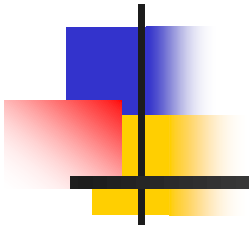


DAE 2012

Optimal experimental design for population PK/PD studies

Sergei Leonov
(*Vertex Pharmaceuticals Inc*)

October 19, 2012

VERTEX

Outline

- Pharmacokinetic/pharmacodynamic models
- Population models and optimal design
- Optimal design software, PkStaMp library
- Information matrix, approximation options

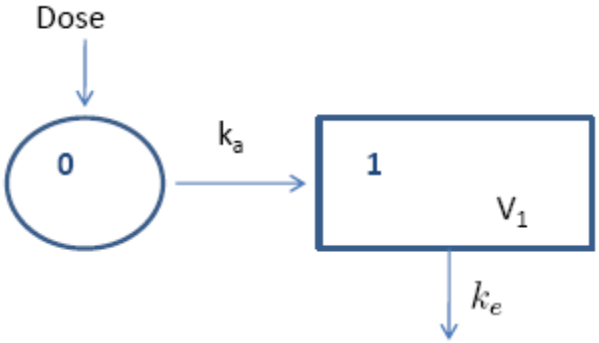
Pharmacokinetics & pharmacodynamics

- PK: what body does to the drug (time-concentration)
  - Compartmental, systems of ordinary differential equations (ODE)
  - Non-compartmental (AUC, T_{max}, C_{max})

Example:

One-compartment model, 1st order absorption and linear elimination

$$\begin{cases} \dot{f}_0(t) = -k_a f_0(t) \\ \dot{f}_1(t) = k_a f_0(t) - k_e f_1(t) \end{cases}$$



- PD: what drug does to the body (concentration → effect)
  - Progression of clinically relevant endpoint or biomarker

Optimal Designs

$$\mathbf{M}(\xi, \boldsymbol{\theta}) = \frac{\mathbf{M}_N(\boldsymbol{\theta})}{N} = \sum_i w_i \boldsymbol{\mu}(\mathbf{x}_i, \boldsymbol{\theta})$$ - normalized information, per observation

$\xi = \{w_i, \mathbf{x}_i\}$ - normalized design; $w_i = n_i/N$ - weights

Locally D-optimal designs: $|\mathbf{M}^{-1}(\xi, \boldsymbol{\theta})| \rightarrow \min_{\xi}$: minimization wrt

- Continuous designs: $0 \leq w_i \leq 1$, $\Sigma_i w_i = 1$,
- Admissible sampling sequences $\mathbf{x}_i \in \mathbf{X}$ - design region.

Equivalence Theorem: *Kiefer, Wolfowitz (1960), Fedorov (1972)* - background for 1-st order optimization algorithms (Fedorov-Wynn)

Mixed effects model model

- $\boldsymbol{\gamma}_i$ - response parameters of patient i (*sampled from population*):

  normal, $\boldsymbol{\gamma}_i \sim N(\boldsymbol{\gamma}^0, \boldsymbol{\Omega})$, or log-normal ($\boldsymbol{\gamma}^0$ - "typical values")

- Data $y(x_{ij}) = \eta(x_{ij}, \boldsymbol{\gamma}_i)\left[1 + \varepsilon_{ij}^p\right] + \varepsilon_{ij}^a, \quad j = 1, \ldots, k_i. \qquad (1)$

  $\varepsilon_{ij}^a \sim N(0, \sigma_a^2), \quad \varepsilon_{ij}^p \sim N(0, \sigma_p^2)$

- Combined vector of parameters: $\boldsymbol{\theta} = (\boldsymbol{\gamma}^0;\ \boldsymbol{\Omega};\ \sigma_A^2,\ \sigma_P^2)$

Example: one-compartment model, single dose D at $x = 0$,

$$\eta(x, \boldsymbol{\gamma}) = \frac{Dk_a}{V(k_a - k_e)}\left(e^{-k_e x} - e^{-k_a x}\right), \quad \boldsymbol{\gamma} = (k_a, k_e, V)^T$$

VERTEX

Information matrix for sequence x

(1) Gaussian $\mathbf{Y}$: $\mathbf{E}[\mathbf{Y}|\mathbf{x}] = \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})$, $\mathbf{Var}[\mathbf{Y}|\mathbf{x}] = \mathbf{S}(\mathbf{x}, \boldsymbol{\theta})$

$\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ - information matrix of a single (k-dimensional) sequence $\mathbf{x}$:

$$\mu_{\alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial \boldsymbol{\eta}}{\partial \theta_\alpha} \mathbf{S}^{-1} \frac{\partial \boldsymbol{\eta}}{\partial \theta_\beta} + \frac{1}{2} \mathrm{tr} \left[\mathbf{S}^{-1} \frac{\partial \mathbf{S}}{\partial \theta_\alpha} \mathbf{S}^{-1} \frac{\partial \mathbf{S}}{\partial \theta_\beta} \right],$$

$\mathbf{S} = \mathbf{S}(\mathbf{x}, \boldsymbol{\theta})$, $\boldsymbol{\eta} = \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})$ [*Muirhead (1982), Magnus and Neudecker (1988)*]

