* Ordinal Toxicity Intervals (BCLM 3.2.5)
* Neuenschwander et al (2008 StatMed)
  - The assumption that a single target toxicity exits and can be reliably identified by the investigator is *unrealistic*.
  - **Toxicity probability intervals:** Replace unrealistic *single* target toxicity by intervals for $p_i = Pr(\text{Tox at dose } i)$, $i = 1, \ldots, m$.
  - Let $\bar{p}_i$ denote the posterior probability of a DLT at dose i. Classify the probability of a DLT into four categories;
    - ** Under-dosing: $\bar{p}_i \in [0, 0.20]$ (I_U)
    - ** Targeted toxicity: $\bar{p}_i \in (0.20, 0.35]$ (I_T)
    - ** Excessive toxicity: $\bar{p}_i \in (0.35, 0.60]$ (I_E)
    - ** Unacceptable toxicity: $\bar{p}_i \in (0.60, 1.00]$ (I_A)

† Statistical Model

- We assume a dose grid, $d = (d_1, d_2, \ldots, d_7)$ with $d^o = d_7$ as a reference dose and $m = 7$.
- **Sampling Model:** probit regression

$$\pi(d) = P(y_i = 1 \mid d_i) = 1 - \Phi\left(-\log(a) - b\log(d_i/d^o)\right).$$

- **Prior:** $(\log(a), \log(b)) \sim \mathrm{N}(\mu, \Sigma)$
  - The original paper has a logistic regression.
  - typo! the book has prior on b, not $\log(b)$.
- **Posterior:**

$$p(a, b \mid \text{data}) \propto p(\log(a), \log(b)) \prod_{i=1}^{n} (\pi(d_i))^{y_i} (1 - \pi(d_i))^{1-y_i}$$

† How to use the posterior to choose a dose for the next cohort?

- **Dose escalation:** based on posterior probabilities $P(p_i \in I_X \mid \mathbf{y})$, $X \in \{U, T, E, A\}$

dose level $i = 1, \ldots, m\ (=7)$

- **Dose:** Let $\mathcal{D} = \{i : p_i(I_E) + p_i(I_A) < 0.25\}$ (safety guard).

  ⋆⋆ Assign for the next patient (cohort) $d^* = d_{i^*}$ with

$$i^* = \arg\max_{i \in \mathcal{D}} \{p_i(I_T)\}$$

  and $i^* = \emptyset$ if no dose is found with $p_i(I_E) + p_i(I_A) < 0.25$.

- **Stop:** if $n > n_{\max}$ or $i^* = \emptyset$

$*$ **Example 3.7:** Dose escalation cancer trial–Neuenschwander et al (2008 StatMed)

– We implement the algorithm to find the MTD for a new cancer drug.

- **Dose grid:** $d = (12.5, 25, 50, 100, 150, 200, 250)$ with $d^o = d_7 = 250$ as a reference dose

- **Prior:** $(\log(a), \log(b)) \sim \mathrm{N}(\mu, \Sigma)$ with $\mu = (2.15, 0.52)$ and $\mathrm{Var}(\log(a)) = 0.84^2$, $\mathrm{Var}(\log(b)) = 0.8^2$ and $\mathrm{Corr}(\log(a), \log(b)) = 0.2$.

  The moments are chosen to best match the 2.5%, 50%, and 97.5% quantiles for the toxicity probabilities that are implied by a one-parameter CRM model.

* **Example 3.7:** – contd

  - **Step 0.** Initialize

    ⋆⋆ Maximum sample size $n_{\max} = 30$ and cohort size $k = 3$. Start with $d = 1$.

    ⋆⋆ Target toxicity interval: $[0.20, 0.35]$

    ⋆⋆ **Simulation truth:** $p^o = (0.11, 0.27, 0.52, 0.76, 0.84, 0.90, 0.94)$.

    ⋆⋆ Under p^o, dose d_2 is the maximum dose with true toxicity within the target interval.

  - **Step 1.** Next cohort: Record responses $y \sim \text{Bin}(k, p^o_d)$. Increment $n \equiv n + k$.

