Optimization of dose selection using multiple surrogates of toxicity as a continuous variable in Phase I cancer trial

Se Yoon Lee

Ph.D. in Statistics

Executive Summary

- We propose a novel fully Bayesian adaptive clinical trial design for Phase I cancer clinical trial to find maximum tolerated dose (MTD).

- This new Bayesian model is called the **2-parameter linear dose finder (2PLD).**

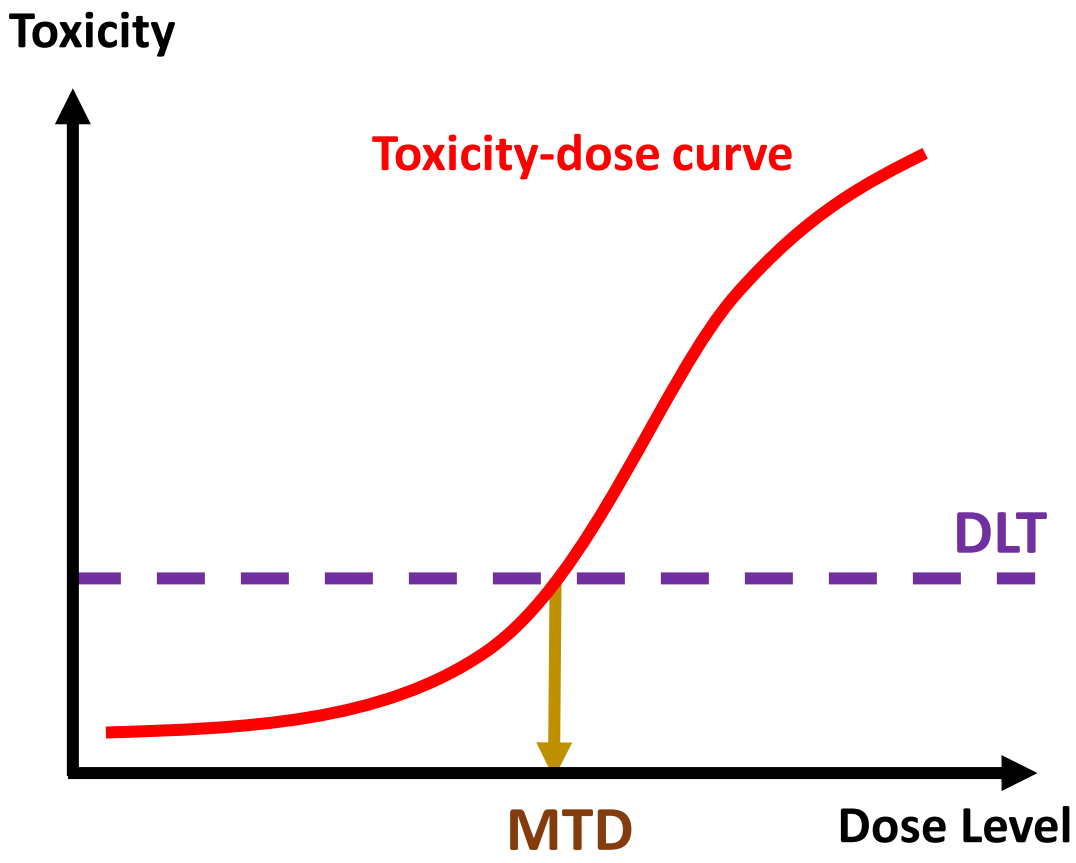
- The 2PLD has been developed to accommodate multiple Adverse Events from Common Toxicity Criteria for Adverse Events (CTCAE) from National Cancer Institute (NCI).

➢ **Keywords:** Maximum Tolerated Dose (MTD); Dose Limiting Toxicity (DLT); Common Toxicity Criteria for Adverse Events (CTCAE); Adaptive Clinical Trial Design