(2) First-order approximation of variance matrix $\mathbf{S}$, model (1): for normal $\boldsymbol{\gamma}$

$$\mathbf{S}(\mathbf{x}, \boldsymbol{\theta}) \simeq \mathbf{F}\, \boldsymbol{\Omega}\, \mathbf{F}^T + \sigma_P^2\, \mathrm{Diag}[\boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})\, \boldsymbol{\eta}^T(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{F}\, \boldsymbol{\Omega}\, \mathbf{F}^T] + \sigma_A^2 \mathbf{I}_k,$$

$$\mathbf{F} = \mathbf{F}(\mathbf{x}, \boldsymbol{\gamma}^0) = \left[\frac{\partial \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\gamma}_\alpha} \right] \Bigg|_{\boldsymbol{\gamma} = \boldsymbol{\gamma}^0} - (k \times m_\gamma) \text{ matrix}$$

Optimal design for population PK/PD models

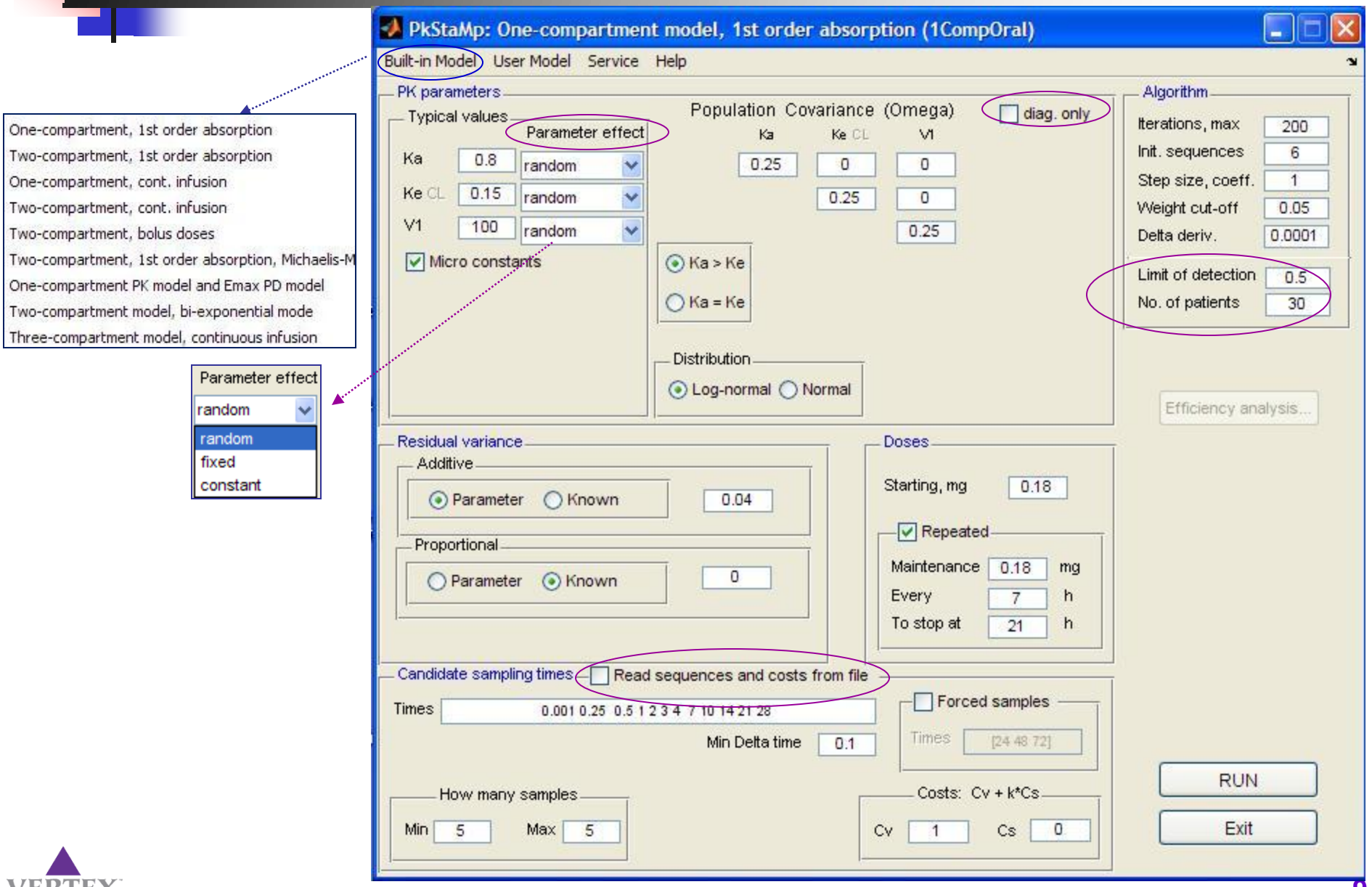
- Annual Population Optimum Design of Experiments (PODE) Workshop created in 2006
  - Optimal design for nonlinear mixed effects models: theory and applications in drug development

- Discussion of population optimal design tools started in 2007
  - PFIM (developed in INSERM, Universitè Paris 7, France)
  - PkStaMp (GlaxoSmithKline/Vertex)
  - PopDes (CAPKR, University of Manchester, UK)
  - PopED (Uppsala University, Sweden)
  - WinPOPT (University of Otago, New Zealand)

PkStaMp library

- Sampling Times Allocation (STand-Alone Application), Matlab Platform
- Why Matlab:
  - Allows for creating executable files (no license required for end-users)
  - Easy to create GUI
- Collection of independent modules created for various GSK projects, development started in 2002-2003
- Last 3+ years: joint work with Dr. Alexander Aliev (Institute for Systems Analysis, Russian Academy of Science, Moscow)
  - Recent addition: *user-defined option*

Typical screen: one-compartment,1st order absorption



Design region $\mathbf{X}$

Option 1: specify set of candidate times $(x_1, x_2, \ldots, x_K)$

- Number of sampling times per patient $k \in [k_{min}, k_{max}]$

- Lag between samples: $x_{i,j+1} - x_{i,j} \geq \Delta$

- $\mathbf{X} = \{$all possible combinations of k times$\}$ $(C_K^k$, *binomial coefficient*$)$

