Using Hamiltonian Monte-Carlo to design longitudinal studies with discrete outcomes accounting for parameter and model uncertainties

Florence Loingeville, Jérémy Seurat, Thu Thuy Nguyen, Marie-Karelle Riviere, France Mentré

IAME, UMR 1137 INSERM - University Paris Diderot, Paris, France

PODE, September 8th, 2017, Basel

Integrated DEsign and AnaLysis of small population group trials

Contents

Designs in pharmacometrics

- Last decades: several methods/software for **maximum likelihood estimation** of population parameters from **longitudinal data** using **nonlinear mixed effect models** (NLMEM)

- Problem beforehand: **choice of "population" design**
  - To obtain precise estimates / adequate power
    - number of individuals?
    - number of sampling times/individual?
    - allocation of sampling times?
    - other design variables (doses, etc.)
  - **Clinical trial simulation (CTS)**: time consuming
  - Asymptotic theory: **expected Fisher Information Matrix** [1] (FIM)

[1] Mentré et al. *Biometrika*, 1997.

Fisher Information Matrix in NLMEM

- **Analytical expression for FIM in NLMEM**
  - Current approach in PFIM [2] and other design software programs [3]:
    first order linearisation of model around the expectation of random effects (FO)
    - Only for continuous data
    - Performs well but has limitations in case of complex nonlinear models and/or large variability
- **FIM for discrete longitudinal data**:
  - Methods based on approximations [4, 5]
  - We propose new approaches for computation of FIM
    - Monte Carlo - Adaptive Gaussian Quadrature (MC-AGQ)[6]
    - Monte Carlo - Hamiltonian Monte Carlo (MC-HMC)[7]

    These approaches:
    - Without model linearisation
    - Evaluated and compared to CTS and Laplace approx. on 4 longitudinal data types: continuous, binary, count, time to event

[2] PFIM group, www.pfim.biostat.fr.

[3] Nyberg et al. *Br J Clin Pharmacol*, 2014.

[4] Waite and Woods. *Biometrika*, 2015.

[5] Ogungbenro and Aarons. *J Pharmacokinet Pharmacodyn*, 2011.

[6] Ueckert and Mentré. *Comput Stat Data Anal*, 2016.

[7] Riviere, Ueckert and Mentré. *Biostatistics*, 2016.

Parameter and model uncertainty in designs

- **Optimal design depends on knowledge on model and parameters**
  - Local planification: given the model m and parameter values Ψ_m^*
  - Widely used criterion: D-optimality [8]
- **Alternative: Robust designs**
  - Taking into account uncertainty on parameters
  - Across a set of candidate models (Example in dose-response study proposed [9, 10] and implemented in MCP-MOD [11])

[8] Atkinson, Donev and Tobias. *Optimum experimental designs, with SAS*, 2009.

[9] Bretz, Pinheiro and Branson. *Biometrics*, 2005.

[10] Pinheiro et al. *Stat Med*, 2014.

[11] Bornkamp et al. cran.r-project.org/web/packages/MCPMod/index.html, 2008.

NLMEM: Notations

For continuous data:
$y_i = f(g(\mu, b_i, z_i, \beta), \xi_i) + \epsilon_i$

For discrete data:
$p(y_i|b_i) = \prod_{j=1}^{n_i} h(y_{ij}, g(\mu, b_i, z_i, \beta), \xi_i)$

with

$y_i = (y_{i1}, \ldots, y_{in_i})^T$ response for individual i $(i = 1, \ldots, N)$

f, h structural model

ξ_i elementary design for subject i

$\theta_i = g(\mu, b_i, z_i, \beta)$ individual parameters vector

μ vector of fixed effects

b_i vector of random effects for individual i, $b_i \sim \mathcal{N}(0, \Omega)$

z_i vector of covariates

β vector of covariate effects

ϵ_i vector of residual errors, $\epsilon_i \sim \mathcal{N}(0, \Sigma)$ and Σ diagonal matrix

Ψ: Population parameters $\{\mu, \Omega, \Sigma, \beta\}$

MC-HMC method to compute the FIM in NLMEM

Population FIM for one group design: $\mathcal{M}(\Psi,\Xi) = N \times \mathcal{M}(\Psi,\xi)$
Population design $\Xi = \{\xi, N\}$ with identical elementary design ξ in all N subjects

Elementary FIM: $\mathcal{M}(\psi,\xi) = E_y\left(\frac{\partial \log(L(y,\psi))}{\partial \psi}\frac{\partial \log(L(y,\psi))}{\partial \psi}^T\right)$

$$\mathcal{M}(\psi,\xi)_{k,l} = E_y\left(\underbrace{\frac{\partial \log(L(y,\psi))}{\partial \psi_k}\frac{\partial \log(L(y,\psi))}{\partial \psi_l}^T}_{D_{k,l}}\right)$$

Monte Carlo - MC

After calculation... $D_{k,l} \Longleftrightarrow$

$$\int_{b_1} \frac{\partial\left(\log(p(y|b_1,\psi)p(b_1|\psi))\right)}{\partial \psi_k} \underbrace{\frac{p(y|b_1,\psi)p(b_1|\psi)}{\int p(y|b,\psi)p(b|\psi)db}}_{\text{conditional density of } b \text{ given } y} db_1 \cdot \int_{b_2} \frac{\partial\left(\log(p(y|b_2,\psi)p(b_2|\psi))\right)}{\partial \psi_l} \underbrace{\frac{p(y|b_2,\psi)p(b_2|\psi)}{\int p(y|b,\psi)p(b|\psi)db}}_{\text{conditional density of } b \text{ given } y} db_2$$

[7] Riviere, Ueckert and Mentré. *Biostatistics*, 2016.

[12] Stan Development Team. Stan: A C++ Library for Probability and Sampling.

[13] Riviere and Mentré. https://cran.r-project.org/web/packages/MIXFIM/index.html, 2015.

MC-HMC method to compute the FIM in NLMEM

$$\mathcal{M}(\psi,\xi) = E_y\left(\frac{\partial \log(L(y,\psi))}{\partial \psi}\frac{\partial \log(L(y,\psi))}{\partial \psi}^T\right)$$

$$\mathcal{M}(\psi,\xi)_{k,l} = E_y\left(\underbrace{\frac{\partial \log(L(y,\psi))}{\partial \psi_k}\frac{\partial \log(L(y,\psi))}{\partial \psi_l}^T}_{D_{k,l}}\right)$$

Monte Carlo - MC

After calculation... $D_{k,l} \Longleftrightarrow$

$$E_b\left(\frac{\partial(\log(p(y|b,\psi)p(b|\psi)))}{\partial \psi_k}\middle|Y\right).E_b\left(\frac{\partial(\log(p(y|b,\psi)p(b|\psi)))}{\partial \psi_l}\middle|Y\right)$$

Markov Chains Hamiltonian Monte Carlo (MC-HMC)

$\Rightarrow$ **Two integrals to compute: w.r.t. y and w.r.t. b**

Use of **MC** and **Hamiltonian Monte Carlo (HMC)** **(in Stan** [12]) [7], implemented in R package MIXFIM [13]

[7] Riviere, Ueckert and Mentré. *Biostatistics*, 2016.

[12] Stan Development Team. Stan: A C++ Library for Probability and Sampling.

[13] Riviere and Mentré. https://cran.r-project.org/web/packages/MIXFIM/index.html, 2015.

MC-HMC method to compute the robust FIM in NLMEM

Robust FIM, assuming a distribution $p(\Psi)$ on the parameters

$$\mathcal{M}_R(\Xi) = E_\Psi(\mathcal{M}(\Psi,\Xi))$$

$$\mathcal{M}_R(\Xi) = E_\Psi(\mathcal{M}(\Psi,\Xi)) = E_\Psi\left(E_y(\underbrace{\frac{\partial \log(L(y,\psi))}{\partial \psi_k}\frac{\partial \log(L(y,\psi))}{\partial \psi_l}^T}_{D_{k,l}})\right) = E_{\Psi,y}(D_{k,l})$$

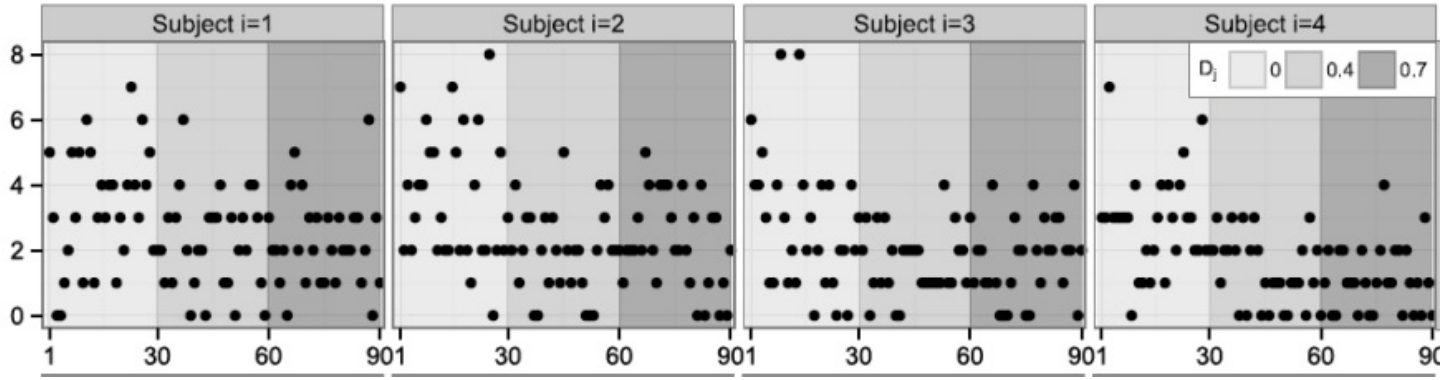
- two integrals **w.r.t. y** and **w.r.t. b** for evaluation of $\mathcal{M}(\Psi,\Xi)$
- one supplementary integral **w.r.t. Ψ** for evaluation of $\mathcal{M}_R(\Xi)$
- Evaluation by **MC-HMC** using Stan (drawing jointly Ψ and y by MC)

Application to robust designs for repeated count data

- Example: Daily count of events that we want to prevent [7]
- Poisson model for repeated count response

$$P(y = k|b) = \frac{\lambda^k exp(-\lambda)}{k!}$$

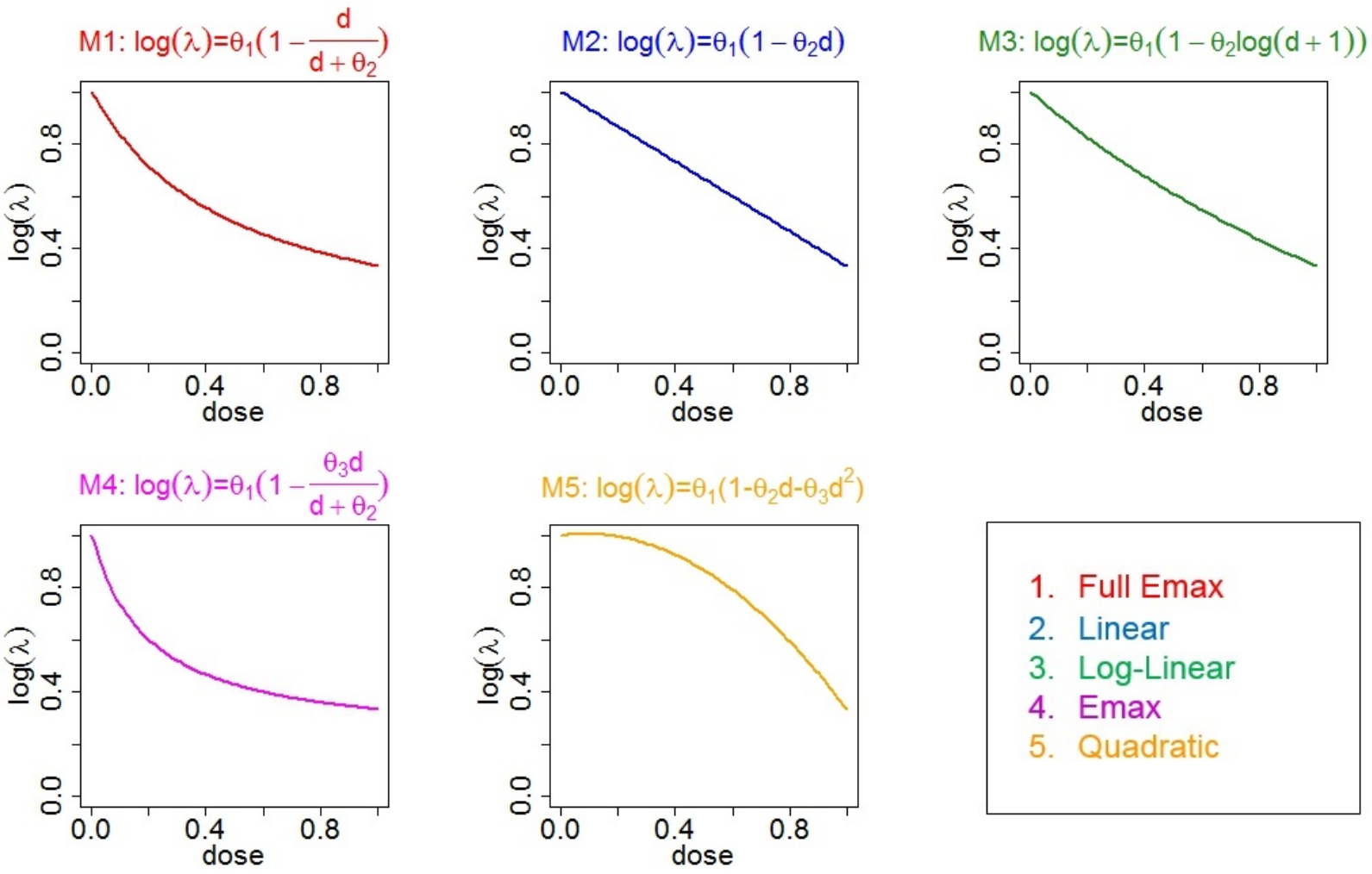
- Each patient observed at 3 dose levels (one placebo) during x days



- Several candidate models for the link between $\log(\lambda)$ and dose
- λ: mean number of events / day

[7] Riviere, Ueckert and Mentré. *Biostatistics*, 2016.

