A Bayesian decision-theoretic approach to incorporating pre-clinical information into phase I trials

Haiyan Zheng, Lisa V. Hampson

Medical and Pharmaceutical Statistics Research Unit
Department of Mathematics and Statistics, Lancaster University

PSI Translational Statistics Meeting
30th March 2017

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567. (IDEAS - http://www.ideas-itn.eu/)

Dose escalations

Consider:

- First-in-man studies $\rightarrow$ limited knowledge about the toxicity to humans
- Binary endpoint: dose-limiting toxicity (DLT) versus no-DLT
- Doses available: $d_1, \ldots, d_J$

Aim:

- to estimate the TDπ, the dose associated with risk of DLT at level π
- Commonly, $\pi \in (0.20, 0.35)$ for oncology trials

Bayesian model-based designs

Existing approaches:

- CRM
- BLRM
- EWOC

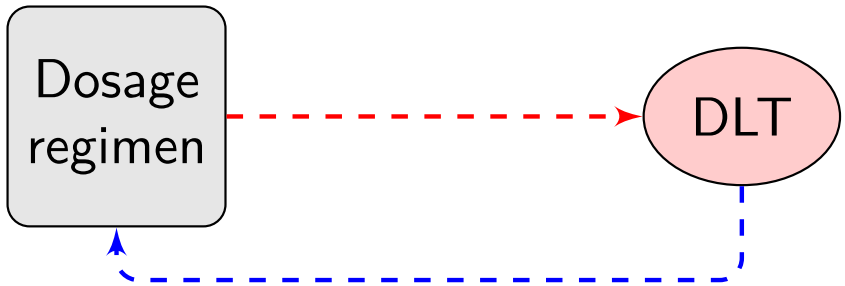
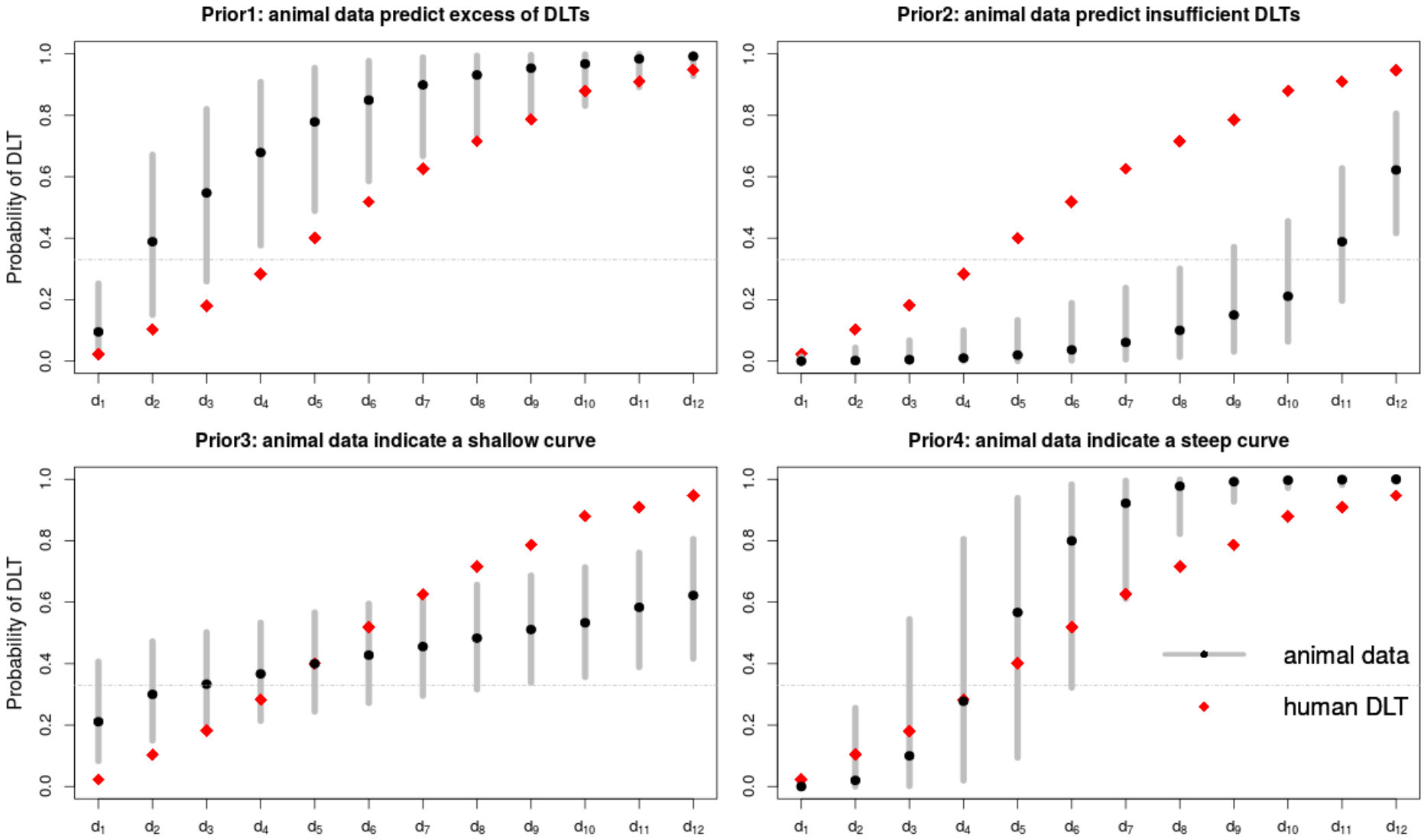


Figure : Modelling the dose-toxicity relationship.