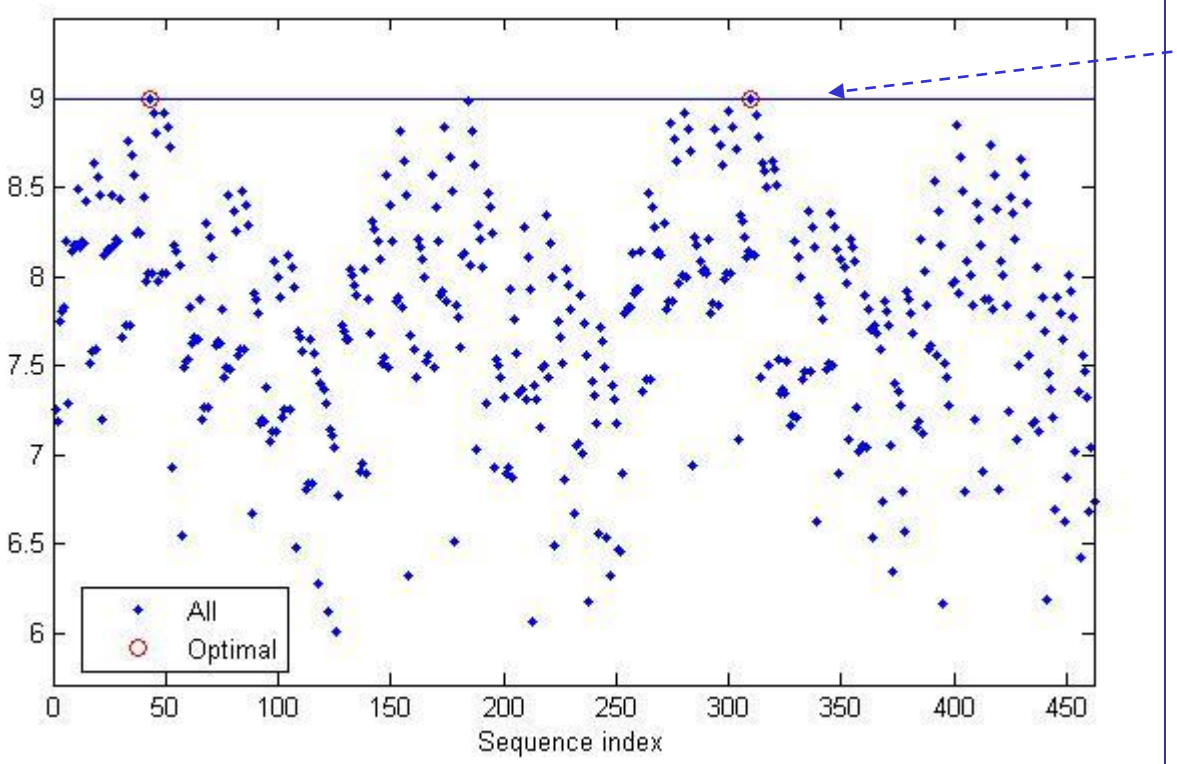
Option 2: specify an arbitrary set of candidate sequences in a file $\Rightarrow$

Design region $\mathbf{X} = \{\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,k_i})\}$

Both options:

- Discrete (finite) design region, by construction
- Forward step of 1st-order optimization algorithm: finite optimization

Diagnostics: whether the algorithm converged

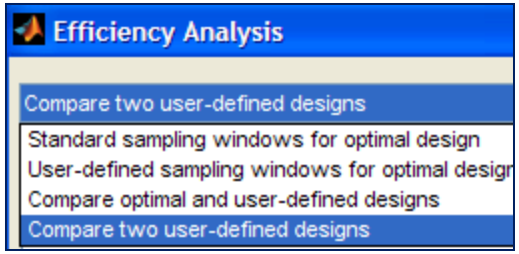


- Sensitivity function, total # of parameters = 9
- Design region: dimension = # of samples
- X-axis: sequence index
- Red circles denote optimal points (sequences)

Efficiency analysis

Goals:

- Compare optimal design with alternative designs: *benchmarking*
- Test robustness of the optimal design (*sampling windows*)



Cost-based designs

Measurements associated with cost $c(\mathbf{x_i})$ $\Rightarrow$ normalize $\mathbf{M}$ by total cost $\mathcal{C}$

$$\sum_i n_i c(\mathbf{x}_i) \le \mathcal{C} \Longrightarrow \mathbf{M}_C(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{n_i}{\mathcal{C}} \boldsymbol{\mu}(\mathbf{x_i}, \boldsymbol{\theta}) = \sum_i \tilde{w}_i \, \tilde{\boldsymbol{\mu}}(\mathbf{x_i}, \boldsymbol{\theta}),$$

$$\tilde{w}_i = n_i c(\mathbf{x_i})/\mathcal{C}; \quad \tilde{\boldsymbol{\mu}}(\mathbf{x_i}, \boldsymbol{\theta}) = \boldsymbol{\mu}(\mathbf{x_i}, \boldsymbol{\theta})/c(\mathbf{x_i}) \Longrightarrow \underline{\text{same framework}},$$

same algorithms

Costs in design problems: *Elfving (1952), Cook, Fedorov (1995), Mentré, Mallet, Baccar (1997), Fedorov, Gagnon, Leonov (2002)*