Five models of effect of dose on decreasing Poisson parameter



$$\theta_p = \mu_p exp(b_p);\ b_p \sim \mathcal{N}(0, \omega_p^2)$$

Constraints

	Count example
N	60 subjects
n_{rep}	10 replications
n	3 doses
Choice of doses	$d_1 = 0$ (placebo) d_2, d_3 from 0.1 to 1 ($step = 0.1$, no repetition)

a priori parameter values / parameter distributions:

	ψ_m^*					$p_m(\psi_m)$				
	μ_1^*	μ_2^*	μ_3^*	ω_1^*	ω_2^*	μ_1	$\boldsymbol{\mu_2}$	μ_3	ω_1	$\boldsymbol{\omega_2}$
M_1	1.00	0.50		0.30	0.30	1.00	$\mathcal{LN}(-0.89, 0.63)$		0.30	$\mathcal{LN}(-1.50, 0.77)$
M_2	1.00	0.67		0.30	0.30	1.00	$\mathcal{LN}(-0.60, 0.63)$		0.30	$\mathcal{LN}(-1.50, 0.77)$
M_3	1.00	0.96		0.30	0.30	1.00	$\mathcal{LN}(-0.24, 0.63)$		0.30	$\mathcal{LN}(-1.50, 0.77)$
M_4	1.00	0.20	0.80	0.30	0.30	1.00	$\mathcal{LN}(-1.81, 0.63)$	0.80	0.30	$\mathcal{LN}(-1.50, 0.77)$
M_5	1.00	0.80	0.13	0.30	0.30	1.00	$\mathcal{LN}(-0.60, 0.63)$	0.13	0.30	$\mathcal{LN}(-1.50, 0.77)$

Optimality criteria

Parameters \ Model	Given model m	Set of candidate models $m = 1, \ldots, M$
Given parameter values ψ_m^*	**D-optimality** [8] $\Phi_{D,m}(\Xi) = \det(\mathscr{M}(\Psi_m^*, \Xi))^{1/P_m}$	**Compound-D optimality** [8, 14, 15] $\Phi_{CD}(\Xi) = \prod_{m=1}^{M} \Phi_{D,m}(\Xi)^{\mathrm{w}_m}$
a priori distribution on parameters $p_m(\psi_m)$	**DE-optimality** [8] $\Phi_{DE,m}(\Xi) = \det(\mathscr{M}_R(\Xi))^{1/P_m}$	**Compound-DE optimality** $\Phi_{CDE}(\Xi) = \prod_{m=1}^{M} \Phi_{DE,m}(\Xi)^{\mathrm{w}_m}$

- P_m: number of population parameters of model m
- w_m: weight quantifying the balance between the M models, $\sum_m \mathrm{w}_m = 1$

[8] Atkinson, Donev and Tobias. *Optimum experimental designs, with SAS*, 2009.

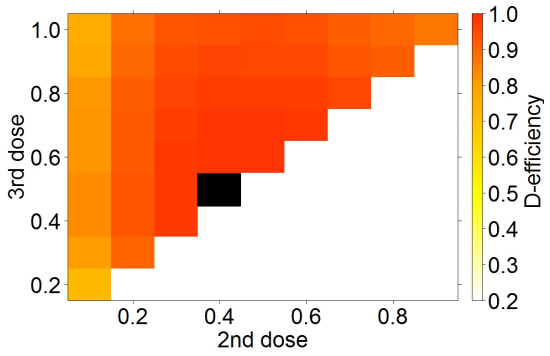
[14] Atkinson et al. *J Stat Plan Inference*, 2008.

[15] Nguyen et al. *Pharm Stat*, 2016.

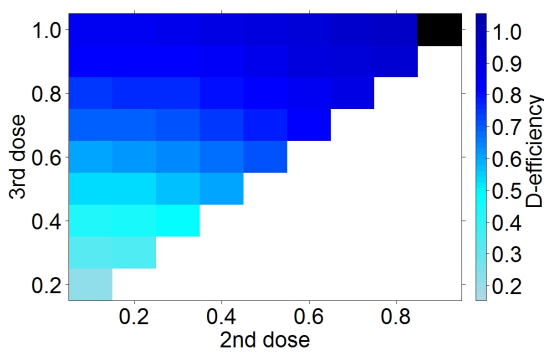
Design optimisation: robustness on model

		Count example
Combinatorial optimisation	Evaluation of FIM for all 45 possible designs	5000 MC 200 HMC
	For each model	D-criterion on FIM
	Over 5 models	Compound D-criterion (averaging for uncertainty on models)

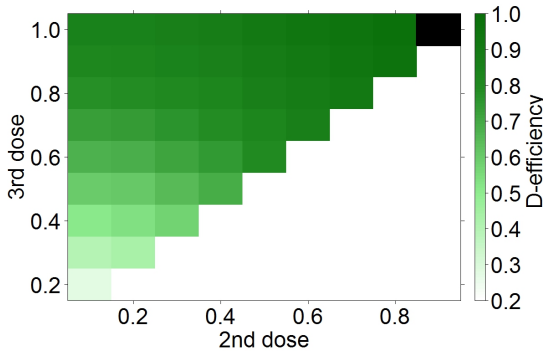
Results: Robust design w.r.t model



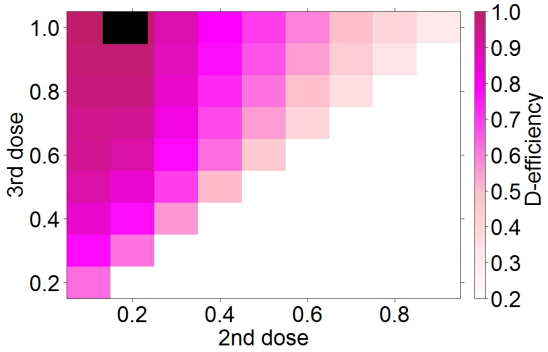
Optimal doses: $\xi_{D,1} = \{0, \mathbf{0.4, 0.5}\}$.



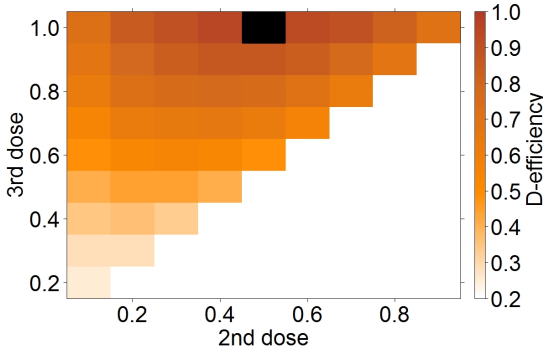
Optimal doses: $\xi_{D,2} = \{0, \mathbf{0.9, 1}\}$.



Optimal doses: $\xi_{D,3} = \{0, \mathbf{0.9, 1}\}$.



Optimal doses: $\xi_{D,4} = \{0, \mathbf{0.2, 1}\}$.



Optimal doses: $\xi_{D,5} = \{0, \mathbf{0.5, 1}\}$.

1. Full Emax
2. Linear
3. Log-Linear
4. Emax
5. Quadratic

$$\text{D-eff}(\Xi) = \frac{\Phi_D(\Xi)}{\Phi_D(\Xi_D)}$$

Results: Robust design w.r.t model

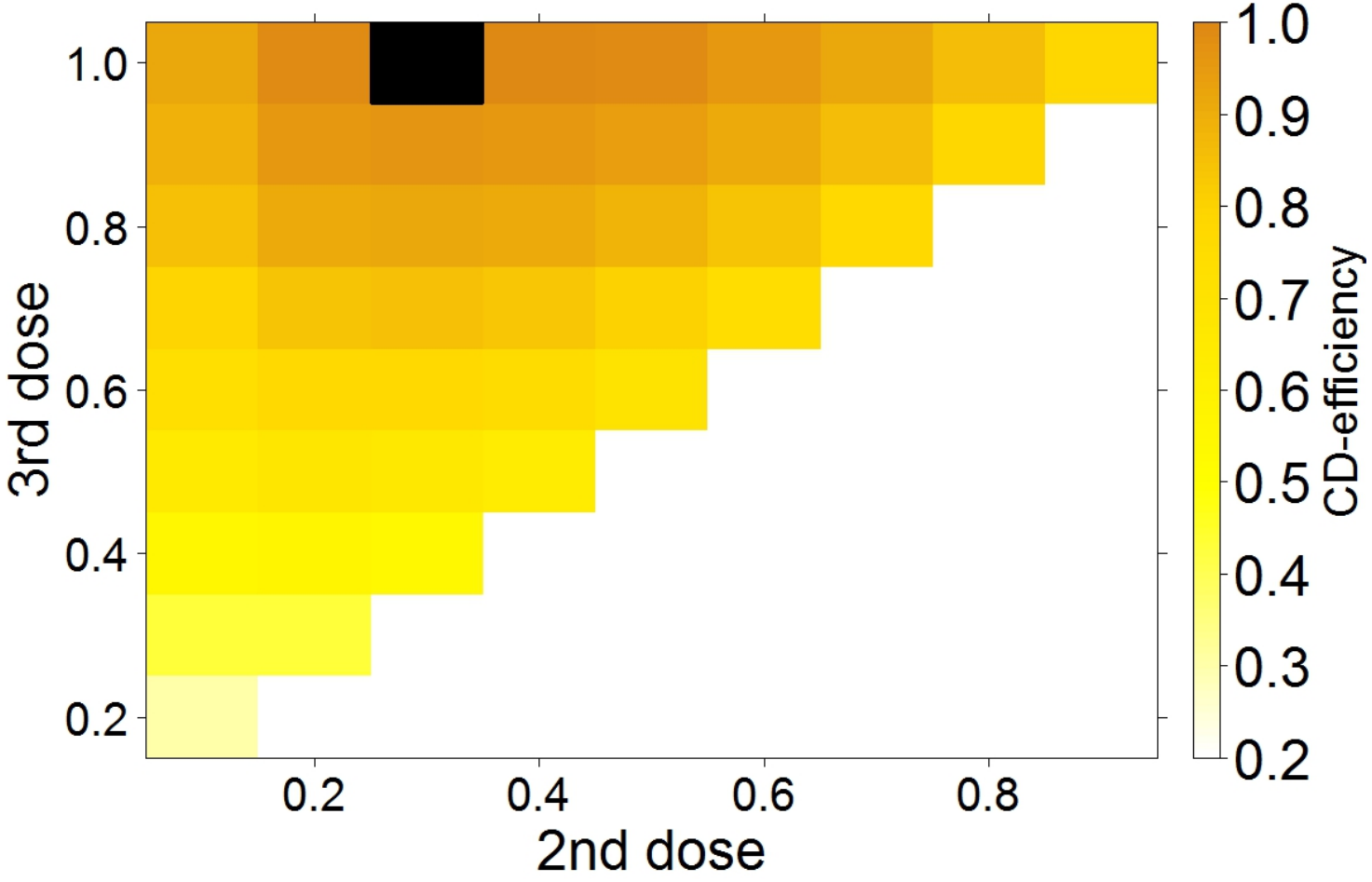
	M_1 Full Emax	M_2 Linear	M_3 Log-linear	M_4 Emax	M_5 Quadratic
$\Xi_{D,1}$ $\{N = 60, \xi = (0, 0.4, 0.5)\}$	100%	60.8%	68.9%	50.3%	27.7%
$\Xi_{D,2}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	87.0%	100%	100%	30.8%	67.2%
$\Xi_{D,3}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	87.0%	100%	100%	30.8%	67.2%
$\Xi_{D,4}$ $\{N = 60, \xi = (0, 0.2, 1)\}$	88.4%	85.7%	85.4%	100%	85.6%
$\Xi_{D,5}$ $\{N = 60, \xi = (0, 0.5, 1)\}$	94.6%	89.9%	91.7%	69.9%	100%

- Important loss of efficiency in some scenarios where the model is not correctly pre-specified

Results: Robust design w.r.t model

Compound D-optimal design: $\xi_{CD} = (0, \mathbf{0.3}, \mathbf{1})$.

$$\text{CD-eff}(\Xi) = \frac{\Phi_{CD}(\Xi)}{\Phi_{CD}(\Xi_{CD})}$$



Results: Robust design w.r.t model

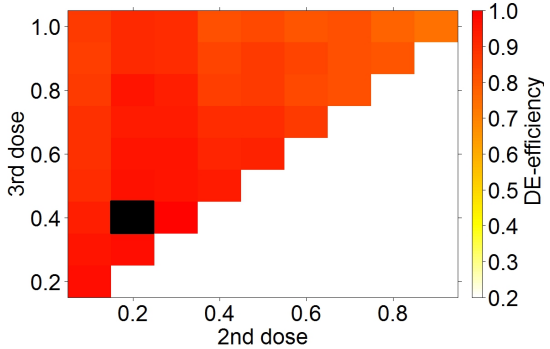
	M_1 Full Emax	M_2 Linear	M_3 Log-linear	M_4 Emax	M_5 Quadratic
$\Xi_{D,1}$ $\{N = 60, \xi = (0, 0.4, 0.5)\}$	100%	60.8%	68.9%	50.3%	27.7%
$\Xi_{D,2}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	87.0%	100%	100%	30.8%	67.2%
$\Xi_{D,3}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	87.0%	100%	100%	30.8%	67.2%
$\Xi_{D,4}$ $\{N = 60, \xi = (0, 0.2, 1)\}$	88.4%	85.7%	85.4%	100%	85.6%
$\Xi_{D,5}$ $\{N = 60, \xi = (0, 0.5, 1)\}$	94.6%	89.9%	91.7%	69.9%	100%
Ξ_{CD} $\{N = 60, \xi = (0, 0.3, 1)\}$	94.1%	88.1%	88.5%	79.7%	93.1%

- Good performance of the compound D-optimal design

Design optimisation: robustness on parameters and model

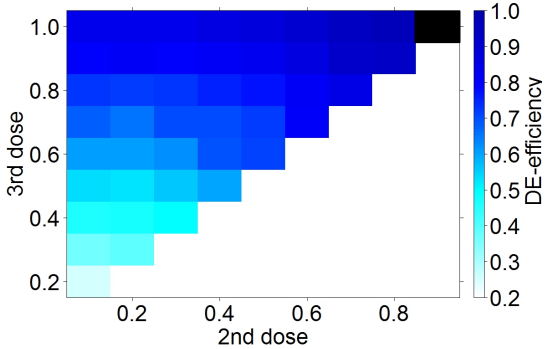
		Count example
Combinatorial optimisation	Evaluation of FIM for all 45 possible designs	5000 MC 200 HMC
	For each model	DE-criterion on FIM
	Over 5 models	Compound DE-criterion (averaging for uncertainty on models)

Results: robust optimal design w.r.t. parameters for each model

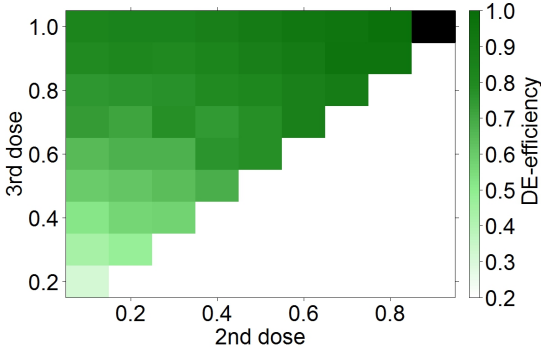


Optimal doses: $\xi_{DE,1}$=(0, **0.2, 0.4**).
$\xi_{D,1}$=(0, **0.4, 0.5**).

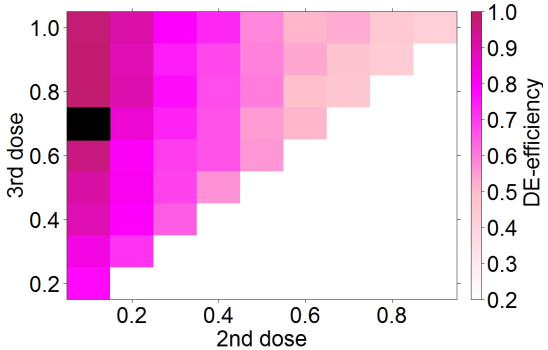
DE-eff($\xi_{D,1}$) = 94.1%



Optimal doses: $\xi_{DE,2}$=(0, **0.9, 1**).
$\xi_{D,2}$=(0, **0.9, 1**).

