Bayesian Dose-Finding in Two Treatment Cycles based on the Joint Utility of Efficacy and Toxicity

Peter F. Thall
Department of Biostatistics
University of Texas, M.D. Anderson Cancer Center

In the session "Multi-Stage Randomized Clinical Trials for the Development of Dynamic Treatment Regimes"

Annual Meeting of the Society for Clinical Trials, Boston

May 22, 2013

Collaborators

Juhee Lee : Model formulation, programming, heavy lifting - - and she laughs at all my jokes

Peter Mueller: Vast wisdom, Austro-Bayesian perspective

Yuan Ji: Knowledge of dose-finding and mathematics

Phase I-II Dose-Finding in Two Cycles

Goal: Develop a practical phase I-II trial design to adaptively optimize each patient's doses, d_1 and d_2, in two cycles of therapy, using binary (Y, Z) = (**Toxicity** , **Efficacy**) in each cycle.

Methodology: Base cycle-specific actions on numerical utilities, $U(y, z)$, for (y, z) = (0,0), (1,0), (0,1), or (1,1)

1. **Action in each cycle** : Treat with the "optimal" dose, or do not to treat, NT.
2. **Dose-outcome model**: Bayesian hierarchical
3. **Optimization of** (a_1, a_2): Apply Bellman (1957), using posterior means of a model-based objective function
4. **Safety**: Include additional dose acceptability rules

Doses, Outcomes, and Actions

Dose set $\{1, \cdots, m\}$ Action set $\mathscr{A} = \{1, \cdots, m\} \cup \{NT\}$

$d_{i,c}$ = dose, $Y_{i,c}$ = I(Toxicity), $Z_{i,c}$ = I(Efficacy) of pat. i in cycle c

$\boldsymbol{Y}_i = (Y_{i,1}, Y_{i,2}), \quad \boldsymbol{Z}_i = (Z_{i,1}, Z_{i,2}), \quad \boldsymbol{d}_i = (d_{i,1}, d_{i,2})$

$\mathscr{X}_t = \{(\boldsymbol{Y}_i, \boldsymbol{Z}_i, \boldsymbol{d}_i) : i = 1, ..., n_t\}$ = current data at trial time t

$a_{i,c}$ = action taken for patient i in cycle c = 1 or 2

$a_{i,1}$ maps $\mathscr{X}_t$ to $\mathscr{A}$ $\Rightarrow$ Adaptive between patients

$a_{i,2}$ maps $[\, \mathscr{X}_t, \text{cycle 1 data} = (d_{i,1}, Y_{i,1}, Z_{i,1}) \,]$ to $\mathscr{A}$

$\Rightarrow$ Adaptive both between and within patients

Actions versus Doses

Bellman's Idea: First find a_2^{opt} by considering all possibilities, then work backwards to find a_1^{opt}, assuming that a_2^{opt} will be taken.

Finding $\mathbf{a}^{opt} = (a_1^{opt}, a_2^{opt})$ is not the same thing as optimizing doses separately in each cycle, $\mathbf{d}^{opt} = (d_1^{opt}, d_2^{opt})$.

Example: (d_1^{opt}, d_2^{opt}) = (3,2), but $\mathbf{a}^{opt} = (3, a_2^{opt})$ with

a_2^{opt} = 3 if (Y_1, Z_1) = (0,1) = No Tox + Eff

a_2^{opt} = 4 if (Y_1, Z_1) = (0,0) = No Tox + No Eff

a_2^{opt} = 1 if (Y_1, Z_1) = (1,1) = Tox + Eff

$a_2^{opt} = NT$ if (Y_1, Z_1) = (1,0) = Tox + No Eff

Joint distribution for $[\boldsymbol{Y}_i, \boldsymbol{Z}_i \mid \boldsymbol{d}_i]$

An Ancient Technique: Use 4 continuous normal latent variables to define 4 discrete observed variables. Induces association among the discrete variables, facilitates posterior computation.