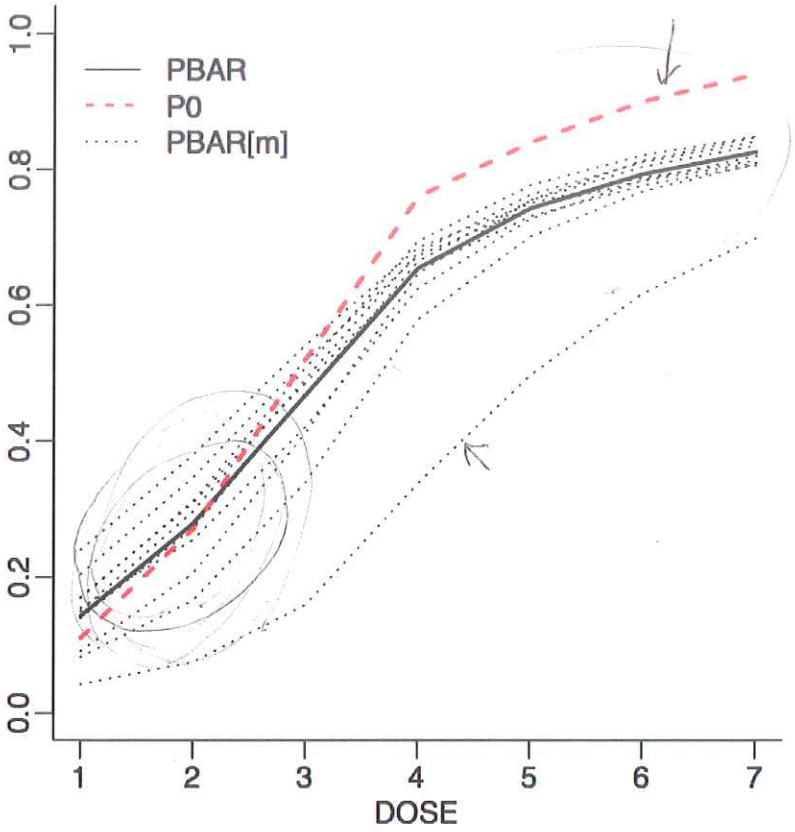
  - **Steps 2 & 3.** Update posterior and find the next dose $d^\star = d_{i^\star}$.

$*$ **Example 3.7:** – contd

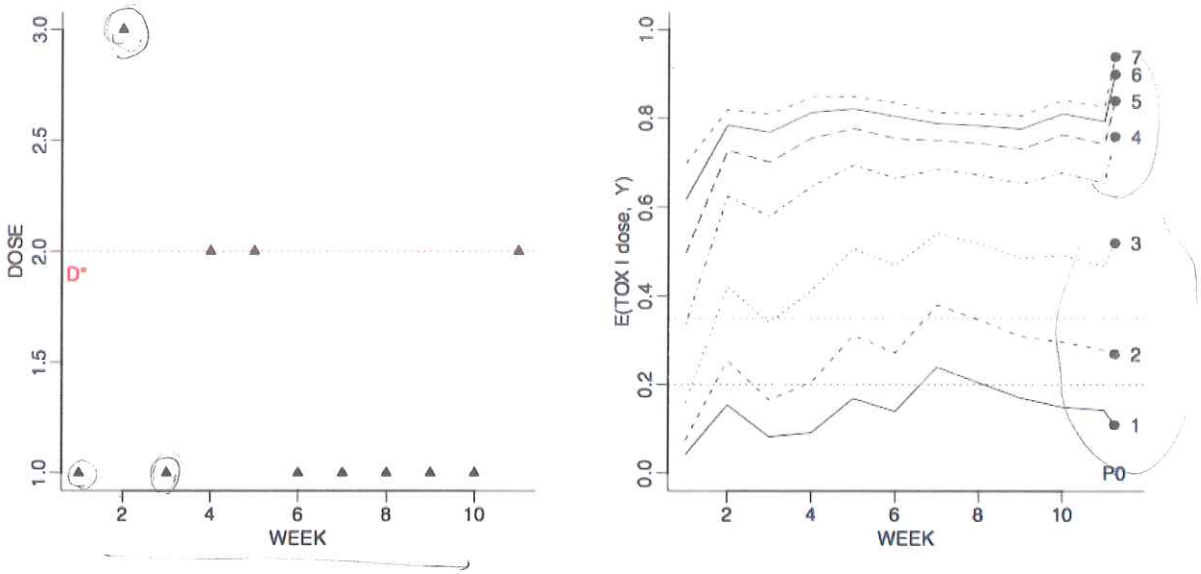
- **Step 4.** Stopping: if $n > n_{\max}$ or $i^\star = \emptyset$. Report the last assigned dose $d = d^\star$ as optimal dose. Otherwise continue with Step 1.

* Example - Results

** Assumed truth p^o (dashed line) and posterior estimated toxicity $E(p_i \mid \mathbf{y})$ at the end of a simulated trial history (soild line, each of 11 cohorts).