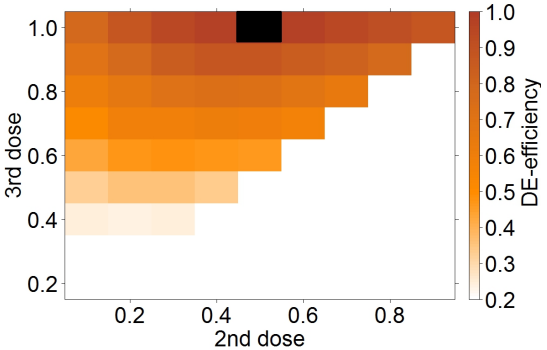


Optimal doses: $\xi_{DE,3}$=(0, **0.9, 1**).
$\xi_{D,3}$=(0, **0.9, 1**).



Optimal doses: $\xi_{DE,4}$ =(0, **0.1, 0.7**).
$\xi_{D,4}$ =(0, **0.2, 1**).

DE-eff($\xi_{D,4}$) = 84.6%



Optimal doses: $\xi_{DE,5}$ =(0, **0.5, 1**).
$\xi_{D,5}$ =(0, **0.5, 1**).

1. Full Emax
2. Linear
3. Log-Linear
4. Emax
5. Quadratic

$$\text{DE-eff}(\Xi) = \frac{\Phi_{DE}(\Xi)}{\Phi_{DE}(\Xi_{DE})}$$

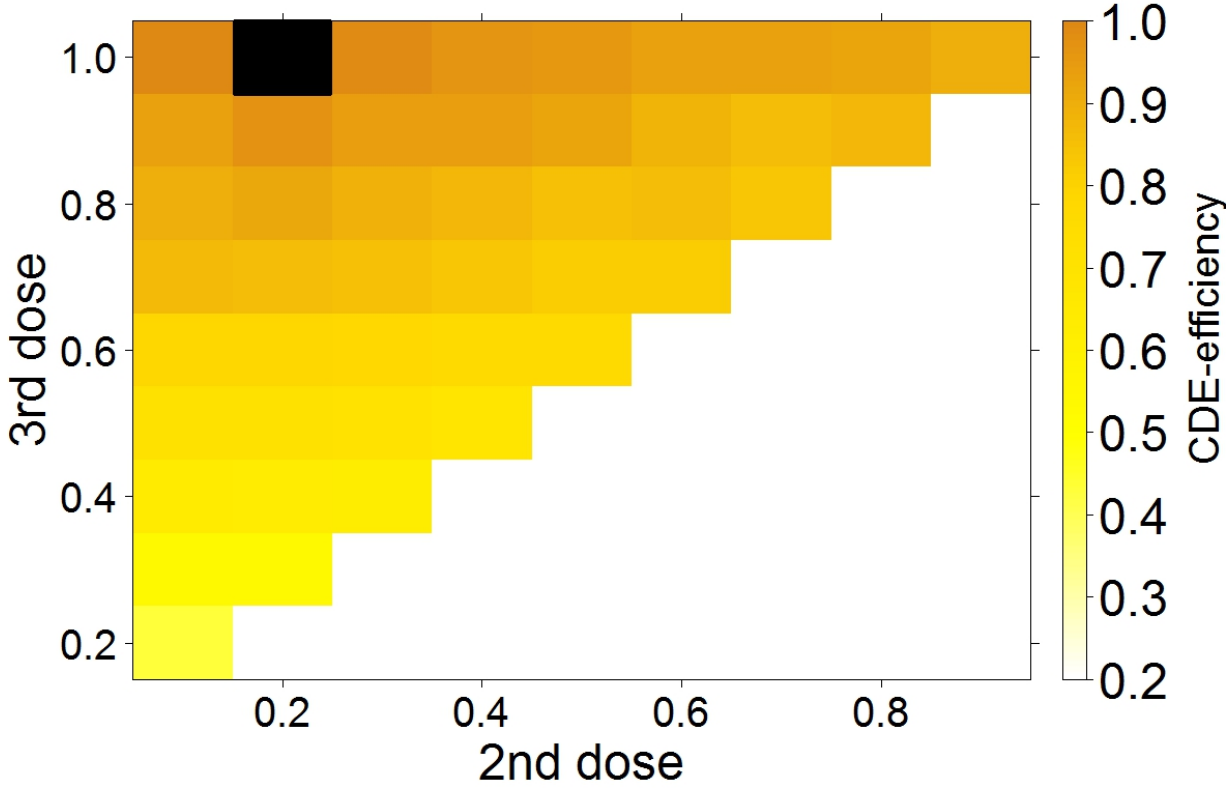
Results: robust optimal design w.r.t. parameters and model

	M_1 Full Emax	M_2 Linear	M_3 Log-linear	M_4 Emax	M_5 Quadratic
$\Xi_{DE,1}$ $\{N = 60, \xi = (0, 0.2, 0.4)\}$	100%	49.9%	56.7%	77.5%	23.6%
$\Xi_{DE,2}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	73.3%	100%	100%	43.5%	87.1%
$\Xi_{DE,3}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	73.3%	100%	100%	43.5%	87.1%
$\Xi_{DE,4}$ $\{N = 60, \xi = (0, 0.1, 0.7)\}$	89.1%	68.1%	73.9%	100%	51.4%
$\Xi_{DE,5}$ $\{N = 60, \xi = (0, 0.5, 1)\}$	83.1%	87.8%	89.6%	58.5%	100%

Results: robust optimal design w.r.t. parameters and model

Compound DE-optimal design: $\xi_{CDE} = (0, \mathbf{0.2}, \mathbf{1})$.
Compound D-optimal design: $\xi_{CD} = (0, \mathbf{0.3}, \mathbf{1})$.

$$\text{CDE-eff}(\Xi) = \frac{\Phi_{CDE}(\Xi)}{\Phi_{CDE}(\Xi_{CDE})}$$



Results: robust optimal design w.r.t. parameters and model

	M_1 Full Emax	M_2 Linear	M_3 Log-linear	M_4 Emax	M_5 Quadratic
$\Xi_{DE,1}$ $\{N = 60, \xi = (0, 0.2, 0.4)\}$	100%	49.9%	56.7%	77.5%	23.6%
$\Xi_{DE,2}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	73.3%	100%	100%	43.5%	87.1%
$\Xi_{DE,3}$ $\{N = 60, \xi = (0, 0.9, 1)\}$	73.3%	100%	100%	43.5%	87.1%
$\Xi_{DE,4}$ $\{N = 60, \xi = (0, 0.1, 0.7)\}$	89.1%	68.1%	73.9%	100%	51.4%
$\Xi_{DE,5}$ $\{N = 60, \xi = (0, 0.5, 1)\}$	83.1%	87.8%	89.6%	58.5%	100%
Ξ_{CDE} $\{N = 60, \xi = (0, 0.2, 1)\}$	90.9%	83.8%	83.9%	84.6%	82.8%

- Efficiency greater than 80% for all models

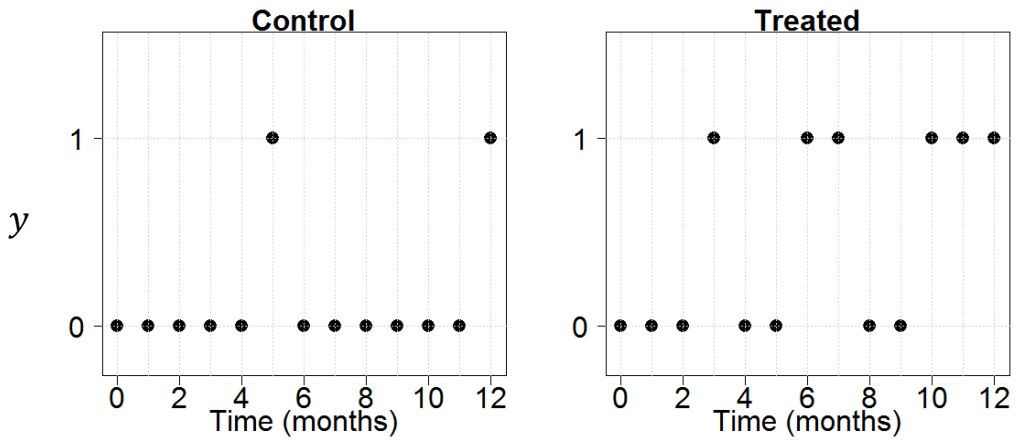
Conclusion

- MC-HMC method for computation of FIM [7] enables applications to design optimisation for count data
- Extension of this method to propose robust optimal designs accounting for uncertainty w.r.t. parameters and/or models

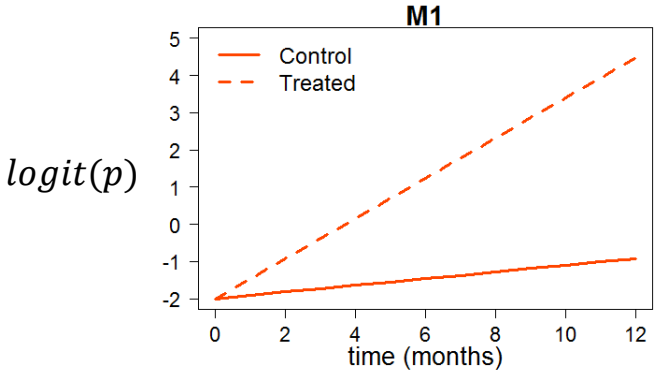
[7] Riviere, Ueckert and Mentré. *Biostatistics*, 2016.

Example

- Repeated binary data
  2 treatment groups : Control *vs.* Treated



- Logistic models to describe the probability p of response y over time



Linear[24]

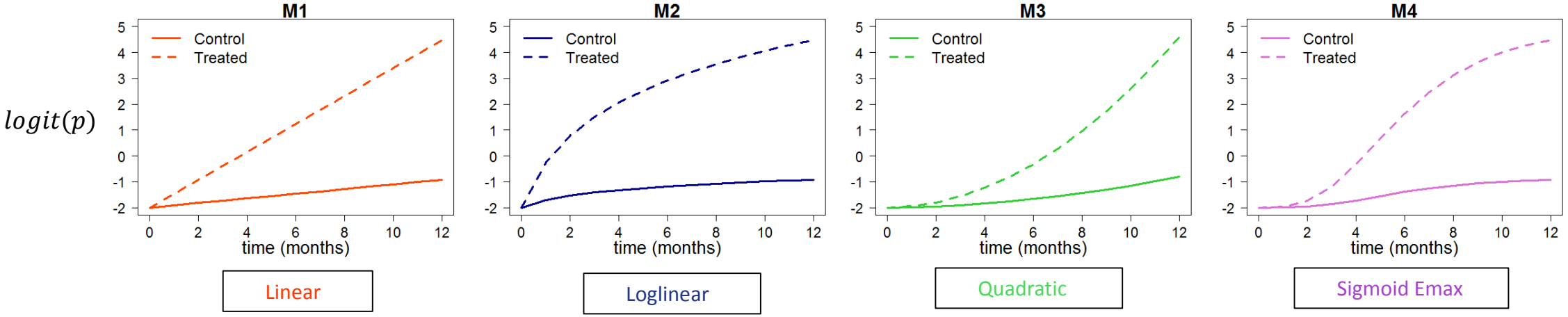
$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t$

24: Lestini G., Ueckert S. and Mentré F., PODE, Sweden, 2016

Example

- Repeated binary data
  2 treatment groups : Control *vs.* Treated

- Logistic models to describe the probability p of response y over time



$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t$

$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)\log(t + 1)$

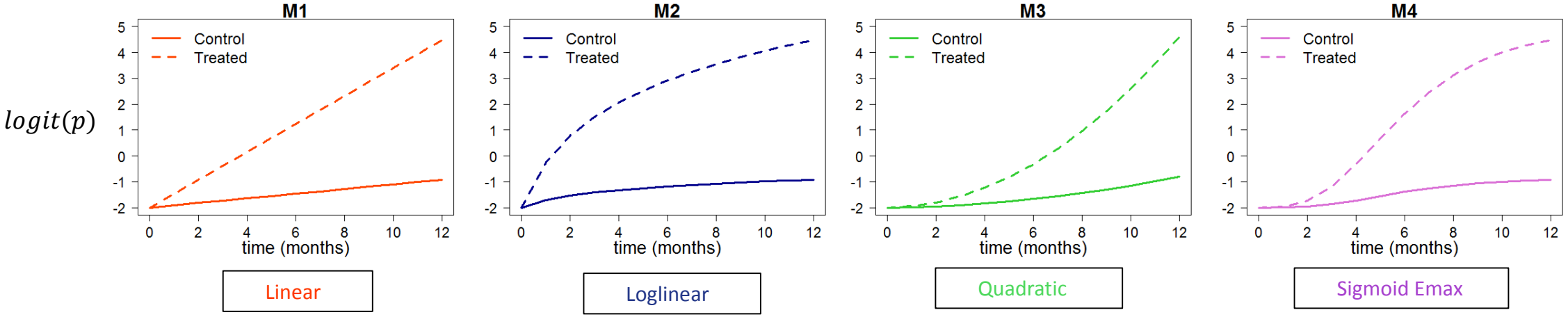
$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t^2$

$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t^3/(\theta_3{}^3 + t^3)$

Example

- Constraints

Number of individuals	N = 100 (50 per treatment group)
Number of samples	n = 4 by individual (from 0 to 12 months)
Sampling times $\xi=(t_1,t_2,t_3,t_4)$	$t_1 = 0$, $t_4 = 12$ months (fixed) **t_2, t_3 optimized among 11 possible times from 1 to 11**



$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t$

$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)\log(t + 1)$

$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t^2$

$\text{logit(p)} = \theta_1 + \theta_2(1 + \beta \times 1_T)t^3/(\theta_3{}^3 + t^3)$

Optimality criteria

Parameters of interest	Given model *m*	Set of candidate models *m* = 1,…,*M*
All the population parameters	D-optimality [8] $\Phi_{D,m} = Det(\mathcal{M}(\psi^*, \Xi))^{\frac{1}{P_m}}$	**CD-optimality** [8,14,15] $\boldsymbol{\Phi_{CD} = \prod_{m=1}^{M} \Phi_{D,m}^{w_m}}$

P_m : number of parameters

$$\sum_{m=1}^{M} w_m = 1$$

$$w_1 = w_2 = w_3 = w_4 = 1/4$$

8: Atkinson AC. Donev A. and Tobias R., Optimum experimental designs, with SAS, 2009
14: Atkinson AC. et *al*., J Stat Plan inference, 2008
15: Nguyen TT. et *al*., Pharm Stat, 2016

