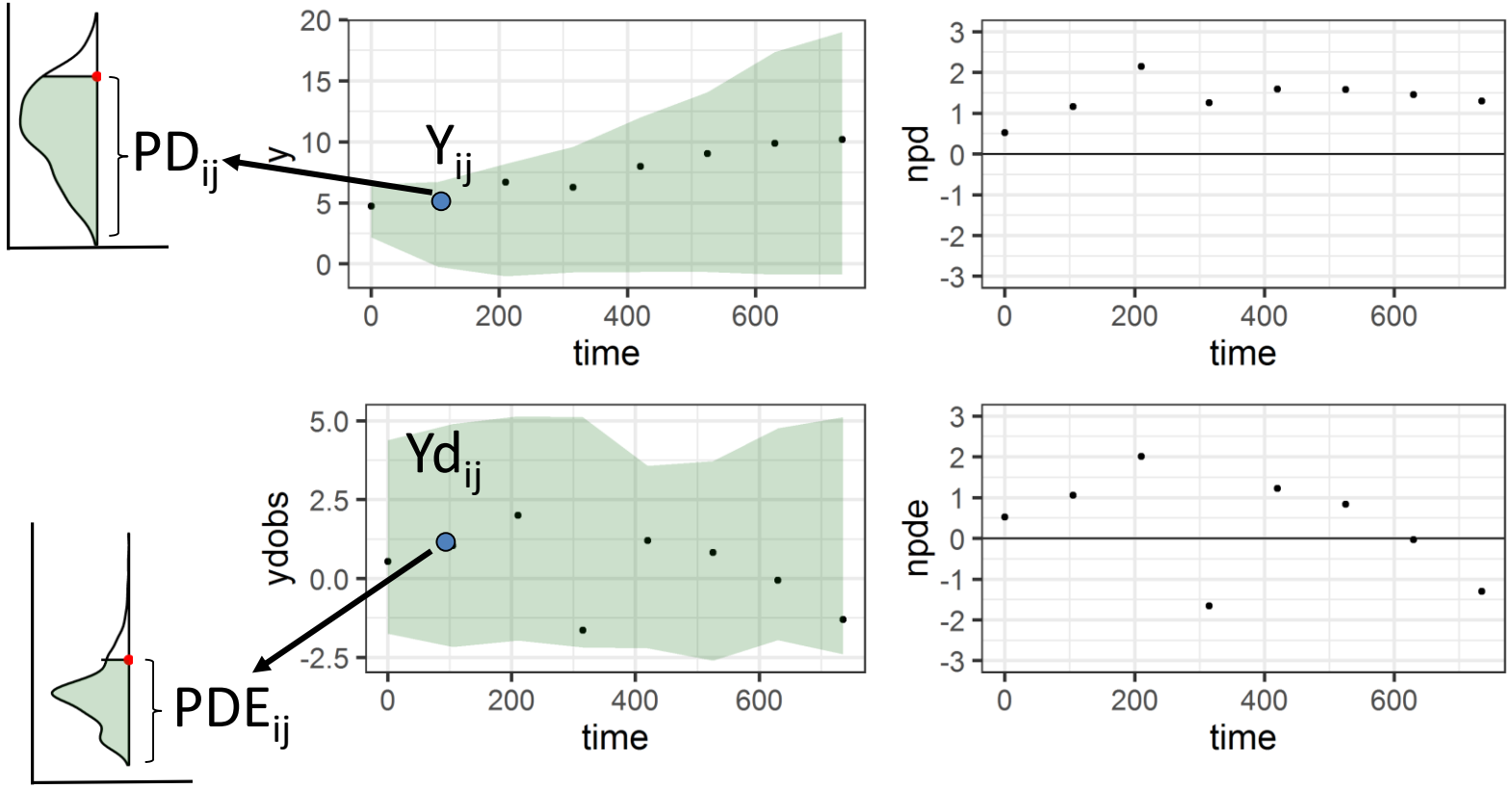
Statistical model

- y_{ij} is the j[th] continuous observation for subject i at time t_{ij}
- T_i is the time to first event
- Model for continuous data
  - $y_{ij} = \mathbf{f}(\boldsymbol{\theta_i}, \boldsymbol{t_{ij}}) + g(\theta_i, t_{ij}, \sigma)e_{ij}$, with $e_{ij} \sim \mathcal{N}(0,1)$ and g the error model
  - $\theta_i = \mathcal{D}(\mu_L, \eta_i)$ with μ_L the fixed effects and η_i the random effects ($\eta_i \sim \mathcal{N}(0, \Omega)$)
- Dependency between observations: conditional **independence** with respect to random effects
- Model for TTE data
  - $\boldsymbol{h_i(t|\theta_i)} = h_0(t) \times \exp\left(\beta_L . \boldsymbol{l}(\boldsymbol{\theta_i}, \boldsymbol{t})\right)$
  - with β_S the vector of parameters of the baseline hazard function h_0, and β_L which represents the strength of the link between $\boldsymbol{l}(\boldsymbol{\theta_i}, t)$ and the hazard
  - $\mu_{TTE} = \{\beta_S, \beta_L\}$
- $\Psi = \{\mu_L, \mu_{TTE}, \Omega, \sigma\}$

Development of npde for continuous data [1]

$$pd_{ij} = F_{ij}(y_{ij}) = \int_0^{y_{ij}} p_i(y, \Psi)dy = \int^{y_{ij}} \int p(y|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dy$$

- prediction discrepancies **pd**: quantile of an observation in its predictive distribution
- prediction distribution error **pde**: quantile of a decorrelated observation in its decorrelated predictive distribution
- normalised prediction distribution **npd**: normalization of pd
- normalised prediction distribution error **npde**: normalization of pde

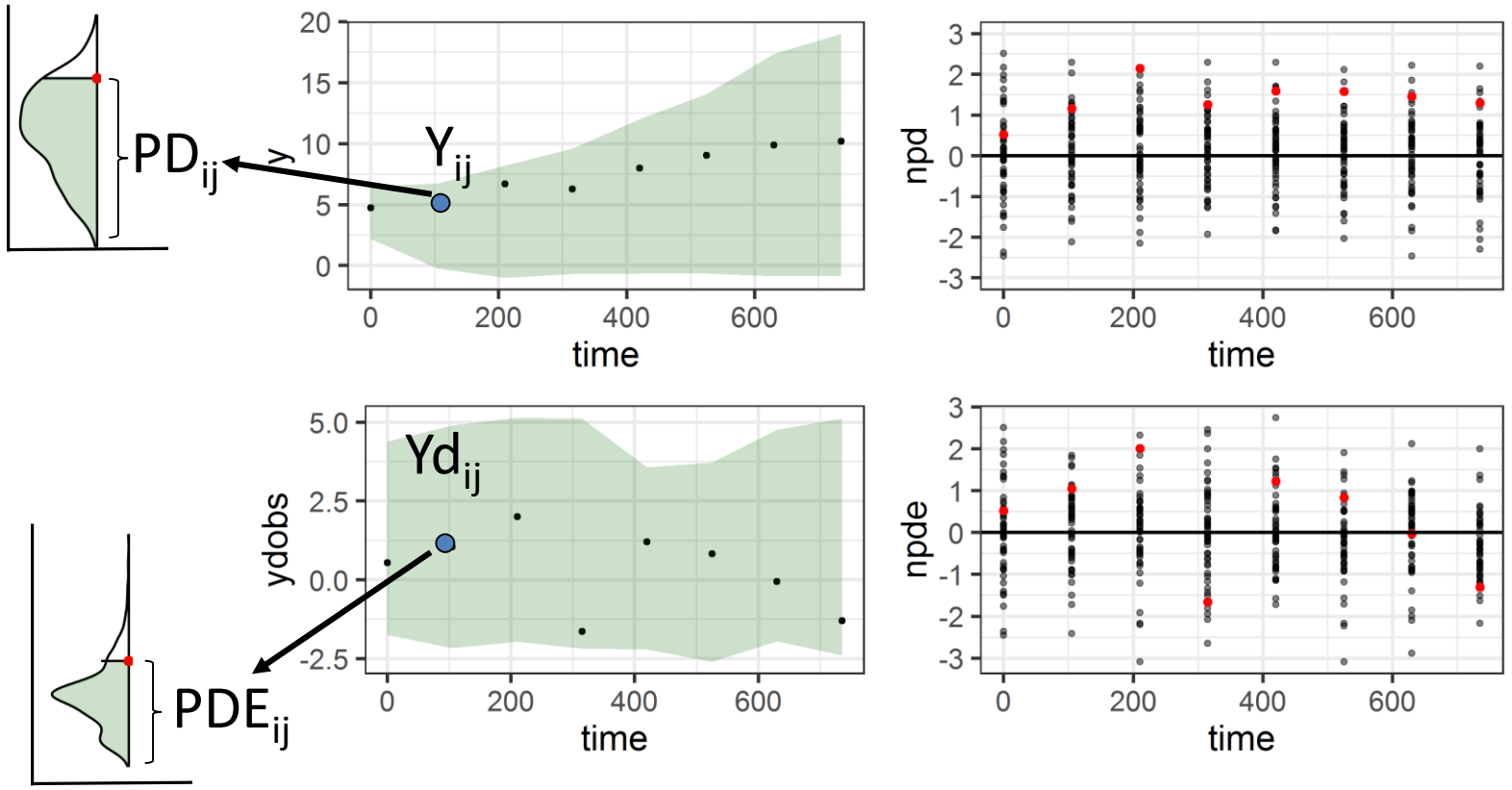


[1] Brendel K et al. Pharm Res. 2006

Development of npde for continuous data [1]

$$pd_{ij} = F_{ij}(y_{ij}) = \int_0^{y_{ij}} p_i(y, \Psi)dy = \int^{y_{ij}} \int p(y|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dy$$