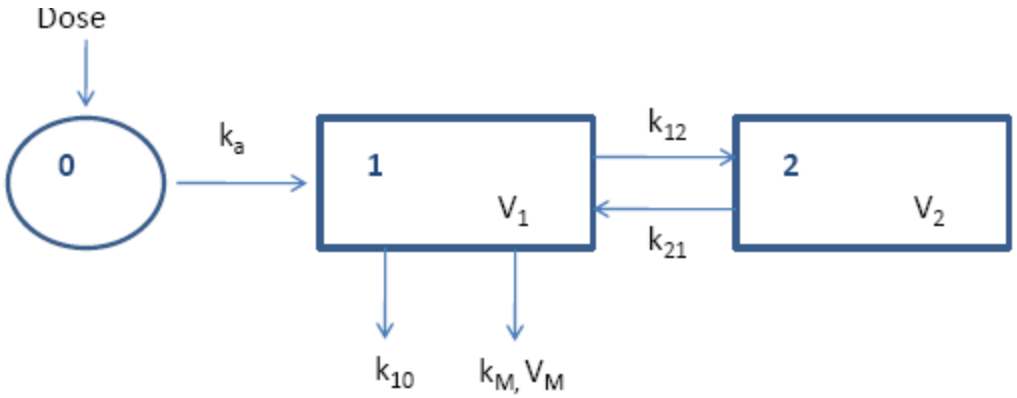
In PkStaMp: (a) Cost $c(\mathbf{x})$ proportional to # of samples in sequence **x**, or
(b) Entered by user for each candidate sampling sequence

More complex models: nonlinear kinetics

Two-compartment model, 1st order absorption,

Michaelis-Menten elimination: no analytical solution (ODE solver)

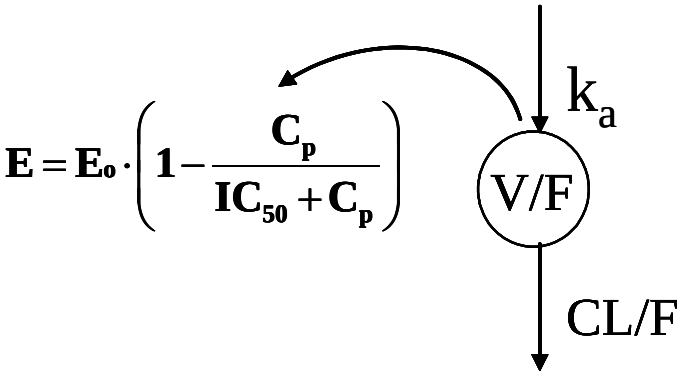
$$\begin{cases} \dot{f}_0(t) = -k_a f_0(t) \\ \dot{f}_1(t) = k_a f_0(t) - (k_{12} + k_e) f_1(t) - \frac{(V_m/V) f_1(t)}{k_m + f_1(t)/V} + k_{21}\ f_2(t) \\ \dot{f}_2(t) = k_{12}\ f_1(t) - k_{21}\ f_2(t), \end{cases}$$



More complex models (2): combined PK/PD

One-compartment PK and Emax PD model

Final PK/PD Model



k_a: first-order absorption rate constant (h^{-1})
V/F: apparent volume of distribution (L)
CL/F: apparent systemic clearance (L/h)
E_o: PD endpoint at baseline (nM/min/mL)
IC_{50}: Drug X plasma concentration causing 50% inhibition of PD endpoint (ng/mL)

PK and PD compartments may be measured at different times

More complex models (3): HCV

Combination drug for treating chronic hepatitis C (HCV) infection
Neumann et al. (1998), *Mentré et al.* (2011)

$$\begin{cases} \dot{f}_0(t) &= -k_a f_0(t) &&+ r(t) \\ \dot{f}_1(t) &= k_a f_0(t) &- k_e f_1(t) & \\ \eta_1(t) &= f_1(t)/V_1 & & \end{cases}$$

PK: parameters (k_a, k_e, V_1), response η_1 (continuous infusion term $r(t)$)

$$\begin{cases} \dot{g}_1(t) &= -C_2 g_1(t) - C_1 g_1(t) g_3(t) + C_3 \\ \dot{g}_2(t) &= -\delta g_2(t) + C_1 g_1(t) g_3(t) \\ \dot{g}_3(t) &= C_4 \left[1 - \frac{1}{1+(EC_{50}/\eta_1)^n}\right] g_2(t) - c g_3(t) \\ \\ \eta_2(t) &= \log_{10} g_3(t) \end{cases}$$

$g_1(t)$ - "target cells", $g_2(t)$ - infected cells, $g_3(t)$ - viral particles (load)

