PHARMACOKINETIC SIMILARITY ANALYSIS USING NONLINEAR MIXED EFFECTS MODELS

Anne Dubois[(1)], Thu Thuy Nguyen[(1)], Etienne Pigeolet[(2)] ,
Sandro Gsteiger[(2)], France Mentré[(1)]

(1) UMR 738 INSERM - University Paris Diderot, Paris, France
(2) Novartis Pharma AG, Basel, Switzerland

Non Clinical Statistics Conference

Lyon, France, September 29th 2010

Context - Background

- **Pharmacokinetic (PK) similarity/equivalence studies**
  - Compare different formulations of the same drug
  - Performed on animals or humans
- **FDA[1] and EMEA[2] guidelines**
  - Two-way crossover trials (two periods, two treatments and two sequences)
  - Equivalence test on AUC and C_{max} using non compartmental analysis (NCA)
- **Non compartmental analysis**
  - Few hypotheses
  - More than 10 samples per subject
  - Not appropriate for nonlinear PK or complex models
  - Omit data below quantification limit
- **Nonlinear mixed effects models (NLMEM)**
  - Simultaneous data analysis for all subjects
  - Few samples per subject

[1] FDA Guidance, ucm070244, 2001.

[2] EMEA Guidance, CPMP/EWP/QWP/1401/98 Rev.1, 2010.

Context - Objectives

- Propose a PK similarity approach using NLMEM and mimicking standard statistical method used for NCA
- Evaluate the proposed approach by simulation
- Illustrate the approach on a real PK dataset in monkeys
- Design a crossover trial analysed by NLMEM

Equivalence tests - Tests based on NCA

- **Equivalence tests based on individual parameters**[1,2]
  - AUC and C_{max} estimated by NCA
  - β_T: treatment effect on log parameters
  - $\beta_{T,AUC}$ ($\beta_{T,C_{max}}$) estimated by linear mixed effects model (LMEM)
    - Taking into account period and sequence effects as recommended in the guidelines[1,2]
- **Schuirmann's test**[3]
  - $H_0 : \{\beta_T < -\delta \text{ or } \beta_T > +\delta\}$ vs. $H_1 : \{-\delta \leq \beta_T \leq +\delta\}$ (usually $\delta = 0.2$)
  - Two one-sided tests (TOST)
    - Reject H_0: reject $H_{0,-\delta} : \{\beta_T < -\delta\}$ and $H_{0,+\delta} : \{\beta_T > +\delta\}$
    - Equivalent to $CI_{1-2\alpha}(\widehat{\beta}_T) \subseteq [-\delta; +\delta]$ where α: type I error

[1] FDA Guidance, ucm070244, 2001.
[2] EMEA Guidance, CPMP/EWP/QWP/1401/98 Rev.1, 2010.
[3] Schuirmann. *J Pharmacokinet Biopharm*, 1987.

Equivalence tests - Statistical model

- Concentration y_{ijk}, subject $i = 1, \cdots, N$, sampling time $j = 1, \cdots, n_{ik}$, period $k = 1, \cdots, K$

  $y_{ijk} = f(t_{ijk}, \theta_{ik}) + \epsilon_{ijk}$

- θ_{ik}: individual parameters

  $\log(\theta_{ik}) = \log(\mu) + \eta_i + \kappa_{ik}$

  with $\begin{cases} \log(\mu) = \log(\lambda) + \beta_T T_{ik} + \beta_P P_k + \beta_S S_i \\ \eta_i \sim N(0, \Omega) \text{ between-subject variability (BSV)} \\ \kappa_{ik} \sim N(0, \Gamma) \text{ within-subject variability (WSV)} \end{cases}$

- ϵ_{ijk}: residual error $\quad \epsilon_{ijk} \sim N(0, \sigma)$ with $\sigma = a + b f(t_{ijk}, \theta_{ik})$

- NLMEM parameters: λ, β_T, β_P, β_S, Ω, Γ, a, b
- Maximum likelihood estimation
  - Linearisation algorithms: FO[1], FOCE[2]
  - Exact algorithms: adaptive gaussian quadrature[3], SAEM[4]

[1] Beal, Sheiner. *Crit Rev Biomed Eng*, 1982.
[2] Lindstrom, Bates. *Biometrics*, 1990.
[3] Pinheiro, Bates. *Comput Graph Stat*, 1995.
[4] Kuhn, Lavielle. *ESAIM P& S*, 2004.

