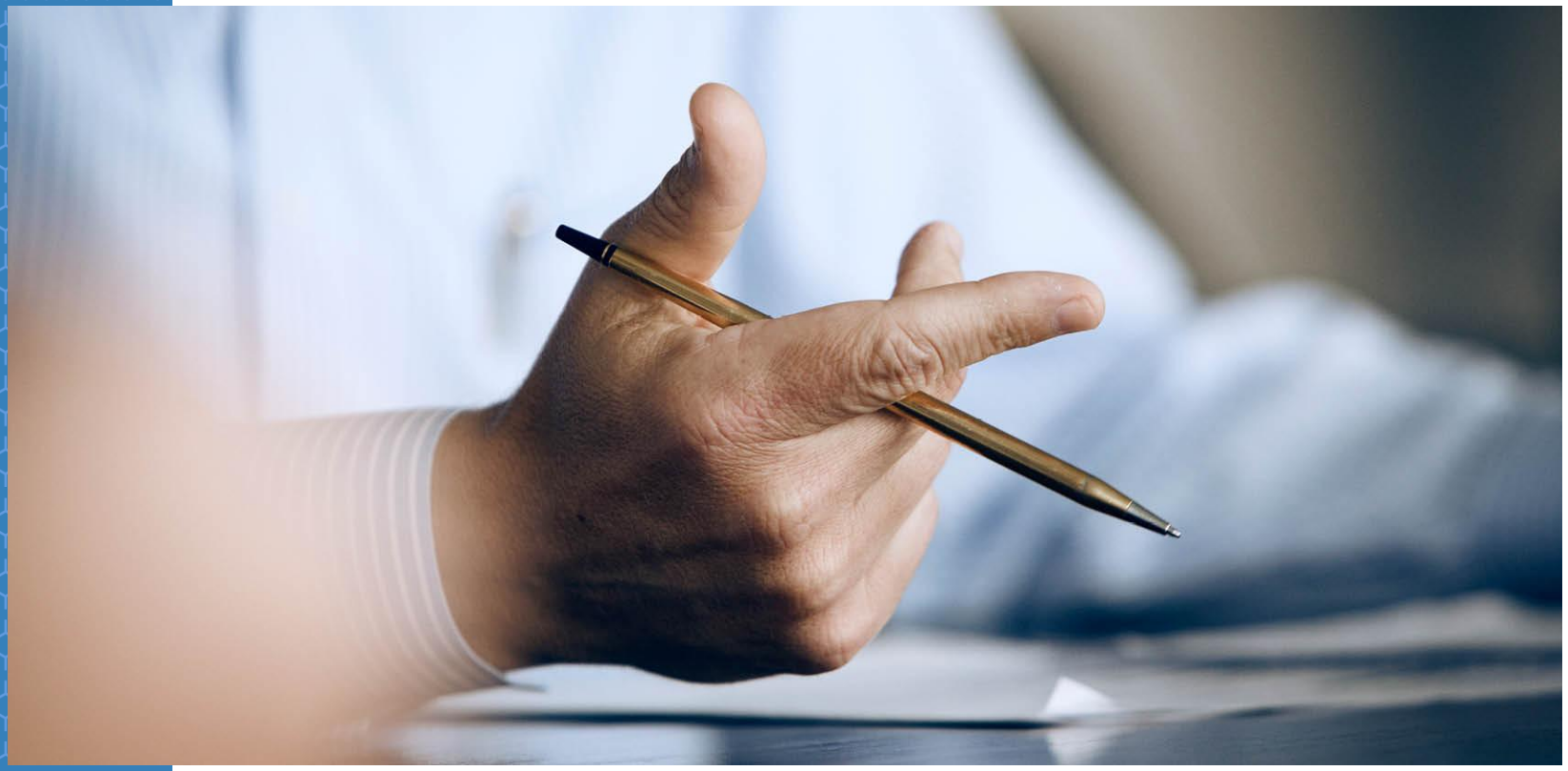


Optimal designs for dose response curves with common parameters

Chrystel Feller
Basel
8.09.2017

NOVARTIS

Acknowledgements

- The work presented here has been done by
  - Kirsten Schörning,
  - Holger Dette,
  - Georgina Bermann
  - Björn Bornkamp
  - Chrystel Feller
- Accepted by Annals of Statistics

Submitted to the Annals of Statistics
arXiv: 1603.04500

OPTIMAL DESIGNS FOR DOSE RESPONSE CURVES WITH COMMON PARAMETERS

*Novartis Pharma AG * and Ruhr-Universität Bochum †*

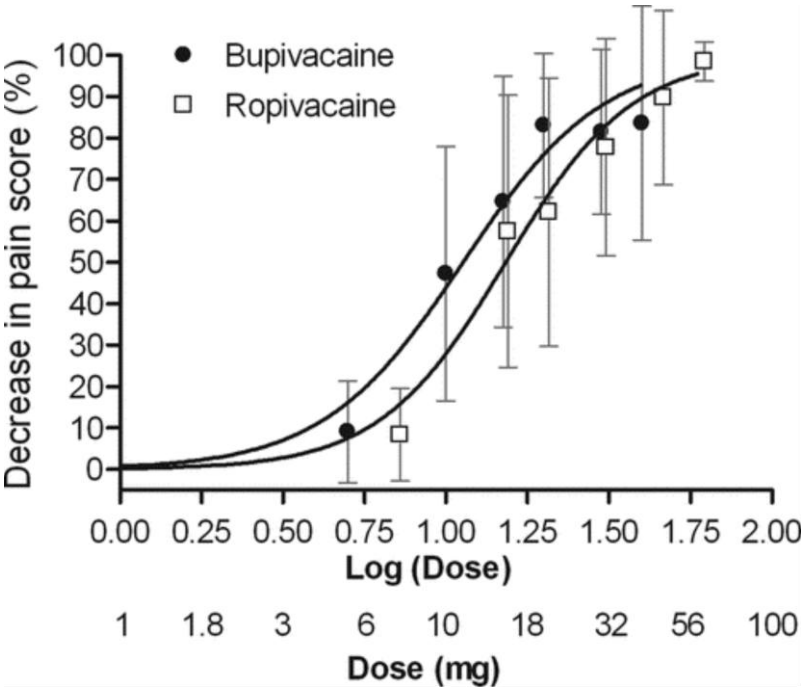
BY CHRYSTEL FELLER*, KIRSTEN SCHORNING†, HOLGER DETTE†, GEORGINA BERMANN* AND BJÖRN BORNKAMP*