Optimality criteria

Parameters of interest	**Given model *m***	**Set of candidate models *m* = 1,…,*M***
All the population parameters	D-optimality [8] $\Phi_{D,m} = Det(\mathcal{M}(\psi^*, \Xi))^{\frac{1}{P_m}}$	**CD-optimality [8,14,15]** $\boldsymbol{\Phi_{CD} = \prod_{m=1}^{M} \Phi_{D,m}{}^{w_m}}$
Subset of parameters of interest (treatment effect)	D_S-optimality [8] $\Phi_{D_S,m} = \left(\frac{Det(\mathcal{M}(\psi^*, \Xi))}{Det(\mathcal{M}_t(\psi^*, \Xi))}\right)^{\frac{1}{S_m}}$	**CD_S-optimality** $\boldsymbol{\Phi_{CD_S} = \prod_{m=1}^{M} \Phi_{D_S,m}{}^{w_m}}$

P_m : number of parameters
S_m : number of parameters of interest
$\mathcal{M}_t$: truncated FIM (without information on parameters of interest)

$$\sum_{m=1}^{M} w_m = 1$$

$$w_1 = w_2 = w_3 = w_4 = 1/4$$

8: Atkinson AC. Donev A. and Tobias R., Optimum experimental designs, with SAS, 2009
14: Atkinson AC. et *al*., J Stat Plan inference, 2008
15: Nguyen TT. et *al*., Pharm Stat, 2016

Optimality criteria

Parameters of interest	Given model *m*	Set of candidate models $m = 1,\ldots,M$
All the population parameters	D-optimality [8] $\Phi_{D,m} = Det(\mathcal{M}(\psi^*,\Xi))^{\frac{1}{P_m}}$	**CD-optimality** [8,14,15] $\Phi_{CD} = \prod_{m=1}^{M} \Phi_{D,m}^{w_m}$
Subset of parameters of interest (treatment effect)	D_S-optimality [8] $\Phi_{D_S,m} = \left(\frac{Det(\mathcal{M}(\psi^*,\Xi))}{Det(\mathcal{M}_t(\psi^*,\Xi))}\right)^{\frac{1}{S_m}}$	**CD_S-optimality** $\Phi_{CD_S} = \prod_{m=1}^{M} \Phi_{D_S,m}^{w_m}$
Compromise	DD_S-optimality [8,16] $\Phi_{DD_S,m} = \left(Det(\mathcal{M}_t(\psi^*,\Xi))\right)^{\frac{1-\alpha_m}{P_m - S_m}} \left(\frac{Det(\mathcal{M}(\psi^*,\Xi))}{Det(\mathcal{M}_t(\psi^*,\Xi))}\right)^{\frac{\alpha_m}{S_m}}$	**CDD_S-optimality** $\Phi_{CDD_S} = \prod_{m=1}^{M} \Phi_{DD_S,m}^{w_m}$

P_m : number of parameters
S_m : number of parameters of interest
$\mathcal{M}_t$: truncated FIM (without information on parameters of interest)
α_m: interest for S, $0 \leq \alpha \leq 1$

$$\sum_{m=1}^{M} w_m = 1$$

$$w_1 = w_2 = w_3 = w_4 = 1/4$$

8: Atkinson AC. Donev A. and Tobias R., Optimum experimental designs, with SAS, 2009
14: Atkinson AC. et *al*., J Stat Plan inference, 2008
15: Nguyen TT. et *al*., Pharm Stat, 2016
16: Atkinson AC. and Bogacka B., Technometrics, 1997

Design optimization accounting for model uncertainty

		Binary example
Combinatorial optimisation	Evaluation of FIM by MC/HMC [7] for all 55 possible designs ξ	5000 MC 200 HMC
	For each model	D-, D_S- and DD_S-criteria
	Averaging over 4 models	Compound D-, D_S- and DD_S-criteria

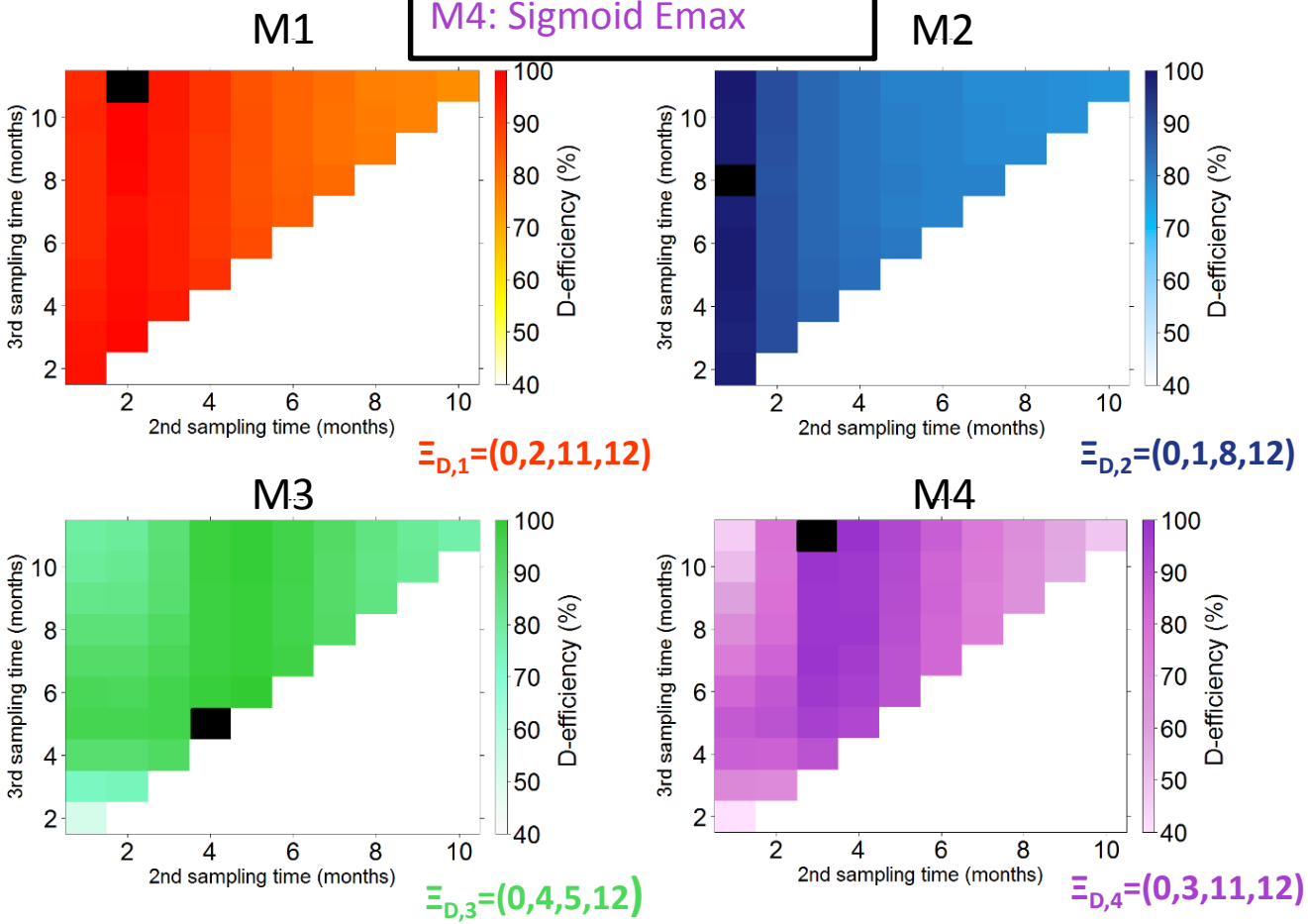
- Choice of α_m in $\Phi_{DD_S,m}$ for each model
  - $\Phi_{DD_S,m}$ computed for values of α between 0 and 1 with a step of 0.05
  - α_m chosen to maximize [D-eff × D_S-eff]
- Predictions from FIM
  - Relative standard errors (RSE) for population parameters of each model
  - Power of the Wald test to detect a significant effect β (with each model [17] and average)
  - Number of subjects needed (NSN) to obtain a $\pi_{average}$ of 80%

$$\pi_{average} = \sum_{m=1}^{M} w_m \pi_m$$

7: Riviere M-K. Ueckert S. and Mentré F., Biostatistics, 2016
17: Retout S. *et al.*, Stat Med, 2007

Results: D-efficiencies for each model

M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax



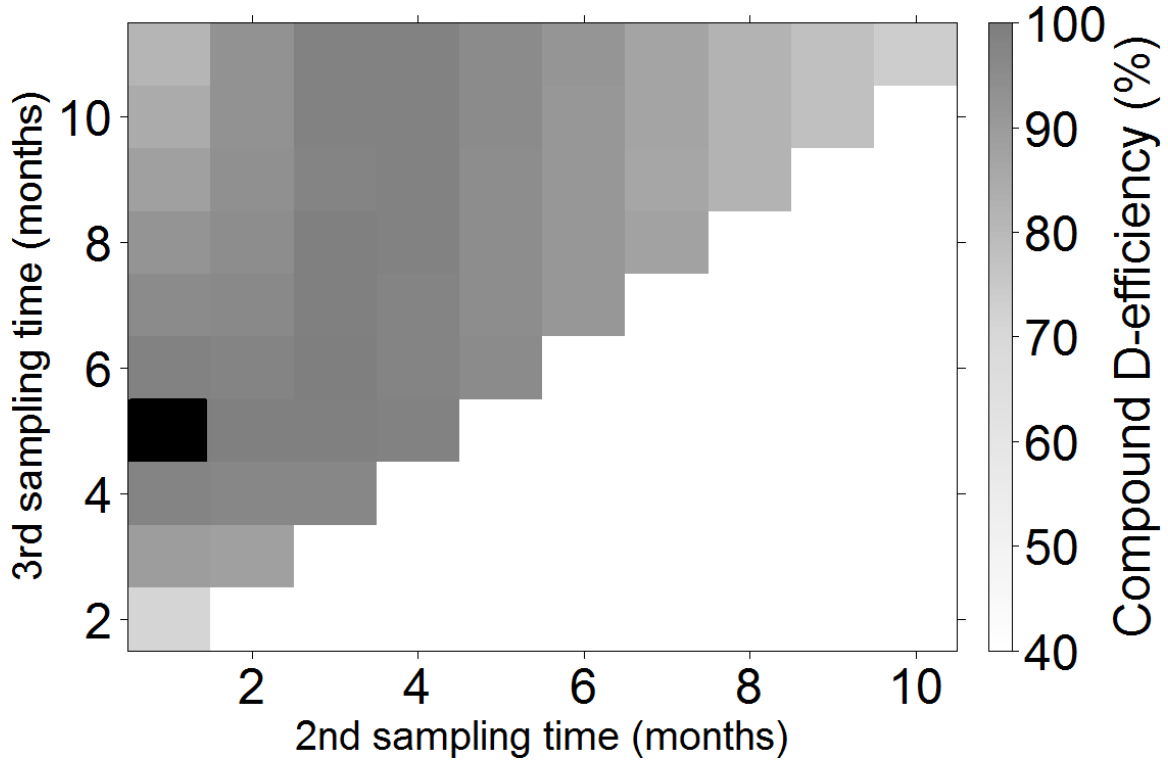
$$\text{D-eff}(\Xi) = \frac{\Phi_D(\Xi)}{\Phi_D(\Xi_D)}$$

Ξ_D : D-optimal design

D-efficiency

D-optimal design for a pre-specified model \ Final model	M1 Linear	M2 Loglinear	M3 Quadratic	M4 Sigmoid Emax
$\Xi_{D,1}$=(0,2,11,12)	100%	89.8%	81.2%	79.1%
$\Xi_{D,2}$=(0,1,8,12)	93.2%	100%	87.5%	68.0%
$\Xi_{D,3}$=(0,4,5,12)	91.6%	83.4%	100%	93.1%
$\Xi_{D,4}$=(0,3,11,12)	96.3%	85.2%	88.4%	100%

CD-efficiencies



Ξ_{CD}=(0,1,5,12)

D-efficiency

Final model / D-optimal design for a pre-specified model	M1 Linear	M2 Loglinear	M3 Quadratic	M4 Sigmoid Emax
$\Xi_{D,1}$=(0,2,11,12)	100%	89.8%	81.2%	79.1%
$\Xi_{D,2}$=(0,1,8,12)	93.2%	100%	87.5%	68.0%
$\Xi_{D,3}$=(0,4,5,12)	91.6%	83.4%	100%	93.1%
$\Xi_{D,4}$=(0,3,11,12)	96.3%	85.2%	88.4%	100%
Ξ_{CD}=(0,1,5,12)	**94.0%**	**98.9%**	**95.1%**	**86.9%**

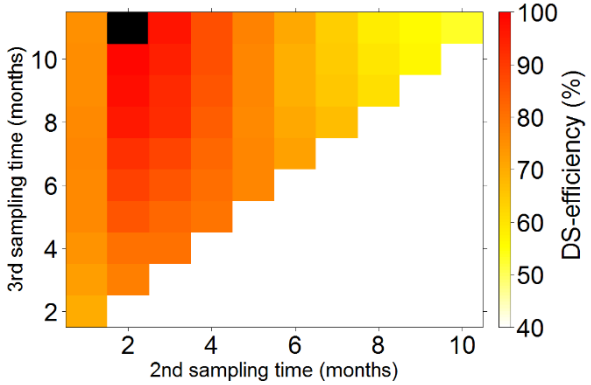
$$\text{CD-eff}(\Xi) = \frac{\Phi_{CD}(\Xi)}{\Phi_{CD}(\Xi_{CD})}$$

Ξ_D : CD-optimal design

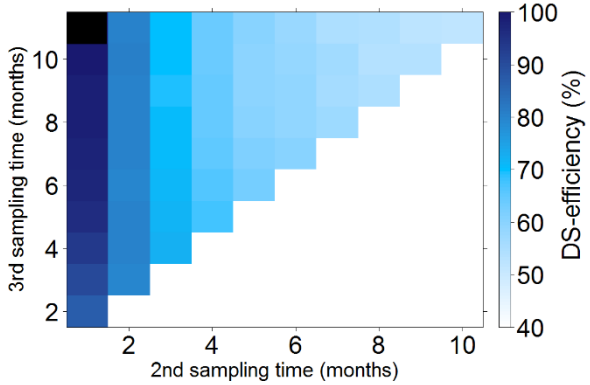
D_S-efficiencies for each model

M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax

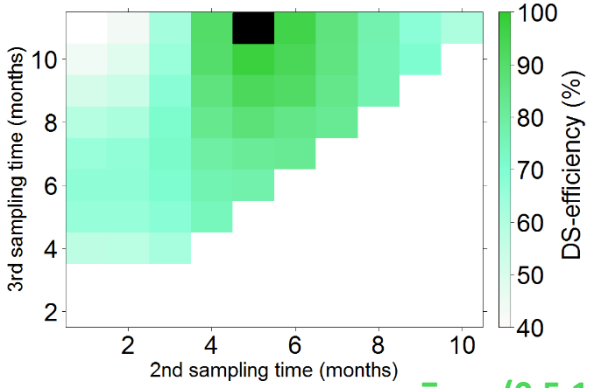
M1



M2

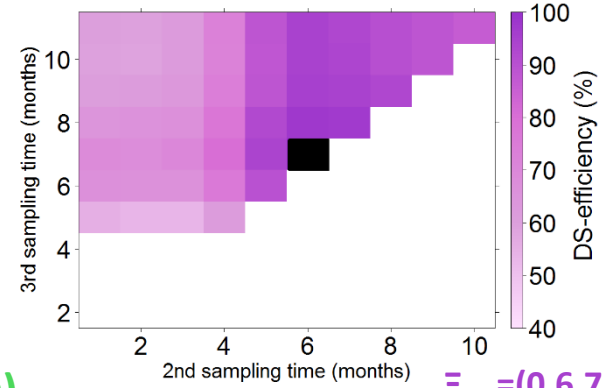


M3 $\Xi_{Ds,1}$=(0,2,11,12)