Equivalence tests - "NLMEM-based" tests

- **Wald test**
  - TOST on estimate of β_T and its standard error (SE)
  - For parameters in the model[1] (e.g. AUC)
  - For secondary parameters (e.g. C_{max})
    - Derivation of SE by delta method or simulation
- **Likelihood ratio test (LRT)**
  - For parameters in the model only
  - Complete model: log-likelihood L_{all}
  - Model with β_T fixed to $\pm\delta$: log-likelihood $L_{-\delta}$ and $L_{+\delta}$
  - Reject H_0 if:

$$-2(L_{-\delta} - L_{all}) \geq \chi^2_1(1-2\alpha)$$
$$-2(L_{+\delta} - L_{all}) \geq \chi^2_1(1-2\alpha)$$
$$-\delta \leq \widehat{\beta}_T \leq +\delta$$

[1] Panhard, Taburet, Piketti and Mentré. *Stat Med*, 2007.

Evaluation - Simulation study

- **One-compartment PK model**

parameters: $k_a, V/F, CL/F$

- **Two sampling designs** with N=40 subjects
  - "Usual" rich design, n=10 samples per subject and period
  - Very sparse design, n=3
- **Variability**
  - For random effects

	BSV	WSV
Low - S_l	10% for V/F 20% for k_a and CL/F	BSV/2
High - S_h	50%	15%

  - For error model
    $a = 1\ mg/l$ and $b = 0.1$
- **Crossover trials**: 1000 simulated datasets
  - For each null hypothesis ($H_{0,-\delta}$ and $H_{0,+\delta}$)
    - $\delta = \log(1.25)$
    - Treatment effect on CL/F and V/F
  - For each sampling design
  - For each variability setting