Agenda

1. Motivation
2. Definition of D-optimal design
3. Results for 1 and 2 groups (regimen like monthly, weekly,...)
4. Application to an example

Motivation

- Classical set-up and question for Phase II clinical trial:
  - Set-up: N patients assigned to doses (with maximal dose d_{max}) or placebo
  - Objective: estimate the dose-response relationship in order to find a therapeutic dose for Phase III
  - Question: determine the appropriate number and actual levels of the doses to be administered to patients, as well as their relative sample size allocations (for e.g to minimize the error in estimating the dose-response curve)
- Large amount of literature discussing the problem of constructing optimal designs for regressions models

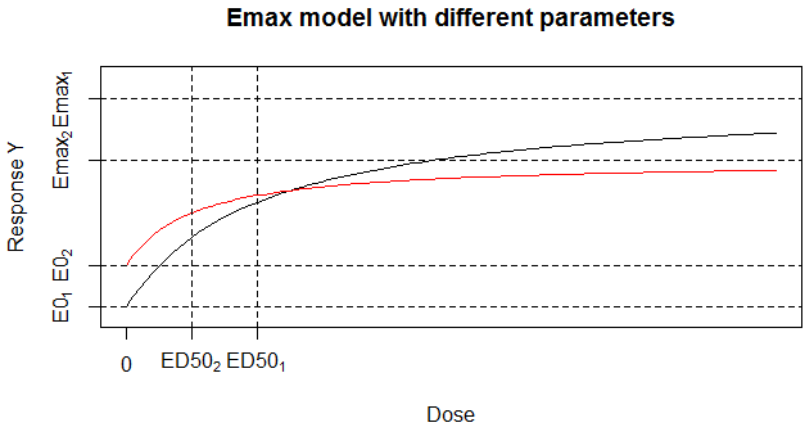


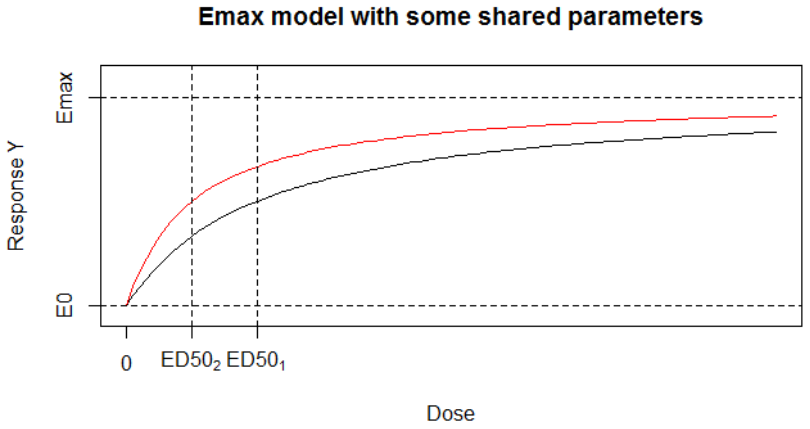
Example of dose-response relationship (Ngan et al. 2010)

NOVARTIS

Motivation (2)

- Generalize framework:
  - N patients, assigned to doses or placebo
  - but assume patients differ not only in terms of “dose” but also in terms of another aspect (e.g. treatment frequency)
- Approach:
  - assume that the same dose-response models holds overall (e.g. formulated in terms of total per time unit),
  - but some parameters are shared while others are not shared between the different groups





Notation

- N patients
- Dose range (per unit of time): $0 \leq d \leq d_{max}$
- M different regimens (like weekly, monthly,…)
- Response: Y
- Dose response model:

$$Y_{ijl} = f\left(d_j, i, \vartheta\right) + \varepsilon_{ijl},\ \varepsilon_{ijl} \sim N(0, \sigma_i^2)$$

$$\mathrm{i} = 1, \ldots, M, \mathrm{j} = 1, \ldots, \mathrm{k_i}, \mathrm{l} = 1, \ldots, \mathrm{n}_{ij}$$

  - Emax: $f(d, i, \vartheta) = {\vartheta_0}^{(i)} + {\vartheta_1}^{(i)} \frac{d}{{\vartheta_2}^{(i)} + d}$
  - Sigmoid Emax: $f(d, i, \vartheta) = {\vartheta_0}^{(i)} + {\vartheta_1}^{(i)} \frac{d^{\gamma}}{({\vartheta_2}^{(i)})^{\gamma} + d^{\gamma}}$

Notation (2)