$\Xi_{Ds,3}$=(0,5,11,12)

M4 $\Xi_{Ds,2}$=(0,1,11,12)



$\Xi_{D,4}$=(0,6,7,12)

$$D_S\text{-eff}(\Xi) = \frac{\Phi_{Ds}(\Xi)}{\Phi_{Ds}(\Xi_{Ds})}$$

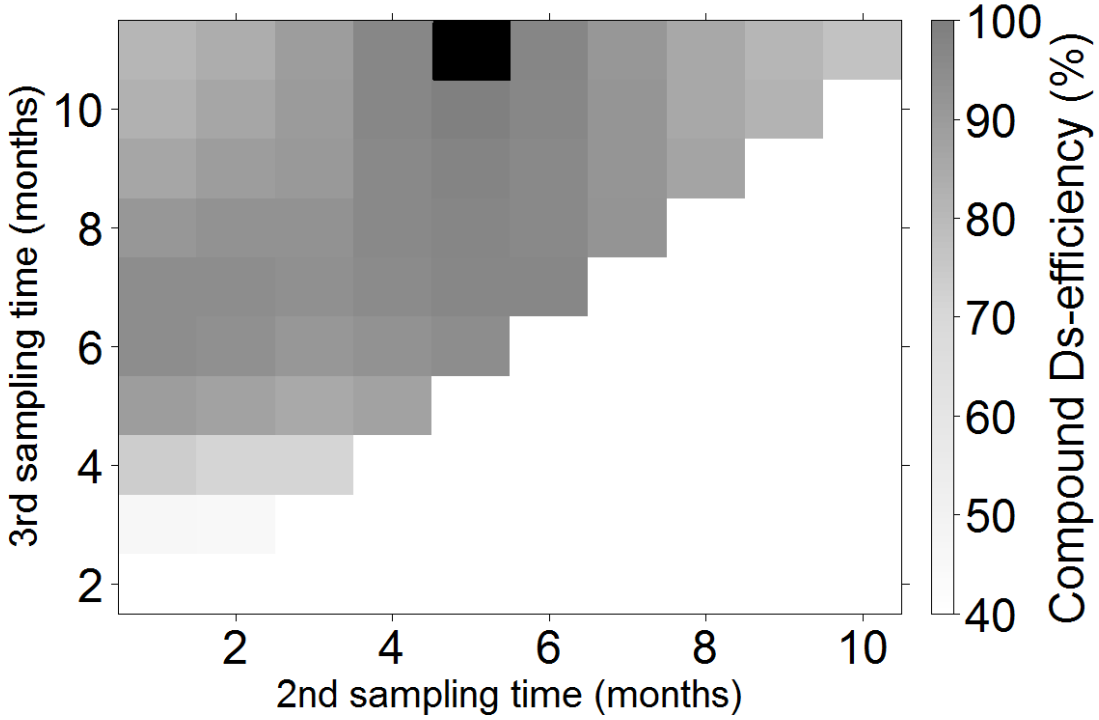
Ξ_{Ds} : D_S-optimal design

D_S-efficiency

Final model / D_S-optimal design for a pre-specified model	M1 Linear	M2 Loglinear	M3 Quadratic	M4 Sigmoid Emax
$\Xi_{Ds,1}$=(0,2,11,12)	100%	80.6%	**43.4%**	59.5%
$\Xi_{Ds,2}$=(0,1,11,12)	75.1%	100%	**38.9%**	60.0%
$\Xi_{Ds,3}$=(0,5,11,12)	77.4%	60.0%	100%	89%
$\Xi_{Ds,4}$=(0,6,7,12)	72.0%	60.2%	82.3%	100%

$\Xi_{D,1}$=(0,2,11,12)
$\Xi_{D,2}$=(0,1,8,12)
$\Xi_{D,3}$=(0,4,5,12)
$\Xi_{D,4}$=(0,3,11,12)

CD_S-efficiencies



Ξ_{CDs}=(0,5,11,12)

D_S-efficiency

Final model / D_S-optimal design for a pre-specified model	M1 Linear	M2 Loglinear	M3 Quadratic	M4 Sigmoid Emax
$\Xi_{Ds,1}$=(0,2,11,12)	100%	80.6%	**43.4%**	59.5%
$\Xi_{Ds,2}$=(0,1,11,12)	75.1%	100%	**38.9%**	60.0%
$\Xi_{Ds,3}$=(0,5,11,12)	77.4%	60.0%	100%	89%
$\Xi_{Ds,4}$=(0,6,7,12)	72.0%	60.2%	82.3%	100%
Ξ_{CDs}=(0,5,11,12)	**77.4%**	**63.7%**	**100%**	**89%**

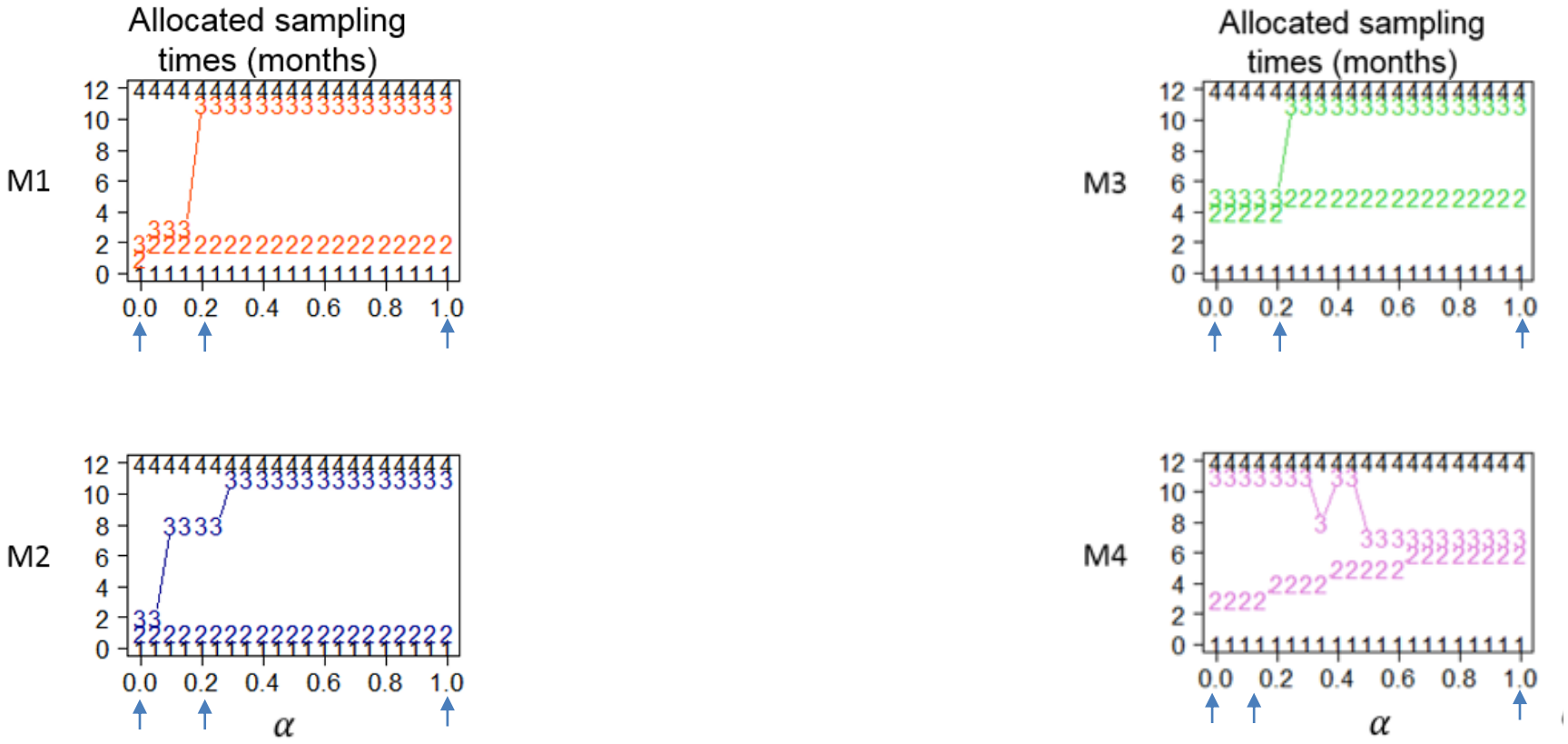
Ξ_{CD}=(0,1,5,12)

$$\text{CDs-eff}(\Xi) = \frac{\Phi_{CDs}(\Xi)}{\Phi_{CDs}(\Xi_{CDs})}$$

Ξ_{CDs} : CDs-optimal design

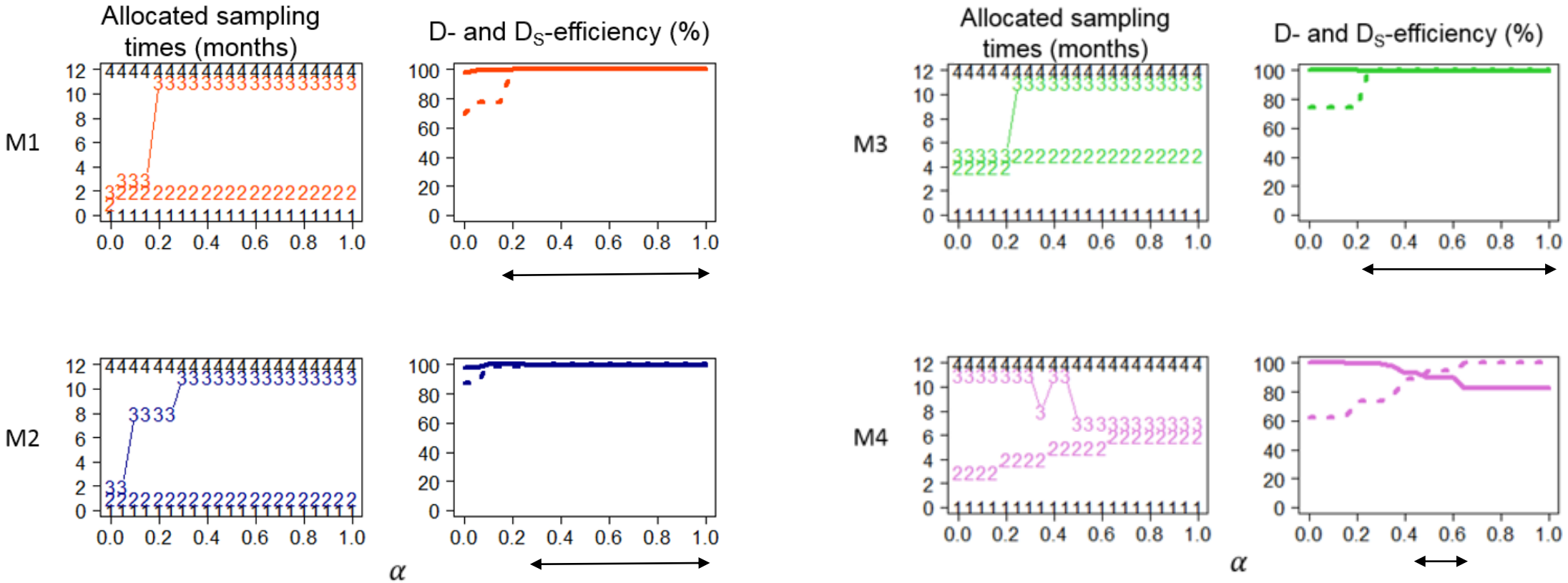
Choice of α_m for optimal $DD_{S,m}$ design

M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax



Choice of α_m for optimal $DD_{S,m}$ design

M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax

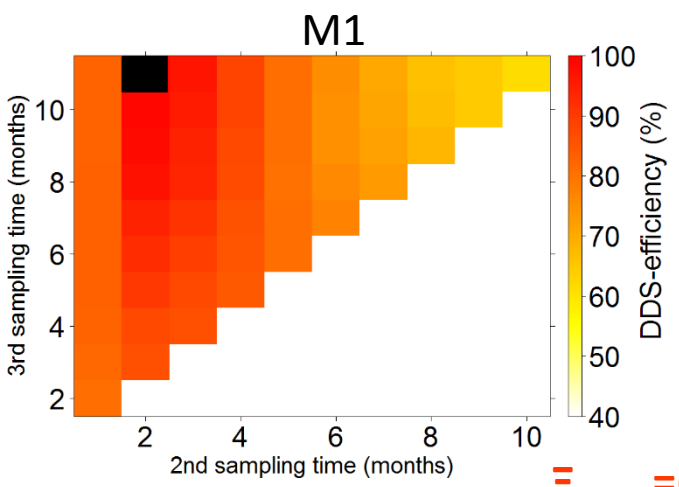


↔ α_m values maximizing [D-eff × D_S−eff]

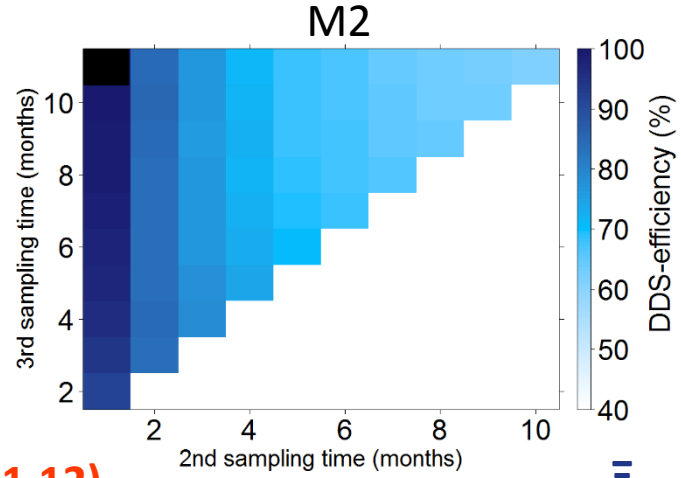
➡ DD_S- and CDD_S-optimality computed with α = 0.6

DD_S-efficiencies for each model (α = 0.6)

