Web-based Supplementary Materials for: A robust Bayesian meta-analytic approach to incorporate preclinical animal data into phase I oncology trials

by Haiyan Zheng, Lisa V. Hampson, Simon Wandel

A. LOGISTIC REGRESSION CURVES WITH VARYING INTERCEPTS OR SLOPES

In Section 2, we introduce a translation parameter, denoted by $\delta_{\mathcal{A}_i}$, as a multiplicative factor due to one underlying assumption generally made in contemporary translational sciences that the dose-toxicity curves across species are more similar in terms of their shapes rather than locations [1]. Figure S1 displays a group of logistic regression curves, which suggest the shape is dominated largely by the slope.

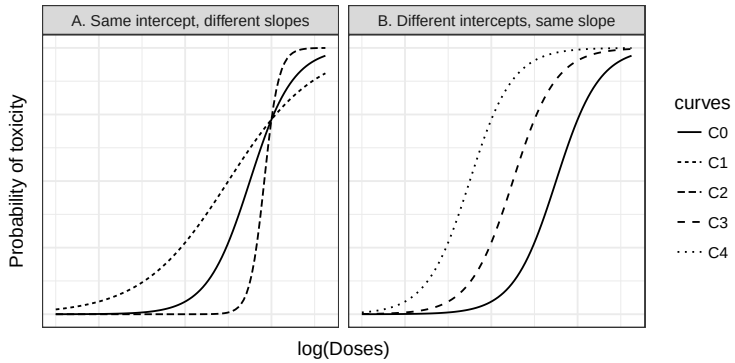

Figure S1: A group of logistic regression curves C0-C4 to graphically illustrate how the slope and the intercept would modify the shape, respectively. The solid curve C0 plotted in both panels is intended to serve as a benchmark for comparison.

B. THE HYPOTHETICAL DOG DATA USED IN THE PAPER

In Section 4, we suppose that prior to the phase I first-in-man trial, historical data are available from three hypothetical preclinical toxicology studies in dogs. Figure S2 shows the binomial data that we have used in Section 4, with the prior effective sample size described in Table S1.

Table S1: Summaries of marginal predictive priors derived from the dog data setting $w_R = 0.3$. Also reported are the parameters of the Beta(a, b) approximtes used for ESS calculation

	d_{i^*1} 2	d_{i^*2} 4	d_{i^*3} 8	d_{i^*4} 16	d_{i^*5} 22	d_{i^*6} 28	d_{i^*7} 40	d_{i^*8} 54	d_{i^*9} 70
Prior means	0.057	0.085	0.135	0.223	0.282	0.338	0.431	0.509	0.573
Prior std dev.	0.118	0.133	0.153	0.176	0.186	0.195	0.212	0.222	0.226
ESS	3.0	3.4	4.0	4.6	4.9	4.8	4.4	4.1	3.8
a	0.2	0.3	0.5	1.0	1.4	1.6	1.9	2.1	2.2
b	2.8	3.1	3.5	3.6	3.5	3.2	2.5	2.0	1.6

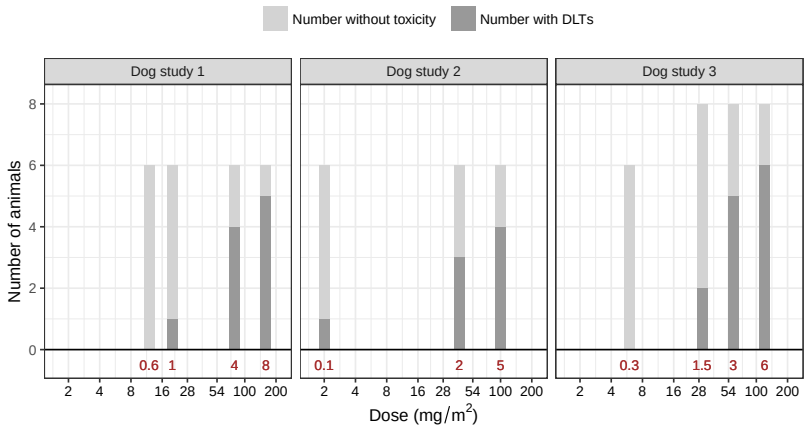

Figure S2: Preclinical data from three hypothetical studies in dogs. The height of the bar represents the number of dogs studied, and the height of the dark grey segment counts the number experiencing an ocular toxicity. Doses listed in brown are the doses (mg/kg) administered to dogs. Doses listed in black are the human-equivalent doses (mg/m^2). Projections are made by scaling animal doses using the prior median of δ_{Dog}.