- Design: M=2

$$\xi = \begin{pmatrix} 1 & \dots & 1 & 2 & \dots & 2 \\ d_1^{(1)} & \dots & d_{k_1}^{(1)} & d_1^{(2)} & \dots & d_{k_2}^{(2)} \\ \omega_1^{(1)} & \dots & \omega_{k_1}^{(1)} & \omega_1^{(2)} & \dots & \omega_{k_2}^{(2)} \end{pmatrix}$$

$$=: \begin{pmatrix} \delta_1 & \dots & \delta_k \\ d_1 & \dots & d_k \\ \omega_1 & \dots & \omega_k \end{pmatrix} \text{ with } k = k_1 + k_2.$$

δ_i= group number of patients receiving dose d_i

ω_i =proportion of patients receiving dose d_i in group δ_i

(with $\sum_{i=1}^{k} \omega_i$=1)

D-optimal design

- Concept: derive local optimal designs that minimize the volume of the confidence ellipsoid for the model parameters under a particular dose–response model (Atkinson et al., 2007; Fedorov and Leonov, 2013)
- Closed form expressions can often be derived by standard arguments using the equivalence theorem for D-optimality, such as for the linear and the Emax model (Dette et al 2010).
- Otherwise, D-optimal designs have to be determined numerically.

D-optimal design

Optimization problem

- True model parameter set: $\bar{\vartheta}$
- Least square estimator: $\widehat{\vartheta_N}$
- Asymptotic normality:

$$\sqrt{N}(\hat{\theta}_N - \bar{\theta}) \to^d z \sim \mathcal{N}(0, \sigma^2 \left(M_\xi(\bar{\theta})\right)^{-1}), \quad N \to \infty$$

where

$$M_\xi(\theta) = \sum_{j=1}^{k} \omega_j \frac{\partial f(d_j, \delta_j, \theta)}{\partial \theta} \frac{\partial f(d_j, \delta_j, \theta)'}{\partial \theta}.$$

- $\Psi(\xi) = \ln(\det(M_\xi(\theta)^{-1}))$ → Min → $\xi^* = \begin{pmatrix} \delta_1 & \dots & \delta_k \\ d_1 & \dots & d_k \\ \omega_1 & \dots & \omega_k \end{pmatrix}$

$$\sum_{j=1}^{k} \omega_j = 1, \qquad \omega_k \in [0,1]$$

$$d_j \in [0, d_{max,i}] \text{ if } \delta_j = i$$

$$i = 1, \dots, M$$

D-optimal design

result for one group (M=1)

- For the Emax model:

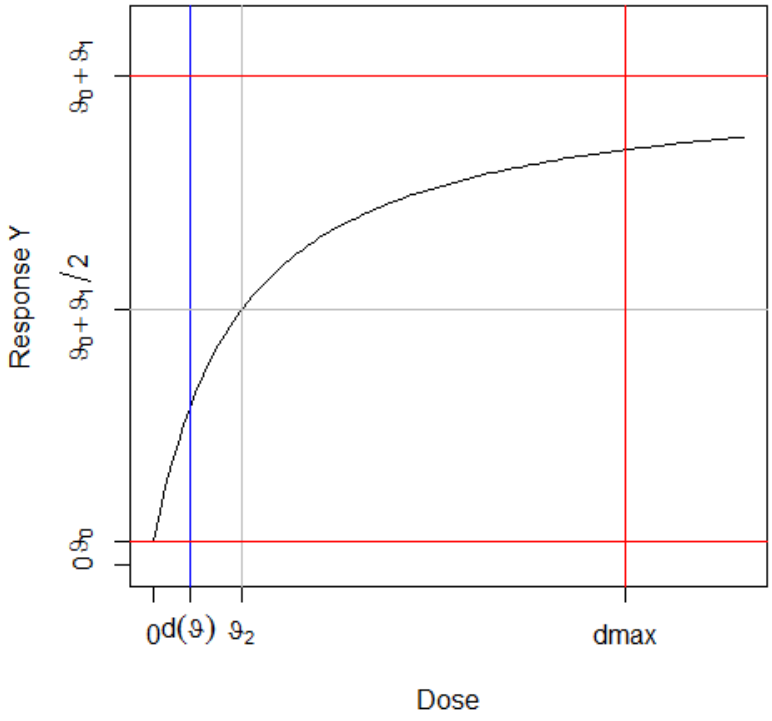
$$\xi_D^* = \begin{pmatrix} 0 & d(\vartheta) & d_{max} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$d(\vartheta) = \frac{\vartheta_2 d_{max}}{2\vartheta_2 + d_{max}}$$

For large d_{max}, $d(\vartheta)$ tends towards ϑ_2

- Results available for other model like the linear in log model:

$$f(d, i, \vartheta) = \vartheta_0 + \vartheta_1 \log\left(\frac{d}{\vartheta_2} + 1\right)$$