Real-valued, cycle-specific latent variables:

$\boldsymbol{\xi}_i = (\xi_{i,1}, \xi_{i,2})$ for $\boldsymbol{Y}_i$ and $\boldsymbol{\eta}_i = (\eta_{i,1}, \eta_{i,2})$ for $\boldsymbol{Z}_i$

Observables: $Y_{i,c} = \mathrm{I}(\xi_{i,c} > 0)$ and $Z_{i,c} = \mathrm{I}(\eta_{i,c} > 0) \Rightarrow$

$p(\boldsymbol{Y}_i, \boldsymbol{Z}_i \mid \boldsymbol{d}_i)$ is induced by $p(\boldsymbol{\xi}_i, \boldsymbol{\eta}_i \mid \boldsymbol{d}_i)$

$(\boldsymbol{\xi}_i, \boldsymbol{\eta}_i) \mid \boldsymbol{d}_i \sim$ multiv. normal, means vary with $\boldsymbol{d}_i$ and patient random effects, $(\boldsymbol{u_i}, \boldsymbol{v_i})$ that induce association

Priors of the Hierarchical Model

Level 1 Priors on the Latent Variables $(\boldsymbol{\xi}_i, \boldsymbol{\eta}_i)$

For patient i in cycle c given dose $d_{i,c} = d$,

$$\xi_{i,c} \mid \boldsymbol{u_i}, \bar{\xi}_{c,d}, \sigma^2_\xi \sim \mathrm{N}(\bar{\xi}_{c,d} + \boldsymbol{u_i},\ \sigma^2_\xi)$$

$$\eta_{i,c} \mid \boldsymbol{v_i}, \bar{\eta}_{c,d}, \sigma^2_\eta \sim \mathrm{N}(\bar{\eta}_{c,d} + \boldsymbol{v_i},\ \sigma^2_\eta)$$

Level 2 Priors on the Random Patient Effects $(\boldsymbol{u_i}, \boldsymbol{v_i})$

$$\boldsymbol{u_i}, \boldsymbol{v_i} \mid \rho,\ \tau^2 \overset{iid}{\sim} \mathrm{MVN}_2(\mathbf{0}_2, \Sigma_{\boldsymbol{u},\boldsymbol{v}})$$

where $\Sigma_{\boldsymbol{u},\boldsymbol{v}}$ has $\sigma^2_{\boldsymbol{u}} = \sigma^2_{\boldsymbol{v}} = \tau^2$ and covariances $\rho\tau^2$.

Priors of the Hierarchical Model

Level 2 Priors on the Mean Dose Effects $(\bar{\xi}, \bar{\eta})$

$$p(\bar{\xi}_{c,d}|\bar{\boldsymbol{\xi}}_{c,-d}) \propto \phi(\bar{\xi}_{c,d}|\xi_{c,0}, \sigma^2_{\xi_{c,0}}) 1(\bar{\xi}_{c,d-1} < \bar{\xi}_{c,d} < \bar{\xi}_{c,d+1})$$

$$p(\bar{\eta}_{c,d}|\bar{\boldsymbol{\eta}}_{c,-d}) \propto \phi(\bar{\eta}_{c,d}|\eta_{c,0}, \sigma^2_{\eta_{c,0}}) 1(\bar{\eta}_{c,d-1} < \bar{\eta}_{c,d} < \bar{\eta}_{c,d+1})$$

Level 2 prior means : $\boldsymbol{\xi}_0 = (\xi_{1,0}, \xi_{2,0}), \quad \boldsymbol{\eta}_0 = (\eta_{1,0}, \eta_{2,0})$

Level 2 prior variances : $\boldsymbol{\sigma}^2_{\xi_0} = (\sigma^2_{\xi_{1,0}}, \sigma^2_{\xi_{2,0}}), \boldsymbol{\sigma}^2_{\eta_0} = (\sigma^2_{\eta_{1,0}}, \sigma^2_{\eta_{2,0}})$

All fixed prior parameters : $\tilde{\boldsymbol{\theta}}^{12\times 1} = (\boldsymbol{\xi}_0, \boldsymbol{\eta}_0, \boldsymbol{\sigma}^2_{\xi_0}, \boldsymbol{\sigma}^2_{\eta_0}, \sigma^2_{\xi}, \sigma^2_{\eta}, \tau^2, \rho)$

Distributions of (ξ_i, η_i) and (Y, Z)

Latent variable distribution, after averaging over $(\boldsymbol{u_i}, \boldsymbol{v_i})$:

$$\boldsymbol{\xi}_i, \boldsymbol{\eta}_i \mid \boldsymbol{d}_i, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\eta}}, \tilde{\boldsymbol{\theta}} \overset{iid}{\sim} \mathrm{MVN}_4\left(\boldsymbol{\mu}_{\boldsymbol{d}_i}, \Sigma_{\xi,\eta}\right)$$