PD: parameters (δ, EC_{50}, n, c), response η_2

HCV example: user-defined option

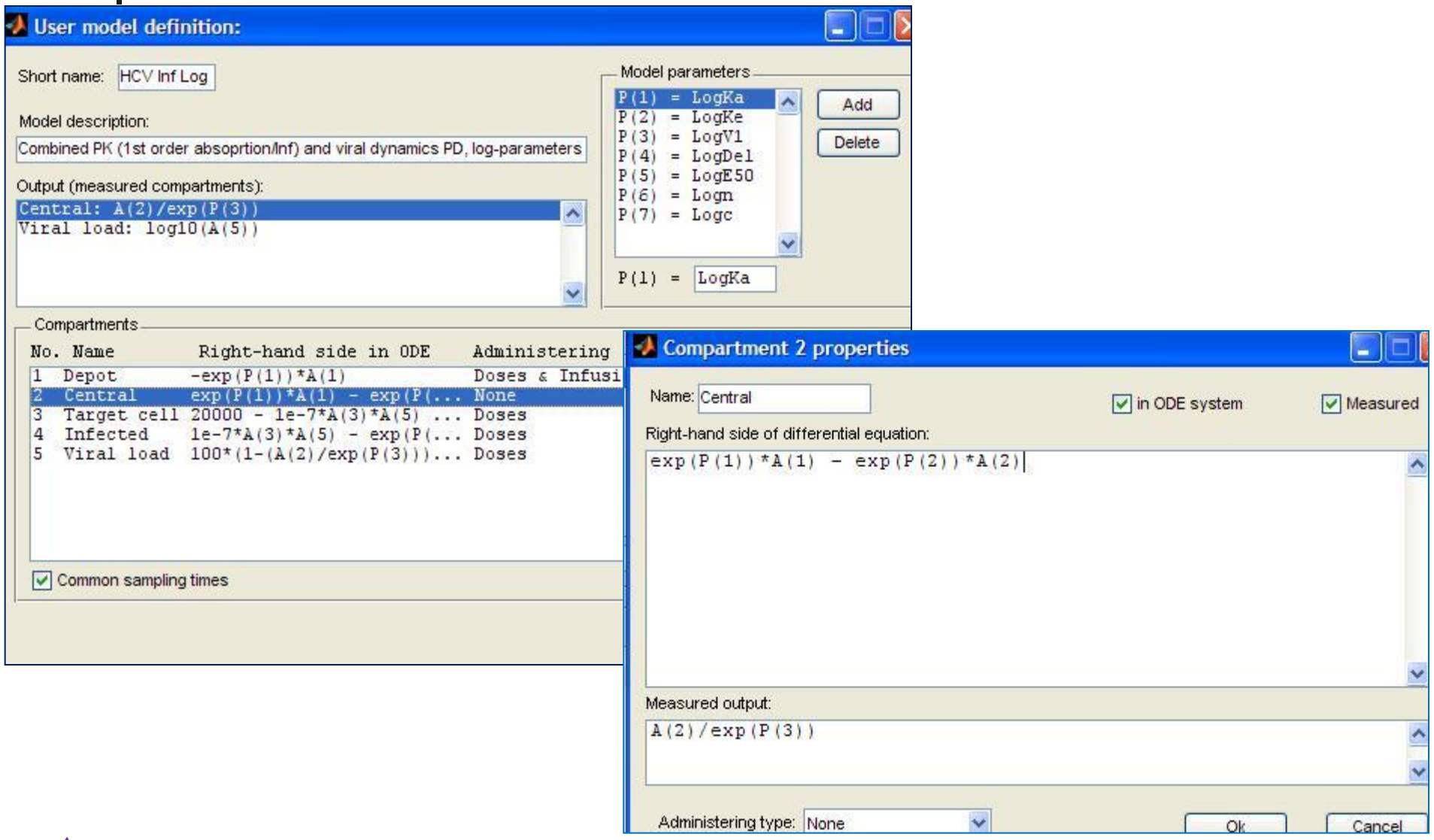
Parameterization

- Log-parameters
- Normal population distribution

User-defined option:

- “Arbitrary” system of ODE, and/or
- “Arbitrary” closed-form solution
- “Arbitrary” number of compartments

Model specification



Software comparison (PODE 2009-11)

One-compartment, 1st order absorption, single dose D

$$\eta(x,\boldsymbol{\gamma}) = \frac{Dk_a}{V(k_a-k_e)}\left(e^{-k_e x} - e^{-k_a x}\right).$$

Response parameters $\boldsymbol{\gamma} = (k_a, CL, V),\ k_e = CL/V$

Individual parameters

$$\boldsymbol{\gamma}_i = \boldsymbol{\gamma}^0 e^{\xi_i},\ \xi_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}),$$

$$\boldsymbol{\gamma}^0 = (1, 0.15, 8),\ \boldsymbol{\Omega} = \mathrm{Diag}(0.6, 0.07, 0.02)$$

Measurements:

$$y_{ij} = \eta(x_{ij}, \boldsymbol{\gamma}_i)\ (1 + \varepsilon_{M,ij}), \qquad (1)$$

$\{x_{ij}\} \equiv \mathbf{x} = (0.5, 1, 2, 6, 24, 36, 72, 120)$ h.

$\varepsilon_{M,ij} \sim \mathcal{N}(0, \sigma_M^2),\ \sigma_M^2 = 0.01$

Parameter $\boldsymbol{\theta} = (k_a^0,\ CL^0,\ V^0;\ \ \omega_{k_a}^2,\ \omega_{CL}^2,\ \omega_V^2;\ \ \sigma_M^2)$

VERTEX

Software comparison (cont.)

Information matrix $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$: block form, *Retout and Mentré (2003)*

$$\boldsymbol{\mu} = \left\{ \begin{array}{cc} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{array} \right\},$$

$\mathbf{A} = \mathbf{F}^T\, \mathbf{S}^{-1}\, \mathbf{F} + \frac{1}{2}\,\mathrm{tr}$ (derivatives wrt γ_α)
$\mathbf{C} = \frac{1}{2}\,\mathrm{tr}$ (mixed derivatives wrt γ_α and $[\omega_\beta^2, \sigma_M^2]$)
$\mathbf{B} = \frac{1}{2}\,\mathrm{tr}$ (derivatives wrt $[\omega_\beta^2, \sigma_M^2]$)

$\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ - information matrix of a single (k-dimensional) sequence $\mathbf{x}$:

$$\mu_{\alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial \boldsymbol{\eta}}{\partial \theta_\alpha}\, \mathbf{S}^{-1}\, \frac{\partial \boldsymbol{\eta}}{\partial \theta_\beta} + \frac{1}{2}\,\mathrm{tr}\left[\mathbf{S}^{-1}\, \frac{\partial \mathbf{S}}{\partial \theta_\alpha}\, \mathbf{S}^{-1}\, \frac{\partial \mathbf{S}}{\partial \theta_\beta}\right],$$