C. ADDITIONAL SIMULATION RESULTS

C.1 Numerical results of all evaluated scenarios

The performance of trials using BLRM-guided dose-escalation under Models A – D are compared with that of the optimal non-parametric benchmark design [2]. The optimal design is defined using the 'complete' toxicity profile of each patient, created by assuming there are $J_{i^\star}$ clones of a patient given doses spanning the dosing set $\mathcal{D}_{i^\star}$. A toxicity tolerance thereshold ϵ_n is generated from $U[0,1]$ for the nth patient, which determines the corresponding toxicity outcome at the jth dose as

$$R_{jn} = \mathbb{1}(\epsilon_n \leq p_{i^\star j}), \quad 1 \leq n \leq N, \quad 1 \leq j \leq J_{i^\star},$$

where $\mathbb{1}(\cdot)$ is the indicator function. An unbiased estimate for $p_{i^\star j}$ is thus $\bar{R}_j(N) = \frac{1}{N}\sum_{n=1}^{N} R_{jn}$ for a trial of which the maximum sample size is N. Consequently, the estimated MTD under the benchmark design is

$$\hat{d}_{\text{M}}^{\text{opt}} = \arg \min_{j=1,\ldots,J_{i^\star}} |\bar{R}_j(N) - 0.25|.$$

Improvements beyond this bound are not possible unless strong parametric assumptions are made about dose-response relationships. In our context, we wish to quantify the gains that can be made over the benchmark designs, in part due to borrowing strength from the preclinical data.

Two variants of Model A – C are evaluated, either treating δ_{Dog} as a random variable with a log-normal prior distribution, or a fixed constant taking the median of the log-normal prior. Table S2 provides a complete listing of all simulation results for analysis models defined in Section 5 and the optimal benchmark design.

C.2 Improved estimation precision

In our way of declaring a dose to be MTD, precision of the posterior estimate of the DLT risk is decisive. We thus evaluate different analysis models by examining the bias, mean squared error (MSE) and coverage probability (CP) of the central 95% credible interval of the posterior median of the DLT risk at the true MTD. Figure S3 visualised the comparison in terms of these metrics. As

illustrated, inference based on the analysis Models A-C reports smaller bias and MSE than Model D, except the Scenarios 1 and 8.