Key features:

- Probabilistic inference about $p(d) \implies$ dose-escalation decisions
- Adopt uninformative, operational priors
- **Incorporating pre-clinical information?** Often informally
- **Challenges?**

Commensurability issues



Goal

To establish a **formal incorporation** of pre-clinical info into phase I trials

- represent the information in a prior for parameters of the dose-toxicity model
- discount it quickly if a **prior-data conflict** emerges anytime during the trial

Problem formulation

- Dose-toxicity model: $\log\left\{\frac{p(d)}{1-p(d)}\right\} = \theta_1 + \exp(\theta_2)\log d$
- Bivariate normal prior for $\boldsymbol{\theta} = (\theta_1, \theta_2)$
  - operational prior
  - **informative** prior, formulated using pre-clinical data
  - mixture prior

$$f(\boldsymbol{\theta}) = \omega \times \underbrace{g(\boldsymbol{\theta})}_{\text{pre-clinical data}} + (1-\omega) \times \underbrace{h(\boldsymbol{\theta})}_{\text{operational prior}},$$

Q1: How to derive $g(\boldsymbol{\theta})$ with pre-clinical toxicology information?

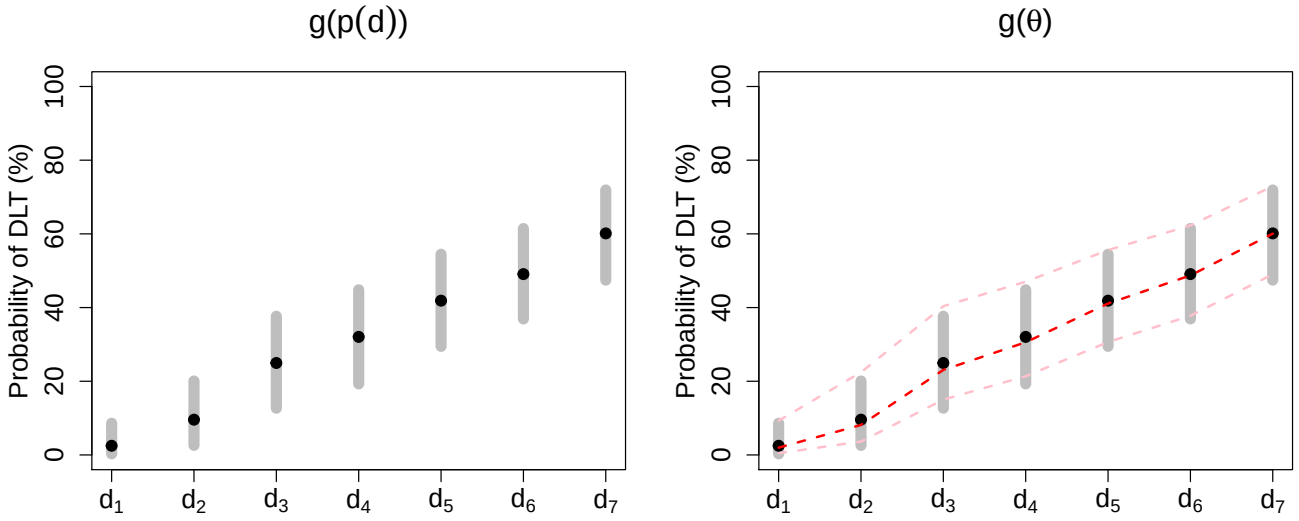
Q2: How to quantify the mixture weight ω?

On deriving the $g(\boldsymbol{\theta})$

Available pre-clinical information $\rightarrow$ informative component $g(\boldsymbol{\theta})$:

1. Summarise pre-clinical information as pseudo-data on the lowest and highest doses d_{-1} and d_0
2. This specifies **independent beta distributions** for $p(d_{-1})$ and $p(d_0)$
3. Given $\text{logit}\{p(d)\} = \theta_1 + \exp(\theta_2)\log d$, derive the **prior distributions** for $p(d_j)$ and their $2.5^{\text{th}}, 50^{\text{th}}$ and 97.5^{th} percentiles
4. Find the bivariate normal prior for $\boldsymbol{\theta} = (\theta_1, \theta_2)$, which approximately agrees with the exact summaries [A STOCHASTIC OPTIMISATION PROBLEM]

$$p(d_{-1}) \sim Beta(t_{-1}, u_{-1})$$
$$p(d_0) \sim Beta(t_0, u_0)$$



Choosing the mixture weight ω

Challenge: difficult to test the prior-data conflict and to quantify the degree of **commensurability**, since phase I trials are typically small

Define the mixture prior for the k^{th} cohort as

$$f_k(\boldsymbol{\theta}) = \omega_k \times \underbrace{g(\boldsymbol{\theta})}_{\text{pre-clinical data}} + (1-\omega_k) \times \underbrace{h(\boldsymbol{\theta})}_{\text{operational prior}},$$

- ω_k is dynamically determined at each interim analysis
  - small value when evidenced by prior-data conflict
  - large value when animal and human data appear commensurate
- A Bayesian decision-theoretic approach to measuring the commensurability
  - How accurate are predictions of human responses based on pre-clinical data?
  - Penalise harshly when they underestimate risks of toxicity in humans

Measuring the prior-data conflict

Fouskakis and Draper (2002), Vehtari and Ojanen (2012)

Let Y denote the response of a human patient receiving a specific dose.