Software comparison (cont.)

$$\mu = \left\{ \begin{array}{cc} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{array} \right\},$$

$$\mathbf{A} = \mathbf{F}^T\ \mathbf{S}^{-1}\ \mathbf{F} + \tfrac{1}{2}\ \mathrm{tr}\ \ (\text{derivatives wrt } \gamma_\alpha)$$
$$\mathbf{C} = \tfrac{1}{2}\ \mathrm{tr}\ \ (\text{mixed derivatives wrt } \gamma_\alpha \text{ and } [\omega_\beta^2, \sigma_M^2])$$
$$\mathbf{B} = \tfrac{1}{2}\ \mathrm{tr}\ \ (\text{derivatives wrt } [\omega_\beta^2, \sigma_M^2])$$

- $\boldsymbol{D}_a = [\boldsymbol{\mu}(\boldsymbol{x},\boldsymbol{\theta})]^{-1}$: identical results for all tools under the same assumptions: *Mentré et al. (2011), Leonov and Aliev (2012)*

- Compared $\boldsymbol{D}_a$ and $\boldsymbol{D}_e$ (empirical variance-covariance matrix: Monte Carlo + estimation in NONMEM/Monolix):
  - *Reduced option*: block $\mathbf{C}$ "excluded" ($\mathbf{C} = \mathbf{0}$), 2nd term in $\mathbf{A}$ removed→ $\boldsymbol{D}_a$ and $\boldsymbol{D}_e$ are very close
  - *Full option*: block $\mathbf{C}$ and 2nd term in $\mathbf{A}$ are both kept → visible difference for some elements of $\boldsymbol{D}$

Approximation options

Individual parameters, log-normal distribution:

$$\gamma_i = e^{\xi_i},\ \xi_i \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega}),$$

- *1st-order approximation*, $\mathbf{E}\xi_i = 0,\ \mathbf{Var}(\xi_i) = V \implies$

$$\mathbf{E}_\xi(e^{\xi_i}) \simeq 1, \qquad \mathbf{Var}_\xi(e^{\xi_i}) \simeq V$$

- *Exact moments*: $\mathbf{E}_\xi(e^{\xi_i}) = e^{V/2},\ \mathbf{Var}_\xi(e^{\xi_i}) = e^V(e^V - 1).$

- $V = 0.6 \implies \mathbf{E}_{1st} = 1, \mathbf{E}_{exact} = 1.35; \mathbf{Var}_{1st} = 0.6, \mathbf{Var}_{exact} = 1.50$

Parameter k_a

Approximation options (cont.)

2nd - order approximation for mean/variance

$$\mathbf{E}_{\boldsymbol{\theta}}[\eta(x,\gamma_i)] \;\approx\; \eta(x,\boldsymbol{\gamma}^0) \;+\; \frac{1}{2}\,\mathrm{tr}\left[\mathbf{H}(\boldsymbol{\gamma}^0)\boldsymbol{\Omega}\right] \;,$$

$$\mathbf{H}(\gamma^0) \;=\; \left[\frac{\partial^2\eta(x,\boldsymbol{\gamma})}{\partial\gamma_\alpha\,\partial\gamma_\beta}\right]\Bigg|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^0} \quad etc \quad \Longrightarrow$$

- All derivatives calculated numerically (central differences)
- Derivatives of variance $\boldsymbol{S}$ require second derivatives of η
- With 2nd order approximation: fourth derivatives…..

Numerically rather cumbersome…

Approximation options: Monte Carlo

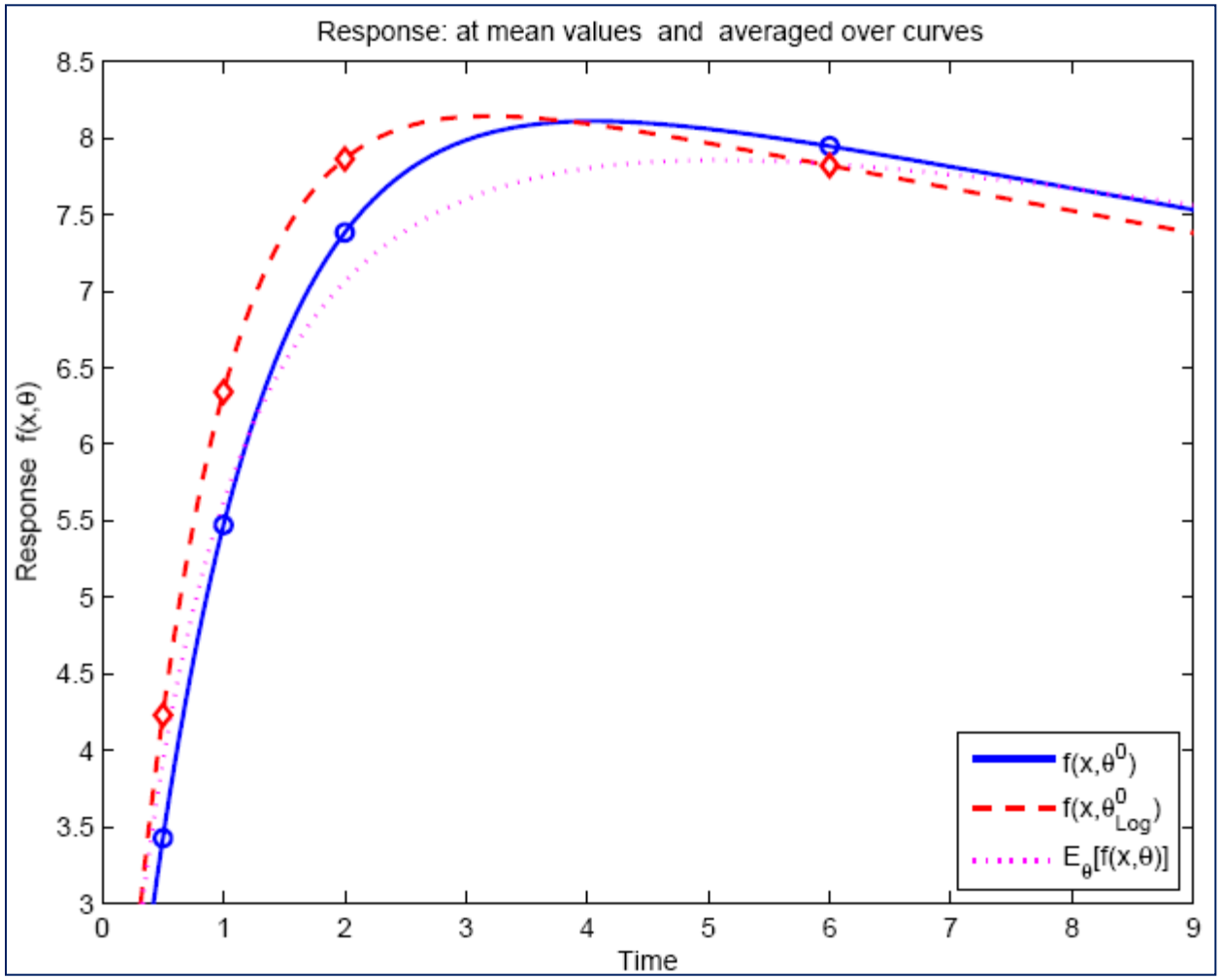
Generate L "patients" according to standard model (1), $\mathbf{Y}_i = \{y_{ij}\}$