- prediction discrepancies **pd**: quantile of an observation in its predictive distribution
- prediction distribution error **pde**: quantile of a decorrelated observation in its decorrelated predictive distribution
- normalised prediction distribution **npd**: normalization of pd
- normalised prediction distribution error **npde**: normalization of pde

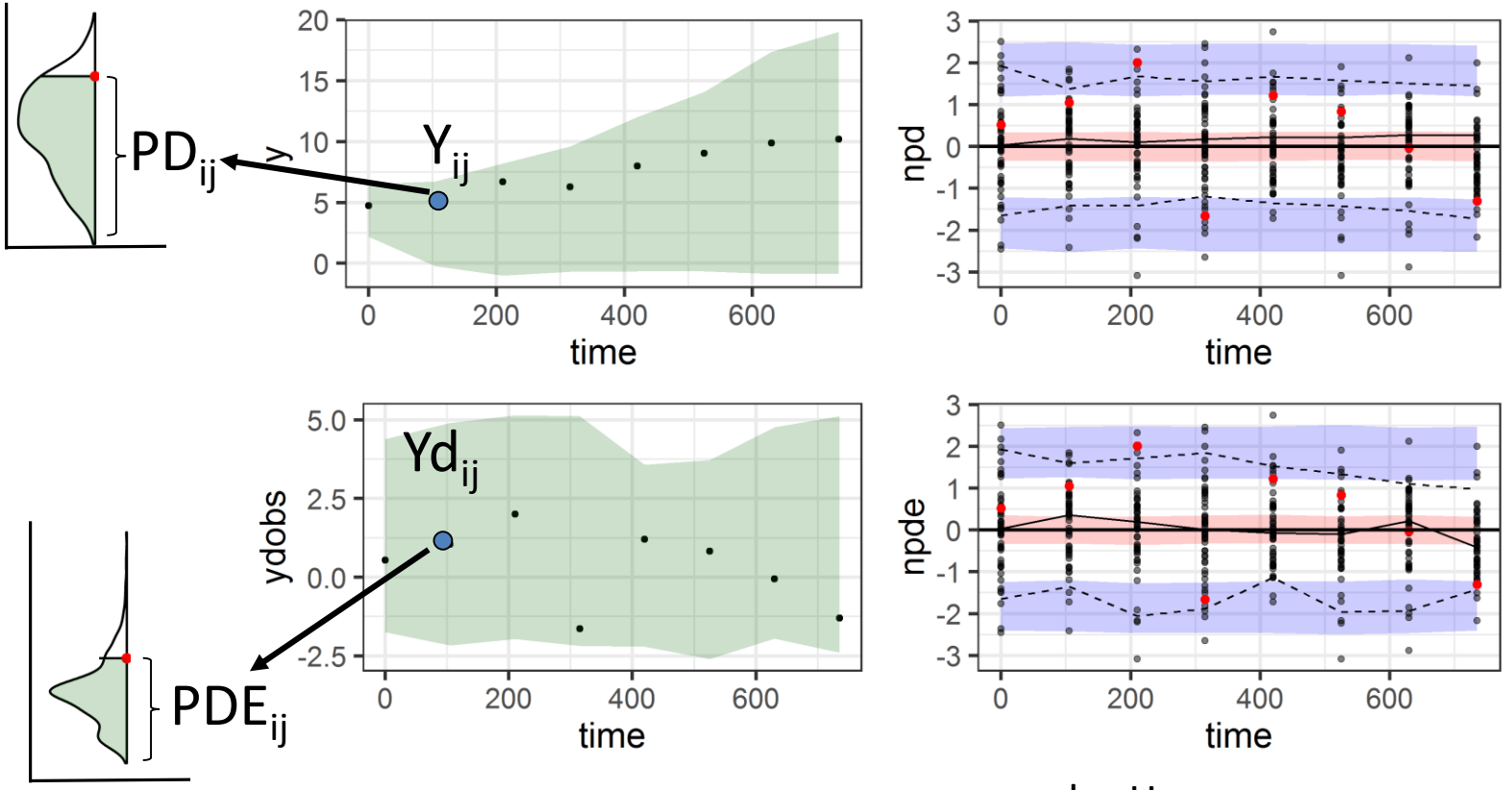


[1] Brendel K et al. Pharm Res. 2006

Development of npde for continuous data [1]

$$pd_{ij} = F_{ij}(y_{ij}) = \int_0^{y_{ij}} p_i(y, \Psi)dy = \int^{y_{ij}} \int p(y|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dy$$

- prediction discrepancies **pd**: quantile of an observation in its predictive distribution
- prediction distribution error **pde**: quantile of a decorrelated observation in its decorrelated predictive distribution
- normalised prediction distribution **npd**: normalization of pd
- normalised prediction distribution error **npde**: normalization of pde

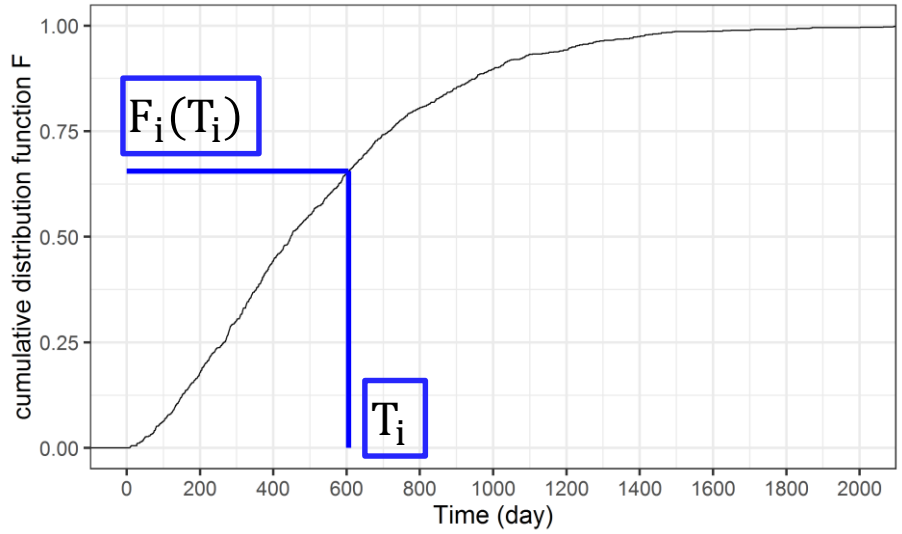


- under H_0:
  $pd/pde \sim U(0,1)$
  $npd/npde \sim \mathcal{N}(0,1) \rightarrow$ test

[1] Brendel K et al. Pharm Res. 2006

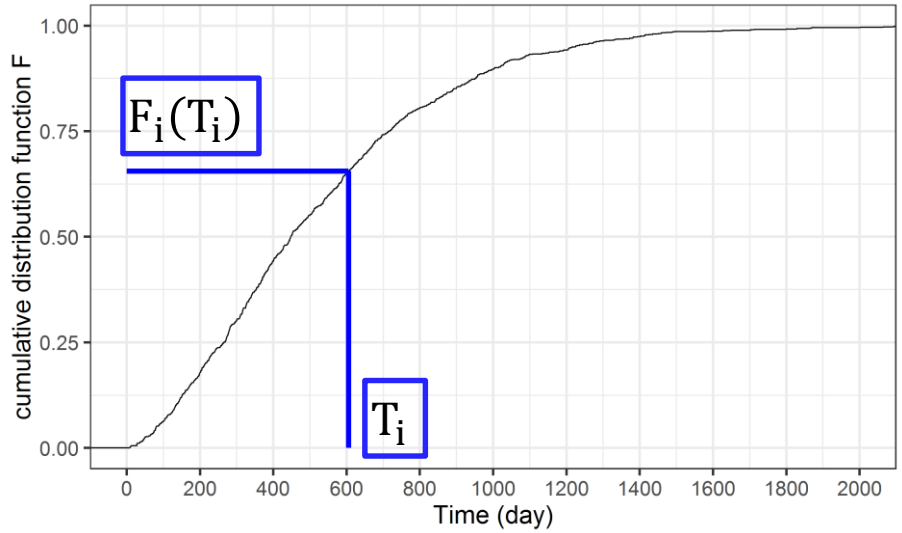
Development of npd for TTE data

- Time to event (T_i) is continuous
- pd for observed event time:
  - $pd_i = F_i(T_i) = \int_0^{T_i} p_i(t, \Psi)dt = \int_0^{T_i} \int p(t|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dt$



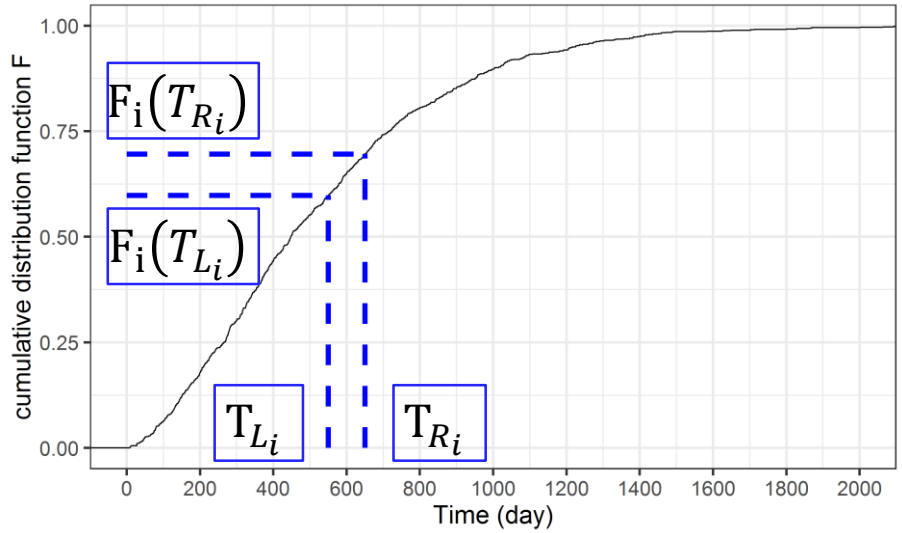
Development of npd for TTE data

- Time to event (T_i) is continuous
- pd for observed event time:
  - $pd_i = F_i(T_i) = \int_0^{T_i} p_i(t, \Psi)dt = \int_0^{T_i} \int p(t|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dt$
- Time to event (T_i) can be censored (right / left / interval censored event)
  - How to deal with censoring ?