1. Derive **prior predictive distributions** $\mathcal{P}\{Y = \tilde{y}\}$ from animal data
2. Derive **optimal prediction** for Y as

$$\hat{\eta} = \arg \max_{\eta \in \{0,1\}} \sum_{\tilde{y}} u(\tilde{y}, \eta) \mathcal{P}\{Y = \tilde{y}\}, \ \tilde{y} \in \{0, 1\}$$

where $u(\tilde{y}, \eta)$ is the utility function that rewards predictions of $\tilde{y}$ as η:

$$u(\tilde{y}, \eta) = \begin{cases} 0, & \text{if } \eta = 0 \text{ while } \tilde{y} = 1 \text{ (incorrectly predict as no-DLT)} \\ s, & \text{if } \eta = 1 \text{ while } \tilde{y} = 0 \text{ (incorrectly predict as DLT)} \\ 1, & \eta = \tilde{y} \text{ (correct prediction)} \end{cases}$$

Note that $0 < s < 1$.

Measuring the prior-data conflict (*Cont'd*)

$f_k(\theta) = \omega_k \times g(\theta) + (1 - \omega_k) \times p(\theta)$

3. Compare optimal prior predictions versus observed human responses for each dose d_j prior to the k^{th} cohort

		Rewards and Penalties		Cell counts	
		Observation (y)			
		No-DLT	DLT		
Prior prediction ($\hat{\eta}$)	No-DLT	u_{00} (1)	u_{10} (0)	n_{00}	n_{10}
	DLT	u_{01} (s)	u_{11} (1)	n_{01}	n_{11}

4. Derive the predictive utility of the animal data for the observed human toxicity data on dose d_j as $U_j^k = \sum_{l=0}^{1} \sum_{m=0}^{1} u_{lm} n_{lm}$
5. Measure commensurability of animal and human data by taking the average of predictive accuracy across doses used so far

$$\bar{a}_k = \frac{1}{J} \sum_{j=1}^{J} \frac{U_j^k}{\sum_{l=0}^{1} \sum_{m=0}^{1} n_{lm}}$$

Measuring the prior-data conflict (*Cont'd*)

$f_k(\theta) = \omega_k \times g(\theta) + (1 - \omega_k) \times p(\theta)$

6. In our investigation, we define

$$\omega_k = \bar{a}_k^{\lambda_k},$$

where λ_k can reflect

- the relative variability:

$$\lambda_k = \frac{s.d.(\bar{a}(y_k, \hat{\eta}_k | \boldsymbol{x_k}))}{s.d.(\bar{a}(\underbrace{y_k, \ldots, y_N}_{\text{simul future obs.,}}, \underbrace{\hat{\eta}_k, \ldots, \hat{\eta}_N}_{\text{optimal pred.}} | \boldsymbol{x_k}))},$$

Notations

$\boldsymbol{x_k}$: phase I trial data

$y_k, \ldots, y_N$: possible outcomes of future patients that receive the dose recommended based on the current best understanding

$\hat{\eta}_k, \ldots, \hat{\eta}_N$: corresponding optimal predictions

Interim dose recommendations

Whitehead and Williamson (1998), Babb et al. (1998)

For the k^{th} cohort, $k = 1, 2, \ldots, N$

- Compare prior animal data with observed human data to derive ω_k
- Update the mixture prior $f_k(\boldsymbol{\theta}) = \omega_k \times g(\boldsymbol{\theta}) + (1 - \omega_k) \times h(\boldsymbol{\theta})$ to derive posterior $f_k(\boldsymbol{\theta}|\boldsymbol{x}_k)$
- Use the accumulated data $\boldsymbol{x}_k$ to recommend a dose for the $(k+1)^{\text{th}}$ cohort according to the patient gain criterion

$$\mathcal{G} = (\tilde{p}(d_j) - \pi)^{-2},$$

where $\tilde{p}(d_j)$ is the implied probability of toxicity at dose d_j and π is the target level

Practical considerations:
1) Effective sample size of the $g(\boldsymbol{\theta})$
2) Run-in period for the incorporation of pre-clinical info

Safety constraint

Throughout the trial, the probability of toxicity is considered to be **excessively high** if

$$\int_{\gamma}^{1} g(p(d_j))\mathrm{d}p(d_j) \geq \delta,$$

where γ is some threshold and δ is the pre-define level.

This naturally specifies an **early stopping rule**:

- Stop when none of the doses available satisfy the safety constraint;

i.e., early stopping for safety, if the lowest dose d_1 is found excessively toxic:

$$\int_{\gamma}^{1} g(p(d_1))\mathrm{d}p(d_1) \geq \delta$$

Note

In our simulations, we set $\gamma = 0.45$ and $\delta = 0.25$.