NOVARTIS

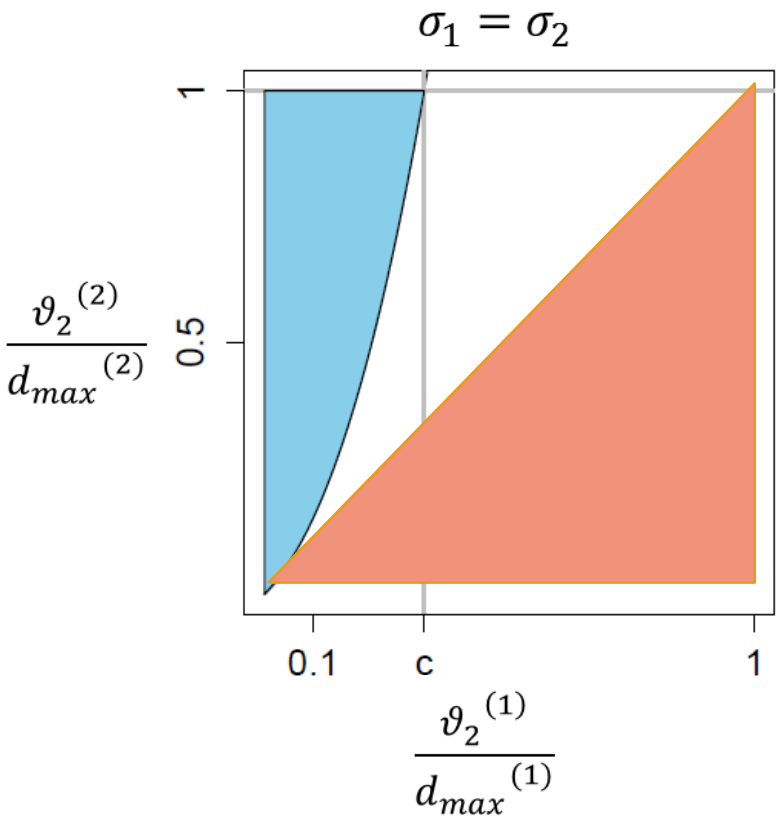
D-optimal design

2 groups (M=2), different slope, same variance

- Model: $f(d,i,\vartheta) = \vartheta_0 + \vartheta_1 \frac{d}{{\vartheta_2}^{(i)}+d}$
- Assume $0 < \frac{{\vartheta_2}^{(1)}}{{d_{max}}^{(1)}} < \frac{{\vartheta_2}^{(2)}}{{d_{max}}^{(2)}} < 1$ (i.e. not in the red region i.e. group 1=smallest normalized ED50) without restriction of generality
- For parameters in the blue region, the locally D-optimal design is:

$$\xi_D^* = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & d(\vartheta^{(1)}) & {d_{max}}^{(1)} & {\vartheta_2}^{(2)} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

$$d(\vartheta^{(i)}) = \frac{{\vartheta_2}^{(i)} {d_{max}}^{(i)}}{2{\vartheta_2}^{(i)} + {d_{max}}^{(i)}}$$



- For parameters in the white region, the D-optimal design has 4 or 5 doses and contains at least ${d_{max}}^{(1)}$ or ${d_{max}}^{(2)}$ -> need to be computed numerically

D-optimal design

2 groups (M=2), different slope, different variance

For parameters in the grey region (next slide), the locally D-optimal design is:

1. If $r = \frac{\sigma_1^2}{\sigma_2^2} \leq 1$ ($group$ 1: $smallest$/equal $variance$):

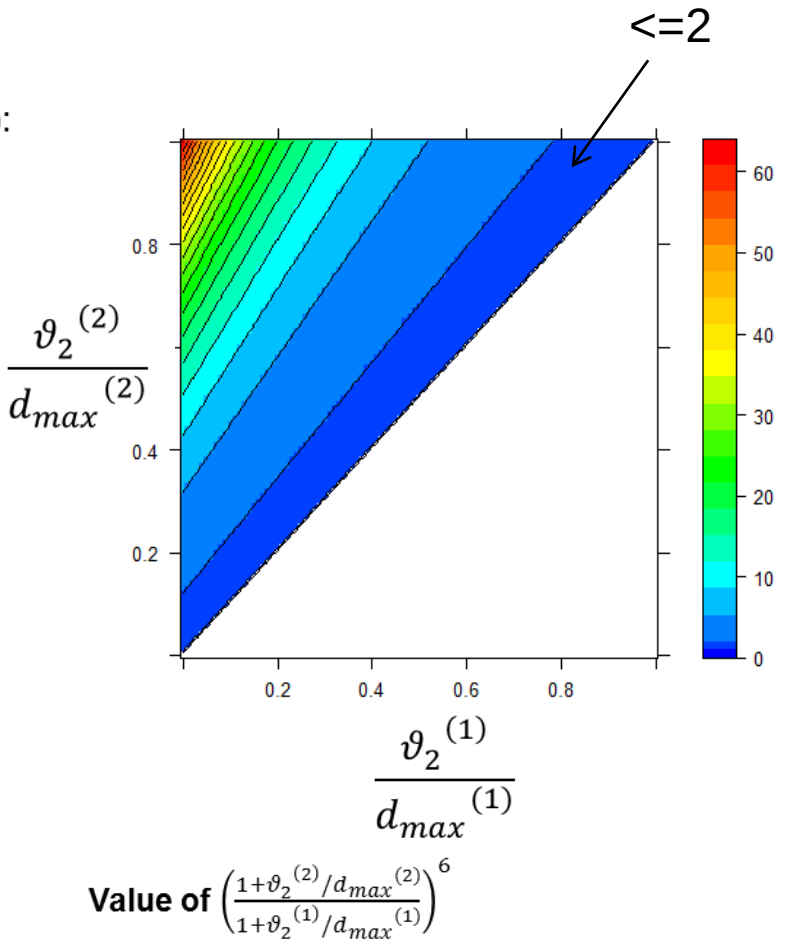
$$\xi_D^* = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & d(\vartheta^{(1)}) & {d_{max}}^{(1)} & {\vartheta_2}^{(2)} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

2. If $1 < r \leq \left(\frac{1+{\vartheta_2}^{(2)}/{d_{max}}^{(2)}}{1+{\vartheta_2}^{(1)}/{d_{max}}^{(1)}}\right)^6$:

$$\xi_D^* = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & d(\vartheta^{(1)}) & {d_{max}}^{(1)} & {\vartheta_2}^{(2)} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

