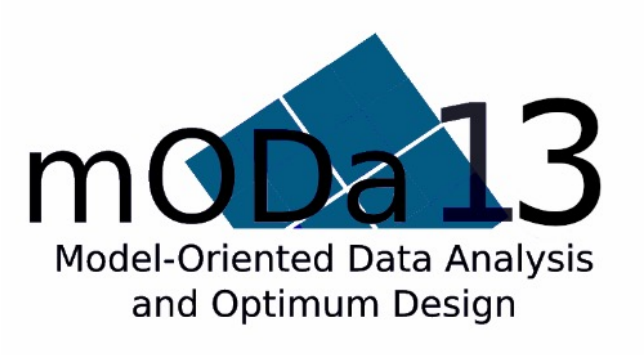


Longitudinal Model for a Dose-Finding Study for a Rare Disease Treatment

Sergei Leonov

Statistical Innovation, CSL Behring

July 13, 2023

Sergei.Leonov@cslbehring.com

Biotherapies for Life™ CSL Behring

Outline

- Background
- Longitudinal model
- Dose selection rule, criteria of optimality
- Trend analysis
- Summary

Joint work with Younan Chen and Mike Fries; see Chen et al. (2023)

Background

- Dose-finding study to identify the most appropriate therapeutic dose
  - A hypothesis that a dose higher than the approved may be more efficacious
  - A longitudinal study, multiple measurements per subject
  - Ethical concerns over the use of placebo (approved treatment exists)

- Statistical challenges
  - Limited sample size, rare disease
  - Large within-subject and between-subject variability

Model

$$Y_{ij} = R_i t_{ij} + \varepsilon_{ij}, \quad R_i = E_{0,i} + \frac{E_{max,i} d_i}{E d_{50,i} + d_i}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2),$$

R_i - monthly rate of subject i, t_{ij} - time of j-th observation of subject i,

E_0 - monthly rate at the minimal dose $\widetilde{d}_{min}$, E_{max} - maximal effect,

$$d_i = \widetilde{d}_i - \widetilde{d}_{min}, \quad E d_{50} = E D_{50} - \widetilde{d}_{min},$$

d_i - "adjusted" doses reduced by $\widetilde{d}_{min}$ from nominal doses $\widetilde{d}_i$,

ED_{50} - dose with half of the maximal effect.

Assumptions

- All subjects measured at the same times $t_{ij} = t_j \in \{1, 2, \ldots, K\}$
- Randomness in E_0 only: $E_{0,i} = E_0 + \eta_i, \quad \eta_i \sim \mathcal{N}(0, \omega^2)$,

$$Y_{ij} = f(d_i, \boldsymbol{\theta}_i) t_j + \varepsilon_{ij}, \quad f(d, \boldsymbol{\theta}) = \theta_1 + \frac{\theta_2 d}{\theta_3 + d}, \quad \boldsymbol{\theta}_i = (E_{0,i}, E_{max}, E d_{50})^T.$$

Sources of variability: population η_i, measurement ε_{ij} *(between/within-subject)*

Population model

Main model: analog of population PK/PD models *(compartmental models)*:

$$y_{ij} = f(x_{ij}, \boldsymbol{\theta}_i) + \varepsilon_{ij}, \ \boldsymbol{\theta}_i \sim \mathcal{N}(\boldsymbol{\theta}, \boldsymbol{\Omega}), \ \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2), \ j = 1, \ldots, k_i,$$

- $f(x, \boldsymbol{\theta})$: response function
- $\boldsymbol{\theta}_i$: individual parameters of subject i (rate constants, volume of distribution)
- x_{ij}: sampling times for subject i
- k_i: number of distinct sampling times for subject i.

Key: derive/approximate Fisher information matrix (FIM) $\boldsymbol{\mu}(\mathbf{x}_i, \boldsymbol{\theta})$ for vector $\mathbf{Y}_i = (y_{i1}, \ldots, y_{ik})$ at times $\mathbf{x}_i = (x_{i1}, \ldots, x_{ik})$.