$\boldsymbol{\mu}_{\boldsymbol{d}_i} = (\bar{\xi}_{1,d_{i,1}}, \bar{\xi}_{2,d_{i,2}}, \bar{\eta}_{1,d_{i,1}}, \bar{\eta}_{2,d_{i,2}})$, depends on dose levels, but not numerical dose values. **Suppressing i, this induces**

$$p(\boldsymbol{y}, \boldsymbol{z} \mid \boldsymbol{d}, \boldsymbol{\theta}) = \Pr(Y_1 = y_1, Y_2 = y_2, Z_1 = z_1, Z_2 = z_2 \mid \boldsymbol{d}, \boldsymbol{\theta})$$

as a 4-dimensional integral of $\phi(\boldsymbol{\xi}, \boldsymbol{\eta} | \boldsymbol{\mu}_{\boldsymbol{d}}, \Sigma_{\xi,\eta})$ **with mean dose effects $\theta = (\bar{\xi}, \bar{\eta})$.**

Establishing a Prior

Obtain a non-informative prior $\pi(\boldsymbol{\theta} \mid \tilde{\boldsymbol{\theta}})$ by using prior expected sample size (ESS, Morita, et al., 2008) to calibrate $\tilde{\boldsymbol{\theta}}$

Prior informativeness, quantified by ESS, is implied by $\tilde{\boldsymbol{\theta}}$

Prototype Example: $Be(a, b)$ with mean $\mu = a/(a+b)$ and variance $\mu(1-\mu)/(a+b+1)$ has exact ESS= $a+b$.

More Complex Models: ESS is not at all obvious, so apply various algorithms to compute exact or approximate prior ESS

Prior Calibration Algorithm

1. Fix preliminary $\tilde{\boldsymbol{\theta}} - (\boldsymbol{\xi}_0, \boldsymbol{\eta}_0) = (\boldsymbol{\sigma}^2_{\xi_0}, \boldsymbol{\sigma}^2_{\eta_0}, \sigma^2_{\xi}, \sigma^2_{\eta}, \tau^2, \rho)$.
2. Solve for $(\boldsymbol{\xi}_0, \boldsymbol{\eta}_0)$ using the elicited prior probabilities for a set of doses and assumed hyperparameters.
3. Assume a vague, proper prior on $\boldsymbol{\theta}$ with large $\boldsymbol{\sigma}^2_{\xi_0}$, $\boldsymbol{\sigma}^2_{\eta_0}$.
4. Simulate a large pseudo-sample of $\boldsymbol{\theta} = (\bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\eta}})$ values
5. Compute several probabilities of interest, $q(\boldsymbol{\theta})$. E.g. $q(\boldsymbol{\theta}) = \Pr(Y_1 = 1 \mid d_1, \theta)$, $\Pr(Y_1 = 2 \mid d_1, Y_1, Z_1 \mid \theta)$, etc.
6. Approximate the distribution of the simulated sample of each $q(\boldsymbol{\theta})$ with a Be(a, b), match means and variances, and set ESS = $a + b$ for the prior of $q(\boldsymbol{\theta})$.
7. Calibrate $\tilde{\boldsymbol{\theta}}$ by repeating Steps 1–6 until ESS values in the range 0.5 to 2.0 are obtained for all $q(\boldsymbol{\theta})$.

Objective Function

Utilities $U(y_c, z_c)$ of outcomes in each cycle c = 1 or 2.

		Efficacy (Z_c)	
		No	Yes
Toxicity (Y_c)	No	35	100
	Yes	0	65

Mean utility of a_2, given $(d_1, Y_1, Z_1, \boldsymbol{\theta})$:

$$Q_2(a_2, d_1, Y_1, Z_1, \boldsymbol{\theta}) \;=\; E\{U(Y_2, Z_2) \mid a_2, d_1, Y_1, Z_1, \boldsymbol{\theta}\}$$

Cycle 2 Objective Function :

$$q_2(a_2, d_1, Y_1, Z_1, \mathcal{X}) \;=\; E\{Q_2(a_2, d_1, Y_1, Z_1, \boldsymbol{\theta}) \mid a_2, d_1, Y_1, Z_1, \mathcal{X}\}.$$