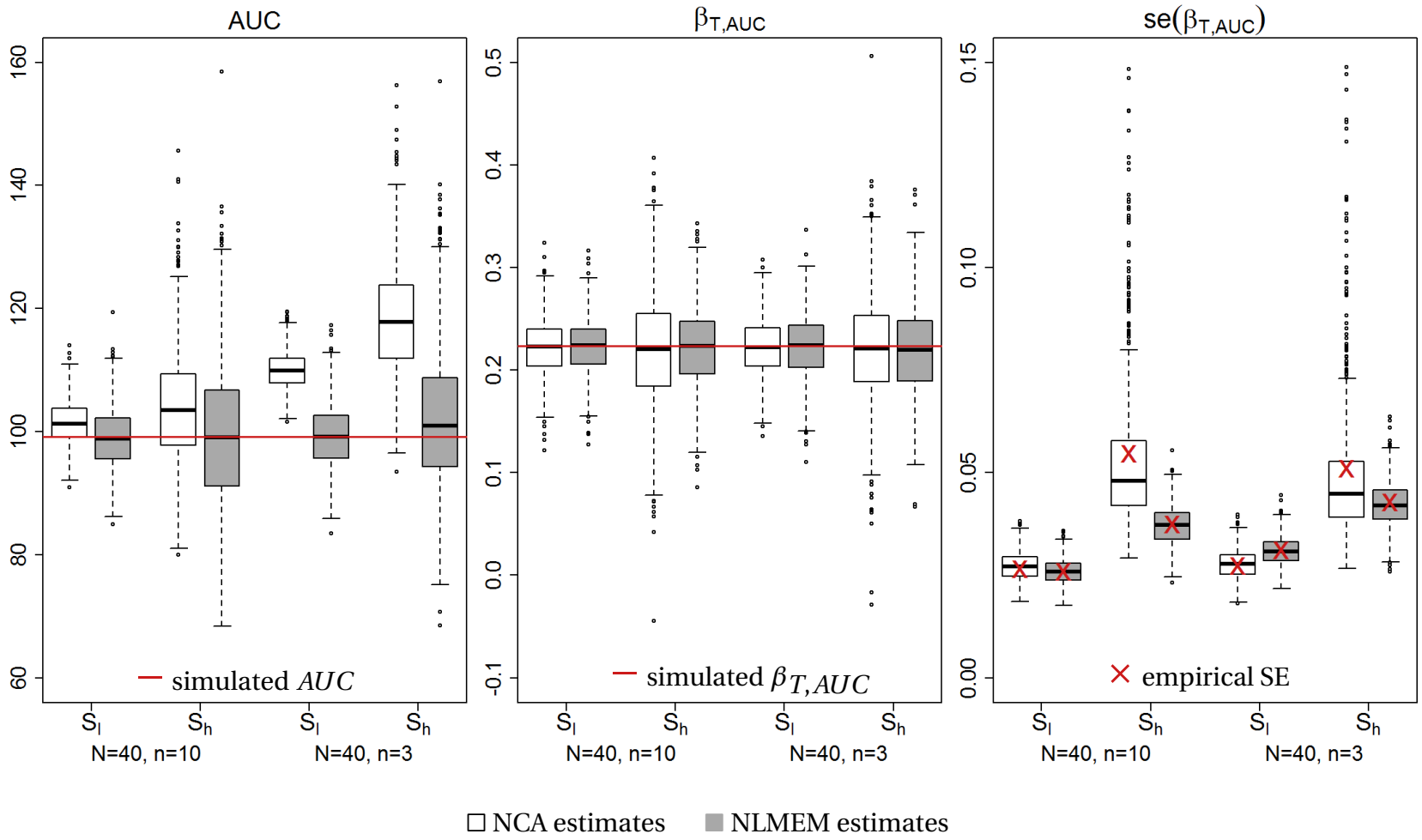
Evaluation - Simulation study

- **Parameter estimation**
  - NCA
    - Geometric means of AUC and C_{max} for each treatment
    - $\beta_{T,AUC}$ and $\beta_{T,C_{max}}$ estimated by LMEM with their SE
  - NLMEM
    - All parameters using SAEM[1] in MONOLIX 3.1[2]
    - AUC, $\beta_{T,AUC}$ and SE($\beta_{T,AUC}$) obtained directly from clearance parameters
    - C_{max}, $\beta_{T,C_{max}}$ computed from fixed effects and SE($\beta_{T,C_{max}}$) using delta method
  - Empirical SE: standard deviation of 1000 treatment effect estimates
- **Evaluation of the estimates for** $H_{0,-\delta}$
  - Means of AUC and C_{max} for reference treatment compared to simulated value
  - $\beta_{T,AUC}$ and $\beta_{T,C_{max}}$ compared to simulated value
  - Estimated SE compared to empirical SE
- **Equivalence test evaluation** ($\alpha = 5\%$)
  - NCA: test on the treatment effect obtained from LMEM
  - NLMEM: Wald test (using estimated and empirical SE) and LRT (for AUC only)
  - Evaluation of the type I error: proportion of simulated trials where H_0 rejected