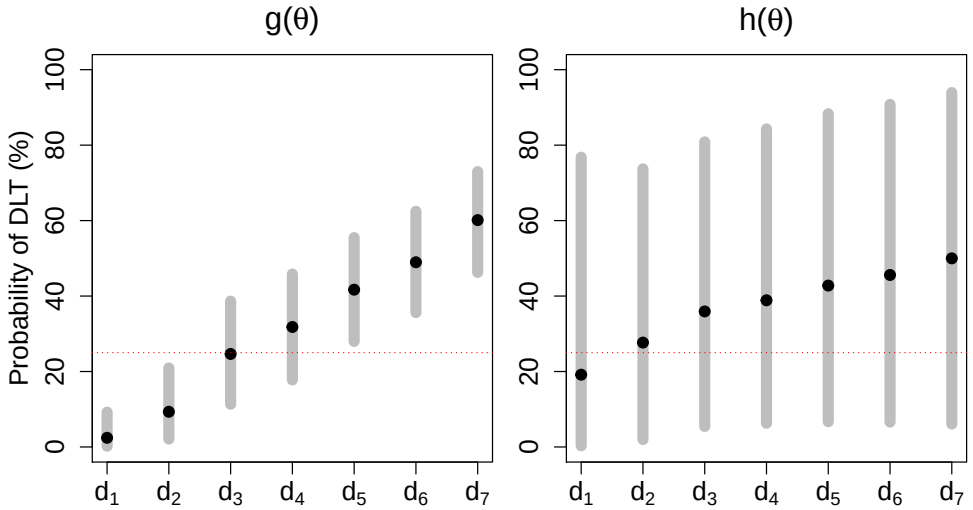
Simulations

From the **pre-clinical studies**:
$p(d_{-1}) \approx 0.03;\ p(d_0) \approx 0.60 \rightarrow$ worth $n_{-1} = n_0 = 60$

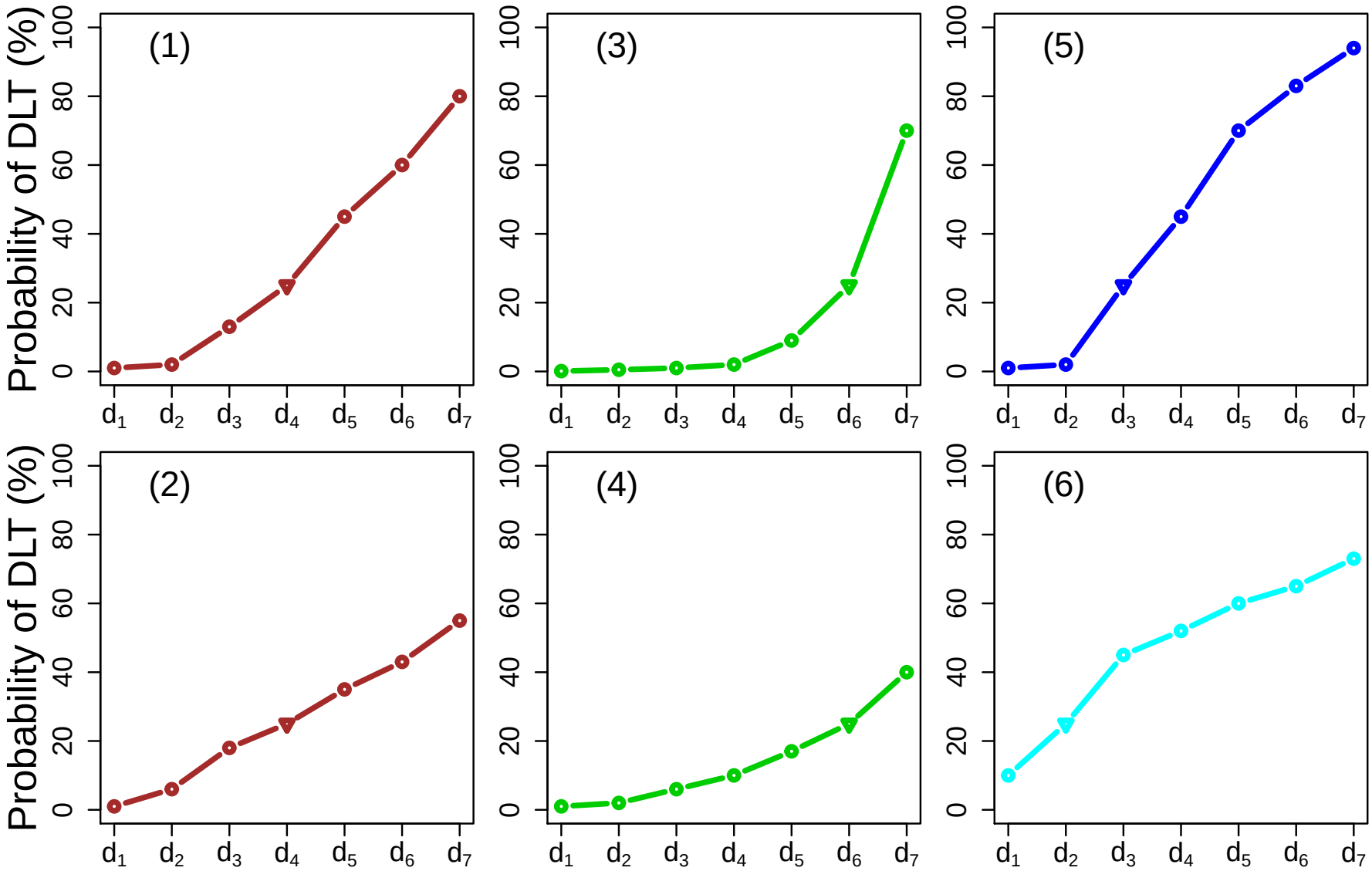
Settings for the **first-in-man trial**:

- Cohort size $c = 1$
- Max. ss $N = 24$
- Number of doses $J = 7$
- Target level $\pi = 0.25$
- Early stopping for accuracy is not considered
- Results based on 1000 simulated trials