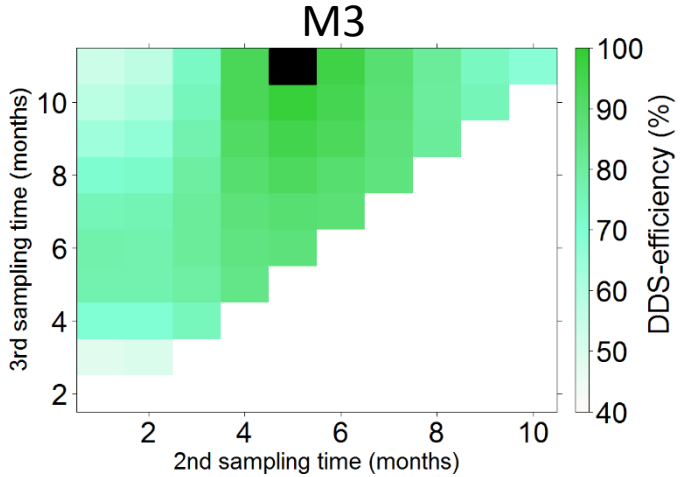
M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax



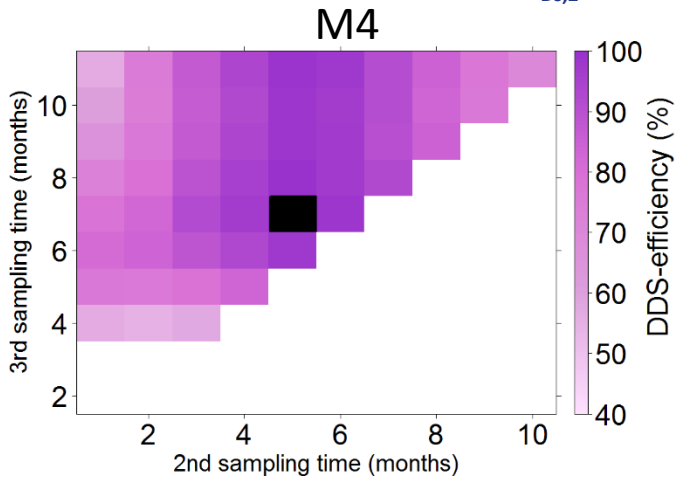
$\Xi_{DDs,1}$=(0,2,11,12)
$\Xi_{D,1}$=(0,2,11,12), DD_S−eff($\Xi_{D,1}$) = 100%
$\Xi_{Ds,1}$=(0,2,11,12), DD_S−eff($\Xi_{Ds,1}$) = 100%



$\Xi_{DDs,2}$=(0,1,11,12)
$\Xi_{D,2}$=(0,1,8,12), DD_S−eff($\Xi_{D,2}$)= 99.3%
$\Xi_{Ds,2}$=(0,1,11,12), DD_S−eff($\Xi_{Ds,2}$) = 100%



$\Xi_{DDs,3}$=(0,5,11,12)
$\Xi_{D,3}$=(0,4,5,12), DD_S−eff($\Xi_{D,3}$)= 84.7%
$\Xi_{Ds,3}$=(0,5,11,12), DD_S−eff($\Xi_{Ds,3}$) = 100%

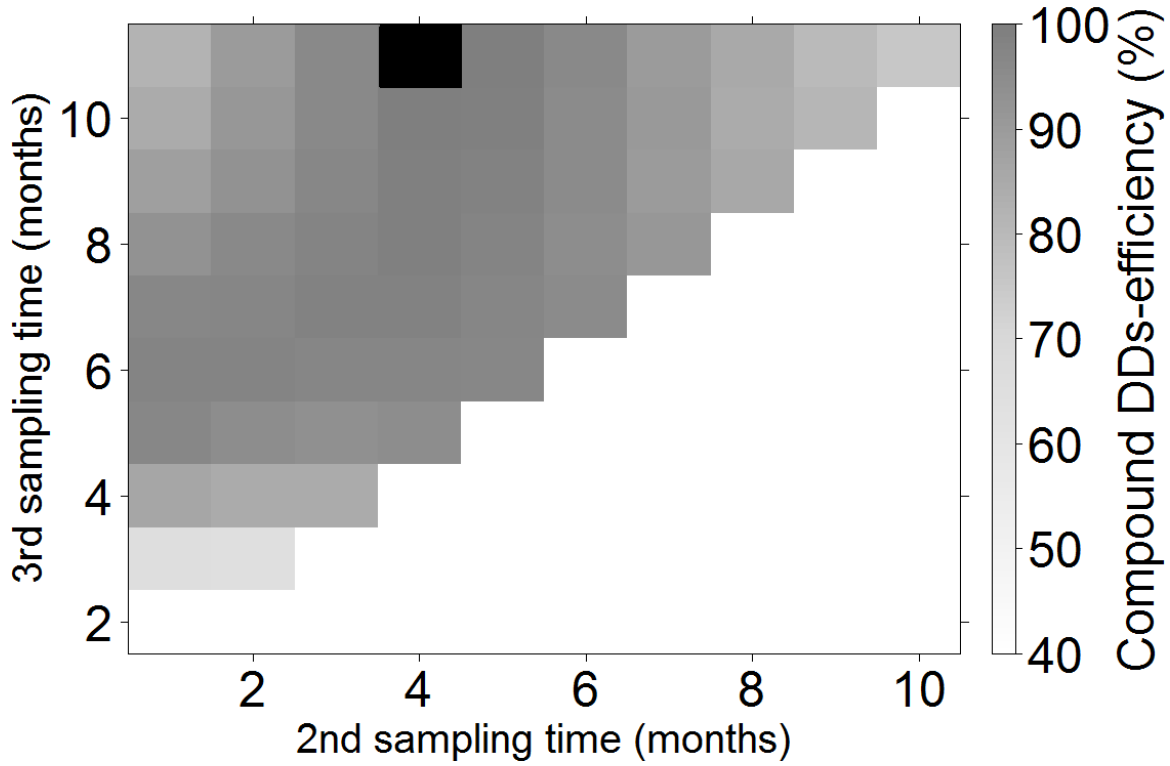


$\Xi_{DDs,4}$=(0,5,7,12)
$\Xi_{D,4}$=(0,3,11,12), DD_S−eff($\Xi_{D,4}$)= 86.8%
$\Xi_{Ds,4}$=(0,6,7,12), DD_S−eff($\Xi_{Ds,4}$) = 98.5%

$$DD_S\text{−eff}(\Xi) = \frac{\Phi_{DDs}(\Xi)}{\Phi_{DDs}(\Xi_{DDs})}$$

Ξ_{DDs} : DDs-optimal design

CDD_S-efficiencies (α = 0.6)



Ξ_{CDDs}=(0,4,11,12)

Final model / DDs-optimal design for a pre-specified model	M1 Linear	M2 Loglinear	M3 Quadratic	M4 Sigmoid Emax
$\Xi_{DDs,1}$=(0,2,11,12)	100%	84.5%	**57.2%**	75.6%
$\Xi_{DDs,2}$=(0,1,11,12)	82.5%	100%	**53.4%**	**56.9%**
$\Xi_{DDs,3}$=(0,5,11,12)	81.0%	68.3%	100%	99.4%
$\Xi_{DDs,4}$=(0,5,7,12)	80.7%	69.2%	88.8%	100%
Ξ_{CDDs}=(0,4,11,12)	88.4%	71.5%	93.9%	93.7%

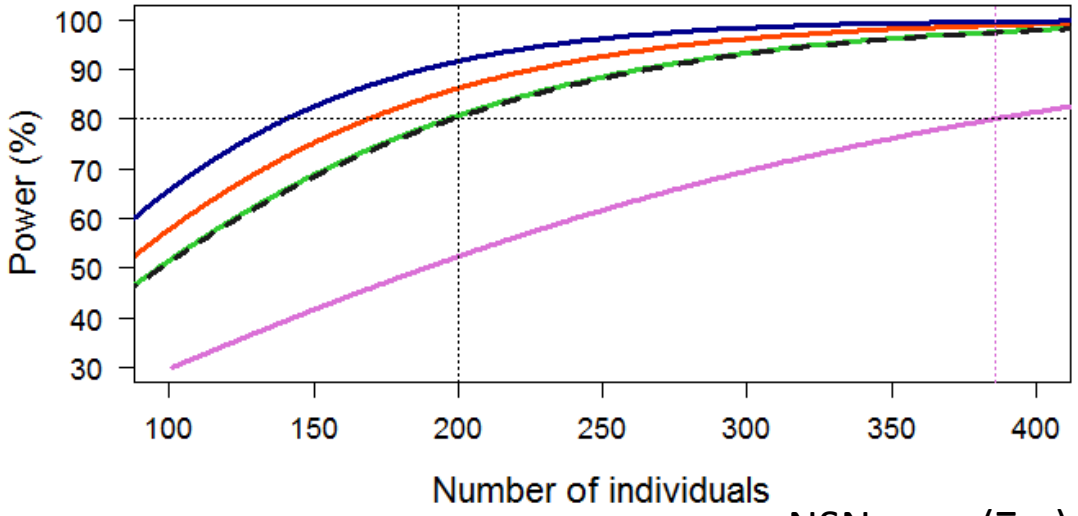
$$CDD_S\text{-eff}(\Xi) = \frac{\Phi_{CDDs}(\Xi)}{\Phi_{CDDs}(\Xi_{CDDs})}$$

Ξ_{CDDs} : CDDs-optimal design

Number of subjects needed to reach an average power of 80 %

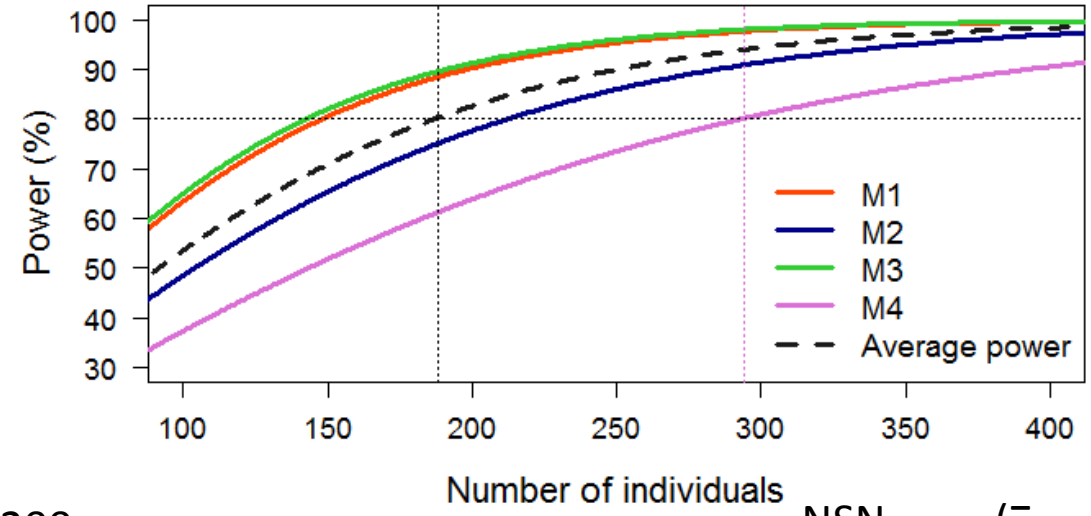
M1: Linear
M2: Loglinear
M3: Quadratic
M4: Sigmoid Emax

Ξ_{CD}=(0,1,5,12)



$NSN_{average}$ (Ξ_{CD}) = 200

Ξ_{CDDs}=(0,4,11,12)



$NSN_{average}$ (Ξ_{CDDs}) = 188

Conclusion

- Robust designs accounting for model uncertainty
- Balance between the parameters precision and the power of the Wald test
- Next steps:
  - To take into account the parameters uncertainty
  - To perform clinical trials simulations

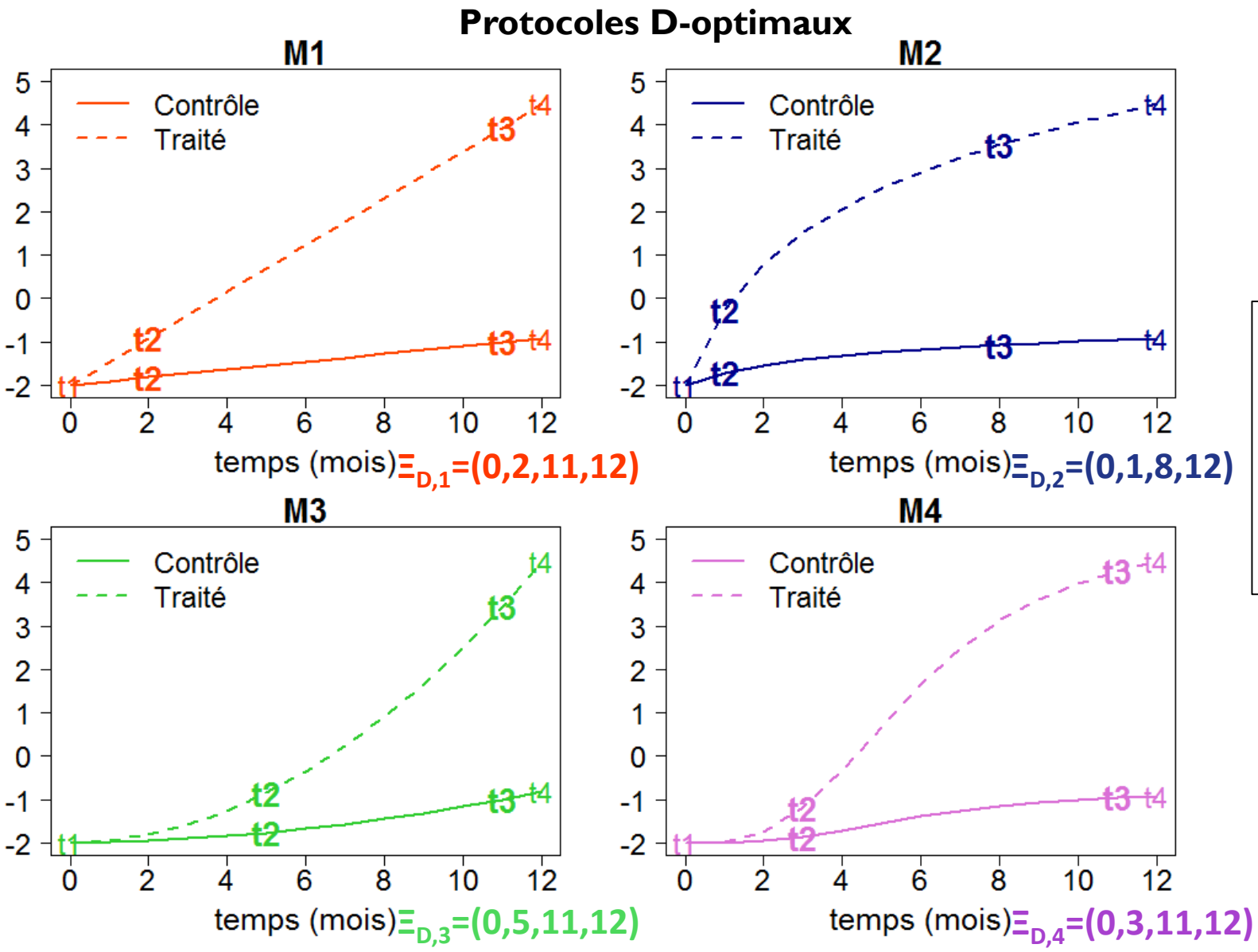
Overall conclusion

- **Discussion**
  - MC-HMC method for computation of FIM [7] enables applications to design optimization for discrete data
  - Extension of this method to propose robust designs accounting for uncertainty w.r.t. parameters and/or models / balance between the parameters precision and the power of the Wald test
  - Computationally challenging, much slower than FO approach
- **Perspectives**
  - Implementation of an optimization algorithm: sparse grids – particle swarm optimization [18]
  - Replacement of MC by more efficient approach: quasi-random sampling [19]
  - Application to other type of data (continuous, time to event)
  - Use in model based adaptive design [20-22]/ model averaging [23]