Objective Function

Mean utility of $a_1 = d_1$ given $\boldsymbol{\theta}$, assuming $a_2^{opt}(d_1, Y_1, Z_1, \mathcal{X})$ will be used in cycle 2 :

$$Q_1(a_1, \boldsymbol{\theta}) \;=\; Q_1(d_1, \boldsymbol{\theta}) \;=\; E\{U(Y_1, Z_1) \mid d_1, \boldsymbol{\theta}\}$$

Cycle 1 Objective Function : Given cycle 2 discount parameter $0 < \lambda < 1$, assume a_2^{opt} will be taken in cycle 2 :

$$\begin{aligned} q_1(d_1, \mathcal{X}) \;&= E\{Q_1(d_1, \boldsymbol{\theta}) \mid \mathcal{X}\} \\ &\quad + \lambda\, E\left[E\{q_2(a_2^{opt}, d_1, Y_1, Z_1, \mathcal{X}) \mid \boldsymbol{\theta}, d_1\} \mid d_1, \mathcal{X}\right] \end{aligned}$$

The optimal cycle 1 action, a_1^{opt}, is the dose maximizing $q_1(d_1, \mathcal{X})$, **unless $q_1(d_1, \mathcal{X}) < U(0, 0)$, which implies $a_1^{opt} = NT$ and stops the trial.**

Additional Dose Acceptability Criteria

Motivation: We do not fully trust our model and decision-making scheme, and want actual oncologists to use this methodology.

1. In each cycle, do not skip untried dose levels when escalating.
2. Do not escalate in cycle 2 if toxicity was observed in cycle 1: $d_2 \leq d_1$ if $Y_1 = 1$, regardless of Z_1.
3. d_1 is unacceptable if $q_1(d_1, \mathcal{X}) < U(0, 0)$
4. d_2 is unacceptable if $q_2(d_2, d_1, Y_1, Z_1, \mathcal{X}) < U(0, 0)$

$\mathcal{A}_1(\mathcal{X})$ = acceptable doses for cycle 1

$\mathcal{A}_{i,2}(d_{i,1}, Y_{i,1}, Z_{i,1}, \mathcal{X})$ = acceptable doses for cycle 2 of patient i

Adaptive Randomization

Motivation: Greedy adaptive algorithms that always take the optimal action risk getting stuck at a suboptimal action.

Define ϵ_i decreasing $\downarrow$ in patient index i, with $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_n)$.

Cycle 1 : The set of ϵ_i-optimal doses for cycle 1 is

$$\{d_1 \in \mathcal{A}_{i,1}(\mathcal{X}) \ : \ q_1(d_1, \mathcal{X}) > q_1(a_{1,i}^{opt}, \mathcal{X}) - \epsilon_i, \}.$$

In words: An ϵ_i-optimal dose has posterior mean payoff near the maximum (optimal) payoff.

Adaptive Randomization

Cycle 2 : The set of $(\epsilon_i/2)$-optimal doses for cycle 2 given $(d_{i,1}, Y_{i,1}, Z_{i,1})$ is all acceptable d_2 having posterior mean utility within $\epsilon_i/2$ of the maximum.

We use $\epsilon_i/2$ because $q_2(a_2, d_1, Y_1, Z_1, \mathcal{X})$ is the posterior expected utility for cycle 2 only.

AR(ϵ)

Randomize the i^{th} patient among the ϵ_i-optimal doses in cycle 1 and $\epsilon_i/2$-optimal doses in cycle 2.

The DTM2 Design: Trial Conduct

For the simulation study, we mimicked a typical phase I-II chemotherapy trial with 5 dose levels, but accounting for 2 cycles of therapy. Maximum sample size = 60, cohort size = 2.

Based on preliminary simulations, we set

ϵ_i = 20 for the first 10 patients

ϵ_i = 15 for the next 10 patients

ϵ_i = 10 for the last 40 patients.

The DTM2 Design: Trial Conduct

1. Cohort 1 is treated at $d = 1$ in cycle 1, their (Y_1, Z_1) are observed, posterior($\boldsymbol{\theta} \mid data$) is computed, and actions taken for cycle 2. When (Y_2, Z_2) are observed from cycle 2, posterior($\boldsymbol{\theta} \mid data$) is updated.