- First-order approximation: need expressions of mean $\mathbf{F}$ and variance-covariance matrix $\mathbf{S}$ of $\mathbf{Y}_i$ (Leonov and Aliev (2013), Nyberg et al. (2015))

Information, variance-covariance matrix

Some algebra/notations:

$$\mathbf{F}(d,\boldsymbol{\theta}) = \mathbf{T}f(d,\boldsymbol{\theta}),\ \mathbf{G} = \mathbf{T}\mathbf{g}(d,\boldsymbol{\theta}),\ \mathbf{T} = \begin{pmatrix} 1 \\ 2 \\ \dots \\ K \end{pmatrix},$$

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \mathbf{g}(d,\boldsymbol{\theta}) = \frac{\partial f}{\partial \boldsymbol{\theta}} = \left[1, \frac{d}{\theta_3 + d}, \frac{-\theta_2 d}{(\theta_3 + d)^2}\right] \implies$$

$$\mathbf{S} = \omega^2 \mathbf{T}\mathbf{T}^T + \sigma^2 \mathbf{I}_K, \boldsymbol{\mu}(d,\boldsymbol{\theta}) = \mathbf{G}^T \mathbf{S}^{-1} \mathbf{G}$$

(Fedorov and Leonov (2013), Chapter 7).

Next steps: from design ξ_N to information matrix $\mathbf{M}(\xi_N, \boldsymbol{\theta})$ to variance-covariance matrix of the MLE $\hat{\boldsymbol{\theta}}$:

$$\xi_N = \{(d_\ell, n_\ell),\ \ell = 1, \dots, L,\ \ N = \sum_{\ell=1}^{L} n_\ell\} \implies$$

$$\mathbf{M}(\xi_N, \boldsymbol{\theta}) = \sum_{\ell=1}^{L} n_\ell \boldsymbol{\mu}(d_\ell, \boldsymbol{\theta}) \to\ \mathbf{D}(\xi_N, \hat{\boldsymbol{\theta}}) \approx \mathbf{M}^{-1}(\xi_N, \boldsymbol{\theta}).$$

Target dose

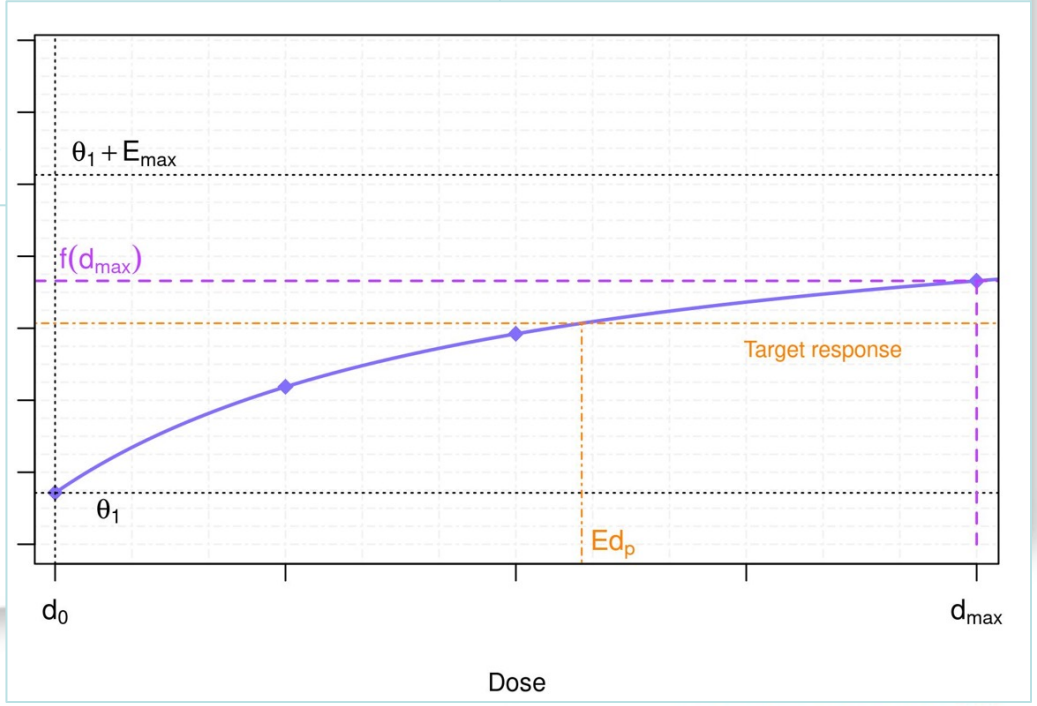
- Traditionally: Ed_p - dose achieving $100p\%$ of the *maximum effect*.