Development of npd for TTE data

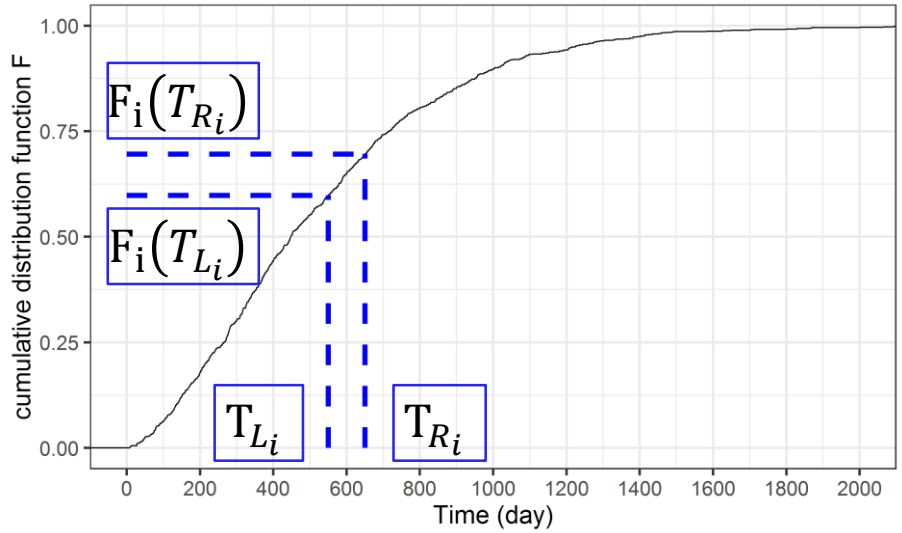
- Time to event (T_i) is continuous
- pd for observed event time:
  - $pd_i = F_i(T_i) = \int_0^{T_i} p_i(t,\Psi)dt = \int_0^{T_i} \int p(t|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dt$
- Time to event (T_i) can be censored (right / left / interval censored event)
  - How to deal with censoring ?
- pd for censored event time: same idea as [1] with npde for BLQ data:
  - T_i lies within the interval $[T_{L_i}; T_{R_i}]$
  - $pd_i = \mathrm{U}\left(F(T_{L_i}), F(T_{R_i})\right)$



[1] Nguyen THT et al. J Pharmacokinet Pharmacodyn. 2012

Development of npd for TTE data

- Time to event (T_i) is continuous
- pd for observed event time:
  - $pd_i = F_i(T_i) = \int_0^{T_i} p_i(t,\Psi)dt = \int_0^{T_i} \int p(t|\boldsymbol{\theta_i})p(\boldsymbol{\theta_i})d\boldsymbol{\theta_i}dt$
- Time to event (T_i) can be censored (right / left / interval censored event)
  - How to deal with censoring ?
- pd for censored event time: same idea as [1] with npde for BLQ data:
  - T_i lies within the interval $[T_{L_i}; T_{R_i}]$
  - $pd_i = \mathrm{U}\left(F(T_{L_i}), F(T_{R_i})\right)$



In practice

- NLMEM: no analytical solution for F
- F approximated by Monte-Carlo integration

[1] Nguyen THT et al. J Pharmacokinet Pharmacodyn. 2012

Computation of the combined test for joint model

- Computing npd / npde
  - **npde** for continuous data obtained as previously by decorrelating the observed and simulated continuous response
  - **npd** for TTE obtained using the inverse normal function of the pd (only one observation)
- Combined statistical test
  - Pr(reject H_0 | **npde**$^{(1)}$) with a Kolmogorov-Smirnov (KS) test of normality $\mathcal{N}(0,1)$
  - Pr(reject H_0 | **npd**$^{(2)}$) with a KS test of normality $\mathcal{N}(0,1)$
    - **Bonferroni correction**: $\Pr(\text{reject } H_0) = \min\left(\Pr\left(\text{reject } H_0 \mid \boldsymbol{npde}^{(1)}\right); \Pr\left(\text{reject H0} \mid \boldsymbol{npd}^{(2)}\right)\right) \times \mathbf{2}$
    - rejection if Pr(reject H_0) < 0.05

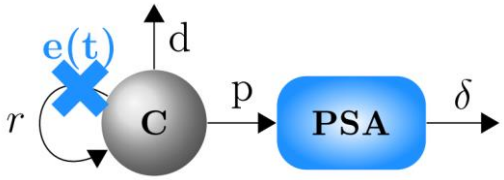
Outline

- Development of npde for the evaluation of joint model with longitudinal and TTE data
- Performance of the statistical test
- How can we visually diagnose model deficiencies ?

Simulation study - Model

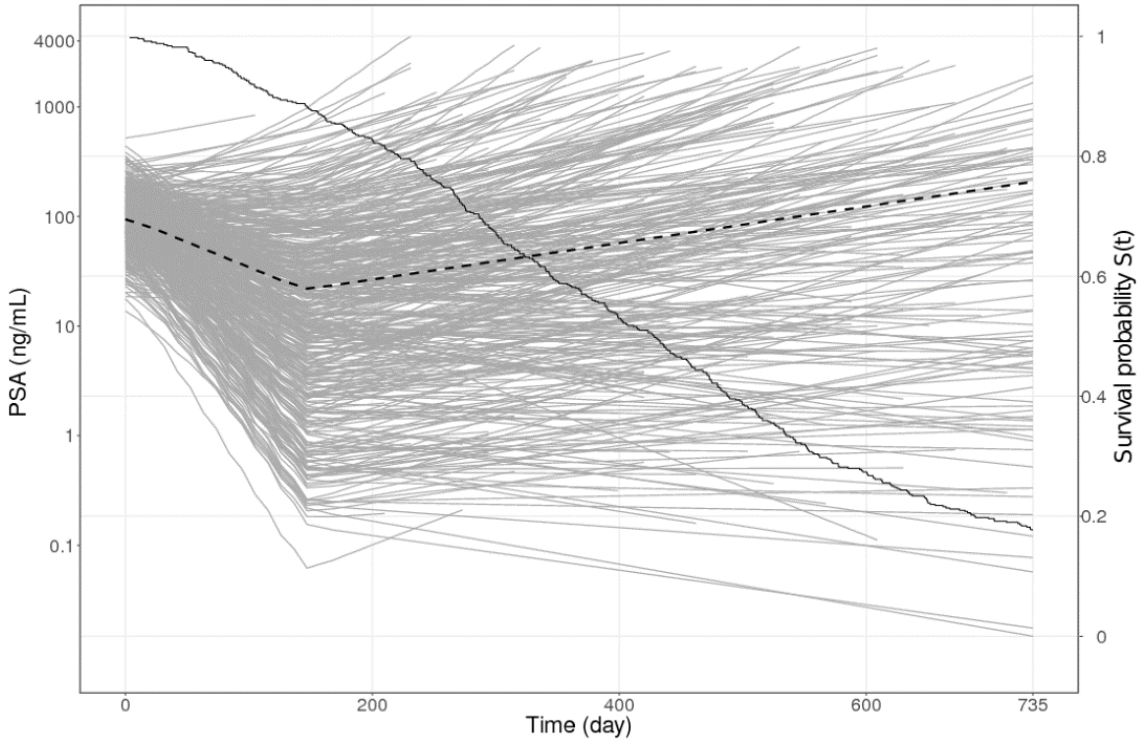
- Inspired from the work of Desmée et al. [1]
  - metastatic castration-resistant prostate cancer patient
  - original design based on a clinical trial
- Data:
  - primary outcome: **survival**
  - biomarker: Prostate Specific Antigen (PSA)

Model for PSA: biexponential



Model for the TTE outcome:

- $h_i(t|\theta_i) = \frac{k}{\lambda}\left(\frac{t}{\lambda}\right)^{k-1} \times \exp\left(\beta . \boldsymbol{PSA(\theta_i, t)}\right)$