3. If r $> \left(\frac{1+{\vartheta_2}^{(2)}/{d_{max}}^{(2)}}{1+{\vartheta_2}^{(1)}/{d_{max}}^{(1)}}\right)^6$:

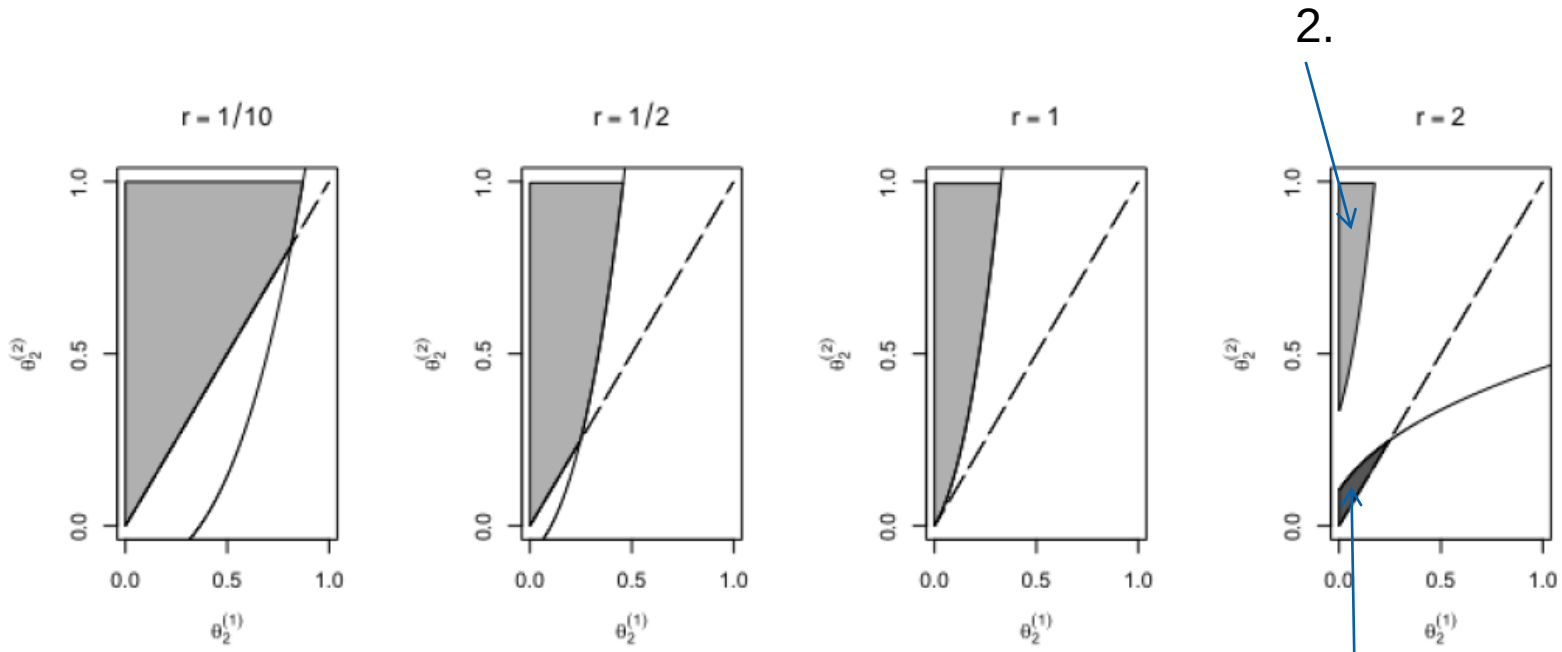
$$\xi_D^* = \begin{pmatrix} 2 & 2 & 2 & 1 \\ 0 & d(\vartheta^{(2)}) & {d_{max}}^{(2)} & {\vartheta_2}^{(1)} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$



D-optimal design

2 groups (M=2), different slope, different variance

Grey region



$$r = \sigma_1^2/\sigma_2^2$$

D-optimal design

M groups, different slope and maximum response

- Model: $f(d,i,\vartheta) = \vartheta_0 + {\vartheta_1}^{(i)} \frac{a}{{\vartheta_2}^{(i)}+d}$
- If group 1 has the smallest variance ($\sigma_1^2 = \min_{i\in\{1,\dots,M\}} {\sigma_i}^2$) then the D-optimal design is

$$\xi_D^* = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & \dots & M & M \\ 0 & d(\vartheta^{(1)}) & {d_{max}}^{(1)} & d(\vartheta^{(2)}) & {d_{max}}^{(2)} & \dots & d(\vartheta^{(M)}) & {d_{max}}^{(M)} \\ \frac{1}{m} & \frac{1}{m} & \frac{1}{m} & \frac{1}{m} & \frac{1}{m} & \dots & \frac{1}{m} & \frac{1}{m} \end{pmatrix}$$

D-optimal design for one group

$$d(\vartheta^{(i)}) = \frac{{\vartheta_2}^{(i)} {d_{max}}^{(i)}}{2{\vartheta_2}^{(i)} + {d_{max}}^{(i)}}$$

$$m = 3 + 2 * M$$

Example

- Phase II study
- Principle of action: increase the level of a biomarker that induces a beneficial clinical effect in patients
- Dosing group: weekly or monthly administration
  - Dose range [0,400] (weekly) and [0,1000] (monthly)
- Objective: characterization of dose-response relationships at a given time-point T
- Design questions:
  - which doses should be studied in each group
  - how to split the total sample size between the two groups