Thus, **priors** derived as



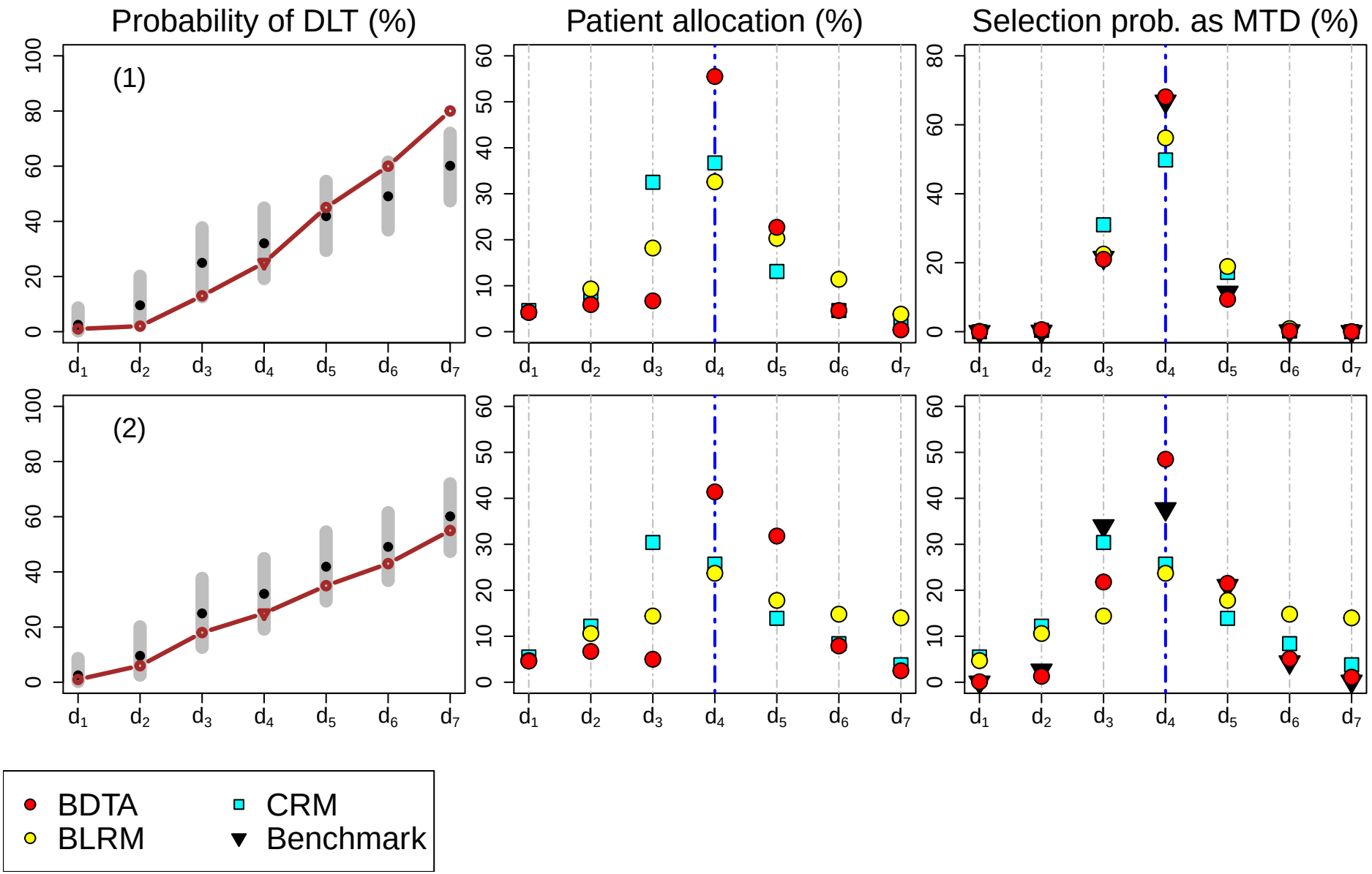
Investigated human toxicity scenarios



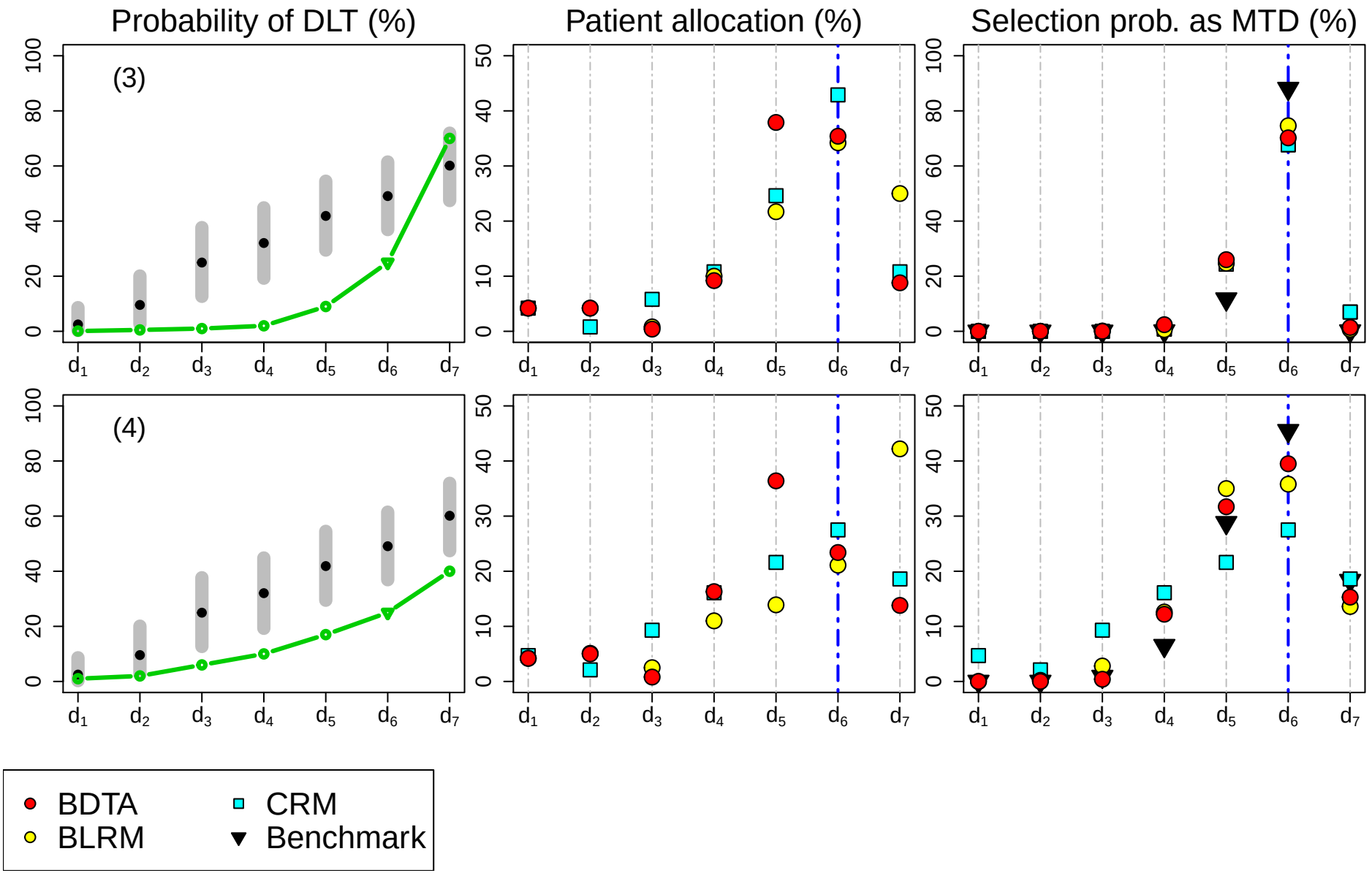
Comparator designs

- **BDTA**