2. Cohort 2 is enrolled after cohort 1 has been evaluated for cycle 1.

3. For cohorts 2, 3, ... , $q_1(d_1, \mathcal{X})$ is computed using $\lambda = 0.8$, $\mathbf{a}^{opt}$ is identified, and AR($\boldsymbol{\epsilon}$) is applied to choose $\mathbf{a}$.

4. Steps 1 – 3 are repeated until the trial is stopped early, or N = 60 is reached and a final optimal two-cycle $\boldsymbol{a}_{sel}$ is chosen.

2-Cycle Comparators

We compared the DTM2 design to 2-cycle extensions of 3+3 algorithms and the continual reassessment method (CRM)

(3+3)a implicitly targets d with $\mathrm{P}(Y_1 = 1 \mid d) \leq 0.17$

(3+3)b implicitly targets d with $\mathrm{P}(Y_1 = 1 \mid d) \leq 0.33$

The extended (3+3) methods both choose d_2 as follows:

If Y_1 = 1, then $d_2 = d_1 - 1$. (Tox in cycle 1 $\Rightarrow$ de-escalate)

If Y_1 = 0, then $d_2 = d_1$. (No Tox in cycle 1 $\Rightarrow$ repeat d_1)

2-Cycle Comparators

CRM model $\Pr(Y_1 = 1|d) = p_d^{\exp(\alpha)}$, with $\alpha \sim \mathrm{N}(0, 2)$ choosing $p_1 < \cdots < p_5$ using "getprior" in "dfcrm" (Cheung, 2011)

Cycle 1 : Each patient's d_1 was chosen
(i) to have posterior mean Pr(Tox) closest to 0.30
(ii) imposing the "do not skip an untried dose" rule.

Cycle 2 : d_2 was chosen using the same rule as for the extended (3+3) methods. Also, d_2 was unacceptable if

$$\Pr\{\Pr(Y_1 = 1 \text{ or } Y_2 = 1) > 0.50|\mathcal{X}, \boldsymbol{d}\} > 0.90.$$

In words: A cycle 2 dose is too high if it makes it likely that the probability of at least one toxicity in two cycles is over 50%.

Simulation Study: Evaluation Criteria

Empirical mean total utility for n patients in the trial :

$$\bar{U} = \sum_{i=1}^{n} \{U(Y_{i,1}, Z_{i,1}) + U(Y_{i,2}, Z_{i,2})\}/n$$

setting $U(Y_{i,2}, Z_{i,2}) = U(0, 0)$ if $a_{i,2} = NT$.

Empirical mean total payoff for all patents :

$$\bar{\bar{U}} = \frac{1}{N} \sum_{r=1}^{N} \bar{U}^{(r)}.$$

indexing simulated trials by $r = 1, ..., N$. This quantifies ethical desirability, given $U(y, z)$.

Simulation Study: Evaluation Criteria

Expected payoff of $a_{1,sel}$, given θ^{true} :

$$Q_{1,sel}(a_{1,sel}) \;=\; E\{U(Y_1, Z_1) \mid a_{1,sel}, \boldsymbol{\theta}^{true}\},$$

Expected payoff of $a_{2,sel}$:

$$Q_{2,sel}(a_{2,sel}) \;=\; E\{U(Y_2, Z_2) \mid a_{1,sel}, a_{2,sel}(y_1, z_1), y_1, z_1, \boldsymbol{\theta}^{true}\}$$

where $E\{U(Y_2, Z_2) \mid a_{1,sel}, a_{2,sel}(y_1, z_1), y_1, z_1, \boldsymbol{\theta}^{true}\} = U(0, 0)$ if $a_{2,sel}(y_1, z_1) = NT$.

Expected payoff to a future patient treated using a_{select} :

$$Q_{sel}(a_{sel}) \;=\; Q_{1,sel}(a_{1,sel}) + \lambda Q_{2,sel}(a_{2,sel}).$$

Simulation Study: Evaluation Criteria

Denote $\delta_{i,2}$ = 1 if patient i received $a_{i,2} = d_2$, and $\delta_{i,2}$ =0 if $a_{i,2} = NT$.