Example

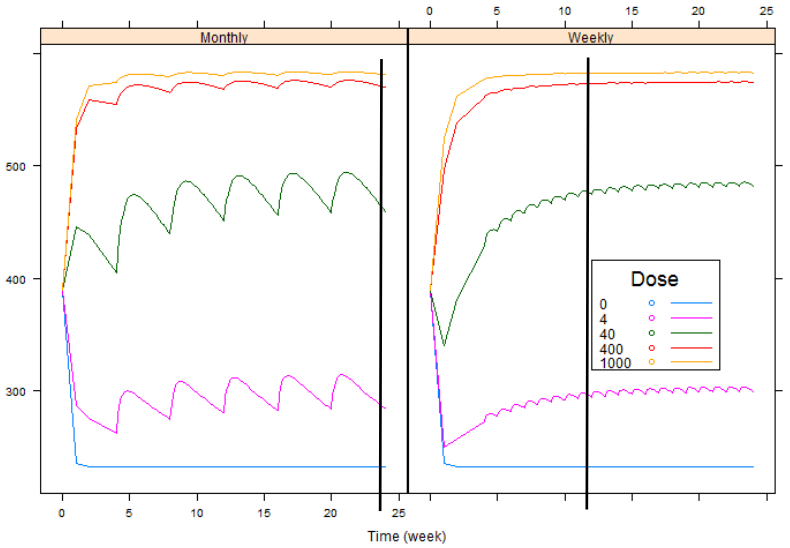
Use information from PK-PD model on Phase IIa data

- As the design computed is D-optimal **locally**, it is important to have a good guess of the underlying true dose-response curve
- Idea: use data from Phase IIa to get an estimation of the dose-response curve

Example

Use PK-PD model to derive an estimate of the dose-response

1. Fit of the PK/PD model
2. Parameter set for each patient
3. Simulate a new trial data
   – 22 arms and 200 patients



4. Fit an Emax model to each regimen on the simulated data at the end of the study

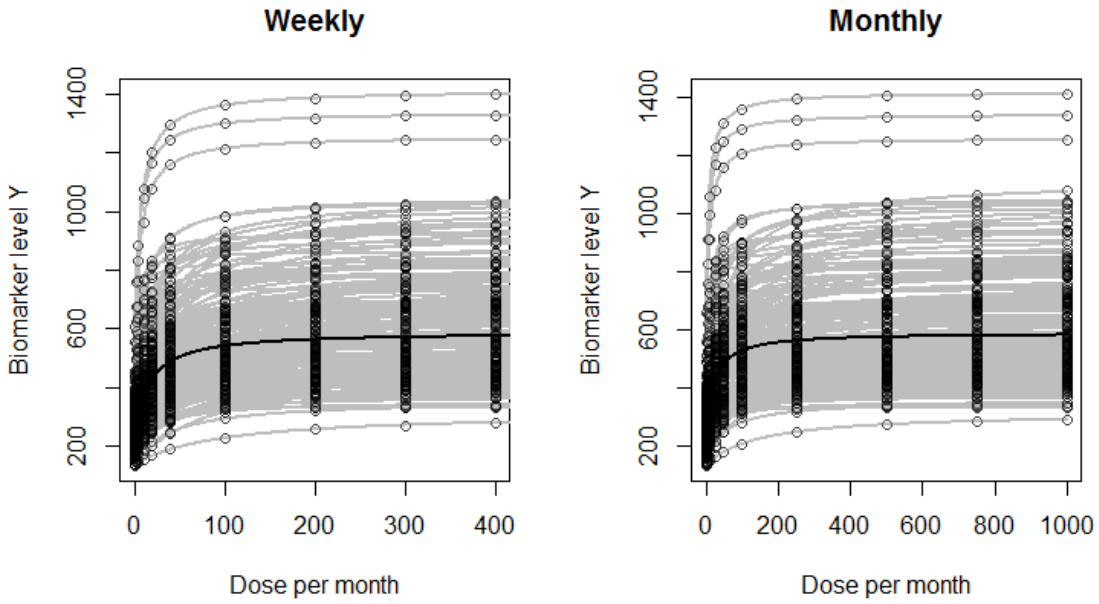
Example

Guess of the dose-response

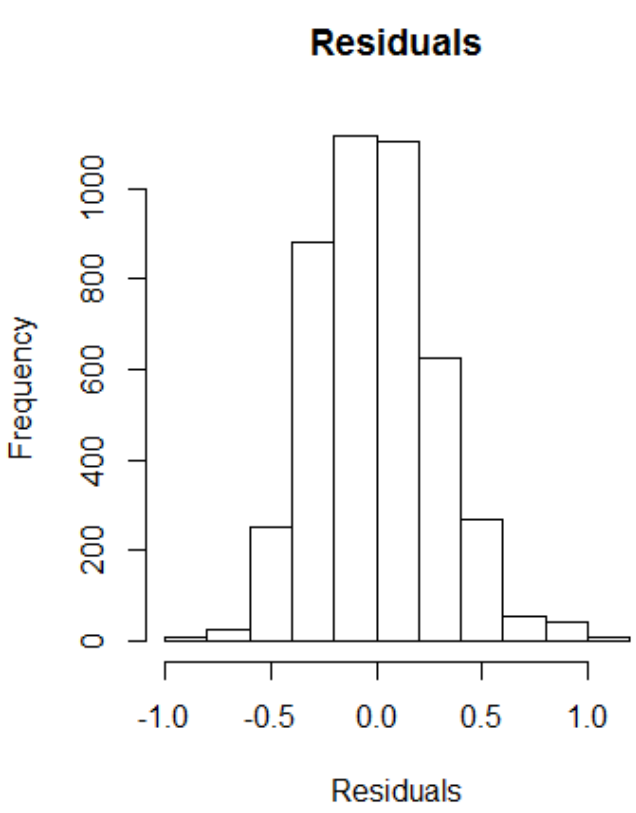
Result of previous slide procedure:

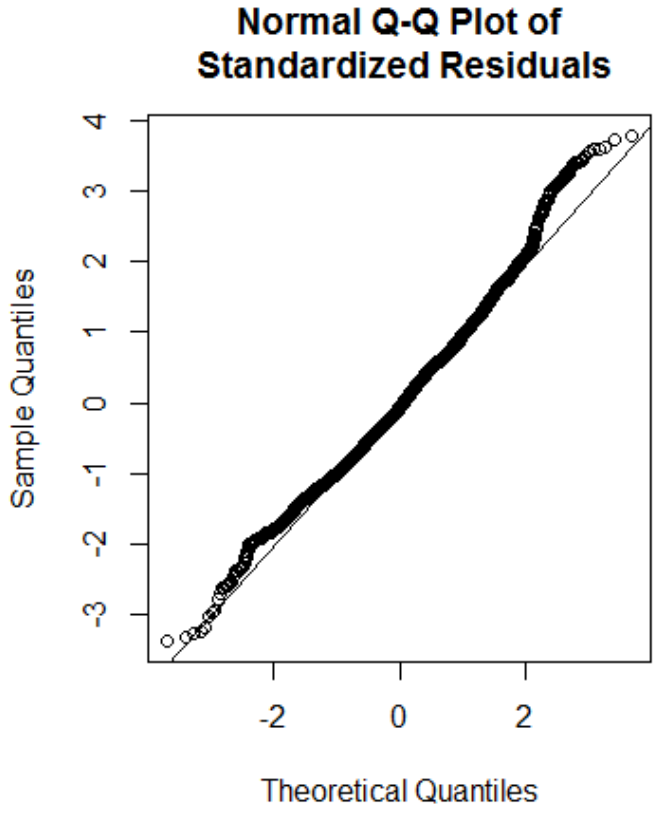
$$f(d, i, \vartheta) = \vartheta_{11} + \vartheta_{12} \frac{d}{\vartheta_2{}^{(i)} + d}$$

$$\vartheta_{11} = 5.48, \vartheta_{12} = 0.90, \vartheta_2{}^{(1)} = 13.82, \vartheta_2{}^{(2)} = 10.46$$



- Residuals


Residuals
Frequency
0
200
400
600
800
1000
-1.0
-0.5
0.0
0.5
1.0
Residuals


Normal Q-Q Plot of
Standardized Residuals
Sample Quantiles
-3
-2
-1
0
1
2
3
4
-2
0
2
Theoretical Quantiles

Example
D-optimal design

- We are in the blue region. The D-optimal design is:

$$\xi_D^* = \begin{pmatrix} \delta_1 & 1 & 1 & 2 \\ 0 & d(\vartheta^{(1)}) & {d_{max}}^{(1)} & {\vartheta_2}^{(2)} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \delta_1 & 1 & 1 & 2 \\ 0 & 13.45 & 1000 & 10.46 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

Example

Robust design

- design derived for a fixed model are rather sensitive with respect to this assumption
  -> construct robust optimal designs that take into account a set of potential dose–response profiles within classes of models are commonly used in drug development practice

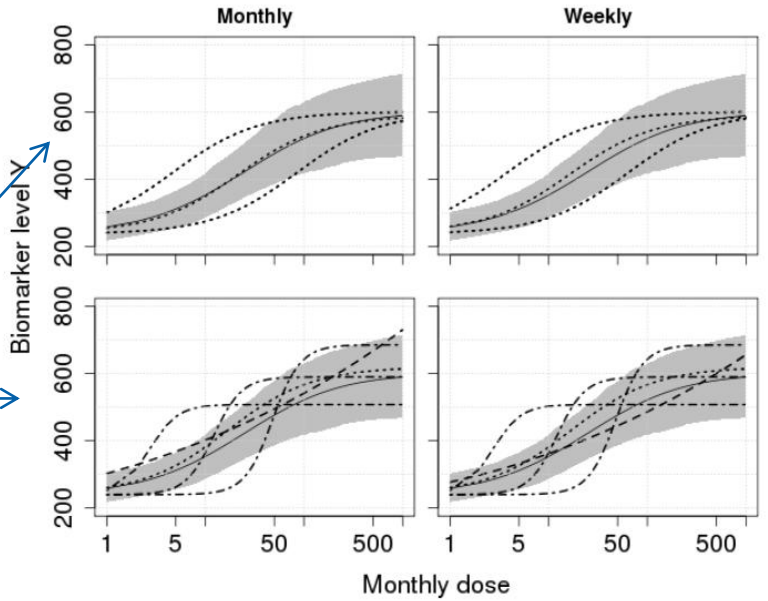
Example

Robust design

Candidate models for the dose-response curve in monthly and weekly group.