* Example - Results (contd)



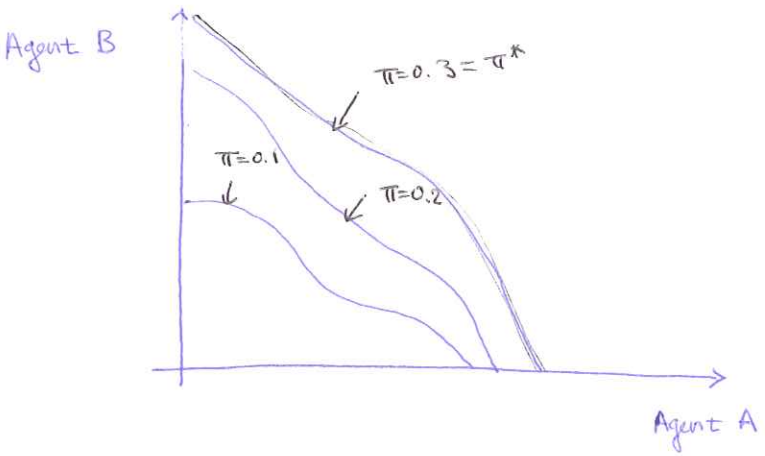
d_i $E(p_i \mid y)$ across weeks

** Assigned doses d_i (left panel) and estimated toxicity probabilities $\bar{p}_i = \mathrm{E}(p_i \mid \mathbf{y})$ against weeks (panel b). For comparison, the final point shows p^o.

$*$ Combination Therapy: 2 Agents (BCLM 3.4.4)

$\star\star$ Thall et al (2004, Biometrics): Phase I oncology trial for the combination of two cytotoxic agents.

- So far dose-finding for a *single* drug
- Investigate the therapeutic effect of *multiple* drugs used in combination. e.g. two agents, A and B.
- Need to discover the dose *combination* (A_j, B_k) having probability of dose-limiting toxicity (DLT) no larger than some prespecified limit $\pi^\star \in (0, 1)$.
- **(strategy)** Consider a contour of MTD values in two-dimensional dose space.



* Statistical Model

- **Data:** binary toxicity response $y_i \in \{0, 1\}$; standardized dose $\mathbf{x}_i = (x_{1i}, x_{2i})$ for agents.

- **Model:** Assume

$$P(y = 1 \mid \mathbf{x} = (x_1, x_2), \theta) = \pi(x_1, x_2, \theta).$$

** Consider a 6-parameter bivariate logistic model – parsimonious & sufficiently flexible),

$$\pi(x_1, x_2, \theta) = \frac{a_1 x_1^{b_1} + a_2 x_2^{b_2} + a_3 (x_1^{b_1} x_2^{b_2})^{b_3}}{1 + a_1 x_1^{b_1} + a_2 x_2^{b_2} + a_3 (x_1^{b_1} x_2^{b_2})^{b_3}},$$

where $\theta = (a_j, b_j, j = 1, 2, 3)$ with $a_j > 0$ and $b_j > 0$ for all j.

* Statistical Model – contd

- **Prior:** independent gamma prior $a_j \sim \text{Gamma}(\alpha_{1j}, \alpha_{2j})$ and $b_j \sim \text{Gamma}(\beta_{1j}, \beta_{2j})$, $j = 1, 2$

  Consider $\log(a_3) \sim \text{N}(\mu_{a3}, \sigma^2_{a3})$ and $\log(b_3) \sim \text{N}(\mu_{b3}, \sigma^2_{b3})$.

- **Posterior:** $\theta = (a_1, a_2, a_3, b_1, b_2, b_3)$

$$p(\theta \mid \mathbf{x}, \mathbf{y}) \propto \prod_{i=1}^{n} \pi(\mathbf{x}_i, \theta)^{y_i} \{1 - \pi(\mathbf{x}_i, \theta)\}^{1-y_i} p(\theta).$$

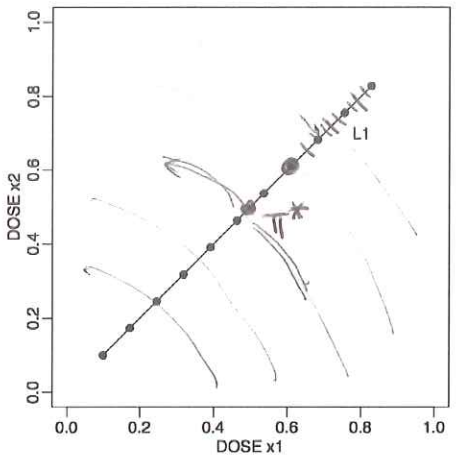
$$\Rightarrow \bar{\pi}_n(\mathbf{x}) = \text{E}(\pi(\mathbf{x}, \theta) \mid \mathbf{y}_n) = \int \pi(\mathbf{x}, \theta) dp(\theta \mid \mathbf{y}_n).$$

$*$ **Two-stage dose finding algorithm:** climb up to reach target toxicity; then adjust dose combinations, keeping toxicity unchanged.

- Stage 1: Dose escalation on a grid of (x_1, x_2) until target toxicity π^* is reached.
- Stage 2: Maintaining toxicity, adjust dose combination to maximize cancer-cell killing and learning.

$**$ We do this based on the currently estimated posterior expected toxicity surface.

* Stage 1 – initial dose escalation



** Define grid $D_1 = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(R)})$ on a fixed line segment.

** Start with the lowest dose $\mathbf{x}^{(1)}$.

* figure from P. Müller.

