Prediction discrepancies (pd) for evaluation of models with data under limit of quantification

Thi Huyen Tram Nguyen, Emmanuelle Comets and France Mentré

UMR 738 INSERM and University Paris Diderot, Paris, France

Introduction

- Nonlinear mixed effect models (NLMEM) are increasingly used for analysis of longitudinal data in clinical trials or cohorts
- Evaluation is an important part of modeling. Simulation-based approaches have been proposed such as VPC, prediction discrepancies (pd) and normalised prediction distribution errors (npde)[1-6]
- Data below the quantification limit (BQL data) are a common challenge for longitudinal data analysis in clinical trials, particularly in HIV clinical trials
  - ➢ appropriate estimation methods have been proposed to take them into account, and have been implemented in reference software (NONMEM, MONOLIX)
  - ➢ however, evaluation methods do not take into account BQL data
- Omitting BQL data for the evaluation plots, as often done, could introduce fake indications of model misspecification if the amount of BQL data is large

Objectives

1. To develop an extension to pd taking into account BQL observations
2. To illustrate the use of this new method on simulated data

Data and model

- Data from the COPHAR 3 – ANRS 134 multicenter clinical trial:
  - ➢ 34 naïve HIV-infected patients treated once daily with atazanavir, ritonavir and tenofovir/emtricitabine during 24 weeks
  - ➢ viral load were measured on the 1st day of treatment and at the 4th, 8th, 12th, 16th (20th) and 24th weeks
  - ➢ limits of quantification of HIV assay are 40 or 50 copies/mL
- A bi-exponential model (Equation 1)[7] was used to describe HIV viral load decrease during treatment. Parameter estimates for the real data were obtained using the SAEM algorithm in MONOLIX 3.2[8]

$$f(\theta,t) = \log_{10}(P_1 e^{-\lambda_1 t} + P_2 e^{-\lambda_2 t}) \qquad \textit{(Equation 1)}$$

Methods

- Notations: N subjects i = 1,…, N

  $y_{ij} = f(q_i, t_{ij}) + g(q_i, t_{ij})\,\varepsilon_{ij}$: observation for individual *i* at time t_{ij}

  f : structural nonlinear model; g : model for residual error

  ε_{ij}: residual errors – $\varepsilon_{ij} \sim N(0,\sigma^2)$

$$\theta_i = \mu \times e^{\eta_i}$$

  - θ_i : vector of the individual parameters for subject *i*
  - μ : vector of the *p* fixed effect parameters
  - η_i : vector of the *q* random effect parameters
  - $\eta_i \sim N(0, \Omega)$: Ω defined as a $q \times q$ - non diagonal matrix
- Prediction discrepancy pd_{ij} for observation y_{ij} above limit of quantification (LOQ)

$$pd_{ij} = F_{ij}(y_{ij}) = \int_{-\infty}^{y_{ij}} p(y|\theta_i)p(\theta_i)d\theta_i dy \approx \frac{1}{K}\sum_{k=1}^{K} 1_{y_{ij}^{sim(k)} < y_{ij}}$$

  F_{ij}: cumulative distribution function (cdf) of the predictive distribution of y_{ij} under tested model obtained by *K* Monte-Carlo simulations
- Prediction discrepancy pd_{ij} for observation y_{ij} below LOQ

  pd_{ij} is randomly sampled from a uniform distribution U[0, Pr$(y_{ij} \le LOQ)$]

$$\text{with } \Pr(y_{ij} \le LOQ) = F_{ij}(LOQ) \approx \frac{1}{K}\sum_{k=1}^{K} 1_{y_{ij}^{sim(k)} < LOQ}$$

- npd = ϕ^{-1}(pd); if the model is correct, npd ~ N(0,1)
  npd are correlated if repeated measurements within subjects
- npde are the decorrelated version of npd, computed as described in [3-6]
  if the model is correct, npde ~ N(0,1).
- Evaluation graph: scatterplot of npd (npde) vs time with the 10th, 50th and 90th percentiles corresponding to observed data. To facilitate model evaluation, the 95% prediction intervals for these selected percentiles of simulated data are added into graph as colored bands[6]
- Tests of npd (npde): Wilcoxon, Fisher and Shapiro Wilks tests are used [3-6]
  - ➢ global p-value is obtained using Bonferroni correction

Simulation study

- Designs: 300 observations at 0, 24, 56, 84, 112, 168 days after initiation of treatment
  - ➢ sparse design: N = 300 subjects, n = 1 observation/subject
  - ➢ rich design: N = 50 subjects, n = 6 observations/subject
- Models: to simulate different validation datasets V
  - ➢ "true" model (H_0) inspired by the real data results
  - ➢ "false" models with modification in fixed (H_{1_mean}) or random effects parameters (H_{1_var})
- LOQ levels: 0, 20 or 50 copies/mL
- Computation of npd using the new approach: K = 1000 MC simulations
- Type I error and power using the global test of npd (or npde): 1000 validation datasets were simulated for each scenario