Parameters	Distribution	Value	i.i.v (ω)
r (day^{-1})	Log-normal	0.05	0.1
PSA_0 ($ng.mL^{-1}$)	Log-normal	80	0.6
ε	Logit-normal	0.3	1.5
Tesc (day)	Log-normal	140	0.6
d (day^{-1})	-	0.046	
δ (day^{-1})	-	0.23	
k	-	1.5	-
λ (day)	-	580	-
β	-	0.001	-

[1] Desmée S et al. AAPS J. 2015

Simulation study

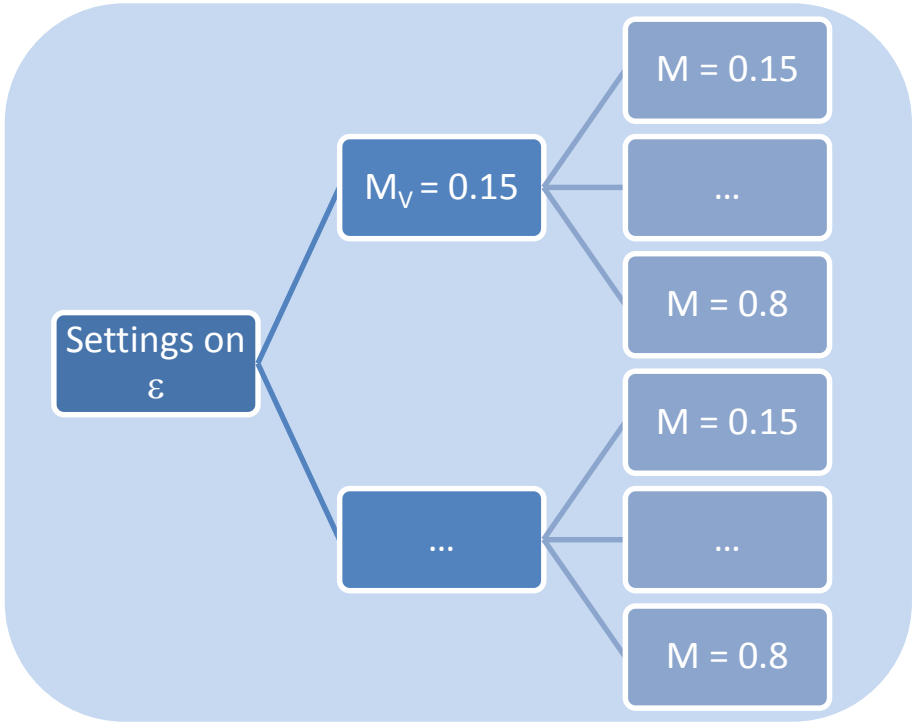
Settings

- Misspecification on **PSA model parameter**
  - ε : {0.15, 0.3, 0.45, 0.6, 0.8}
- Misspecification on **TTE model parameter**
  - β : $\{0, 0.001, 0.003, 0.005\}$
- Design
  - $\{50, 100, 200\}$ patients
  - follow-up censored at 735 days
  - up to 9 measurements

Simulation study

Settings

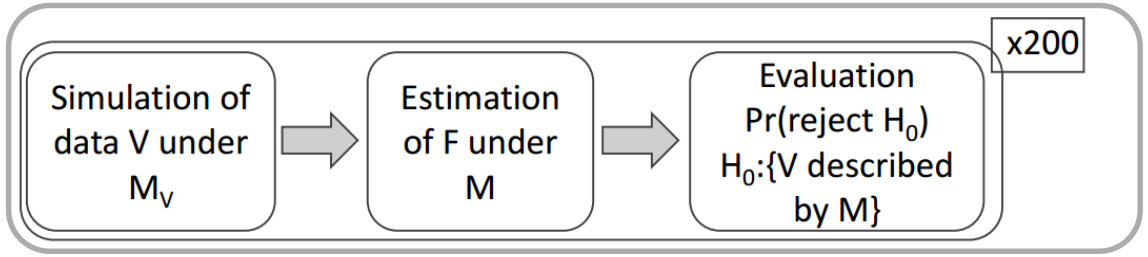
- Misspecification on **PSA model parameter**
  - ε : {0.15, 0.3, 0.45, 0.6, 0.8}
- Misspecification on **TTE model parameter**
  - β : $\{0, 0.001, 0.003, 0.005\}$
- Design
  - $\{50, 100, 200\}$ patients
  - follow-up censored at 735 days
  - up to 9 measurements



Evaluation criteria

- Notation used:
  - V: dataset generated under model M_V (e.g. $\varepsilon = 0.15$)
  - M: model to test (e.g. $\varepsilon = 0.8$)
  - H_0: the data V can be described by the model M
- Performance of npde:

N={50, 100, 200}

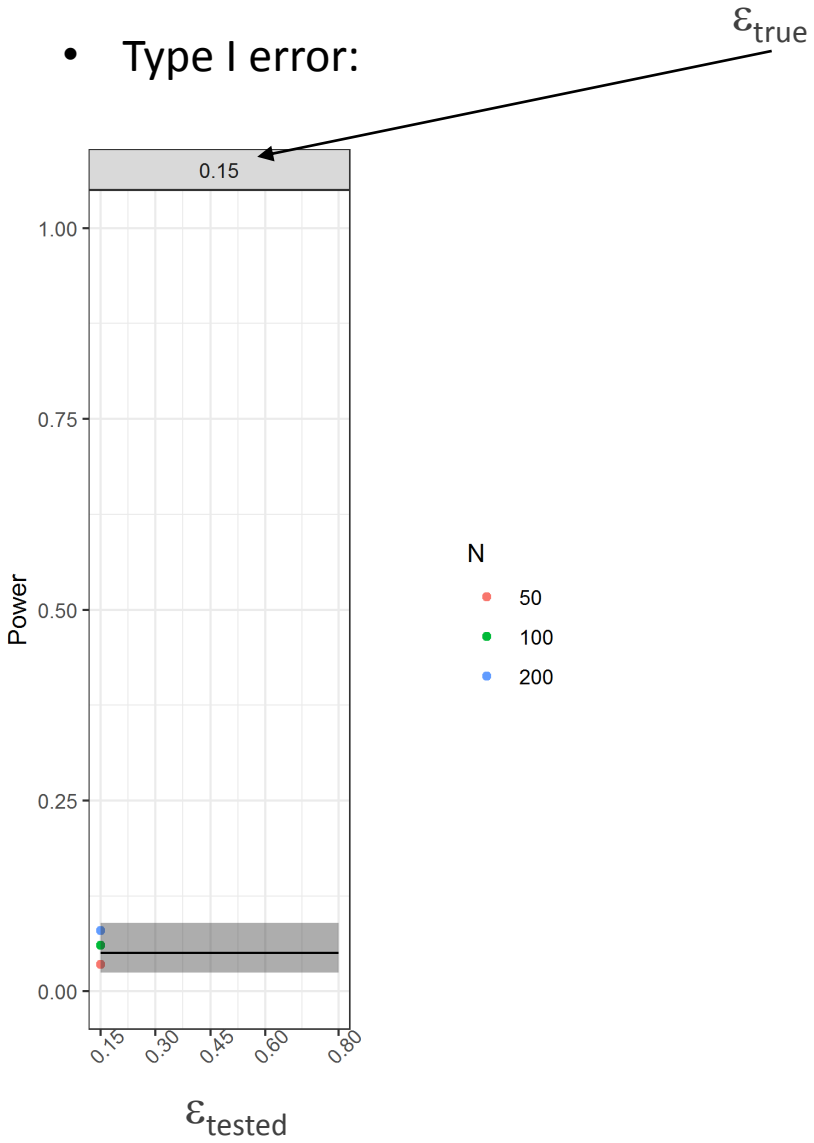


- **type I error**: % of rejection of M under H_0 ($M=M_V$)
- **power**: % of rejection of M under H_1 ($M \neq M_V$)