Set of candidate models used in the robust criterion (5.1)

Model type	id	ϑ_{11}	ϑ_{12}		$\theta_2^{(1)}$	$\theta_2^{(2)}$	γ
Emax	1	5.48	0.90		13.82	10.46	1
Emax	2	5.47	0.93		2.93	2.39	1
Emax	3	5.47	0.93		2.93	40.40	1
Emax	4	5.47	0.93		53.49	2.39	1
Emax	5	5.47	0.93		53.49	40.40	1
Sigmoid Emax	6	5.48	0.90		13.82	10.46	3
Model type	id	θ_1	$\vartheta_1^{(1)}$	$\vartheta_1^{(2)}$	$\vartheta_2^{(1)}$	$\vartheta_2^{(2)}$	γ
Emax	7	5.48	0.85	0.95	13.82	10.46	1
Sigmoid Emax	8	5.48	0.65	0.75	2.93	2.39	3
Sigmoid Emax	9	5.48	0.95	1.05	53.49	40.40	3
Log	10	5.44	0.13	0.14	0.32	0.41	



When $\vartheta_0^{(1)} = \vartheta_0^{(2)}$ **then** $\vartheta_0^{(1)} = \vartheta_0^{(2)} =: \vartheta_{11} = \theta_1$
When $\vartheta_1^{(1)} = \vartheta_1^{(2)}$ **then** $\vartheta_1^{(1)} = \vartheta_1^{(2)} =: \vartheta_{12}$

Solid line: population average of the new trial data generated with the PK/PD model
Grey area: biomarker level between the 25th and 75th quantiles of the patient responses of the new trial data
Dotted curves: Emax models

Example

Robust design

1=weekly
2=monthly

- D-optimal design for the best guess model (model 1):

$$\xi_1^* = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 13.45 & 1000 & 10.46 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

- The compound optimal design* based on the first 5 models:

$$\xi_{c,5}^* = \begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 2 \\ 0 & 3.02 & 43.67 & 1000 & 2.53 & 37.51 \\ 0.17 & 0.16 & 0.17 & 0.17 & 0.16 & 0.17 \end{pmatrix}$$

- The compound optimal design* based on the 10 models:

$$\xi_{c,5}^* = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 0 & 2.90 & 12.98 & 41.91 & 1000 & 3.01 & 13.16 & 49.46 & 400 \\ 0.16 & 0.08 & 0.13 & 0.08 & 0.14 & 0.14 & 0.09 & 0.13 & 0.06 \end{pmatrix}$$

*Design that maximize the mean efficiency: $g_c(\xi, s) = \frac{1}{s}\sum_{i=1}^{s} Eff_i(\xi)$

Where $Eff_i(\xi) = \left(|M_i(\xi, \vartheta_i)|/|M_i(\xi^{*,i}, \vartheta_i)|\right)^{1/m_i}$ and m_i = number of parameters in model i

Example

Efficiency

- Efficiency: $Eff_i(\xi) = \left(|M_i(\xi, \vartheta_i)|/|M_i(\xi^{*,i}, \vartheta_i)|\right)^{1/m_i}$
  m_i = number of parameters in model i

Efficiency $Eff_i(\xi^\star_{c,s})$ of the two compound optimal designs compared to each of the locally D-optimal designs for the 10 models.

	$g_c(\cdot, s)$	1	2	3	4	5	6	7	8	9	10
$\xi^\star_{c,5}$	0.82	0.71	0.84	0.88	0.85	0.85	0.10	0.80	0.93	0.91	0.63
$\xi^\star_{c,10}$	0.75	0.83	0.75	0.78	0.77	0.79	0.75	0.90	0.76	0.75	0.75

Means that if the true dose-response is model 1 then the sample size need to be increase by approx +20% to get the same precision as the optimal design ξ_1^*

Summary

- Described an approach for optimal designs in dose-response situations, when there are groups that differ in dose and one further aspect (e.g. treatment frequency)
  - Basic assumption: Some parameters are shared between the dose-response functions in the different groups
- We provided an upper bound for the number of doses for the locally D-optimal designs and derived a formula for the D-optimal design (for a part of the parameter space)
- In practice it is important to account for possible mis-specification of the parameter guesses used and use robust designs (have to be calculated numerically)

Thank you

NOVARTIS

References

1. Ngan Kee WD, Ng FF, Khaw KS, Lee A, Gin T, **Determination and Comparison of Graded Dose–Response Curves for Epidural Bupivacaine and Ropivacaine for Analgesia in Laboring Nulliparous Women,** 2010, Anesthesesiology 113(2):445-453.
2. Anthony Atkinson, Alexander Donev, and Randall Tobias, **Optimum experimental designs, with SAS**, Oxford University Press, 2007.
3. Valerii V. Fedorov and Sergei L. Leonov, **Optimal design for nonlinear response models**, New York: CRC Press, 2013.
4. Dette, H., Kiss, C., Bevanda, M. and Bretz, F, Optimal designs for the Emax, log-linear and exponential models, 2010, *Biometrika, 97*(2), 513-518

Back-up slides

Example

PK/PD model

- 2 compartments-model (Peripheral, central volume)

$$\frac{dC_c^{(n)}}{dt} = \frac{d}{V_c} e^{-k_a(t-t_n)} - k_{10} C_c^{(n)}(t) - k_{12} C_c^{(n)}(t) + k_{21} C_p^{(n)}(t)$$

$$\frac{dC_p^{(n)}}{dt} = k_{12} C_c^{(n)}(t) - k_{21} C_p^{(n)}(t)$$

After taking the nth dose

$$\frac{dR^{(n)}}{dt} = k_{in} - k_{out}\left(1 - I_{max}\frac{C_c{}^{(n)}(t)}{IC_{50} + C_c{}^{(n)}(t)}\right)R^{(n)}(t)$$

Inhibition of loss