- Let $\bar{\pi}_n$ = posterior mean toxicity surface after n-th cohort.
- Assign $(n+1)-$st cohort at:

$$\mathbf{x} = \arg \min_k |\bar{\pi}_n(\mathbf{x}^{(k)}) - \pi^*|,$$

subject to escalation control (not skipping any untried dose in D_1).

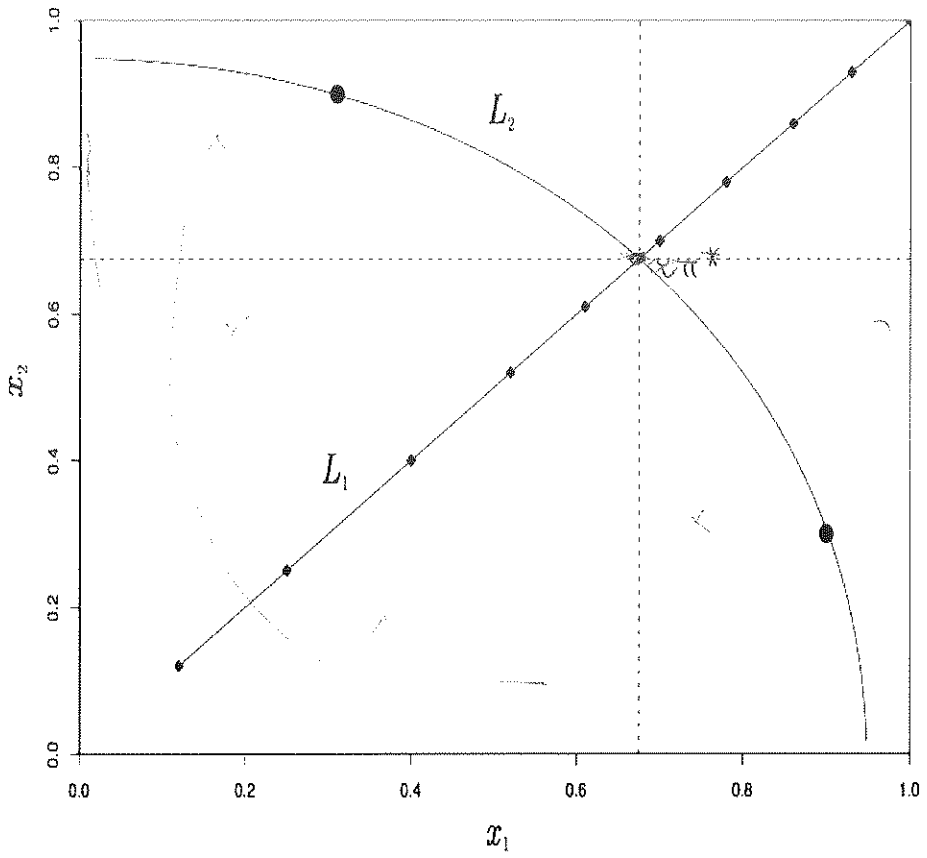
- When the first toxicity is observed at $\mathbf{x}^{(r)}$, refine the grid $\mathbf{x}^{(r)}, \ldots, \mathbf{x}^{(R)}$ by introducing half-steps above $\mathbf{x}^{(r)}$, $\frac{1}{2}(\mathbf{x}^{(k)} + \mathbf{x}^{(k+1)})$, $k = r, \ldots, R-1$ $\Rightarrow$ Consider doses in D_1 and newly defined midway doses.
- Stop after n_1 patients.

* **Stage 2:** explore at equal toxicity level
  - Let L_{2n} = set of doses with $\pi_n(\mathbf{x}) \approx \pi^*$
  - $L_{2n} = L_{2n}^{left} \cup L_{2n}^{right}$ where

    ⋆⋆ $L_{2n}^{left} = \{\mathbf{x} \in L_{2n} \text{ and } x_2 > x_1\}$: the segment of L_{2n} above the 45 degree line.

    ⋆⋆ $L_{2n}^{right} = \{\mathbf{x} \in L_{2n} \text{ and } x_1 > x_2\}$: the segment of L_{2n} below the 45 degree line.
  - Alternate between dose in L_{2n}^{left} and L_{2n}^{right}.
  - Each cohort, assign dose in $L_{2n}^{left/right}$ that maximizes expected learning and cancer cell killing.

* **Stage 2:** explore at equal toxicity level



* figure from P. Müller.

* **Stage 3:** final recommendation

  - Let $\mathbf{x}_\ell^\star$ denote the optimal dose pair in $L = L_2^{left}$. Similarly, $\mathbf{x}_r^\star$ denote the optimal dose combination in $L = L_2^{right}$, and let $\mathbf{x}_m^\star$ denote the optimal dose pair on $L = L_1$.