Empirical Toxicity and Efficacy Probabilities

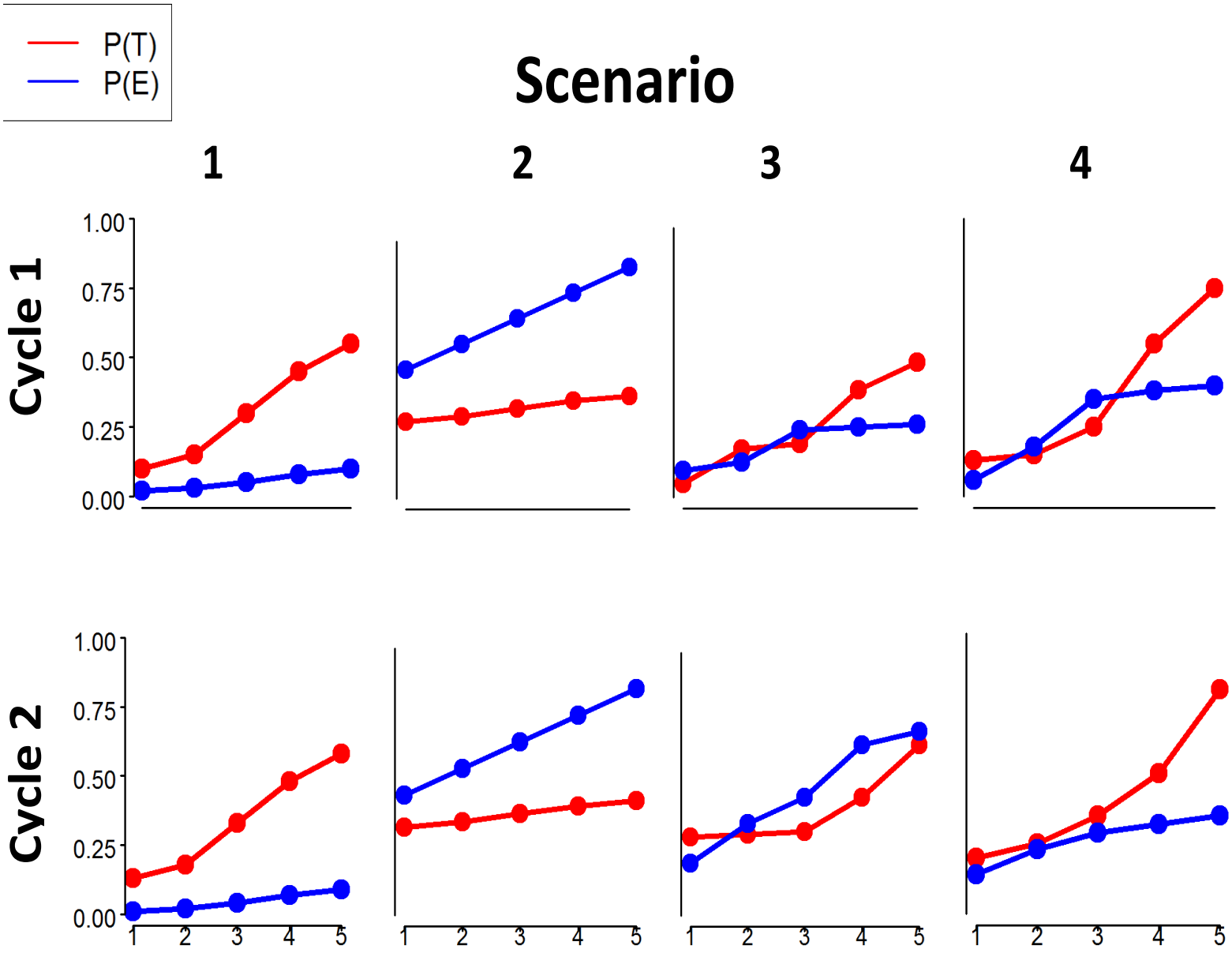
$$\Pr(\text{Tox}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1(Y_{i,1}=1)+\delta_{i,2}1(Y_{i,2}=1)}{1+\delta_{i,2}}$$

$$\Pr(\text{Eff}) = n^{-1}\sum_{i=1}^{n}\frac{1(Z_{i,1}=1)+\delta_{i,2}1(Z_{i,2}=1)}{1+\delta_{i,2}}$$

Simulation Scenarios

True $(p_T, p_E)^{true}$ under the four simulations scenarios.