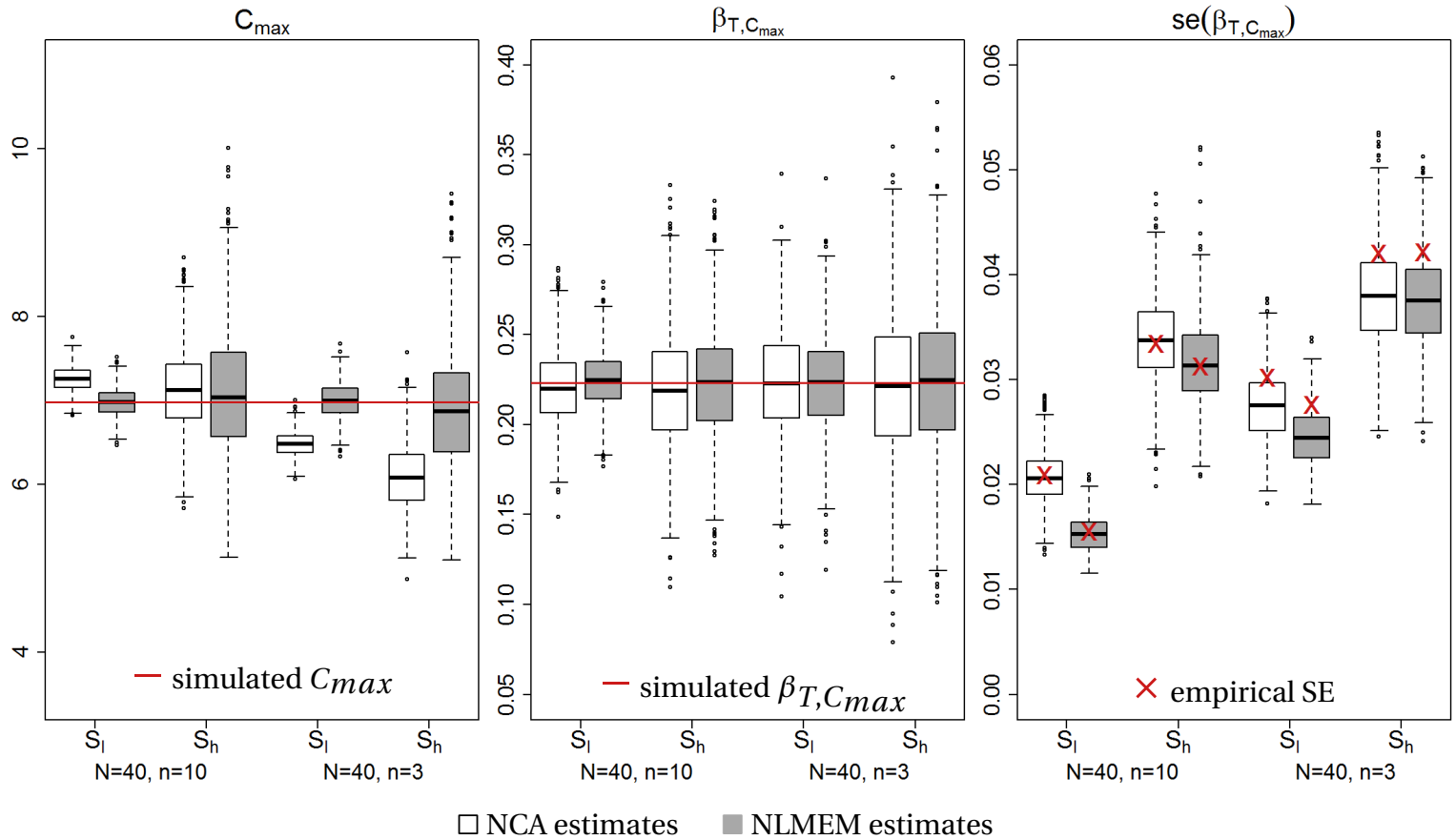
[1] Panhard, Samson. *Biostatistics*, 2009.