Simulation study – Misspecification on PSA model

Parameter: ε

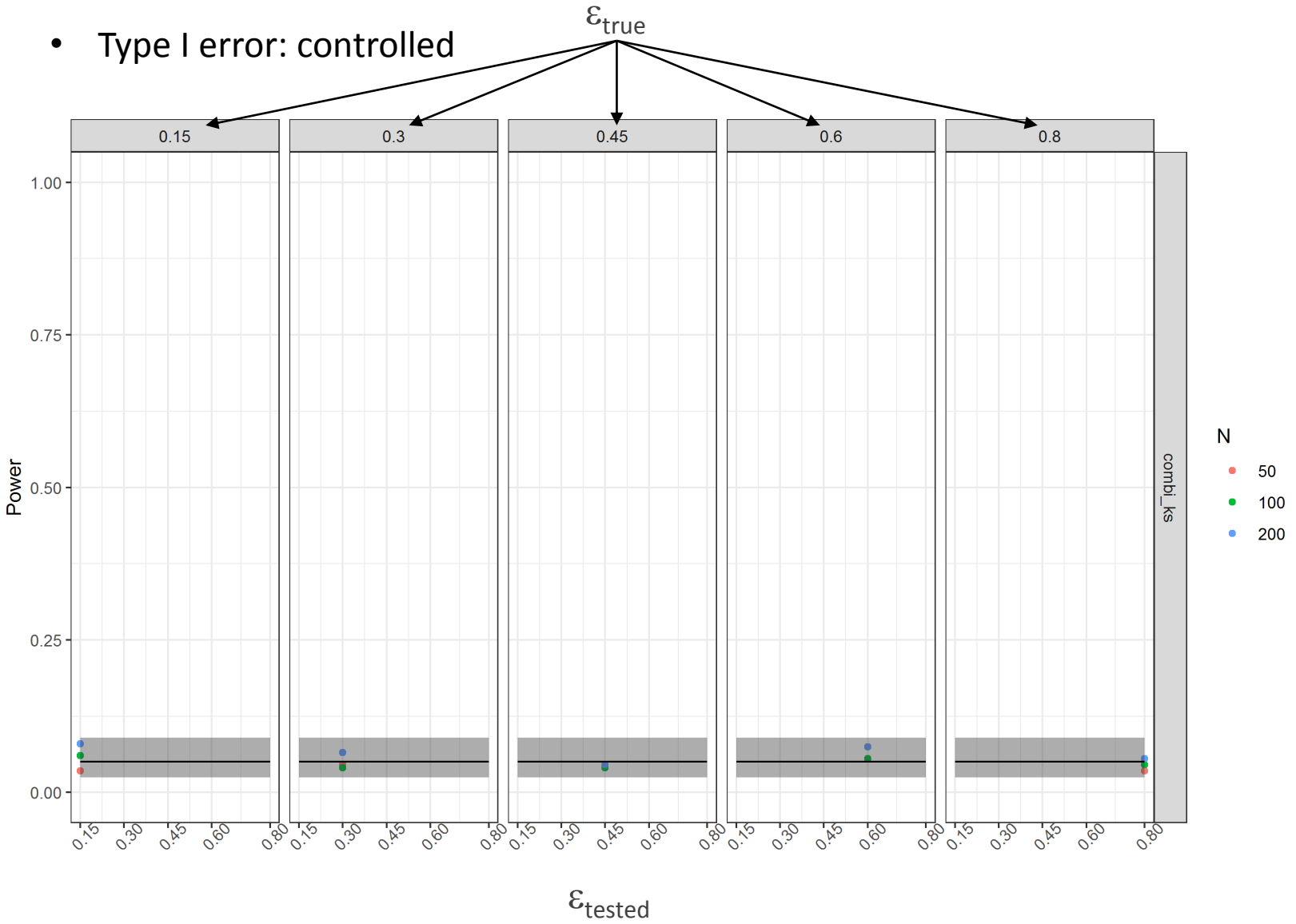
- Type I error:



Simulation study – Misspecification on PSA model

Parameter: ε

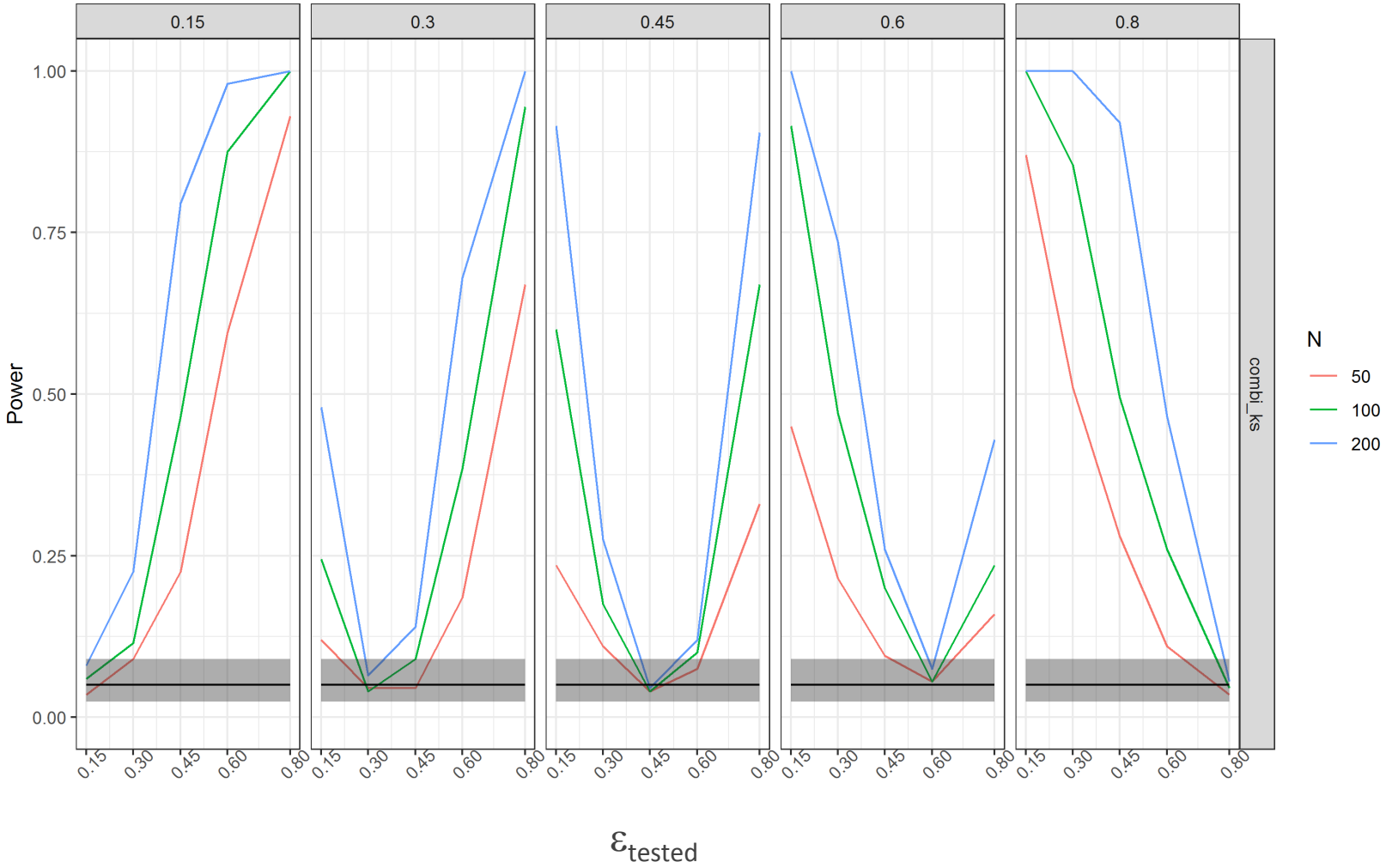
- Type I error: controlled



Simulation study – Misspecification on PSA model

Parameter: ε

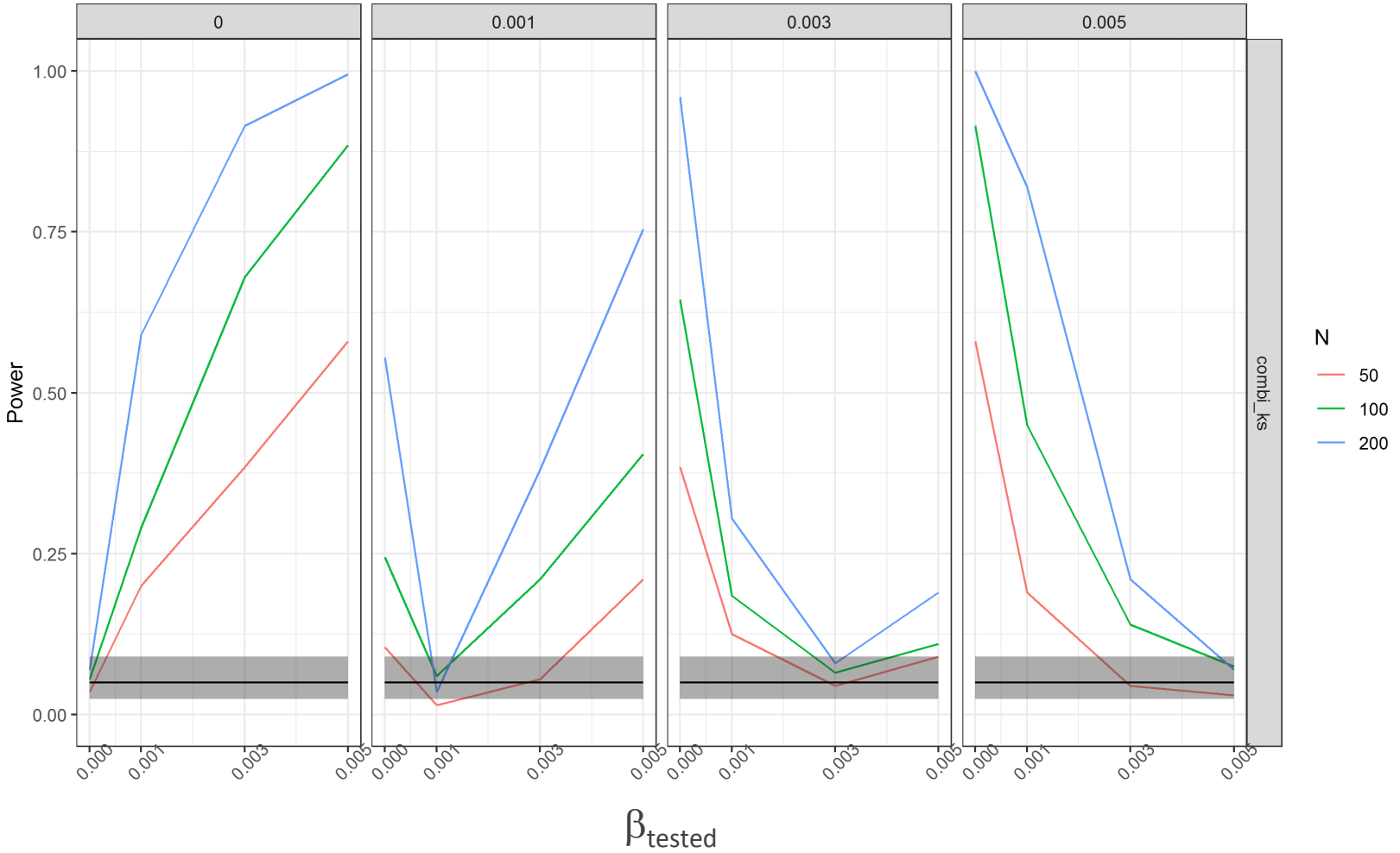
- Type I error: controlled
- Power increases as difference from true value and N increases