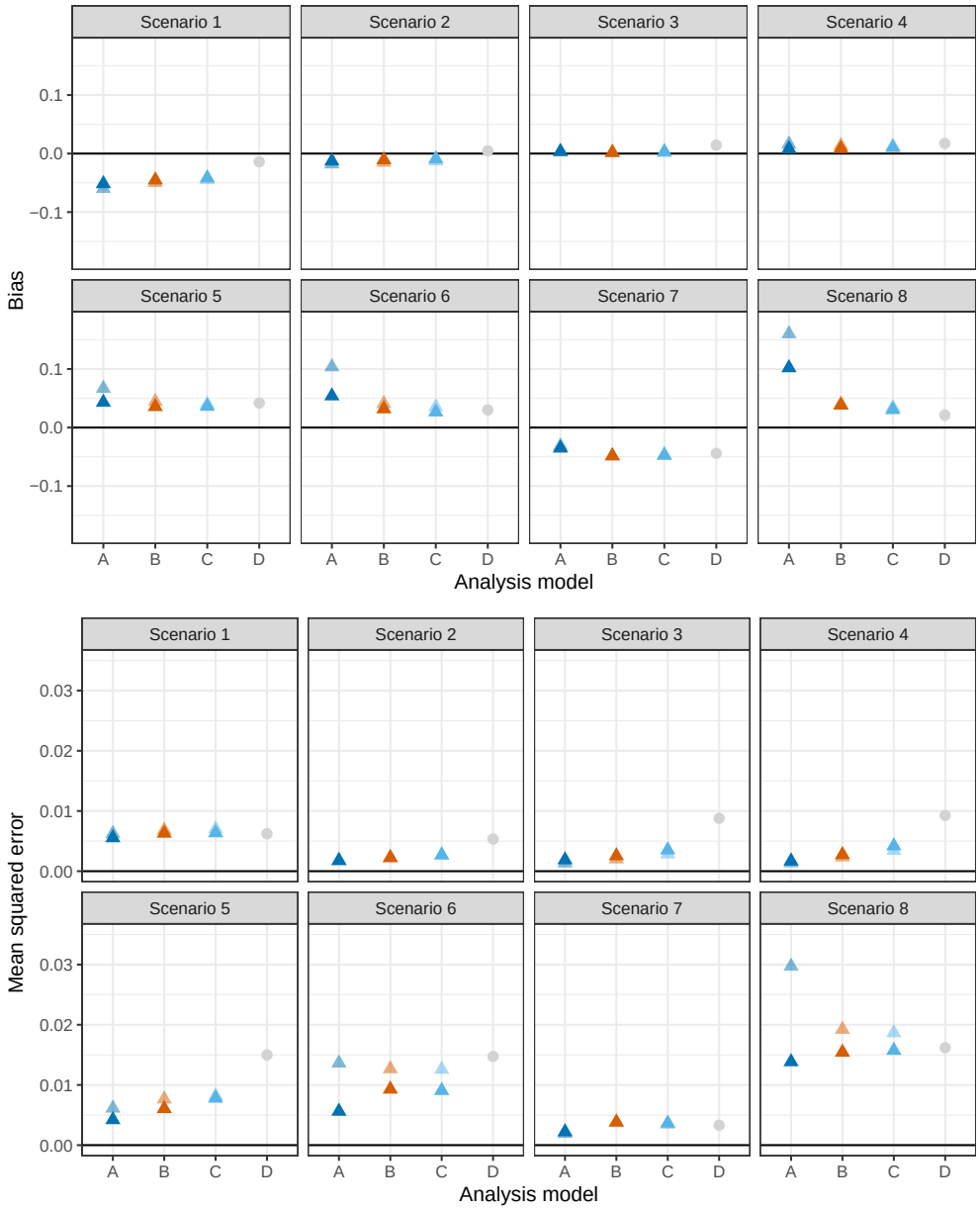

Figure S3: Comparison of the performance in terms of bias and mean squared error of the toxicity rate estimator at the true human MTD, based on different analysis Models A – D. Solid plotting symbols correspond to analysis models with δ_{Dog} defined as a random variable. Transparent ones correspond to the counterparts defined with δ_{Dog} as a fixed constant.

Figure S4 visualised the comparison in terms of the CP of central 95% credible interval at the true MTD by applying analysis Models A-C to incorporate preclinical data versus Model D to entirely discard them. We observe that at least 95% CP was attained for almost all simulation scenarios except Scenario 1. The low convergence probability of analysis models A-C in Scenario 1 is explainable, as risk of toxicity at the dose 16 mg/m^2 tends to be underestimated and thus easier to be concluded as the MTD in humans after synthesising the dog data, which advise a safer toxicity profile of the drug to humans. Across the scenarios considered here, Model A tends

to attain lowest CP for scenarios when there is a discrepancy between the prediction of human toxicity based on Bayesian meta-analysis of the dog data and the true MTD. This is because large weight placed on preclinical data would lead to excessive shrinkage to the animal parameter, although on the equivalent human dosing scale. In addition, a fixed constant of δ_{Dog} in general would produce an estimate of toxicity rate at the target dose with less accurate confidence interval.