[2] http://monolix.org

Evaluation - Results on *AUC* estimates

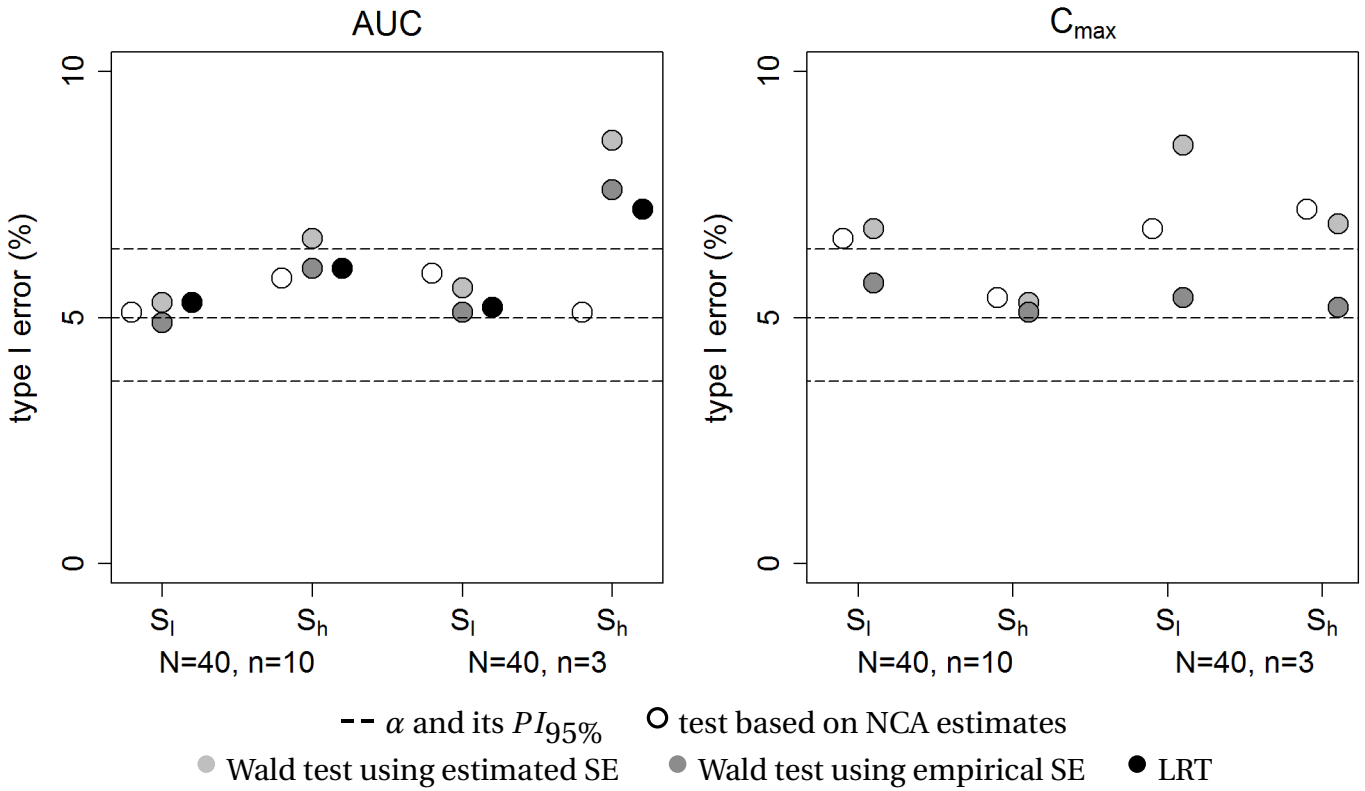


Evaluation - Results on C_{max} estimates



- Evaluation of estimates
  - Biased estimation of means for NCA and sparse design
  - Good estimation of treatment effect for NCA and NLMEM
  - Slight underestimation of SE for sparse design

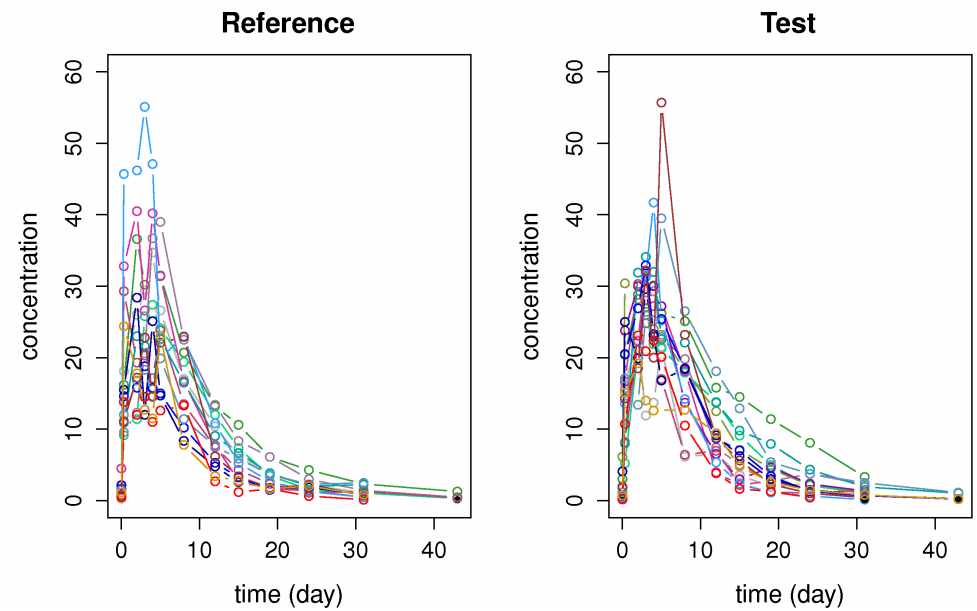
Evaluation - Results on the type I error