- Here: Ed_p - dose achieving $100p\%$ of the *effect at the maximum observed dose*: Dette et al. (2010) $\Longrightarrow$ find $Ed_p = \psi(p, \boldsymbol{\theta})$ from

$$f(d, \boldsymbol{\theta}) - \theta_1 = p[f(d_{max}, \boldsymbol{\theta}) - \theta_1] \implies Ed_p = \frac{p d_{max} \theta_3}{\theta_3 + (1-p) d_{max}} .$$

Variance of the estimator of Ed_p for design ξ_N:

$$\mathbf{Var}_{\xi_N}(\widehat{Ed_p}) = \frac{\partial \psi(p, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \mathbf{D}(\xi_N, \widehat{\boldsymbol{\theta}}) \frac{\partial \psi(p, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} .$$



Optimality and practical design

Design region $[d_{min}, d_{max}]$, single measurement per subject $(K = 1)$

- D-optimality ("generalized" variance): smallest volume of conf. region.
  Optimal design: 3 doses $\{d_{min}, d_{int}(\boldsymbol{\theta}), d_{max}\}$ with equal weights $1/3$.
- Ed_p-optimality: smallest variance of the estimator of the dose Ed_p.
  Same *optimal points*, weights $\{1/4, 1/2, 1/4\}$ (Dette et al., 2010)

Locally optimal designs:

- Depend on $\boldsymbol{\theta}(\theta_3)$, not always practical
- Provide a useful benchmark for other candidate designs

Practical designs: independent of $\boldsymbol{\theta}$, robust across plausible parameter range

Simulation settings

Doses $d_i \in \{0, 30, 60, 120\}$ *units*, $\xi = \{d_i, w_i = 1/4,\ i = 1, \ldots, 4\}$,

75 subjects/arm, 300 subjects total.

Question: Why not a 4-parameter model?

$$f(d, \boldsymbol{\beta}) = \beta_1 + \frac{\beta_2 d^{\beta_4}}{\beta_3^{\beta_4} + d^{\beta_4}}, \quad \boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4),$$

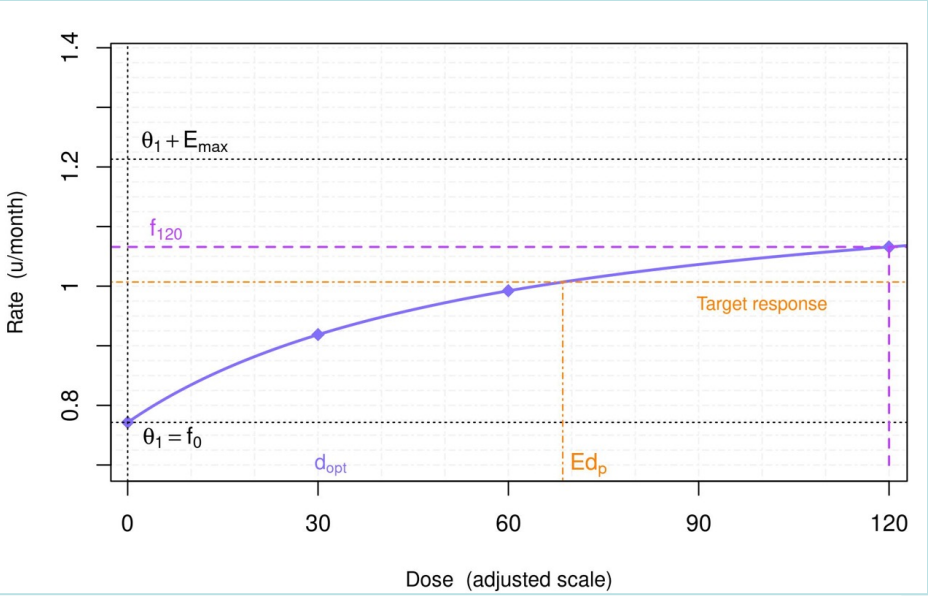
Answer: Slope at $d = 0$ is zero, $f'_d(d, \boldsymbol{\beta}) = 0$ at $d = 0$ for $\beta_4 > 1$