    Report $\{\mathbf{x}_\ell^\star, \mathbf{x}_r^\star, \mathbf{x}_m^\star\}$ as three alternative MTD dose pairs.

$n_1 = 20$, $n_2 = 40$ ($\Rightarrow N = 60$)

$\pi^* = 0.3$

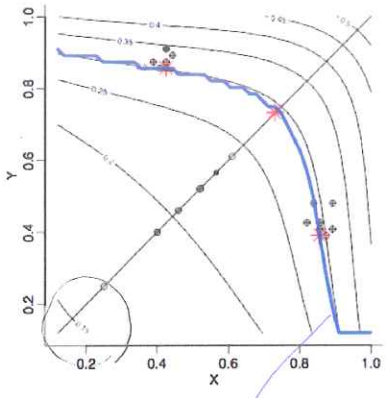
cohort size = 2

* **Example 3.9** - Gemcitibine + Cyclophosphamide
  - Escalation on grid over line segment.
  - Adjust dose combination.

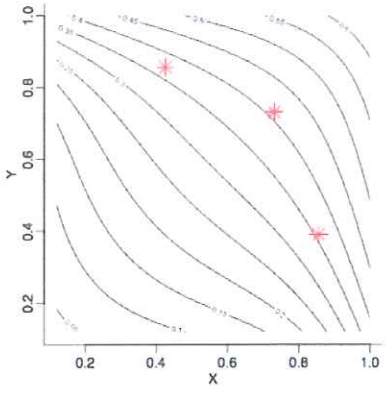
$\hat{\pi} = 0.2$ $\quad \pi^{TRUE} = 0.3$ $\quad$ d

d s.t $\hat{\pi} = 0.3$ $\quad \pi^* = 0.3$

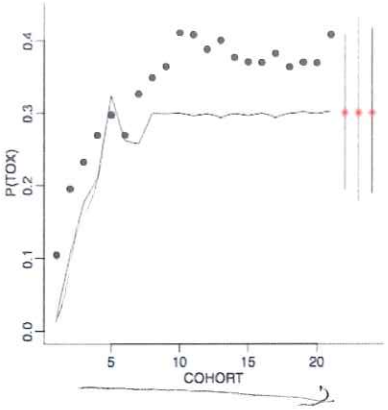
* : MTD



posterior estimate



simulation truth



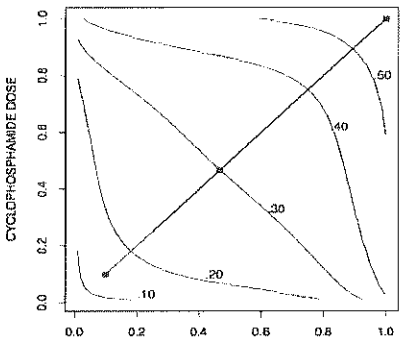
estimated toxicity at allocated dose

Estimated π^*

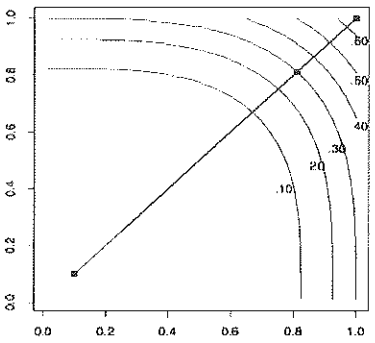
Toxicity Curve

* Simulation Scenarios

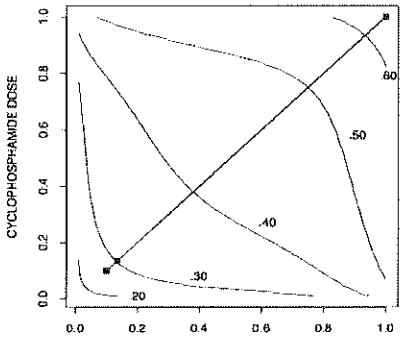
SCENARIO 1



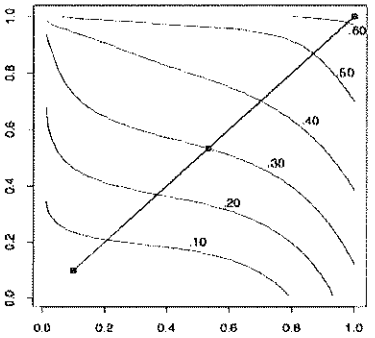
SCENARIO 2



SCENARIO 3

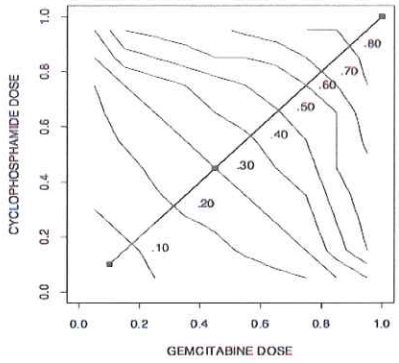


SCENARIO 4

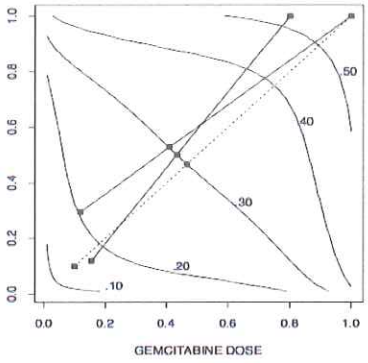


* Simulation Scenarios–contd

SCENARIO 5



SCENARIO 1*



Simulation Truth different from the modeling assumption

What if we change L_1?