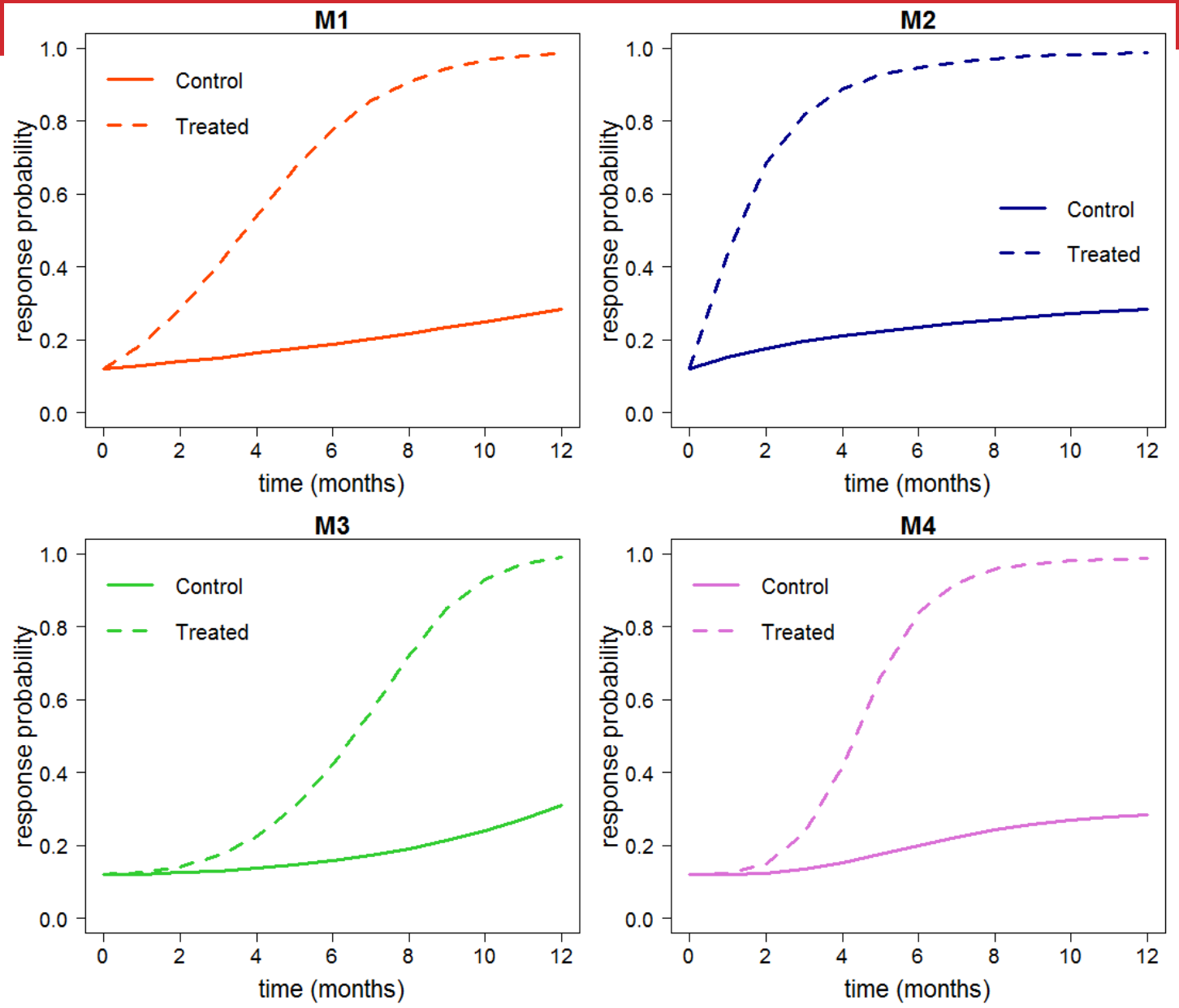
7: Riviere M-K. Ueckert S. and Mentré F., Biostatistics, 2016
18: Shi Y., Zhang Z. and Wong W.K., BIRS Design Workshop, Canada, 2017
19: Ueckert S. and Mentré F., CM Statistics Conference, UK, 2015
20: Lestini G., Dumont C. and Mentré F., Pharm Res, 2015
21: Dumont C. *et al.*, Commun Stat Simul Comput, 2016
22: Strömberg E.A. and Hooker A.C., JPKPD, 2017
23: Buckland S.T., Burnham K.P. and Augustin N.H., Biometrics, 1997

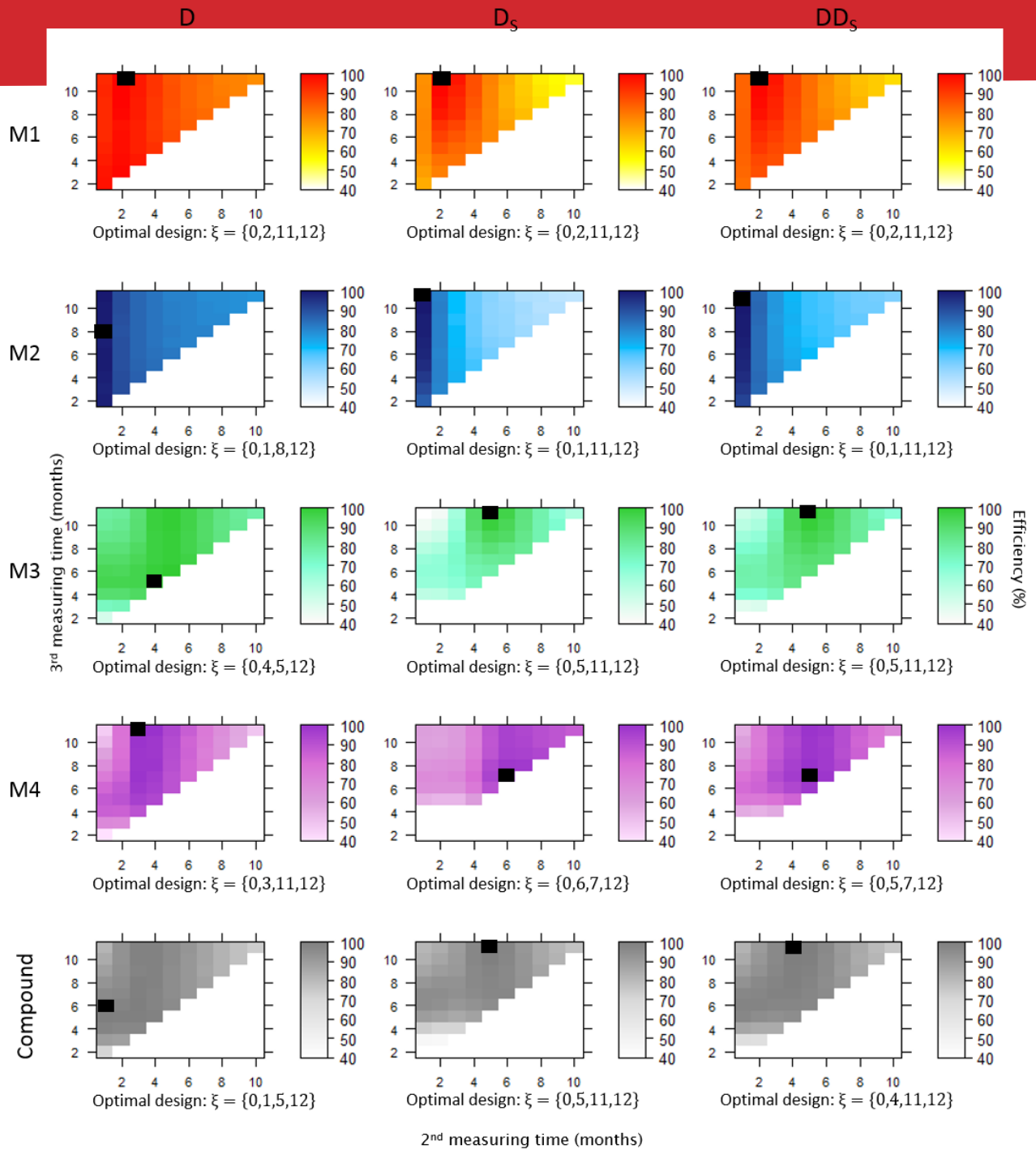
Thank you for your attention

D-optimal times



1. Linear
2. Loglinear
3. Quadratic
4. Sigmoid emax


M1
M2
M3
M4
Control
Treated
response probability
time (months)
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12


D
D_S
DD_S
M1
M2
M3
M4
Compound
Optimal design: ξ = {0,2,11,12}
Optimal design: ξ = {0,2,11,12}
Optimal design: ξ = {0,2,11,12}
Optimal design: ξ = {0,1,8,12}
Optimal design: ξ = {0,1,11,12}
Optimal design: ξ = {0,1,11,12}
Optimal design: ξ = {0,4,5,12}
Optimal design: ξ = {0,5,11,12}
Optimal design: ξ = {0,5,11,12}
Optimal design: ξ = {0,3,11,12}
Optimal design: ξ = {0,6,7,12}
Optimal design: ξ = {0,5,7,12}
Optimal design: ξ = {0,1,5,12}
Optimal design: ξ = {0,5,11,12}
Optimal design: ξ = {0,4,11,12}
3rd measuring time (months)
2nd measuring time (months)
Efficiency (%)

Efficacité des protocoles optimaux selon le modèle final

Protocole robuste CDD_S-optimal

D-efficacité

Modèle final / Protocole D-optimal du modèle *a priori*	M1 Linéaire	M2 Loglinéaire	M3 Quadratique	M4 Emax sigmoïde
$\xi_{D,1}$=(0,2,11,12)	100 %	89,8 %	81,2 %	79,1 %
$\xi_{D,2}$=(0,1,8,12)	93,2 %	100 %	87,5 %	68,0 %
$\xi_{D,3}$=(0,4,5,12)	91,6 %	83,4 %	100 %	93,1 %
$\xi_{D,4}$=(0,3,11,12)	96,3 %	85,2 %	88,4 %	100 %
ξ_{CD}=(0,1,5,12)	**94,0 %**	**98,9 %**	**95,1 %**	**86,9 %**
ξ_{CDDs}=(0,4,11,12)	**91,0 %**	**82,8 %**	**97,7 %**	**99,5 %**

NSN pour une puissance de 80 % du test de Wald

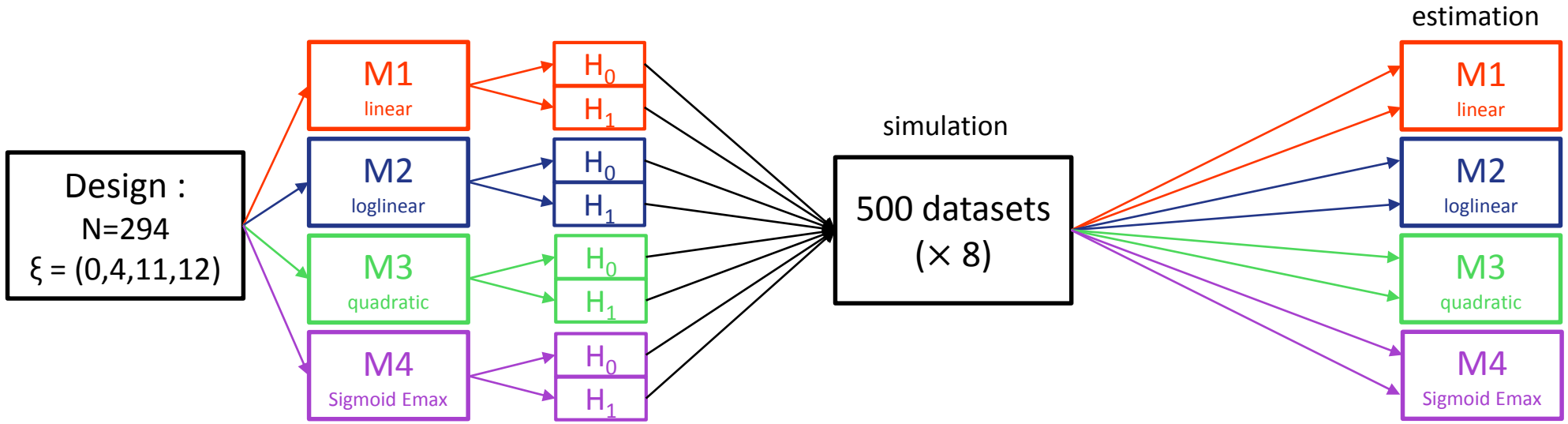
ξ_{CD}=(0,1,5,12)	**170**	**142**	**198**	**386**
ξ_{CDDs}=(0,4,11,12)	**150**	**214**	**144**	**294**

Simulations d'essais cliniques avec le protocole CDD_S-optimal ξ_{CDDs}=(0,4,11,12) et 294 patients

D-efficacité : rapport de la D-optimalité d'un protocole ξ sur celle du protocole D-optimal $\xi_{D,m}$ du modèle m

Clinical trial simulations

- Clinical trial simulations (CTS) : For each candidate model, 500 datasets with a treatment effect (under H_1) and 500 without treatment effect (under H_0)