$$\widehat{\boldsymbol{\eta}} = \widehat{\boldsymbol{\eta}}(\mathbf{x}, \boldsymbol{\theta}) = \widehat{\mathbf{E}}_{\boldsymbol{\theta}}\mathbf{Y} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{Y}_i \ ,$$

$$\widehat{\mathbf{S}}(\mathbf{x}, \boldsymbol{\theta}) = \widehat{\mathbf{Var}}_{\boldsymbol{\theta}}\mathbf{Y} = \frac{1}{L-1}\sum_{i=1}^{L}[\mathbf{Y}_i - \widehat{\boldsymbol{\eta}}][\mathbf{Y}_i - \widehat{\boldsymbol{\eta}}]^T$$

Use $\widehat{\boldsymbol{\eta}}$, $\widehat{\mathbf{S}}$ in the formula for $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$

Approximation options (cont.)



Mean response curves for one-compartment model example

- Solid - 1st order approximation
- Dashed - computed at mean values of log-normal distribution,
- Dotted - Monte Carlo average

Approximation options: Monte Carlo (cont.)

Central difference

$$g'(\theta) = \frac{g(\theta + h) - g(\theta - h)}{2h} + O(h^2) + \frac{rg(\theta)}{h}$$

For model (1) with proportional residual error,

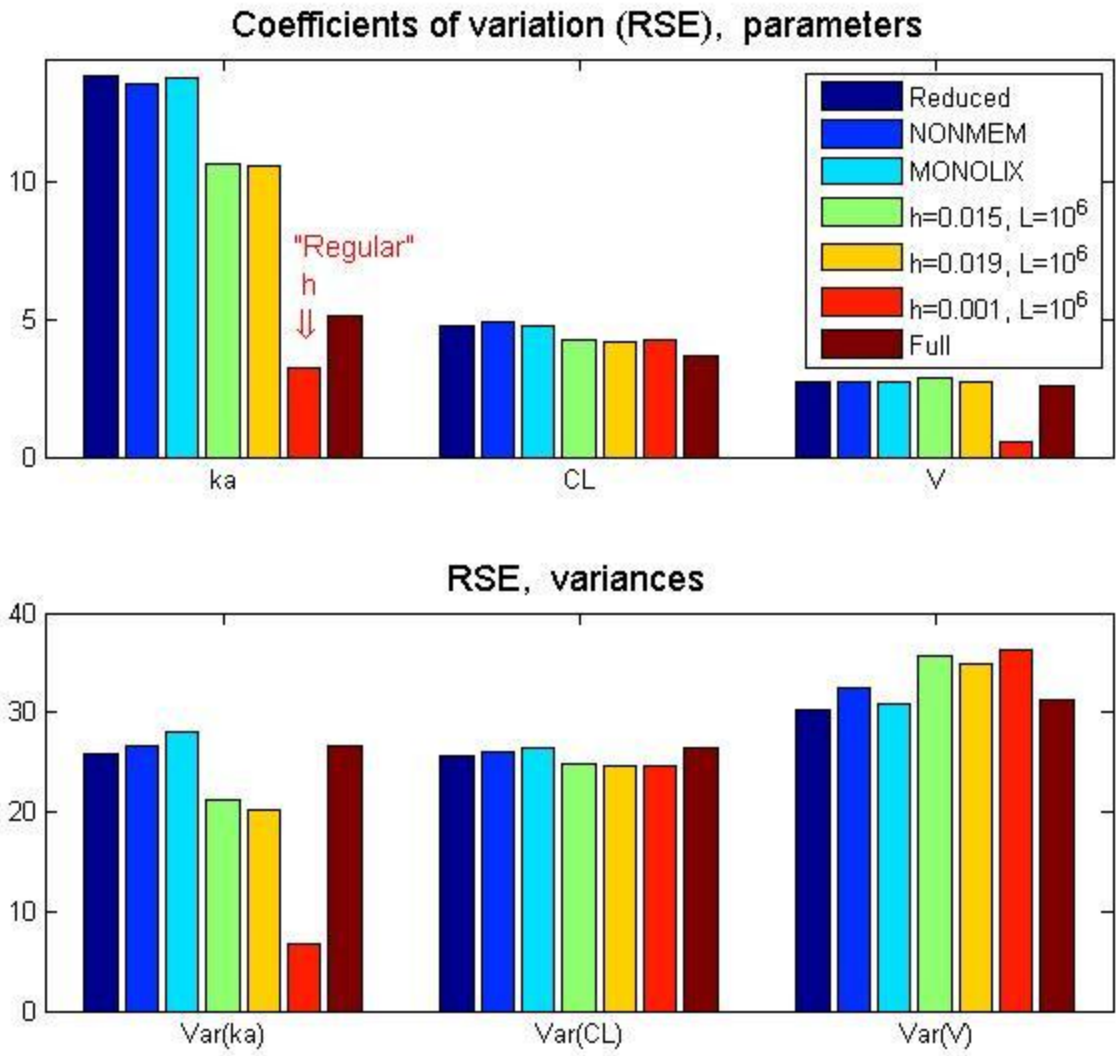
$$Std[\widehat{\eta}(x, \boldsymbol{\theta})] \sim \frac{\sigma_M |\eta(x, \boldsymbol{\theta})|}{\sqrt{L}},$$