Results

1. Parameter estimates for the real data

Table 1. Estimated parameters from real data and parameters for simulation study

	Estimation		Simulation		
	Estimate	RSE (%)	True model	False model "mean"	False model "var"
P1 (cp/mL)	21900	36	25000	25000	25000
P2 (cp/mL)	182	32	250	250	250
λ_1 (day^{-1})	0.205	6	0.2	0.2	0.2
λ_2 (day^{-1})	0.0195	12	0.02	**0.04**	0.02
ω_{P1}	2.07	12	2.1	2.1	2.1
ω_{P2}	1.5	18	1.4	1.4	1.4
$\omega_{\lambda 1}$	0.206	21	0.3	0.3	0.3
$\omega_{\lambda 2}$	0.301	25	0.3	0.3	**0.9**
$\rho(\eta_{P1},\eta_{P2})$	0.856	7	0.8	0.8	0.8
Additive σ	0.15	4	0.15	0.15	0.15

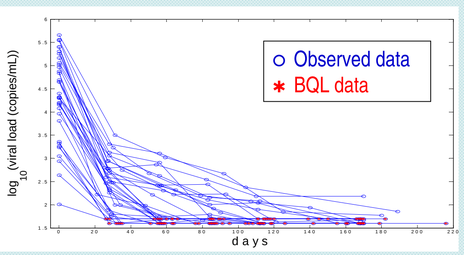

Figure 1. Spaghetti plot of COPHAR 3 viral load of 34 patients

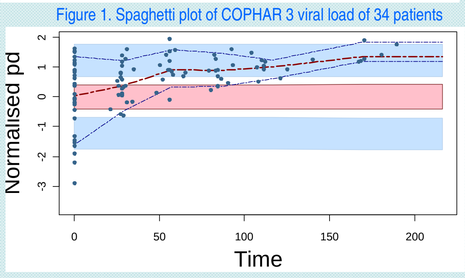

Figure 2. NPD vs time for the COPHAR 3 dynamic model calculated by omitting the BQL observations

- Data : 205 observations in 34 patients with 49.8% BQL data
- BQL data are taken into account in model building step using the extended version of SAEM[8]
- Parameters are well estimated (***Table 1***). The scatterplot of npd vs time computed by omitting the BQL data suggests model misspecifications

2. Graphic illustration (*Figure 3*)

- H_0, omitting BQL data: clear departure of the median of npd from 0
- With new approach:
  - ➢ H_0: allows to select the right model
  - ➢ H_{1_mean}: shift of npd away from 0, becoming less clear as % of BQL data increases
  - ➢ H_{1_var}: model misspecification not apparent when % of BQL data increases