Simulation study – Misspecification on TTE model

Parameter: β

- Type I error: controlled
- Power increases as difference from true value and N increases



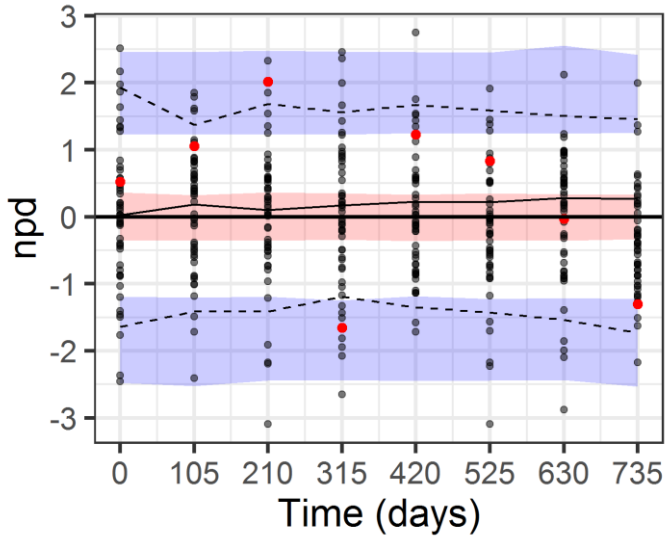
Outline

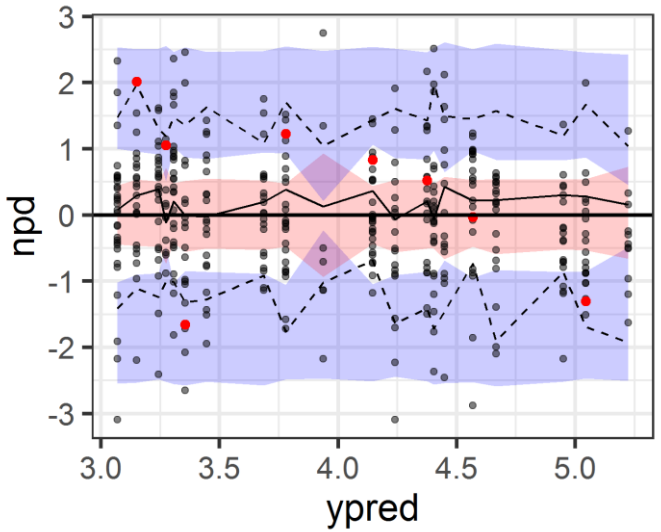
- Development of npde for the evaluation of joint model with longitudinal and TTE data
- Performance of the statistical test
- How can we visually diagnose model deficiencies ?

How can we visually diagnose model deficiencies ?

For continuous outcome:

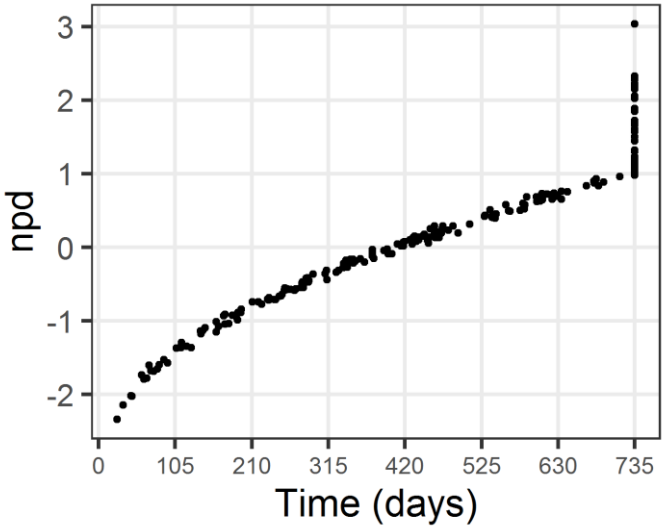
- scatter plot of npd/npde vs time and pred



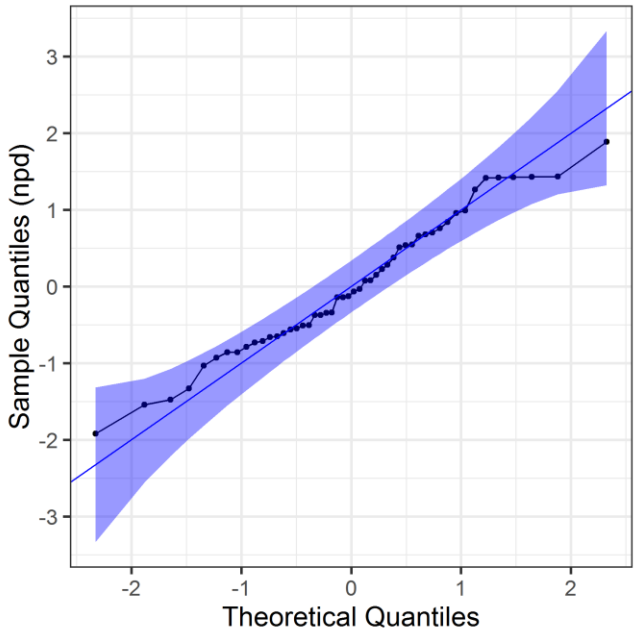


For TTE outcome:

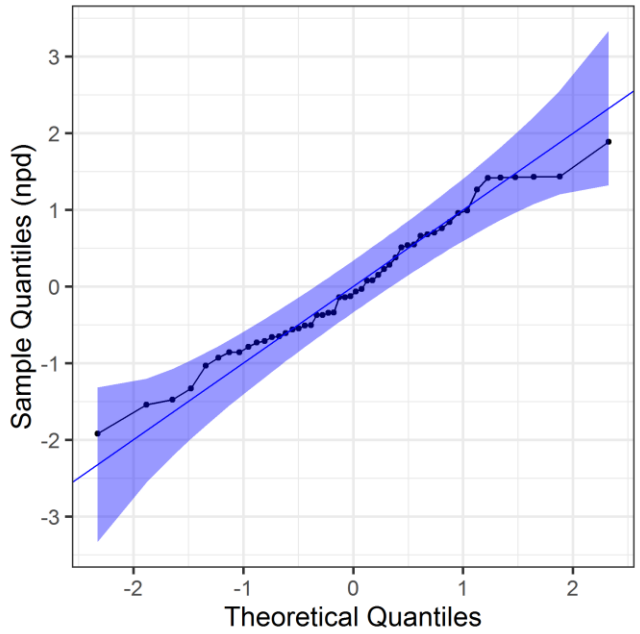
- time is the predictor itself
- if npd vs time: a trend is expected (population residuals vs. individual observations)

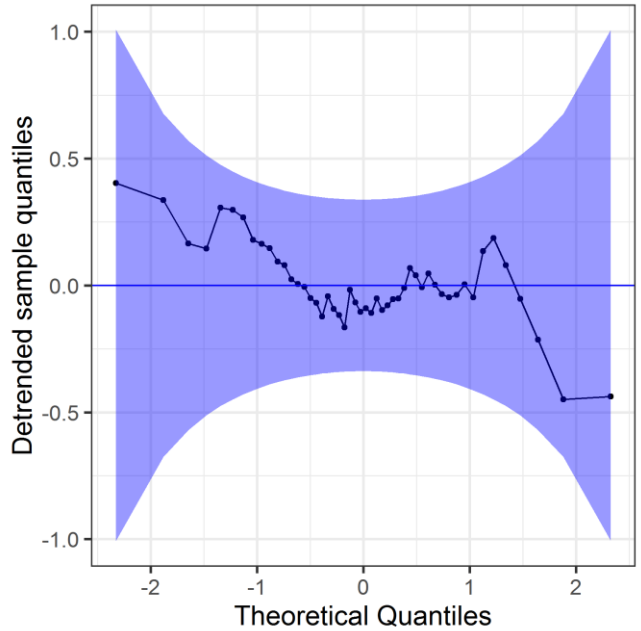


How can we visually diagnose model deficiencies ?