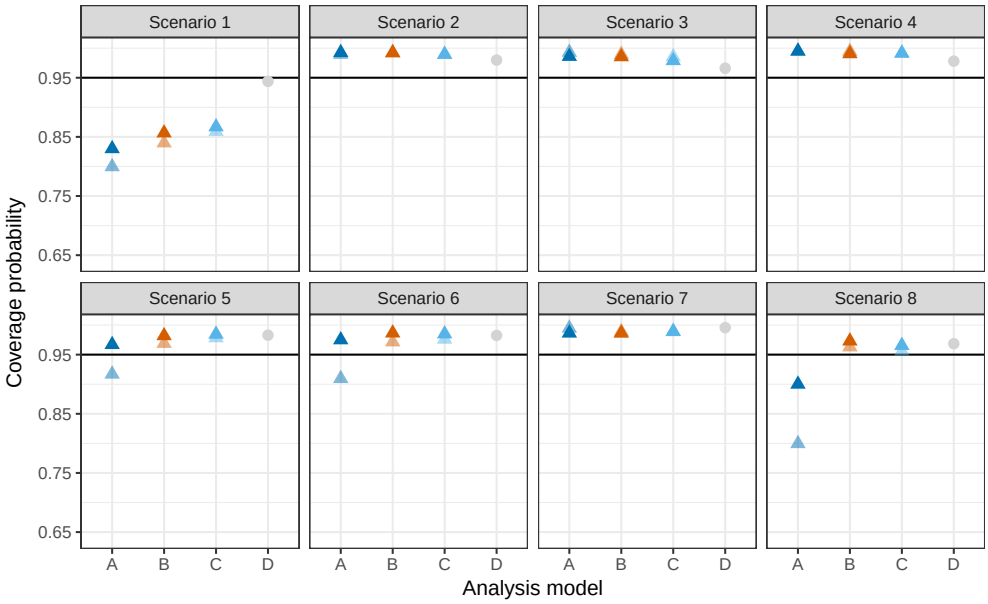

Figure S4: Comparison of the performance in terms of coverage probability of central 95% credible interval of the toxicity rate estimator at the true human MTD, based on different analysis Models A – D. Solid plotting symbols correspond to analysis models with δ_{Dog} defined as a random variable. Transparent ones correspond to the counterparts defined with δ_{Dog} as a fixed constant.

Another interesting comparison is investigated between two variants of Bayesian meta-analytic Models A – C, by treating the translation factor δ_{Dog} as a random variable or fixed constant. As shown in both Figures S3 and S4, it is not surprising that meta-analytic models constrained with fixed δ_{Dog} would lead to increased precision in Scenarios 3 and 4 due to prior-data consistency. In other simulation scenarios, however, those models experience problems because of ineffective down-weighting of the dog data. This is especially true for Scenarios 6 and 8, when the prior predictions mismatch the true human MTD and, meanwhile, incorporating preclinical data that overestimate the human toxicity leads to more conservative dose escalations.

D. COMPARISONS WHEN ADDITIONAL EARLY STOPPING RULES ARE APPLIED

In Figure S5, we show operating characteristics based on 2000 simulated phase I clinical trials under the same escalation rules, defined in Section 4.2, and the same basis of declaring a dose to be MTD as defined in Section 5. The difference is that trials will be terminated earlier if criteria (a)-(c), specified as follows, are satisfied:

(a) There have been at least 8 cohorts (24 patients) recruited

(b) The dose selected for the next cohort is the same as the current dose

(c) At least 6 patients have been treated with this dose

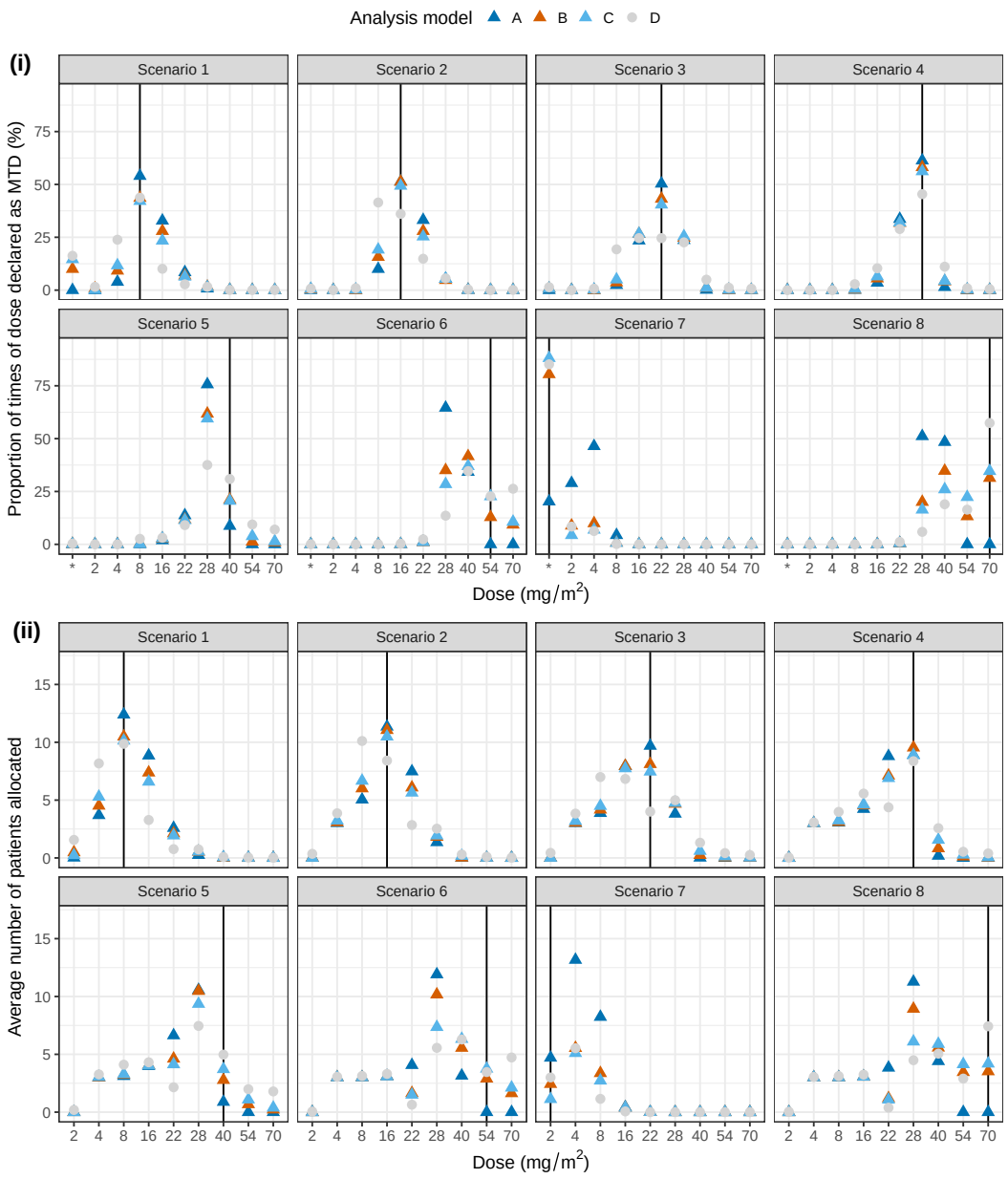

Figure S5: Operating characteristics of BLRM-guided dose-escalation procedures basing inferences on Models A-D, when additional early stopping rules may be applied.

Table S2: Comparison of alternative analysis models in terms of the percentage of selecting a dose as MTD at the end of the trials, percentage of early stopping for safety, average patient allocation, and average number of patients with toxicity.