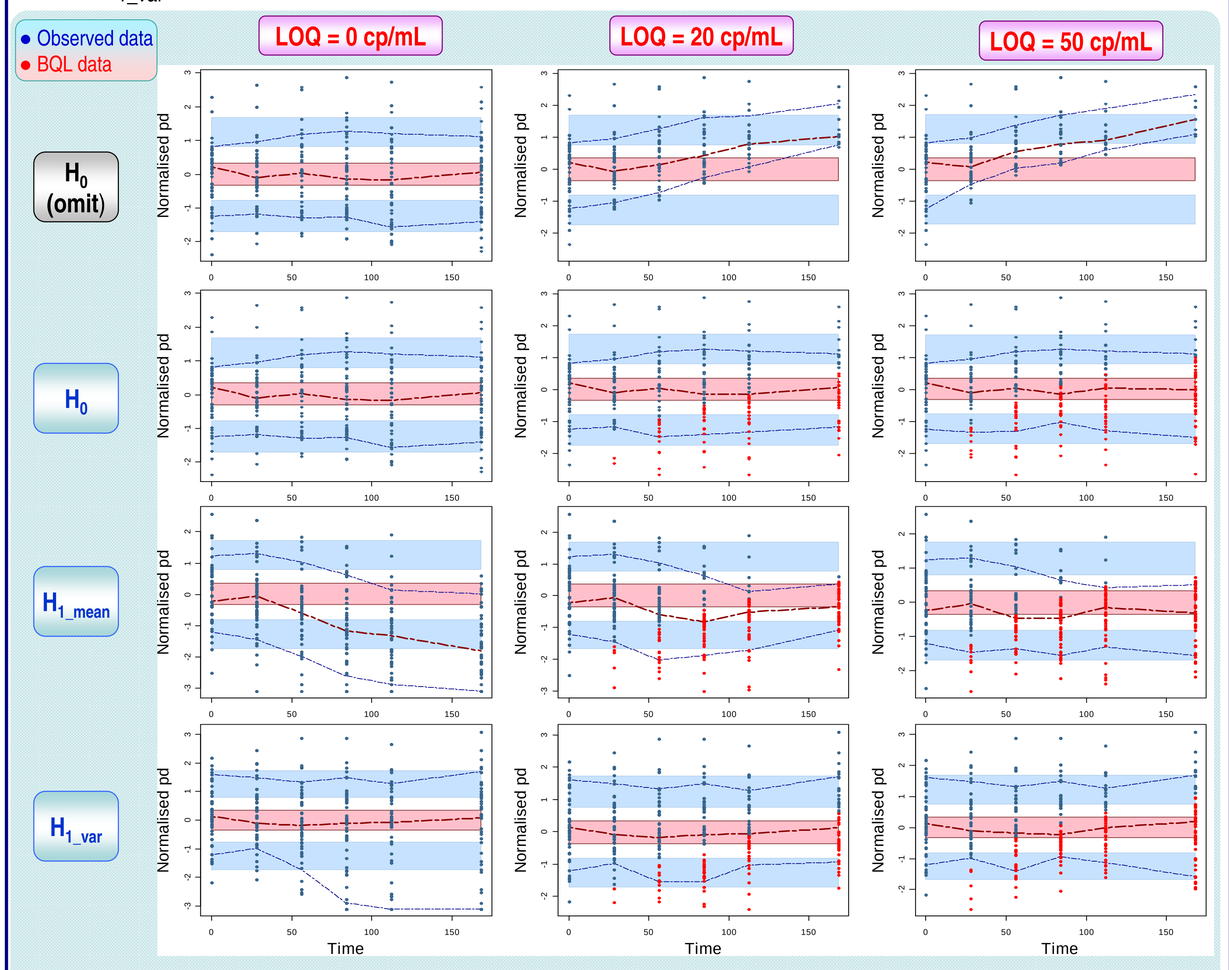

Figure 3. NPD vs time of 1 simulated dataset under different models at 3 LOQ levels (rich design)
H_0(omit): NPD computed by omitting BQL data for the basic model. H_0, H_{1_mean}, H_{1_var}: NPD by new approach counting for BQL data under several models

3. Evaluation by simulation of the npd with BQL

- Design with 1 observation/subject
  - ➢ type I error close to 5% regardless of LOQ (H_0)
  - ➢ high power to detect model misspecification for H_{1_mean} even for large amounts of BQL data
  - ➢ high power to detect model misspecification for H_{1_var} on the full dataset, but quick decrease of power as the % of BQL data increases
- Design with 6 observations/subject, simulation H_0
  - ➢ in the absence of BQL data, large type I error for npd, corrected with npde which take correlations into account
  - ➢ when omitting BQL data, even npde show large type I errors

Table 2. Type 1 error and power under several assumptions of the global test for npd evaluated on 1000 datasets simulated with the sparse design

Assumptions	LOQ (cp/mL)		
	0	20	50
H_0	0.043	0.041	0.041
H_{1_mean}	1.000	1.000	1.000
H_{1_var}	1.000	0.494	0.336

Table 3. Type 1 error under H_0 of the global test for npd and npde computed by omitting BQL data evaluated on 1000 datasets simulated with the rich design

Assumption H_0 (omit)	LOQ (cp/mL)		
	0	20	50
npd	0.643		
npde	0.054	0.257	0.469

Conclusion and perspectives

- Omitting BQL data in model evaluation can lead to misleading conclusion in the presence of large amounts of BQL data
- The new method for computing the prediction discrepancies is a promising approach to take into account BQL data in evaluation graphs
- Intra-subject correlations should be taken into consideration when testing, and a decorrelation method is currently under development in case of BQL data

References

[1] F Mesnil, F Mentré, C Dubruc, JP Thenot, A Mallet (1998). J Pharmacokinet Biopharm, 26: 133-61.
[2] F Mentré, and S Escolano (2006). J Pharmacokinet Pharmacodyn, 33: 345-67.
[3] E Comets, K Brendel, F Mentré (2008). Comput Methods Programs Biomed, 90: 154-66.
[4] K Brendel et al (2006). Pharm Res, 23:2036–49
[5] K Brendel, E Comets, C Laffont, F Mentré (2010). J Pharmacokinet Pharmacodyn, 37: 49-65.
[6] E Comets, K Brendel, F Mentré (2010), Journal de la Société Française de Statistique, 151: 106-128.
[7] A Ding, H Wu (1999), Math. Biosci., 160: 63-82
[8] A Samson, M Lavielle, F Mentré (2006), Computational Statistics & Data Analysis, 51:1562-1574.