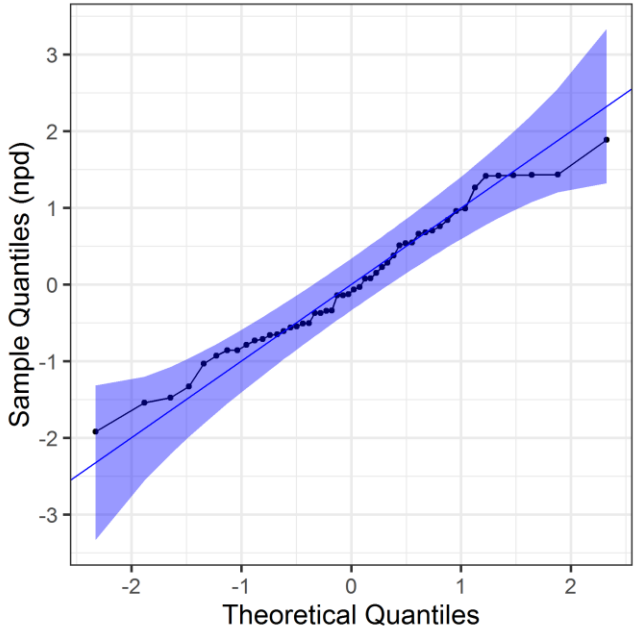


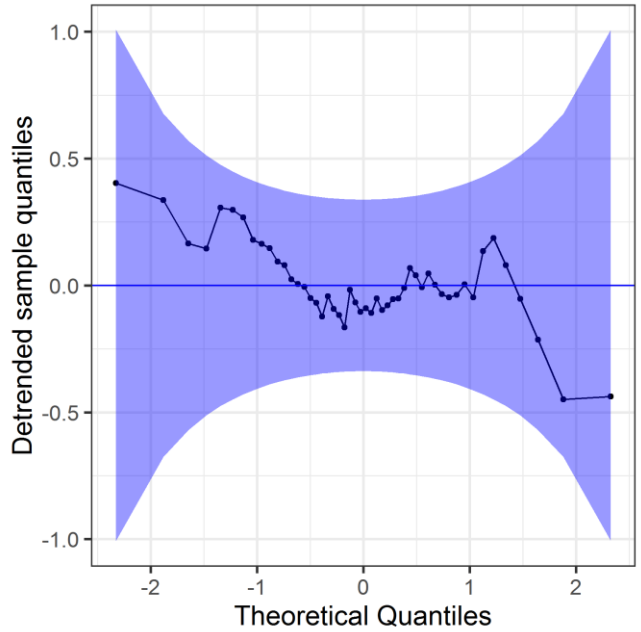
How can we visually diagnose model deficiencies ?

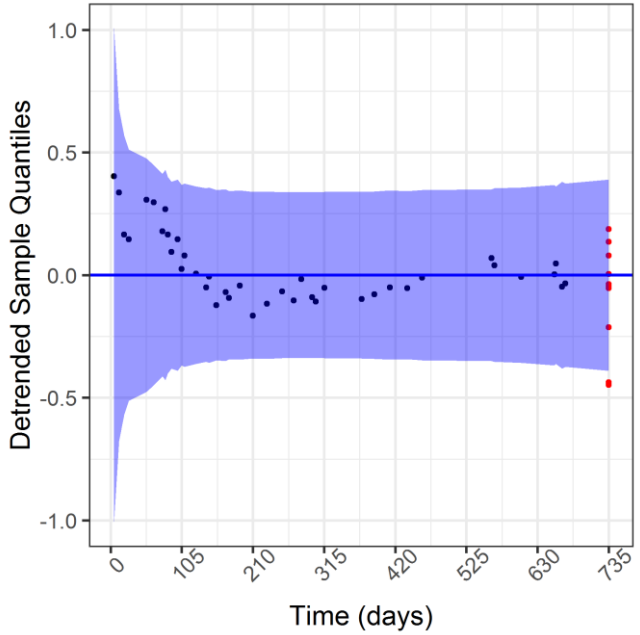




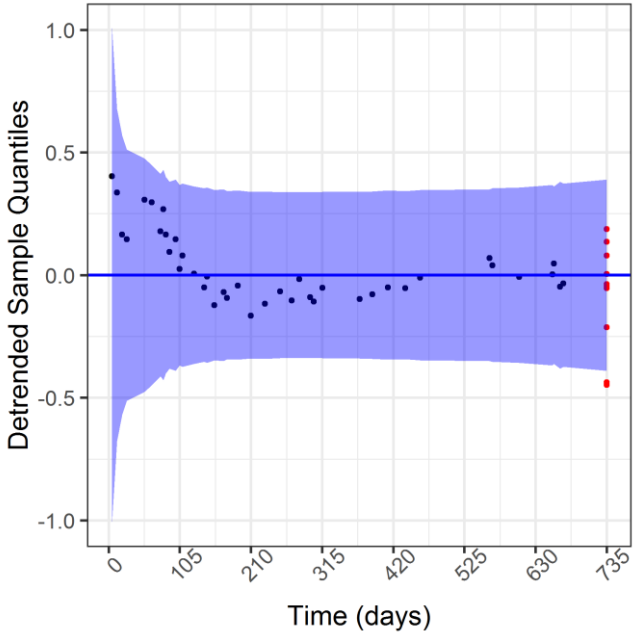
How can we visually diagnose model deficiencies ?





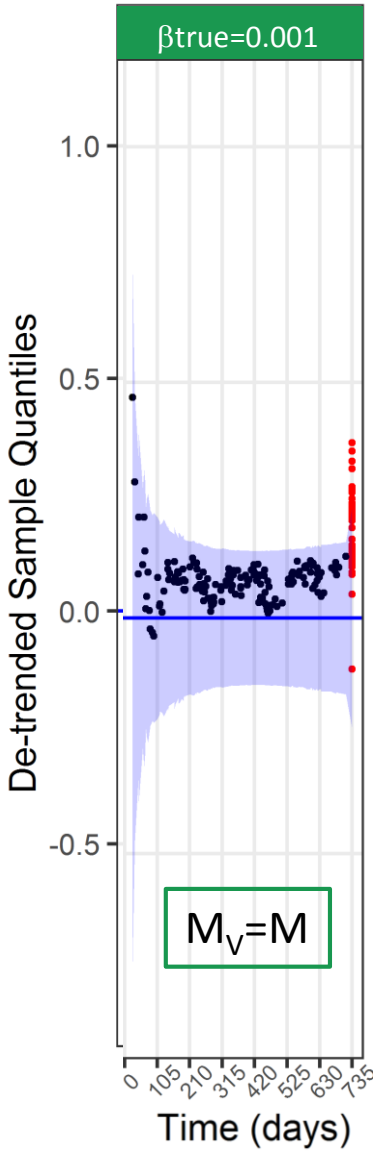


How can we visually diagnose model deficiencies ?