- NCA: 5% type I error except for C_{max} with sparse design
- NLMEM
  - Close results LRT and Wald test for AUC
  - 5% type I error for rich design, inflation for sparse design
  - Correction of the inflation by the use of empirical SE for Wald test

Application - Data

- Two-way crossover trial comparing two formulations of a biologic drug in development at Novartis Pharma AG
  - 16 monkeys
  - 12 sampling times per monkey and per period
- PK similarity analysis using NCA and NLMEM
  - Half-life time, 26 days and wash-out period, 42 days
    ⇒ residual concentration from the first period at the drug administration of the second period for some monkeys
  - Few concentrations below the limit of quantification (LOQ)



Application - Results

- **NLMEM analysis**
  - One-compartment model: k_a, V/F, CL/F
  - Treatment, period and sequence effects + BSV + WSV on all PK parameters (12 fixed effects + 7 variance parameters)
  - Taking into account residual concentration of the first period and LOQ[1]
- **Results of the PK similarity analysis**
  - $\alpha = 5\%$, $\delta = \log(1.25)$

	AUC		C_{max}	
	β_T	$CI_{90\%}$	β_T	$CI_{90\%}$
NCA	0.05	[-0.08; 0.19]	0.07	[-0.09; 0.23]
NLMEM	0.07	[-0.04; 0.18]	0.07	[-0.06; 0.20]

  - NCA: equivalence assessment for AUC only
  - NLMEM: equivalence assessment for AUC and C_{max}

[1] Samson, Lavielle, Mentré. *CSDA*, 2006

Design - Method

- Design PK similarity crossover trials analysed through NLMEM
  - Number of subjects and number of samples per subject, choice of sampling times
  - Impact on the study results (precision of parameter estimates, power of test)
- Design evaluation and optimisation
  - From a model and a priori values of parameters
  - Derivation of the expected population Fisher information matrix (M_F)
    - No analytical expression for M_F

    $\Rightarrow$ First-order Taylor expansion of the structural model around the expectation of the random effects (=0)[1]
- Extension for crossover trial[2,3] using PFIM 3.2[4] (R package, January 2010)
  - Within-subject variability
  - Categorical covariate changing (or not) with time
  - For equivalence Wald test: given β_T, α and δ, using predicted SE to compute
    - Expected power
    - Number of subjects needed (NSN) for a given power

[1] Mentré, Mallet, Baccar. *Biometrika*, 1997.
[2] Nguyen, Bazzoli, Mentré. PAGE conference, 2010.
[3] Nguyen, Bazzoli, Dubois, Mentré. ISCB conference, 2010.
[4] http://www.pfim.biostat.fr

Design - Illustration

- Optimise the design of crossover trial on monkeys using the Fedorov-Wynn algorithm[1]
  - 16 monkeys, 6 samples per period, two periods
  - Equivalence test on the clearance
  - Parameter estimates of previous NLMEM analysis
    - Slight treatment effect: $\beta_{T,CL/F} = -0.05$
    - No period or sequence effects
    - Design taking into account WSV
  - $\alpha = 0.05$, power=0.9, $\delta = 0.2$

Design	Sampling times	Power	NSN
Original	0.01, 0.33, 2, 3, 4, 5, 8, 12, 15, 19, 24, 31 days	0.89	17
Optimal	0.01, 2, 3, 4, 5, 31 days	0.84	20

- 17 monkeys to show the PK similarity on the clearance for both formulations
- Close results between original and optimal design with 1.7 times less samples

[1] Retout, Comets, Samson, Mentré. *Stat Med*, 2007.

Conclusion

- NLMEM-based similarity analysis of two-way crossover trials
  - Can be applied to other crossover study designs
- Equivalence test using NLMEM
  - Wald test for model parameters and secondary parameters
  - LRT for model parameters
- Possible applications
  - "Sparse" sampling design
  - Complex / nonlinear pharmacokinetics (TMDD)
  - Model for pharmacodynamic response
- NLMEM for analysis of crossover trials: many parameters to estimate
  - Need to use a robust algorithm as SAEM in MONOLIX
- Method for design evaluation / optimisation of crossover trials
  - Decrease the number of samples per subject