so "best" stepsize h can be found from

$$\frac{\sigma_M \eta(x, \boldsymbol{\theta})}{\sqrt{L}\, h} \sim h^2, \quad \text{or} \quad h \sim \left[\frac{\sigma_M \eta(x, \boldsymbol{\theta})}{\sqrt{L}}\right]^{1/3}$$

$f \sim 1, \sigma_M = 0.1, L = 10^6 \Rightarrow h \sim 0.05$, not too small!

Approximation options (cont.)



Summary

- Finding most “informative” levels of controls (sampling times)
- Validating standard designs

  (optimal designs as benchmark)
- Test robustness of optimal designs

  (sampling windows)
- Can incorporate costs/penalties
- Reduce # of samples with “minimal” precision loss
  - Example: from 16 sampling times – to 8 most informative

  D-efficiency (8 samples vs 16 samples) = 0.84 (only 16% lost)
  Gagnon, Leonov (2005)

References

- Aliev, A., Fedorov, V., Leonov, S., McHugh, B., Magee, M. (2012). PkStaMp library for constructing optimal population designs for PK/PD studies. *Comm. Statist. Simul. Comp.* **41** (6), 717–729.
- Atkinson, A.C., Donev, A. (1992). *Optimum Experimental Design*. Clarendon Press, Oxford.
- Elfving, G. (1952). Optimum allocation in linear regression theory, *Ann. Math. Statist*., **23**, 255-262.
- Fedorov, V.V. (1972), *Theory of Optimal Experiment*. Academic Press, NY.
- Fedorov, V.V., Gagnon, R., Leonov, S., Wu, Y. (2007). Optimal design of experiments in pharmaceutical applications. In: Dmitrienko, A., Chuang-Stein, C., D'Agostino, R. (Eds), *Pharmaceutical Statistics Using SAS.* SAS Press, Cary, NC, pp. 151-195.
- Fedorov, V.V., Hackl, P. (1997). *Model-Oriented Design of Experiments*. Springer, NY
- Gagnon, R., Leonov, S. (2005). Optimal population designs for PK models with serial sampling. *J. Biopharm. Stat*., **15** (1), 143-163.
- Graham, G., Aarons, L. (2006). Optimum blood sampling windows for parameter estimation in population pharmacokinetics experiments. *Stat. Med*., **25**, 4004-4019.
- Leonov, S., Aliev, A. (2012). Optimal design for population PK/PD models. *Tatra Mountains Math. Publ.* **51**, 115-130.
- Lindstrom, M.J., Bates, D.M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, **46**, 673-687.

References

- Magnus, J.R., Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, NY.
- Mentré, F., Duffull, S., Gueorguieva, I., Hooker, A., Leonov, S., Ogungbenro, K., Retout, S. (2007). Software for optimal design in population pharmacokinetics and pharmacodynamics: a comparison. In: Abstracts of the PAGE Annual Meeting. ISSN 1871-6032, http://www.page-meeting.org/?abstract=1179
- Mentré, F., Mallet, A., Baccar, D. (1997). Optimal design in random-effects regression models, *Biometrika*, **84** (2), 429-442.
- Mentré, F., Nyberg, J., Ogungbenro, K., Leonov, S., Aliev, A., Duffull, S., Bazzoli, C., Hooker, A. (2011). Comparison of results of the different software for design and evaluation in population pharmacokinetics and pharmacodynamics. In: *Abstracts of the Annual Meeting of the Population Approach Group in Europe (PAGE).* ISSN 1871-6032, www.page-meeting.org/?abstract=2066
- Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S. (1998). Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon therapy. *Science*, **282**, 103-107.
- Pázman, A. (1986). *Foundations of Optimum Experimental Design*. Reidel, Dordrecht.
- Retout, S., Mentré, F. (2003). Further developments of the Fisher information matrix in nonlinear mixed effects models with evaluation in population pharmacokinetics. *J. Biopharm. Stat*., **13** (2), 209-227.
- Wynn, H.P. (1970) The sequential generation of D-optimal experimental designs. *Ann. Math. Statist.,* **41**, 1655-1664.