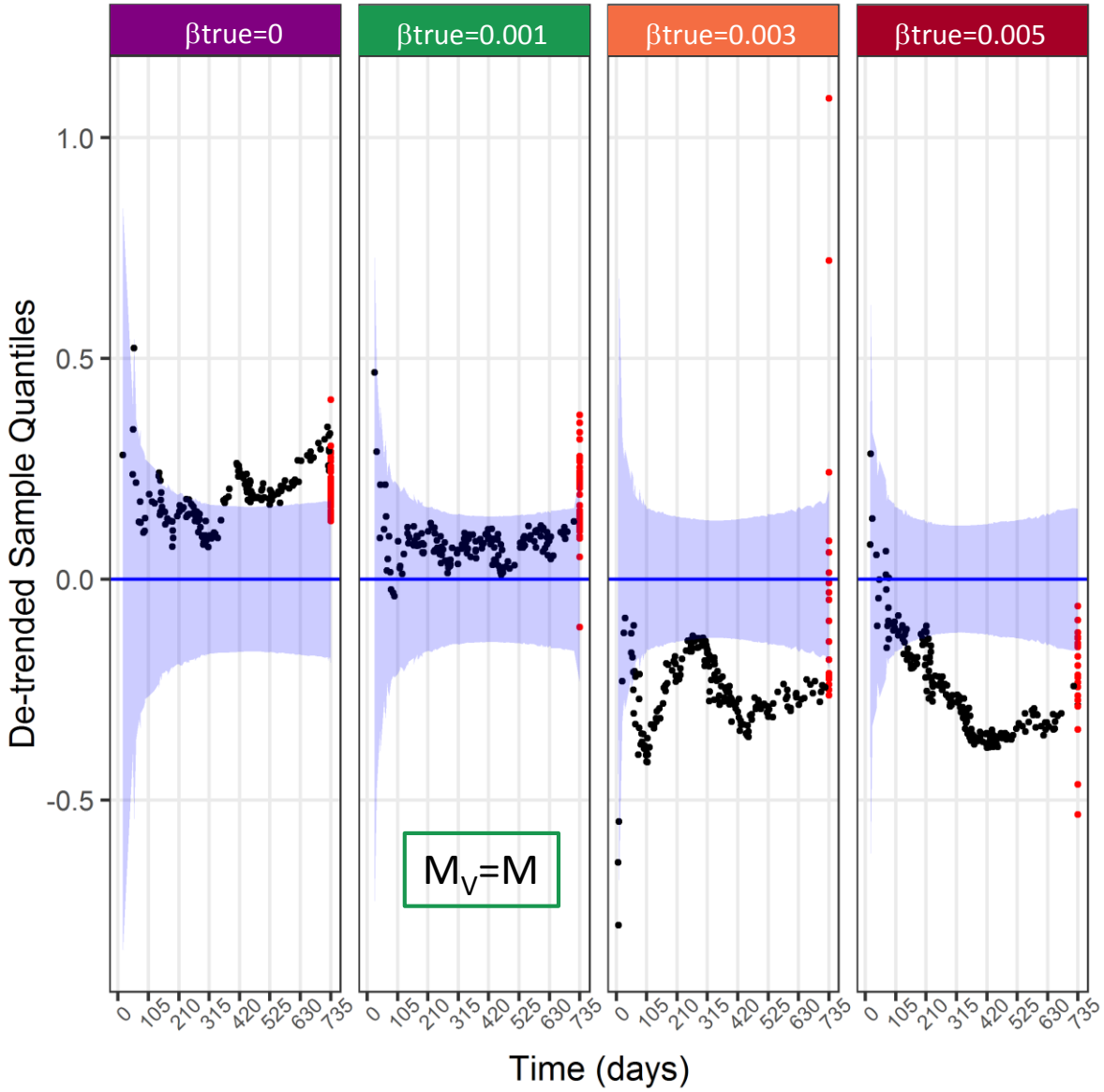
Example of graphs – TTE part

Strength of the impact of PSA on survival (β tested): 0.001



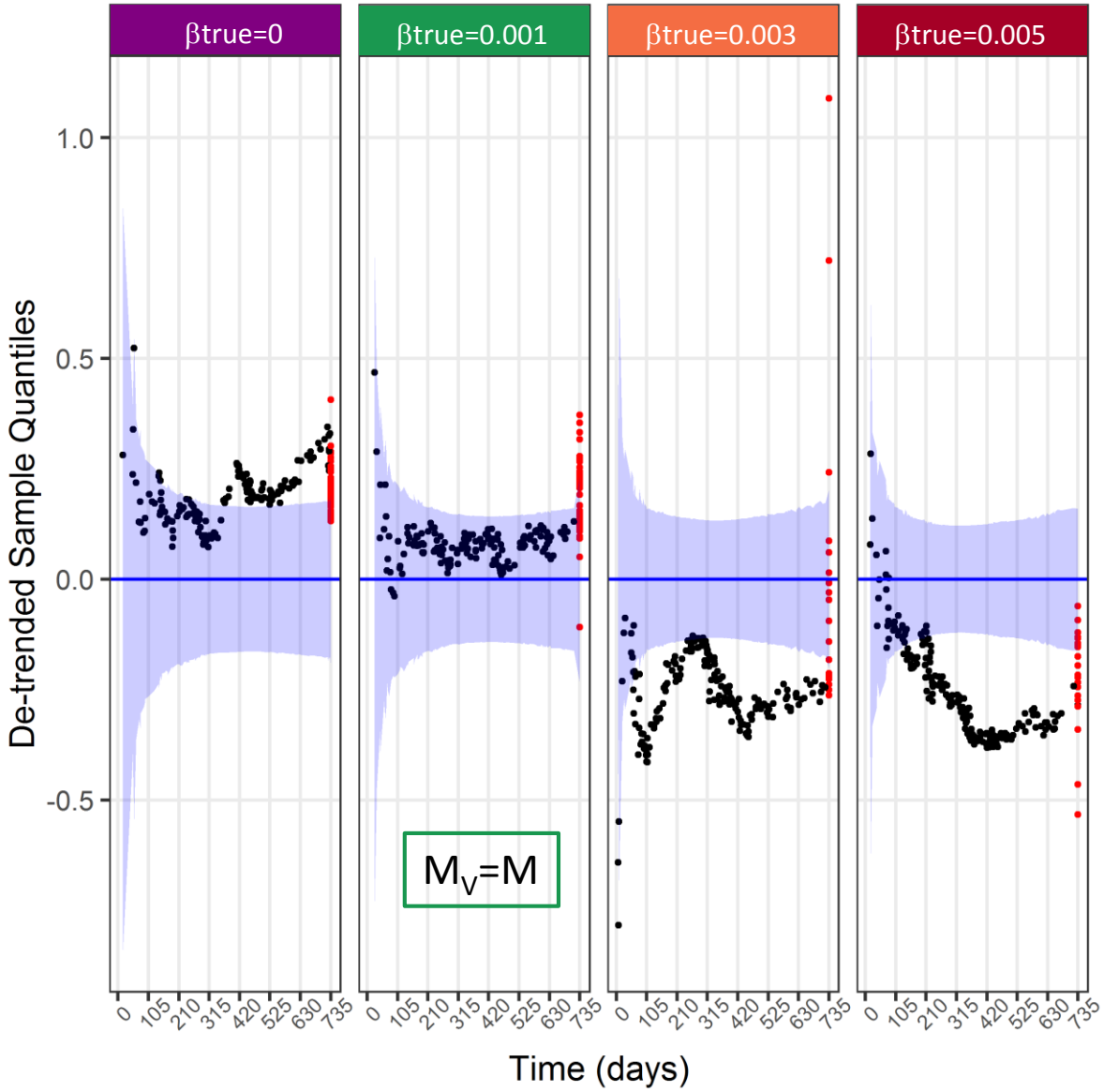
Example of graphs – TTE part

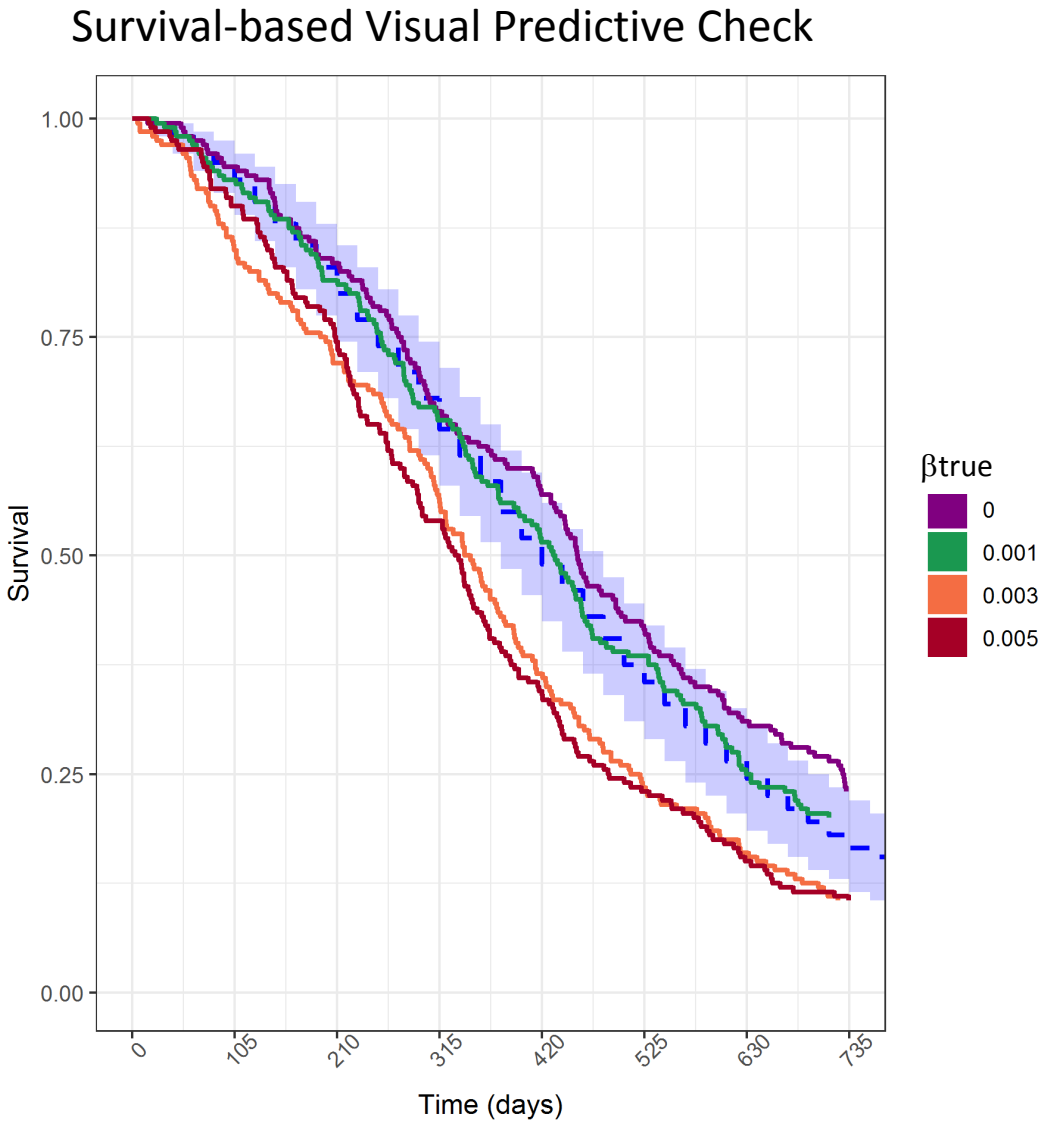
Strength of the impact of PSA on survival (β tested): 0.001



Example of graphs – TTE part

Strength of the impact of PSA on survival (β tested): 0.001





Discussion

- Good performance of the combined test
  - adequate type I error
  - power increases as the difference in shape differs
  - also evaluated in other settings
- Graphs are useful to assess the inadequacies of the model
- More details about the npd for TTE in [1]
  - good performance of npd-TTE evaluated on survival data alone, assuming longitudinal model is correct

[1] Cerou M et al. Pharm Res. 2018

Perspectives

- Application on data from a clinical trial
- Comparison of npd-TTE to other diagnostics for survival-type data:
  - Survival-based VPC
  - Hazard-based VPC [1]
  - Kaplan Meier Mean Covariate [2]
- Extension to repeated TTE

[1] Yeamin H and Hutmacher MM. J Pharmacokinet Pharmacodyn. 2016
[2] Hooker A and Karlsson MO. PAGE 2012

Acknowledgment

- Grant: Servier
- To my PhD supervisor
  - Emmanuelle Comets
  - Marylore Chenel
- To my Inserm colleagues
- To my Servier colleagues