- BLRM with operational prior
- CRM with naïve opinion of incorporating pre-clinical info
- Non-parametric optimal benchmark

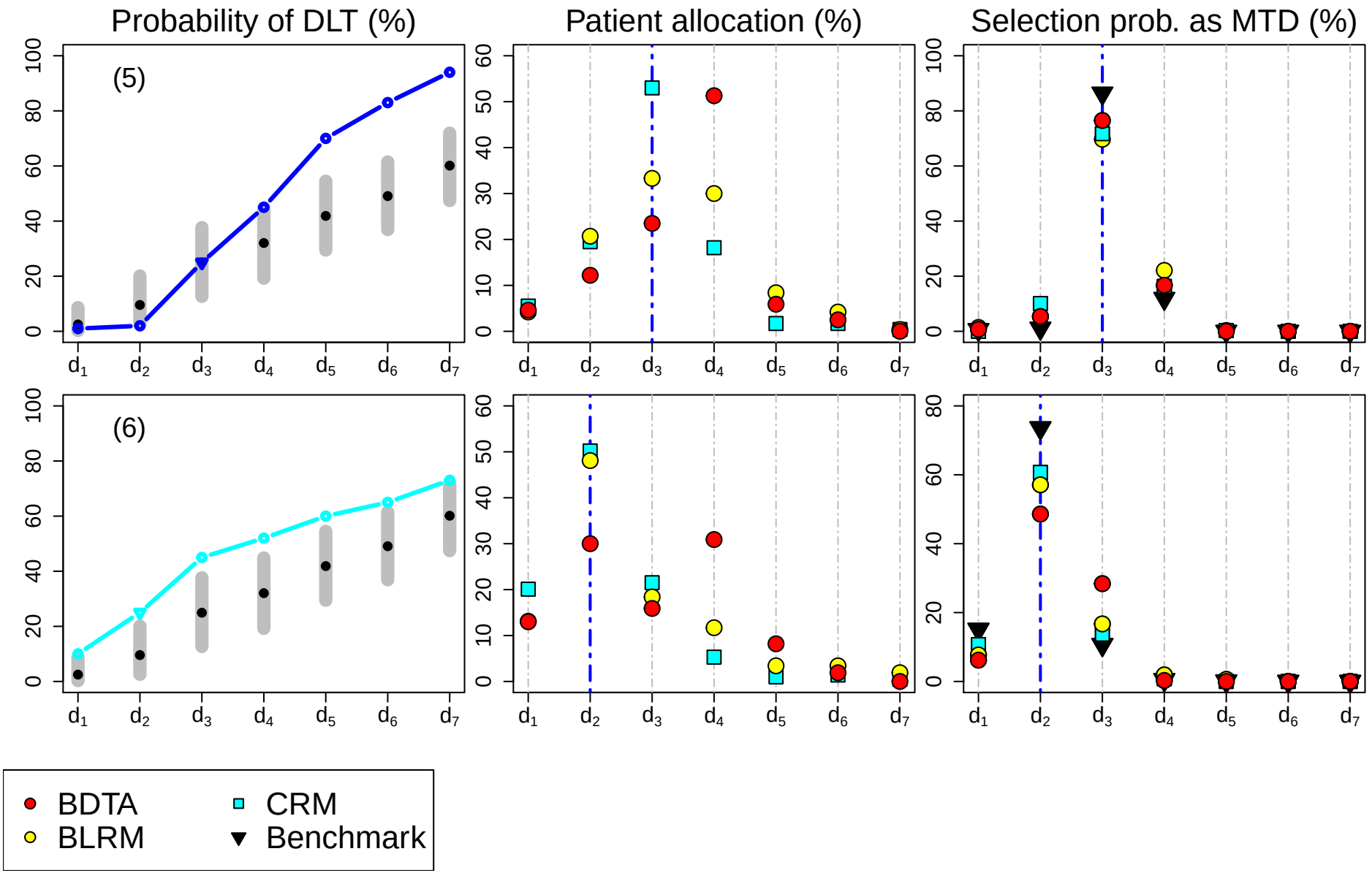
Simulation results (I)