* Simulation Results: dose combinations selected for the Gem/CTX (SD as subsripts)

	$\mathbf{x}_{n_1}^{\text{middle}}$	$\mathbf{x}_N^{\text{middle}}$	$\mathbf{x}_N^{\downarrow\text{right}}$	$\mathbf{x}_N^{\uparrow\text{left}}$
		Scenario 1		
Selected dose $\mathbf{x}$	$(0.47_{0.16}, 0.47_{0.16})$	$(0.50_{0.10}, 0.50_{0.10})$	$(0.66_{0.11}, 0.24_{0.19})$	$(0.34_{0.13}, 0.67_{0.10})$
$\bar{\pi}_n(\mathbf{x})$, true $\pi(\mathbf{x})$	$0.30_{0.05}$, $0.30_{0.06}$	$0.30_{0.01}$, $0.31_{0.04}$	$0.30_{0.01}$, $0.28_{0.07}$	$0.30_{0.01}$, $0.32_{0.04}$
$\text{Var}\{\pi(\mathbf{x}) \mid \mathbf{Z}_n\} \times 10^2$	$0.82_{0.32}$	$0.33_{0.12}$	$0.43_{0.19}$	$0.37_{0.18}$
		Scenario 2		
Selected dose $\mathbf{x}$	$(0.81_{0.04}, 0.81_{0.04})$	$(0.80_{0.03}, 0.80_{0.03})$	$(0.92_{0.05}, 0.72_{0.05})$	$(0.58_{0.12}, 0.86_{0.05})$
$\bar{\pi}_n(\mathbf{x})$, true $\pi(\mathbf{x})$	$0.30_{0.01}$, $0.30_{0.08}$	$0.30_{0.01}$, $0.28_{0.05}$	$0.29_{0.01}$, $0.35_{0.05}$	$0.30_{0.01}$, $0.22_{0.05}$
$\text{Var}\{\pi(\mathbf{x}) \mid \mathbf{Z}_n\} \times 10^2$	$0.78_{0.14}$	$0.29_{0.03}$	$0.41_{0.10}$	$0.38_{0.11}$
		Scenario 3		
Selected dose $\mathbf{x}$	$(0.28_{0.14}, 0.28_{0.14})$	$(0.23_{0.17}, 0.23_{0.17})$	$(0.46_{0.21}, 0.07_{0.17})$	$(0.10_{0.14}, 0.40_{0.22})$
$\bar{\pi}_n(\mathbf{x})$, true $\pi(\mathbf{x})$	$0.37_{0.07}$, $0.36_{0.05}$	$0.30_{0.02}$, $0.33_{0.08}$	$0.29_{0.02}$, $0.29_{0.08}$	$0.30_{0.02}$, $0.31_{0.08}$
$\text{Var}\{\pi(\mathbf{x}) \mid \mathbf{Z}_n\} \times 10^2$	$0.94_{0.28}$	$0.40_{0.22}$	$0.48_{0.23}$	$0.44_{0.26}$
		Scenario 4		
Selected dose $\mathbf{x}$	$(0.57_{0.12}, 0.57_{0.12})$	$(0.56_{0.07}, 0.56_{0.07})$	$(0.70_{0.10}, 0.38_{0.15})$	$(0.39_{0.10}, 0.68_{0.08})$
$\bar{\pi}_n(\mathbf{x})$, true $\pi(\mathbf{x})$	$0.29_{0.05}$, $0.32_{0.07}$	$0.30_{0.01}$, $0.32_{0.04}$	$0.30_{0.005}$, $0.26_{0.08}$	$0.30_{0.01}$, $0.34_{0.04}$
$\text{Var}\{\pi(\mathbf{x}) \mid \mathbf{Z}_n\} \times 10^2$	$0.75_{0.29}$	$0.35_{0.09}$	$0.43_{0.16}$	$0.39_{0.15}$
		Scenario 5		
Selected dose $\mathbf{x}$	$(0.47_{0.11}, 0.47_{0.11})$	$(0.45_{0.09}, 0.45_{0.09})$	$(0.67_{0.13}, 0.17_{0.16})$	$(0.29_{0.13}, 0.63_{0.11})$
$\bar{\pi}_n(\mathbf{x})$, true $\pi(\mathbf{x})$	$0.31_{0.03}$, $0.33_{0.10}$	$0.30_{0.01}$, $0.32_{0.10}$	$0.30_{0.01}$, $0.27_{0.13}$	$0.30_{0.01}$, $0.32_{0.10}$
$\text{Var}\{\pi(\mathbf{x}) \mid \mathbf{Z}_n\} \times 10^2$	$0.92_{0.31}$	$0.38_{0.14}$	$0.50_{0.22}$	$0.43_{0.21}$