Scenario	Design	δ_{S_k}		% dose declared as MTD & average patient allocation										DLT	$\bar{N}$
				d_{i^*1} 2	d_{i^*2} 4	d_{i^*3} 8	d_{i^*4} 16	d_{i^*5} 22	d_{i^*6} 28	d_{i^*7} 40	d_{i^*8} 54	d_{i^*9} 70	None		
1			pTox	0.08	0.16	**0.25**	0.35	0.41	0.45	0.52	0.58	0.63			
	Benchmark		Sel	0.4	19.2	**58.7**	19.3	1.9	0.4	0.1	0	0			
	Model A	Par	Sel	0	4.3	**64.6**	27.6	3.3	0.2	0	0	0	0		
			Pts	0	4.4	22.9	13.7	3.6	0.4	0	0	0		12.9	45.0
		Fix	Sel	0	3.0	52.4	39.4	5.1	0.1	0	0	0	0		
			Pts	0	4.0	17.0	17.9	5.7	0.4	0	0	0		13.7	45.0
	Model B	Par	Sel	0.4	10.2	**52.4**	22.4	4.4	0.2	0	0	0	10.0		
			Pts	0.6	6.2	18.9	11.4	3.1	0.6	0	0	0		11.3	40.8
		Fix	Sel	0.4	10.9	44.8	28.8	4.7	0.2	0	0	0	10.2		
			Pts	0.6	6.2	15.9	12.6	5.0	0.6	0	0	0		11.3	40.9
	Model C	Par	Sel	0.4	12.4	**50.0**	18.6	3.6	0.3	0	0	0	14.7		
			Pts	0.3	7.3	18.0	10.0	2.8	0.6	0.1	0	0		10.6	39.1
		Fix	Sel	0.5	14.0	43.0	22.7	4.8	0.2	0	0	0	14.8		
			Pts	0.4	7.5	15.6	10.8	4.1	0.8	0	0	0		10.6	39.2
	Model D		Sel	0.6	26.0	**47.3**	7.5	1.9	0.4	0	0	0	16.3		
			Pts	1.7	12.5	17.6	4.6	1.1	0.8	0.1	0	0		9.0	38.4
2			pTox	0.01	0.04	0.11	**0.25**	0.35	0.44	0.55	0.65	0.73			
	Benchmark		Sel	0	0	8.3	**70.1**	20.2	1.4	0	0	0			
	Model A	Par	Sel	0	0	9.0	**60.4**	28.1	2.5	0	0	0	0		
			Pts	0	3.0	6.7	20.6	12.5	2.2	0	0	0		11.3	45.0
		Fix	Sel	0	0	5.2	60.4	32.2	2.2	0	0	0	0		
			Pts	0	3.0	4.7	18.0	17.0	2.3	0	0	0		12.1	45.0
	Model B	Par	Sel	0	0	14.5	**56.9**	26.2	2.0	0	0	0	0.4		
			Pts	0	3.1	8.6	19.9	10.8	2.4	0	0	0		10.9	44.8
		Fix	Sel	0	0	10.9	54.9	32.1	1.7	0	0	0	0.4		
			Pts	0	3.1	7.0	17.5	14.7	2.5	0	0	0		11.6	44.8
	Model C	Par	Sel	0	0	20.9	**53.1**	23.2	2.2	0	0	0	0.6		
			Pts	0	3.2	10.3	18.6	9.8	2.6	0.2	0	0		10.0	44.7
		Fix	Sel	0	0	16.9	50.9	29.5	2.1	0	0	0	0.6		
			Pts	0	3.3	8.6	17.0	12.9	3.0	0	0	0		10.8	44.8
	Model D		Sel	0	0.4	42.2	**40.3**	13.9	2.2	0.3	0	0	0.7		
			Pts	0.4	4.0	17.3	14.6	5.0	3.0	0.3	0.1	0		9.1	44.7
3			pTox	0.03	0.05	0.10	0.16	**0.25**	0.32	0.40	0.48	0.55			
	Benchmark		Sel	0	0	1.1	19.5	**50.6**	23.7	4.8	0.3	0			
	Model A	Par	Sel	0	0	1.0	19.7	**55.6**	22.9	0.8	0	0	0		
			Pts	0	3.0	4.1	11.2	18.2	8.3	0.1	0	0		10.0	45.0
		Fix	Sel	0	0	0.4	17.6	60.9	21.1	0	0	0	0		
			Pts	0	3.0	3.4	8.8	21.2	8.6	0	0	0		10.0	45.0
	Model B	Par	Sel	0	0	1.5	20.3	**51.4**	24.9	1.0	0	0	0.8		
			Pts	0	3.0	4.6	11.6	15.8	9.2	0.4	0	0		9.9	44.6
		Fix	Sel	0	0	1.2	17.3	57.6	22.8	0.3	0	0	0.8		
			Pts	0	3.0	4.0	9.0	18.8	9.7	0.2	0	0		9.9	44.7
	Model C	Par	Sel	0	0.1	1.4	22.7	**45.8**	26.0	3.4	0.1	0	0.8		
			Pts	0	3.2	4.5	9.6	17.1	9.9	0.4	0	0		9.7	44.7
		Fix	Sel	0	0.1	2.0	19.1	54.4	22.9	0.6	0	0	0.8		
			Pts	0	3.2	4.5	9.6	17.1	9.9	0.4	0	0		9.7	44.7
	Model D		Sel	0	0.2	13.4	25.1	**32.9**	22.4	3.3	0.9	0.4	1.4		