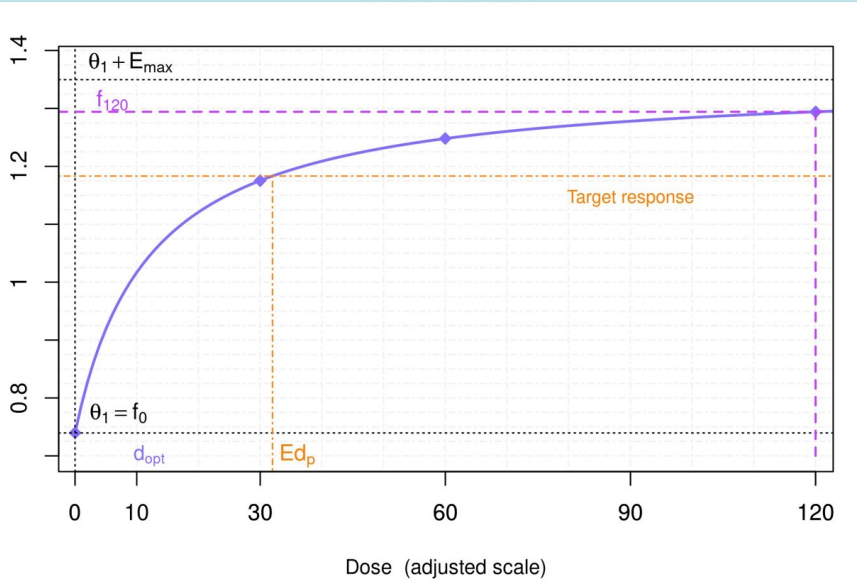
- 3-parameter E_{max} models: parameters selected to "match" 4-parameter E_{max} curves used earlier (doses in $[0, \widetilde{d}_{max}]$, placebo used)
- Variability: $\omega^2 = 1.25$ (*between-subject*), $\sigma^2 = 5$ (*within-subject*)

Examples of 3-parameter E_{max} models

- p=0.8: reasonable balance between efficacy and safety/toxicity
- Two parameter sets used for 3-parameter E_{max} model $f(d, \boldsymbol{\theta})$:

$$\boldsymbol{\theta}_1 = (0.77, 0.44, 60) \qquad \boldsymbol{\theta}_2 = (0.74, 0.61, 12)$$





D- and Ed_p-efficiency (Ed_p-efficiency is *invariant* of p!):

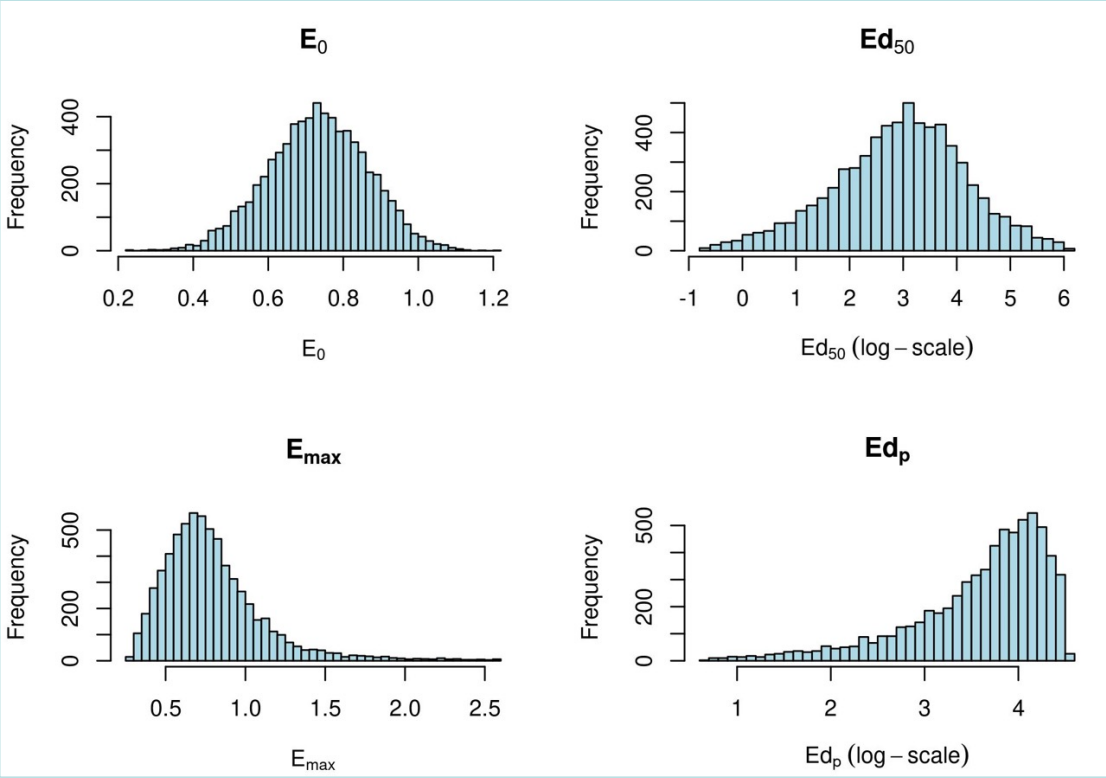
$\boldsymbol{\theta}_1$: efficiency 0.9, 0.79 (D, Ed_p) $\qquad$ $\boldsymbol{\theta}_2$: efficiency 0.67, 0.27