* Summary

- **Phase I:** posterior inference for dose/toxicity curve. Can accommodate essentially arbitrary constraints, structure and goals.
- **Delayed response:** TITE-CRM
- **Target toxicity:** target dose or dose interval
- **Multiple agents:** replace dose/response curve by surface. Same for dose & schedule etc.

- Remove monotocity

cytotoxic agents : usually monotone increasing

biological agents : non-monotonic pattern.

† Do we want to consider toxicity only in a phase I trial? Efficacy vs Toxicity

- After all, drug doses are acceptable only if they are safe **and** efficacious.
- Recently, increasing interest and effort in developing dose finding methods incorporating both toxicity and efficacy endpoints.
- This is called a "phase I-II design" since it combines the goals of conventional phase I and II studies.
- How to consider them together? joint model and define the acceptable doses based on the a trade-off between efficacy and toxicity.
- Read BCLM Section 3.3 for more.

AMS 276
Lecture 9: Phase II Studies

Fall 2016

† What is coming?

We are in DC Chapter 3 and BCLM Chapter 4.

- Intro: **Sequential monitoring trial designs**
- Predictive probability
- Proper Bayes Designs
- Adaptive randomization and dose allocation
- Delayed response
- Hierarchical models
- Decision theoretic designs

† Phase II Studies: a small study of efficacy

- **Aim:** examine whether a drug has the desired efficacy to warrant further development.

  ⋆⋆ A *multi-stage* design with early stopping for futility rules is desirable in phase II settings.

- **Response:** often binary endpoint for "success"
- **Sample size:** still moderate, typically 40-100
- **Phase IIa:** initial efficacy assessment, single-arm, primary endpoint: a binary endpoint of response/no response or success/failure.
- **Phase IIb:** randomized multi-arm study to identify the most promising treatment for phase III, primary endpoint: time-to-event endpoints such as disease-free survival or progression-free survival.

† Monitoring Sequential Trial Designs

Sequential stopping: phase II trials use interim analysis or sequential stopping for futility (or efficacy) to improve efficiency of the study design.

- **Predictive probability designs:** rules are based on predictive probabilities.
- **Proper Bayes designs:** based on posterior probabilities of clinically meaningful events
- **Decision theoretic designs:** uses utility function $u(\cdot)$ (or loss function) in addition to probability model. Rule is based on maximizing utility.

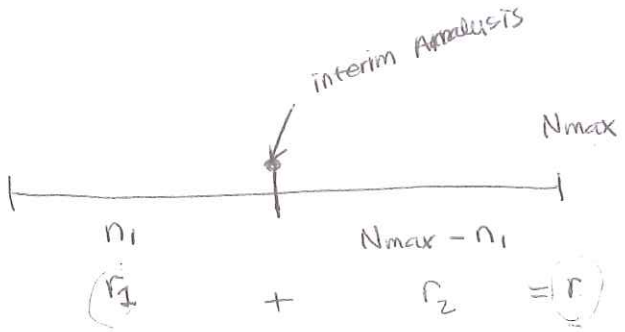
† *What* early stopping? Stop a trial before reaching the maximum number of patients due to futility or efficacy.

† *Why* early stopping?

- Save patients from receiving ineffective treatment
- Save time and resources from developing ineffective treatment
- If the new drug works well, enroll more patients.

  $\Rightarrow$ More patients can benefit from the treatment while the trial continues. & larger sample sizes increase the precision in estimating the response rate.

  $\star\star$ less stronger need for early stop due to efficacy.



† Simon's optimal design

efficacy of the standard trt

target efficacy

Goal: evaluate new drug, testing $H_0 : p \leq p_0$ vs $H_1 : p \geq p_1$ ($p_1 > p_0$)

Construct a design that minimizes the expected sample size under the null hypothesis. e.g.,

- Let $p_0 = 0.1$, $p_1 = 0.3$, and $\alpha = \beta = 0.1$ (type I and II error rates)
- Enroll 12 patients in the first stage.

  $r = 0$ or 1

  ⋆⋆ If no response or only one response is found, the trial stopped and the drug is considered ineffective.

  ⋆⋆ If not, 23 more paitents are enrolled to reach a total of 35 patients.

  ⋆ ⋆ ⋆ At the end of trial, if ≤ 5 responses are observed, the agent is deemed ineffective.

  ⋆ ⋆ ⋆ If ≥ 6 responses are observed, the agent is considered effective.