Table S2 – *Continued.*

Scenario	Design	δ_{S_k}		% dose declared as MTD & average patient allocation										DLT	$\bar{N}$
				d_{i^*1} 2	d_{i^*2} 4	d_{i^*3} 8	d_{i^*4} 16	d_{i^*5} 22	d_{i^*6} 28	d_{i^*7} 40	d_{i^*8} 54	d_{i^*9} 70	None		
			Pts	0.4	3.9	9.8	10.8	8.6	8.2	1.9	0.5	0.3		8.8	44.4
4			pTox	0.001	0.005	0.03	0.10	0.16	**0.25**	0.38	0.50	0.60			
	Benchmark		Sel	0	0	0	1.2	22.2	**62.0**	14.4	0.2	0	0		
	Model A	Par	Sel	0	0	0	1.5	27.4	**65.6**	5.5	0	0	0		
			Pts	0	3.0	3.1	4.5	13.1	20.4	0.9	0	0		8.1	45.0
		Fix	Sel	0	0	0	0.8	33.3	65.0	0.9	0	0	0		
			Pts	0	3.0	3.0	3.7	13.5	21.7	0.1	0	0		8.1	45.0
	Model B	Par	Sel	0	0	0	1.9	29.1	**64.3**	4.3	0.4	0	0		
			Pts	0	3.0	3.1	5.0	12.0	20.3	1.5	0.1	0		8.2	45.0
		Fix	Sel	0	0	0	1.0	32.4	64.8	1.6	0.2	0	0		
			Pts	0	3.0	3.1	4.0	12.5	21.7	0.6	0	0.1		8.2	45.0
	Model C	Par	Sel	0	0	0.2	2.8	28.1	**63.8**	4.9	0.2	0	0		
			Pts	0	3.0	3.3	5.3	11.7	18.8	2.5	0.3	0.1		8.3	45.0
		Fix	Sel	0	0	0.1	2.1	30.4	64.6	2.6	0.2	0	0		
			Pts	0	3.0	3.2	4.5	11.8	21.1	1.2	0.1	0.1		8.3	45.0
	Model D		Sel	0	0	1.4	7.2	32.3	**50.6**	7.6	0.8	0.1	0		
			Pts	0	3.1	4.3	7.0	9.7	16.0	3.8	0.7	0.4		8.5	45.0
5			pTox	0.01	0.02	0.05	0.08	0.11	0.14	**0.25**	0.37	0.47			
	Benchmark		Sel	0	0	0	0.2	2.4	11.9	**67.2**	17.8	0.6	0		
	Model A	Par	Sel	0	0	0	0.3	4.5	48.4	**44.8**	2.0	0	0		
			Pts	0	3.0	3.1	4.1	7.7	20.9	6.1	0.2	0		5.9	45.0
		Fix	Sel	0	0	0	0.2	5.3	75.2	19.1	0.2	0	0		
			Pts	0	3.0	3.0	3.5	8.1	25.8	1.6	0	0		5.4	45.0
	Model B	Par	Sel	0	0	0.2	0.4	4.6	51.0	**38.1**	4.4	1.1	0.2		
			Pts	0	3.0	3.3	4.2	5.7	19.9	7.2	1.2	0.4		6.4	44.9
		Fix	Sel	0	0	0	0.2	5.1	62.8	25.9	4.2	1.6	0.2		
			Pts	0	3.0	3.1	3.6	6.2	23.2	4.5	0.7	0.6		6.1	44.9
	Model C	Par	Sel	0	0	0.2	0.5	4.5	50.3	**36.4**	6.5	1.4	0.2		
			Pts	0	3.0	3.3	4.2	5.2	18.3	8.4	1.8	0.6		6.7	44.8
		Fix	Sel	0	0	0	0.4	4.5	58.2	29.5	4.9	2.3	0.2		
			Pts	0	3.0	3.2	3.7	5.2	21.6	6.0	1.2	1.0		6.5	44.9
	Model D		Sel	0	0	0.7	1.1	6.9	39.6	**34.8**	13.2	3.5	0.3		
			Pts	0.2	3.3	4.4	4.7	3.5	13.5	9.6	3.4	2.4		7.7	45.0
6			pTox	0.003	0.006	0.01	0.02	0.05	0.08	0.15	**0.25**	0.37			
	Benchmark		Sel	0	0	0	0	0	0.4	18.3	**63.8**	17.6			
	Model A	Par	Sel	0	0	0	0	0	7.5	59.9	**30.8**	1.8	0		
			Pts	0	3.0	3.0	3.1	4.1	15.9	13.2	2.6	0.1		4.2	45.0
		Fix	Sel	0	0	0	0	0	31.6	62.3	5.2	0.9	0		
			Pts	0	3.0	3.0	3.0	4.2	24.9	6.7	0.2	0		3.4	45.0
	Model B	Par	Sel	0	0	0	0	0.2	10.3	44.5	**34.1**	10.9	0		
			Pts	0	3.0	3.0	3.1	1.7	13.6	11.8	6.0	2.8		5.4	45.0
		Fix	Sel	0	0	0	0	0.2	19.2	40.2	27.4	13.0	0		
			Pts	0	3.0	3.0	3.0	1.7	18.1	9.1	3.7	3.4		5.0	45.0
	Model C	Par	Sel	0	0	0	0	0.2	10.7	38.2	**38.3**	12.7	0		
			Pts	0	3.0	3.0	3.1	1.6	10.3	12.4	7.8	3.9		6.0	45.0
		Fix	Sel	0	0	0	0	0.2	14.8	38.5	32.2	14.2	0		
			Pts	0	3.0	3.0	3.0	1.2	14.6	9.9	5.4	4.9		5.7	45.0