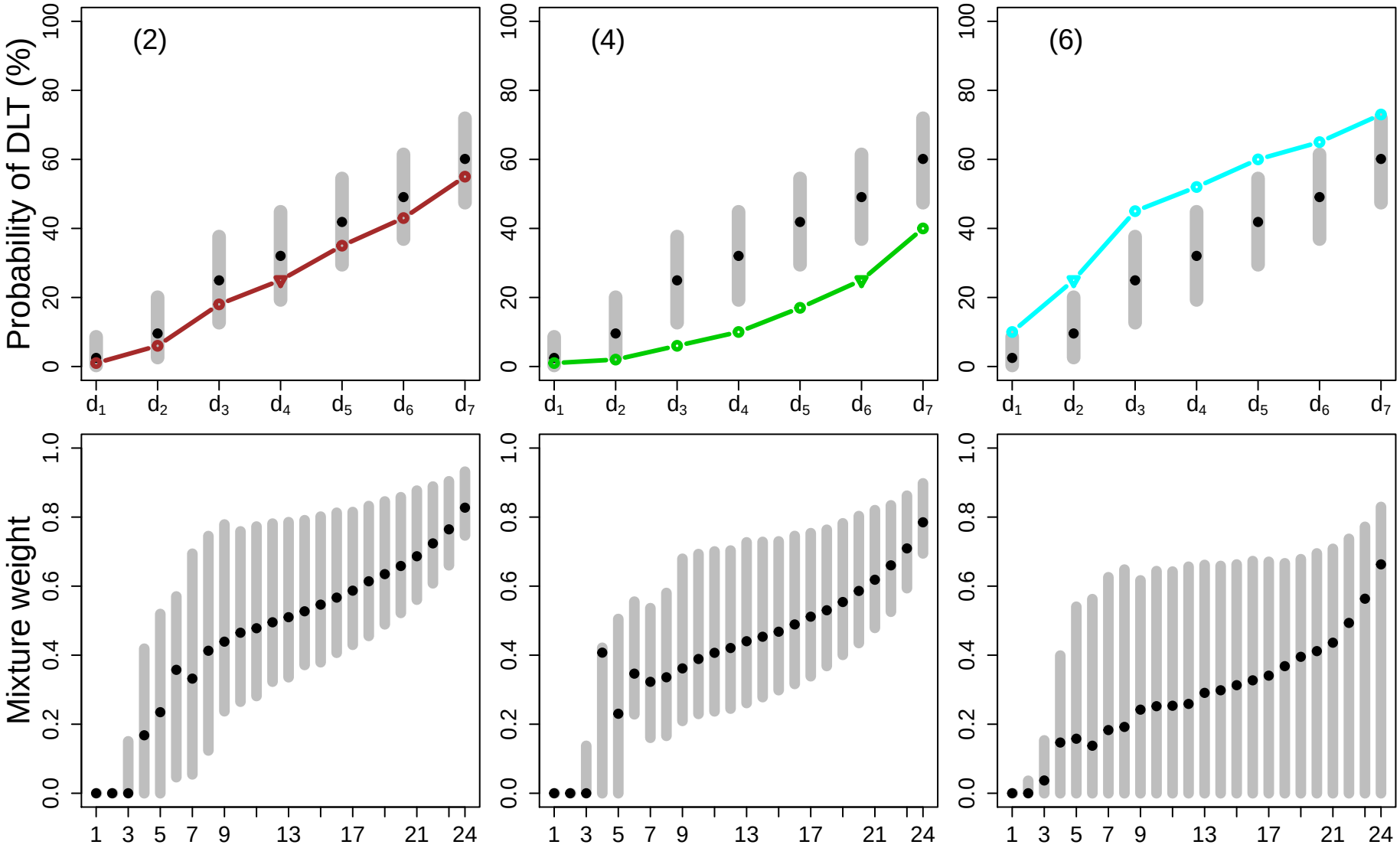
Simulation results (II)