Simulations (Nsim = 10000)

- Data fitting: log-transform doses (better "non-linearity" measures):

$$\theta_3 \to \Theta_3 = \log(\theta_3),\ d \to D = \log(d),\ f(D, \boldsymbol{\theta}) = \theta_1 + \frac{\theta_2}{1 + e^{\Theta_3 - D}}.$$

- R-packages *nlme*, *nlmer* (Pinheiro, Bates (2000), Pinheiro et al. (2020))
- Non-convergent data sets ($\sim 30\%$): relatively large variability ($\omega^2,\ \sigma^2$)

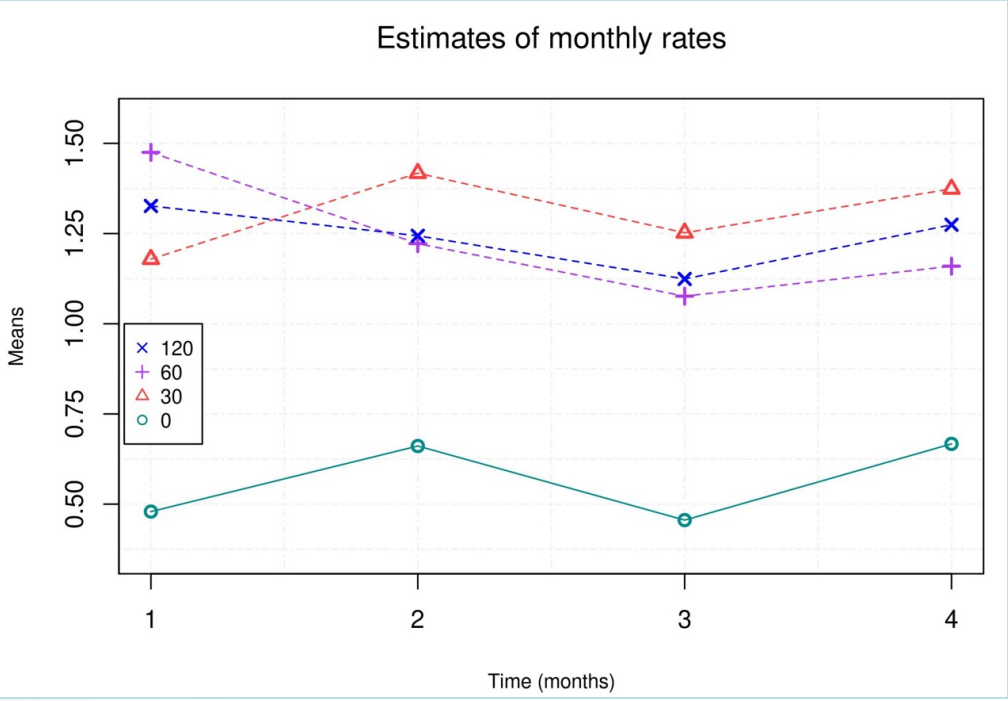


Non-convergent data sets

Let $EY_{\ell,j} = E_{\varepsilon,\theta} Y_{\ell,j} = f(d_\ell, \boldsymbol{\theta}) t_j$

- $f(d, \boldsymbol{\theta})$ is monotonically increasing: $f(d_\ell, \boldsymbol{\theta}) < f(d_{\ell+1}, \boldsymbol{\theta})$
- Reasonable estimate of $f(d_\ell, \boldsymbol{\theta})$: mean of responses at dose ℓ, time t_j:

$$\widehat{f}_j(d_\ell) = \frac{1}{t_j} \sum_{i:\ d_i = d_\ell} Y_{ij}/n, \text{ where } n = \#(\text{subjects on } d_\ell)$$



- Wrt time: ~constant at each dose
- Wrt dose: monotonicity broken

Estimates of monthly rates

Our model:

$$Y_{ij} = [f(d, \boldsymbol{\theta}) + \eta_i]\, t_j + \varepsilon_{ij} \quad \text{for subject } i, \text{ time } j.$$