Scenario	Cycles	Doses				
		1	2	3	4	5
1	1	(0.10, 0.02)	(0.15, 0.03)	(0.30, 0.05)	(0.45, 0.08)	(0.55, 0.10)
	2	(0.13, 0.01)	(0.18, 0.02)	(0.33, 0.04)	(0.48, 0.07)	(0.58, 0.09)
2	1	(0.30, 0.50)	(0.32, 0.60)	(0.35, 0.70)	(0.38, 0.80)	(0.40, 0.90)
	2	(0.33, 0.45)	(0.35, 0.55)	(0.38, 0.65)	(0.41, 0.75)	(0.43, 0.85)
3	1	(0.05, 0.10)	(0.18, 0.13)	(0.20, 0.25)	(0.40, 0.26)	(0.50, 0.27)
	2	(0.30, 0.20)	(0.31, 0.35)	(0.32, 0.45)	(0.45, 0.65)	(0.65, 0.70)
4	1	(0.13, 0.06)	(0.15, 0.18)	(0.25, 0.35)	(0.55, 0.38)	(0.75, 0.40)
	2	(0.20, 0.14)	(0.25, 0.23)	(0.35, 0.29)	(0.50, 0.32)	(0.80, 0.35)


P(T)
P(E)
Scenario
1
2
3
4
Cycle 1
Cycle 2
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5

Scenario Descriptions

Scenario 1: $d = 1, 2, 3$ **safe,** $d = 3, 4$ **too toxic,** but all doses have low efficacy $\Rightarrow$ **The optimal action is to stop the trial early**.

Scenario 2: $0.30 \leq p_T(d_1) \leq 0.40$ **and** $.33 \leq p_T(d_2) \leq 0.43$ **with very high** $p_E(d_1)$ **and** $p_E(d_2) \Rightarrow$ **Big payoff for escalating to higher doses**.

Scenario 3: a_1^{opt} = 3, with $a_2^{opt}(Y_1 = 0)$ = 4 and $a_2^{opt}(Y_1 = 1)$ = 2 $\Rightarrow$ **Accounting for** Y_1 **is very important when choosing** d_2

Scenario 4: **The optimal doses happen to coincide with what the 3+3 and CRM choose while ignoring efficacy. "A stopped clock is right twice a day."**

Optimal Actions Under the 4 Scenarios

Scenario	a_1^{opt}	a_2^{opt}			
		(0,0)	(0, 1)	(1,0)	(1,1)
1	NT	NT	NT	NT	NT
2	5	5	5	4	4
3	3	4	4	2	2
4	3	3	3	NT	2

Percent Completed Trials

Scenarios	DTM2	(3+3)a	(3+3)b	Extended CRM
1	**2.3**	88.6	96.5	99.8
2	**99.4**	39.2	64.7	93.1
3	79.4	99.6	99.2	99.8
4	96.7	83.2	94.7	99.8

In Scenario 1, DTM2 finds that all doses are inefficacious, and stops the trial 97.7% of the time. The other 3 methods all ignore efficacy and continue to treat patients.

Summary of Simulation Results

Scenario		DTM2	(3+3)a	(3+3)b	Extended CRM
2	$\bar{U}$	**136.35**	123.32	117.78	116.38
	Q_{select}	**135.76**	103.85	104.48	103.17
	Pr(Tox)	**0.39**	0.30	0.34	0.37
	Pr(Eff)	**0.72**	0.57	0.55	0.57
3	$\bar{U}$	**94.23**	85.72	85.49	89.29
	Q_{select}	**84.39**	77.97	80.13	78.29
	Pr(Tox)	0.38	0.27	0.28	0.30
	Pr(Eff)	0.38	0.26	0.27	0.33
4	$\bar{U}$	**75.84**	81.83	79.85	84.88
	Q_{select}	**69.49**	74.92	75.76	79.12
	Pr(Tox)	0.51	0.25	0.27	0.29
	Pr(Eff)	0.29	0.22	0.21	0.29

Some Conclusions

1: A "Bayes Meets Bellman" dose-finding method, adaptive between and within patients, with lots of model-method-computing machinery.

2: Very good properties compared to "toxicity only" 2-cycle methods – but not in all cases.

3: Are 2 cycles better than 1 ? Sometimes Yes, sometimes No.

4: Very sensitive to $U(y, z)$ $\Rightarrow$ It is essential to calibrate $U(y, z)$ with the physician/scientists when designing an actual trial.

5: $U(y, z)$ is not enough! Additional safety/efficacy restrictions are needed.