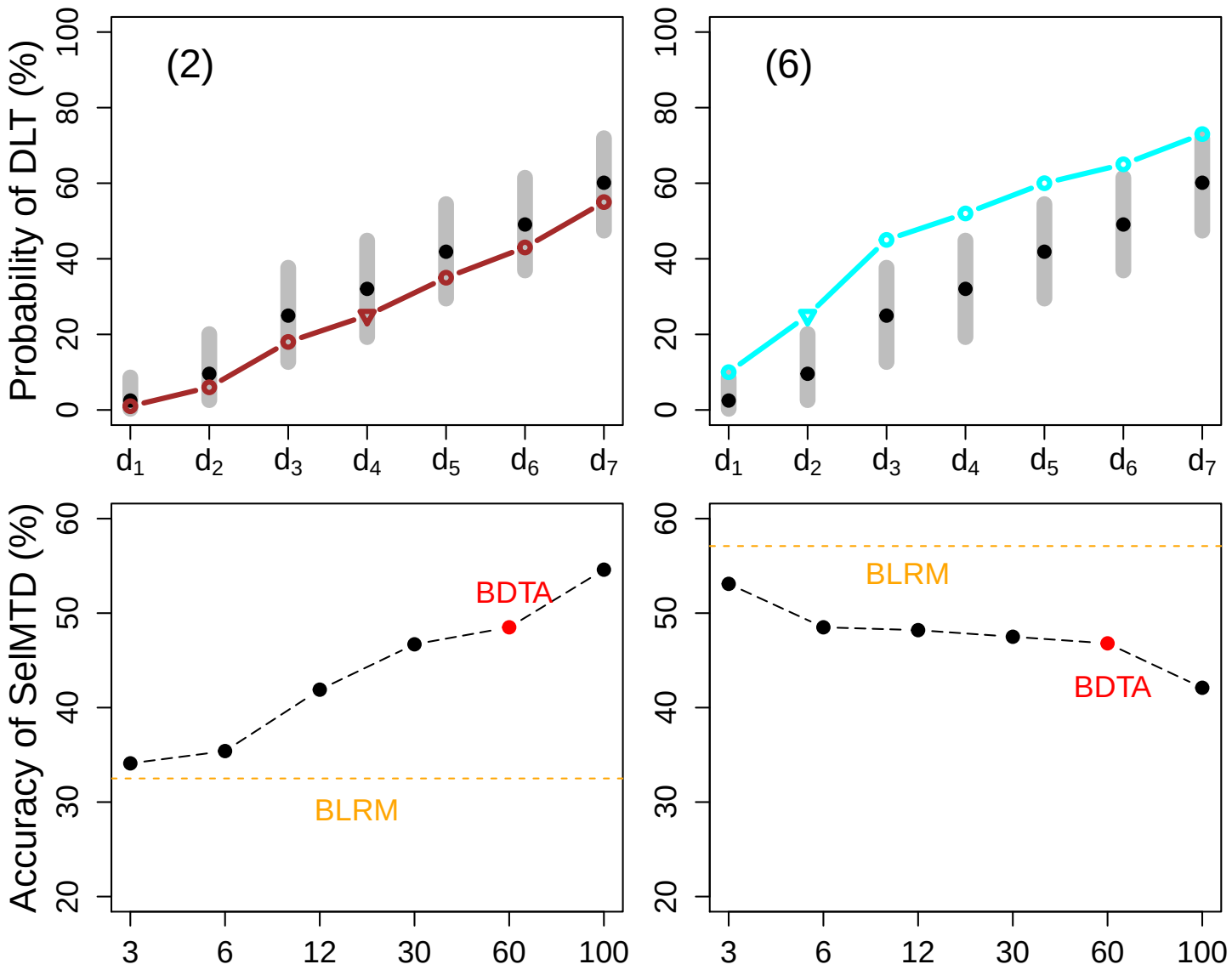
Simulation results (III)



Simulation results (IV)



Sensitivity analysis



Highly toxic & Very safe scenarios

Design		Dose levels 1	2	3	4	5	6	7	None	DLT	$\bar{N}$
	$p(d)$	40	60	80	87	91	93	95			
BDTA	Sel	8.9	1.5	0	0	0	0	0	**89.6**	3.4	6.6
	Pts	3.6	2.2	0.3	0.5	0	0	0			
BLRM	Sel	9.1	0.9	0	0	0	0	0	**90.0**	3.3	6.5
	Pts	3.6	2.4	0.2	0.3	0	0	0			
CRM	Sel	12.7	0.4	0	0	0	0	0	**86.9**	3.4	6.9
	Pts	4.8	1.3	0.7	0.1	0	0	0			
	$p(d)$	0.1	0.2	0.5	2	6	15	25			
BDTA	Sel	0.5	0.1	0.1	1.1	7.3	30.8	**60.7**	0	3.6	24
	Pts	1.0	1.0	0	1.6	4.8	6.2	9.4			
BLRM	Sel	0	0	0	1.9	14.8	17.6	**65.2**	0	4.5	24
	Pts	1.0	1.0	0	1.3	0.9	2.7	17.1			
CRM	Sel	0	0	0	0.3	5.8	39.4	**54.4**	0	3.2	24
	Pts	1.0	1.0	1.0	1.7	4.6	7.1	7.6			

Table : Results for two more extreme cases, i.e., highly toxic and very safe

Conclusions

- Incorporating pre-clinical data will potentially lead to more efficient escalation decision making and greater estimation precision
  - Dose recommendations are robust and competitive
  - Patients have enhanced possibility to receive the target dose
- Pre-clinical information that may undermine the safety of patients can be quickly discounted during the course of the trial

Future work

- Two or more animal species
- Pharmacological information
- Phase I trials with both safety and efficacy endpoints

References


Whitehead J, Williamson D. Bayesian decision procedures based on logistic regression models for dose-finding studies. J Biopharm Med 1998; 8: 445-67.


Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. Stat Med 2008; 27: 2420-39.


O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. Biometrics. 1996; 52: 673-84.


Babb J, Rogatko A, Zacks S. Cancer phase I clinical trials: efficient dose escalation with overdose control. Stat Med 1998; 17:1103-20.


Fouskakis D, Draper D. Stochastic Optimization: a Review. International Statistical Review 2002; 70: 315-49.


Vehtari A, Ojanen J. A survey of Bayesian predictive methods for model assessment, selection and comparison. Statistics Survey 2012; 6: 142-28.


Morita S, Thall PF, Müller P. Determining the effective sample size of a parametric prior. Biometrics 2008; 64: 595-602.