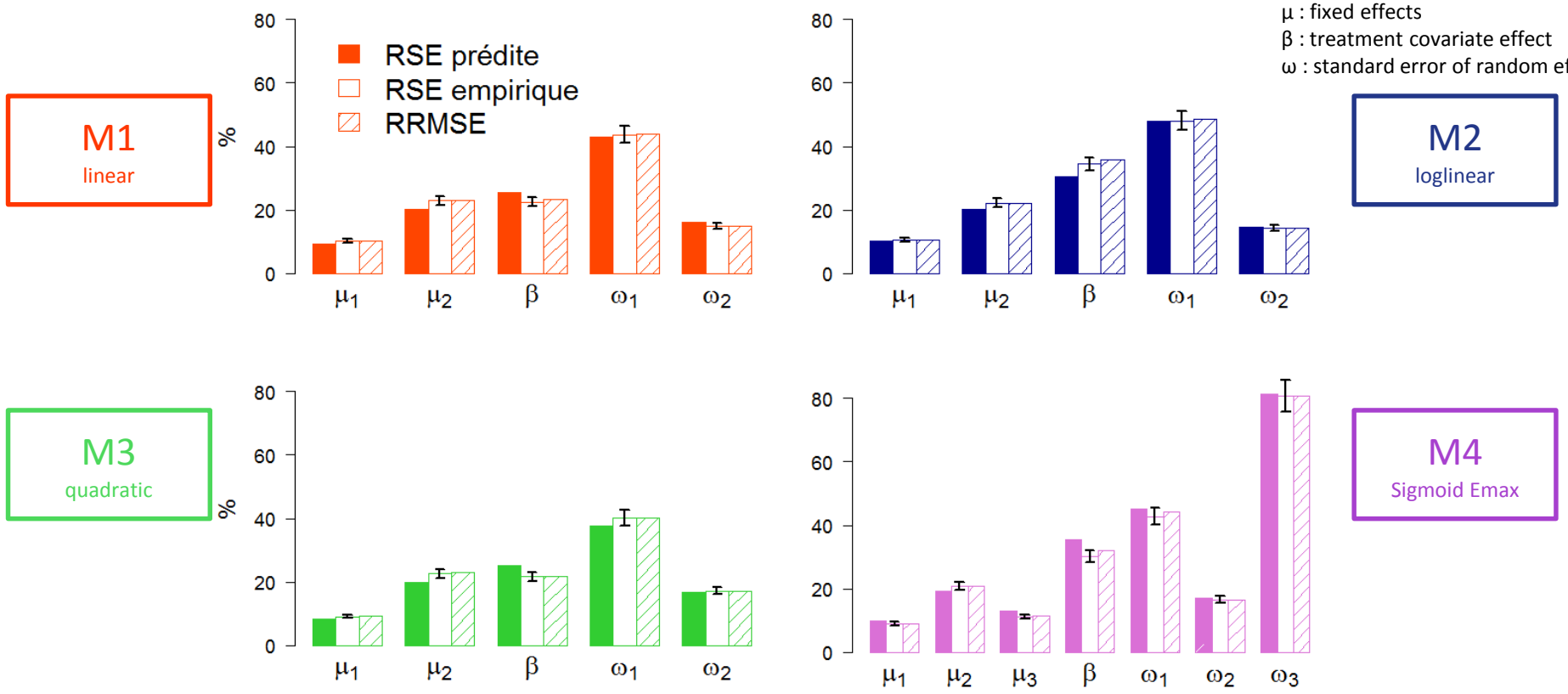
- Estimation of the parameters of each candidate model, by maximum likelihood (SAEM algorithm of MONOLIX 2016R1)
- Comparison between predictions by FIM and results with CTS

Comparison of predictions (FIM) and observations (CTS)

Relative standard errors of parameters

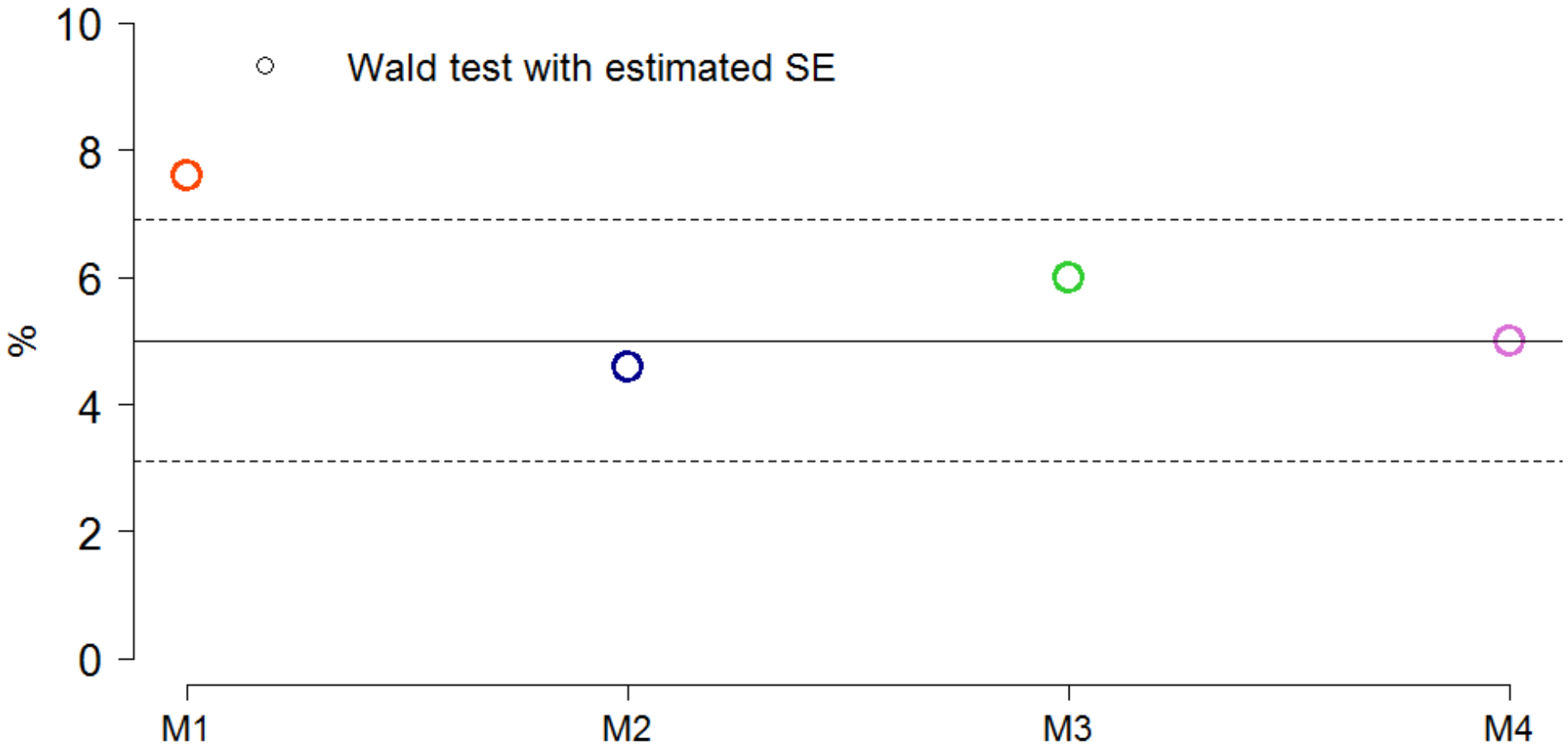
RSE : Relative Standard Error
Predicted RSE by computation of the **FIM**
Empirical RSE from the standard deviation of parameter estimations on **simulations**
RRMSE : Relative root mean square error

Parameters
μ : fixed effects
β : treatment covariate effect
ω : standard error of random effects (interindividual variability)



Comparison of predictions (FIM) and observations (CTS)

Type I error



Comparison of predictions (FIM) and observations (CTS)

Power of the Wald test

	FIM Predicted power (%)	CTS Observed power using estimated SE [CI_{95}] (%)
M1	97.7	99.8 [98.9;100]
M2	90.9	98.4 [96.9;99.3]
M3	97.8	99.6 [98.6;100]
M4	80.1	86.3 [83.0;89.2]

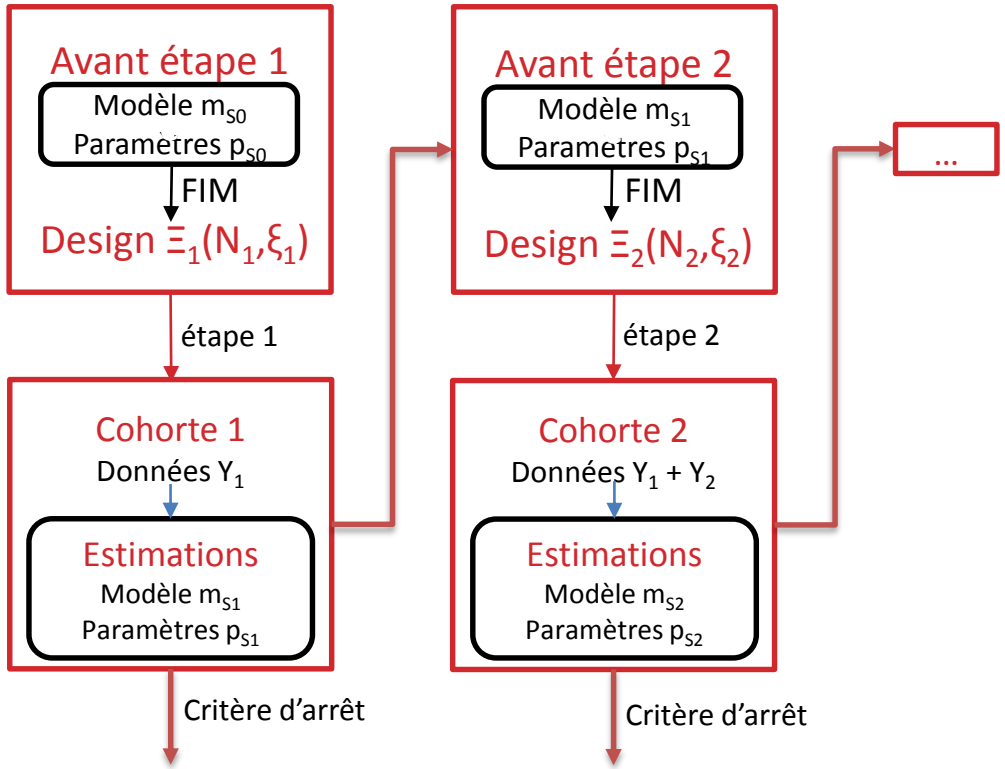
CTS Observed power = Proportion of significant Wald tests with a type I error of 5 %

Travaux de méthodologie : Développer une approche robuste et adaptative tenant compte de l'incertitude du modèle et des valeurs de paramètres

Planification de design prenant en compte l'incertitude du modèle **et des paramètres**

Incorporation de ces méthodes robustes dans la planification adaptative

Schéma adaptatif **non** robuste

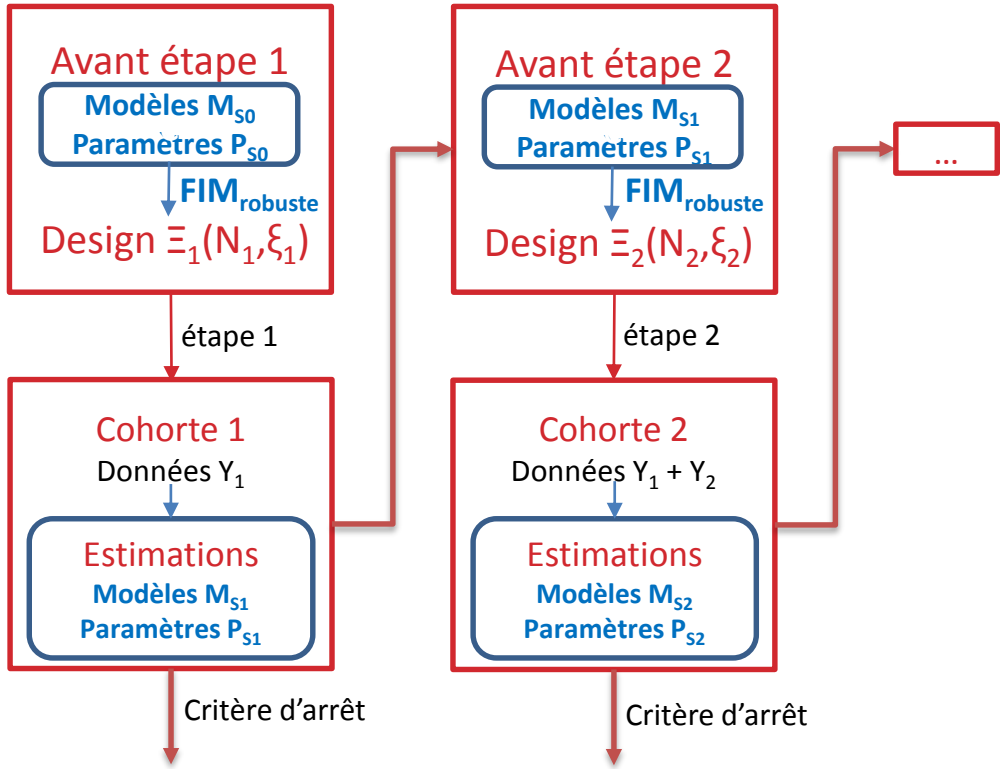


Travaux de méthodologie : Développer une approche robuste et adaptative tenant compte de l'incertitude du modèle et des valeurs de paramètres

Planification de design prenant en compte l'incertitude du modèle **et des paramètres**

Incorporation de ces méthodes robustes dans la planification adaptative

Schéma adaptatif **robuste**



- Extension du schéma de planification non robuste [8]
  - Prise en compte de l'incertitude du modèle et des paramètres à chaque étape
  - Méthode de « model averaging » [9]
  - Influence du nombre d'étapes S et de la taille de chaque cohorte N_s
  - Étude des critères d'arrêt
  - Évaluation par CTS

8 : Lestini G. *et al.*, Pharm Res, 2015
9 : Buckland S.T. *et al.*, Biometrics, 1997

Predicted RSE (N=100)

Table A1. Predicted relative standard errors (%) of candidate models parameters

RSE		μ_1	μ_2	μ_3	β	ω_1^2	ω_2^2	ω_3^2
Model	Optimal design							
M1	$\Xi_{D,1} = (0,2,11,12)$	15	34	-	40	136	56	-
	$\Xi_{D_S,1} = (0,2,11,12)$	15	34	-	40	136	56	-
	$\Xi_{DD_S,1} = (0,2,11,12)$	15	34	-	40	136	56	-
	$\Xi_{CD} = (0,1,5,12)$	14	37	-	46	121	71	-
	$\Xi_{CD_S} = (0,5,11,12)$	16	35	-	46	153	54	-
	$\Xi_{CDD_S} = (0,4,11,12)$	16	35	-	43	147	55	-
M2	$\Xi_{D,2} = (0,1,8,12)$	15	34	-	42	139	56	-
	$\Xi_{D_S,2} = (0,1,11,12)$	15	34	-	42	144	55	-
	$\Xi_{DD_S,2} = (0,1,11,12)$	15	34	-	42	144	55	-
	$\Xi_{CD} = (0,1,5,12)$	15	35	-	42	139	58	-
	$\Xi_{CD_S} = (0,5,11,12)$	18	35	-	54	169	49	-
	$\Xi_{CDD_S} = (0,4,11,12)$	17	35	-	52	164	50	-
M3	$\Xi_{D,3} = (0,4,5,12)$	13	38	-	47	111	80	-
	$\Xi_{D_S,3} = (0,5,11,12)$	15	34	-	41	135	58	-
	$\Xi_{DD_S,3} = (0,5,11,12)$	15	34	-	41	135	58	-
	$\Xi_{CD} = (0,1,5,12)$	13	38	-	50	108	82	-
	$\Xi_{CD_S} = (0,5,11,12)$	15	34	-	41	135	58	-
	$\Xi_{CDD_S} = (0,4,11,12)$	15	34	-	43	130	57	-
M4	$\Xi_{D_4} = (0,3,11,12)$	16	32	24	66	140	59	269
	$\Xi_{D_S,4} = (0,6,7,12)$	18	39	28	52	172	72	290
	$\Xi_{DD_S,4} = (0,5,7,12)$	17	37	25	54	164	71	283
	$\Xi_{CD} = (0,1,5,12)$	14	38	31	70	120	89	266
	$\Xi_{CD_S} = (0,5,11,12)$	18	34	22	55	169	58	283
	$\Xi_{CDD_S} = (0,4,11,12)$	17	33	22	61	155	59	279