Let us drop index i, denote $f = f(d, \boldsymbol{\theta})$, use $t_j = j \implies$

$$Y_j = (f + \eta)j + \varepsilon_j, \ j = 1, \ldots, K.$$

Reasonable *unbiased* estimator of f (monthly rate for a given subject):

$$\widehat{f} = \sum_{j=1}^{K} u_j \frac{Y_j}{j}, \quad 0 \le u_j \le 1, \ \sum_j u_j = 1.$$

$$\mathbf{Var}(\widehat{f}) = E(\widehat{f} - f)^2 = \omega^2 + \sigma^2 \sum_{j=1}^{K} \frac{u_j^2}{j^2} \implies \text{minimize variance:}$$

$$G(\mathbf{u}) = \sum_{j=1}^{K} \frac{u_j^2}{j^2} \to \min_{\{u_j\}} \quad \text{subject to } 0 \le u_j \le 1, \ \sum_j u_j = 1 \qquad (*)$$

Estimates of monthly rates

$$G(\mathbf{u}) = \sum_{j=1}^{K} \frac{u_j^2}{j^2} \to \min_{\{u_j\}} \quad \text{subject to } 0 \le u_j \le 1, \ \sum_j u_j = 1 \qquad (*)$$

Solution of optimization problem (*):

$$\mathbf{u}^* = \left(\frac{1}{S_K}, \frac{2^2}{S_K}, \ldots, \frac{K^2}{S_K} \right), \ S_K = \sum_{j=1}^{K} j^2.$$

$K = 4, \ \omega^2 = 1.25, \ \sigma^2 = 5$:

- $\mathbf{u}^* = \frac{1}{30}(1, 4, 9, 16), \ G(\mathbf{u}^*) = \frac{1}{30} \approx 0.0333, \ \mathbf{Var}(\widehat{f}) = 1.25 + \frac{5}{30} \approx 1.4166.$
- $\mathbf{u} = (0, 0, 0, 1), \ G(\mathbf{u}) = \frac{1}{16} = 0.0625, \ \mathbf{Var}(\widehat{f}) = 1.5625.$

Population studies: relation between within- and between-subject variability is critical!

Nonparametric trend test (Jonckheere-Terpstra)

- Distribution-free test for ordered alternatives (Jonckheere, 1954) to compare hypotheses

$$\mathbf{H}_0: \beta_1 = \beta_2 = \ldots = \beta_K \text{ vs } \mathbf{H}_1: \beta_1 \leq \beta_2 \leq \ldots \leq \beta_K$$

  with at least one strict inequality

- Independent of the underlying dose-response curve
- Provides sufficient power in most plausible scenarios
- Time-normalized means increase power by 4-5%

Summary

- Challenging clinical and statistical issues for a rare disease trial
- Accepted as an overall robust statistical methodology
- Future research topics:
  - Non-convergent models
  - Optimal designs for E_{max} models with multiple measurement per subject

References

- Bretz, F., Dette, H., Pinheiro, J. (2010). Practical considerations for optimal designs in clinical dose finding studies. *Statist. Med*., **29**, 731-742.
- Chen, Y., Fries, M., Leonov, S. (2023). Longitudinal model for a dose-finding study for a rare disease treatment. *Statistical Papers*, DOI: 10.1007/s00362-023-01424-1. Available online at: https://link.springer.com/article/10.1007/s00362-023-01424-1
- Dette, H., Kiss, M., Bevanda, M., Bretz, F. (2010). Optimal designs for the emax, log-linear
- and exponential models. *Biometrika*, **97**(2), 513-518.
- Fedorov, V.V., Leonov, S.L. (2013). *Optimal Design for Nonlinear Response Models*. Chapman & Hall/CRC Biostatistics Series, Boca Raton, FL.
- Jonckheere, A.R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, **41**(1/2), 133-145.
- Nyberg, J., Bazzoli, C., Ogungbenro, K., Aliev, A., Leonov, S., Duffull, S., Hooker, A., Mentre, F. (2015). Methods and software tools for design evaluation in population pharmacokinetics-pharmacodynamics. *Brit. J. Clin. Pharm*., **79**(1), 6-17.
- Pinheiro, J.C., Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer, NY.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team (2020). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-148, https://CRAN.Rproject.org/package=nlme.

Sergei.Leonov@cslbehring.com