Table S2 – *Continued.*

Scenario	Design	δ_{S_k}		% dose declared as MTD & average patient allocation										DLT	$\bar{N}$
				d_{i^*1} 2	d_{i^*2} 4	d_{i^*3} 8	d_{i^*4} 16	d_{i^*5} 22	d_{i^*6} 28	d_{i^*7} 40	d_{i^*8} 54	d_{i^*9} 70	None		
	Model D		Sel	0	0	0	0.2	1.8	8.3	30.4	**41.2**	18.1	0		
			Pts	0.1	3.1	3.1	3.3	1.0	7.3	10.9	8.5	7.7		7.1	45.0
7			pTox	**0.25**	0.42	0.60	0.75	0.82	0.88	0.91	0.94	0.97			
	Benchmark		Sel	**90.5**	9.5	0	0	0	0	0	0	0			
	Model A	Par	Sel	**42.3**	27.5	0.1	0	0	0	0	0	0	**30.2**		
			Pts	10.7	18.9	8.5	0.4	0	0	0	0	0		16.0	38.5
		Fix	Sel	40.7	24.7	0.2	0	0	0	0	0	0	34.4		
			Pts	10.3	16.9	9.3	1.3	0	0	0	0	0		16.2	37.8
	Model B	Par	Sel	**10.8**	3.8	0.1	0	0	0	0	0	0	**85.4**		
			Pts	4.1	6.5	3.4	0.4	0	0	0	0	0		6.1	14.4
		Fix	Sel	9.8	3.4	0	0	0	0	0	0	0	86.8		
			Pts	4.1	6.3	3.4	0.4	0	0	0	0	0		5.9	14.2
	Model C	Par	Sel	**6.6**	2.6	0	0	0	0	0	0	0	**90.8**		
			Pts	2.1	5.8	2.8	0.4	0	0	0	0	0		4.9	11.1
		Fix	Sel	6.8	2.1	0	0	0	0	0	0	0	91.1		
			Pts	2.2	5.7	2.7	0.4	0	0	0	0	0		4.9	11.0
	Model D		Sel	**8.6**	2.8	0	0	0	0	0	0	0	**88.6**		
			Pts	4.3	6.3	1.2	0.1	0	0	0	0	0		4.5	11.9
8			pTox	0.001	0.005	0.01	0.02	0.04	0.05	0.10	0.16	**0.25**			
	Benchmark		Sel	0	0	0	0	0	0	1.6	20.7	**77.8**			
	Model A	Par	Sel	0	0	0	0	0	2.0	41.2	47.9	**8.8**	0		
			Pts	0	3.0	3.0	3.1	3.8	13.1	14.2	4.5	0.3		3.1	45.0
		Fix	Sel	0	0	0	0	0	11.9	71.7	12.5	3.9	0		
			Pts	0	3.0	3.0	3.0	3.9	21.5	10.1	0.5	0		2.4	45.0
	Model B	Par	Sel	0	0	0	0	0.1	4.9	20.5	32.1	**42.4**	0		
			Pts	0	3.0	3.0	3.0	1.2	14.1	8.0	4.3	8.4		4.7	45.0
		Fix	Sel	0	0	0	0	0.1	7.6	21.4	26.9	44.0	0		
			Pts	0	3.0	3.0	3.0	1.2	14.1	8.0	4.3	8.4		4.4	45.0
	Model C	Par	Sel	0	0	0	0	0.1	4.6	15.0	31.5	**48.8**	0		
			Pts	0	3.0	3.0	3.1	1.1	7.4	9.2	8.0	10.2		5.3	45.0
		Fix	Sel	0	0	0	0	0.1	5.5	17.2	25.4	51.8	0		
			Pts	0	3.0	3.0	3.0	0.8	11.0	7.6	5.4	11.2		5.1	45.0
	Model D		Sel	0	0	0	0	0.6	3.5	9.0	28.2	**58.7**	0		
			Pts	0.1	3.0	3.1	3.2	0.5	5.2	6.9	6.7	16.3		6.2	45.0

pTox: true probability of toxicity in humans; **Sel**: proportion of times of declaring a dose as MTD; **Pts**: average number of patients allocated to a dose; **Par**: one variant of the meta-analytic model treating δ_{Dog} as a random variable; **Fix**: another variant of the meta-analytic model treating δ_{Dog} as a fixed constant, i.e., prior median of δ_{Dog} in our implementation.

References

1. Kamrin MA. Toxicology – A Primer on Toxicology Principles and Applications. *CRC Press*, 1988.

2. O'Quigley J, Paoletti X, Maccario J. Non-parametric optimal design in dose finding studies. *Biostatistics*. 2002;3(1):51-56.