$r = 0$ or 1 under H_0: $p = 0.1$

- Under $H_0: \ p \leq p_0 = 0.1$, there is a 66% chance that the trial will be stopped. $\binom{12}{0} 0.9^{12} + \binom{12}{1} 0.9^{11} 0.1^1$
- Under $H_0: \ p \leq p_0 = 0.1$, the expected sample size is

$$12 + (35 - 12)(1 - 0.66) = 19.8.$$

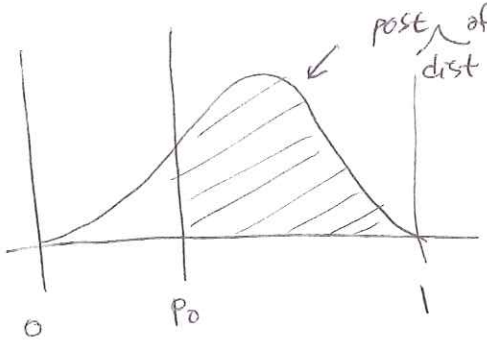
- Recall that the Simon optimal design was constructed to minimize the expected sample size under H_0.
- *Minimax* design: minimize the maximum trial sample size. For details, read Section 4.1.1.
- Multi-stage designs achieve better statistical properties than single-stage designs by utilizing information gained in the interim analysis.
- The frequentist analysis is constrained by the rigid requirement of examining the outcome at the specified sample size at each predetermined stage.

† Predictive Probability Designs

- **Goal:** evaluate new drug, testing $H_0: p \leq p_0$ vs $H_1: p \geq p_1$
- **Model:**
  - ⋆⋆ sampling model: $p(x_i = 1) = p$, $i = 1, \ldots, n$.
  - ⋆⋆ prior: $p \sim \text{Be}(a_0, b_0)$
- **Decision:** assume n observed patients $\mathbf{X} = (x_1, \ldots, x_n)$, and max $m = (N_{\text{max}} - n)$ future patients $\mathbf{Y} = (x_{n+1}, \ldots, n_{n+m})$.

  Decision for continuation:

  If the posterior probability of p exceeding some prespecified level p_0 is greater than some threshold θ_T, we declare efficacy.



$$\Pr(p > p_0 \mid \mathbf{X}, \mathbf{Y}) > \theta_T$$

$p \sim Be(a_0, b_0)$

$S_x \mid p \sim Bin(n, p)$

$\Rightarrow \quad p \mid S_x \sim Be(a_0 + S_x, \; b_0 + n - S_x)$

$\Rightarrow \quad S_y \mid S_x \sim \int \underbrace{p(S_y \mid p, S_x)}_{Bin(m,p)} \underbrace{p(p \mid S_x)}_{Be} \, dp$

$=$ Beta − Binomial

- **Predictive Probability (PP):** Consider future, after $\{x_{n+1}, \ldots, x_{n+m}\}$ will have been recorded.

  ⋆⋆ Let $S_x = \sum_{i=1}^{n} x_i$ and $S_y = \sum_{i=n+1}^{n+m} x_i$

  ⋆⋆ $S_y \sim$ Beta-binomial$(a_0 + S_x + S_y, b_0 + N_{\max} - S_x - S_y)$

  $(m, \; a_0 + S_x, \; b_0 + n - S_x)$ ⇐ corrected

  ⋆⋆ When S_y is recorded, $p \mid S_x, S_y \sim \text{Be}(a_0 + S_x + S_y, b_0 + N_{\max} - S_x - S_y)$.

  ⋆⋆ Predictive prob. of future evidence against H_0 (i.e., the PP of trial success):

$$\text{PP} = \text{E}[Pr(p > p_0 \mid \mathbf{X}, \mathbf{Y}) > \theta_T \mid \mathbf{X}]$$

$P(S_y \mid S_x)$

$p \mid S_x, S_y \sim Be(a_0 + S_x + S_y, \; b_0 + N_{max} - S_x - S_y)$

$$= \sum_{i=0}^{m} P(S_y = i \mid S_x) \, P(p > p_0 \mid S_x, S_y = i) > \theta_T)$$

- Predictive prob. of future evidence against H_0 (i.e., the PP of trial success):

$$\text{PP} = \text{E}[Pr(p > p_0 \mid \mathbf{X}, \mathbf{Y}) > \theta_T \mid \mathbf{X}]$$

⋆⋆**Case 1:** A **high** PP means that the treatment is likely to be efficacious by the end of the study given S_x.

$\Rightarrow$ The trial should be stopped early due to efficacy

⋆⋆**Case 2:** A **small** PP means that the treatment may not have sufficient activity.

$\Rightarrow$ The trial should be stopped early due to futility

⋆⋆**Case 3:** A **not high or small** PP means that the treatment may not have sufficient activity.

$\Rightarrow$ The trial should be continued because the current data are not yet conclusive.