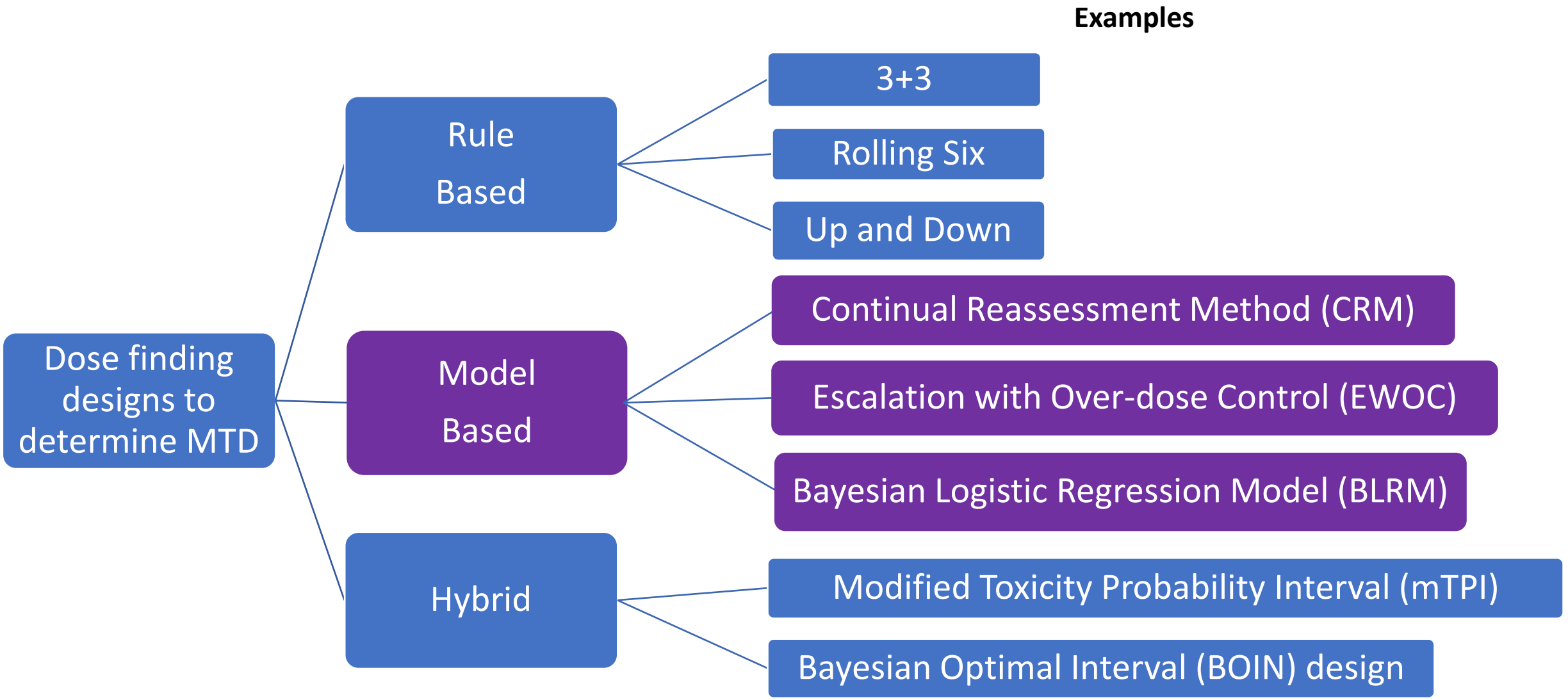
Primary Purpose of Phase I Cancer Clinical Trial

- **Key assumption of anticancer agents :** Toxicity level and therapeutic effect simultaneously increase with dose.

- **Goal :** to find the Maximum Tolerated Dose (MTD), the highest dose associated with an acceptable level of toxicity, called the Dose Limiting Toxicity (DLT).



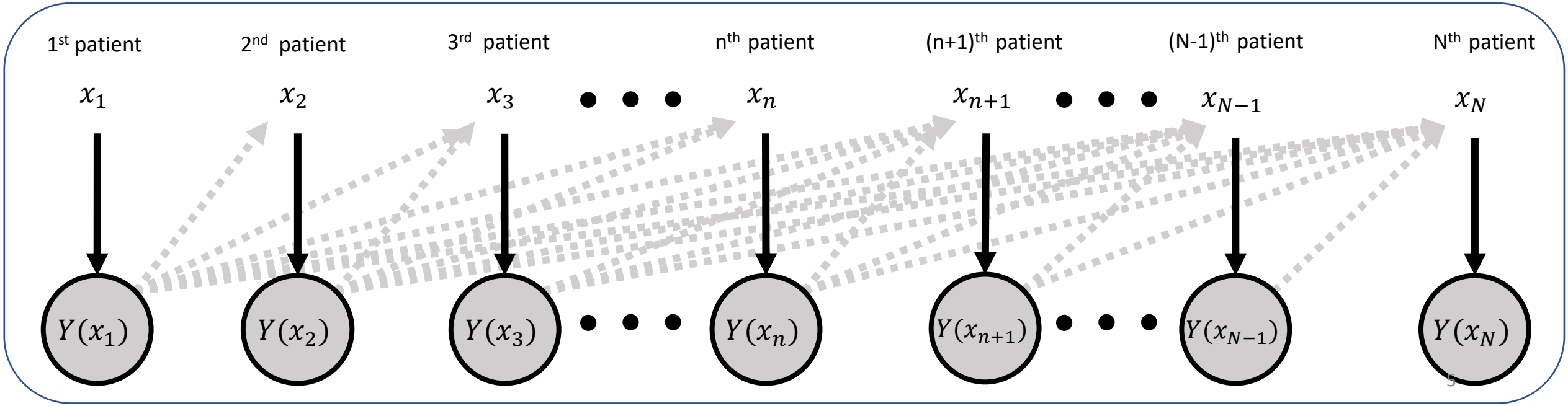
Categories of Dose Finding Designs in Phase I cancer clinical trials



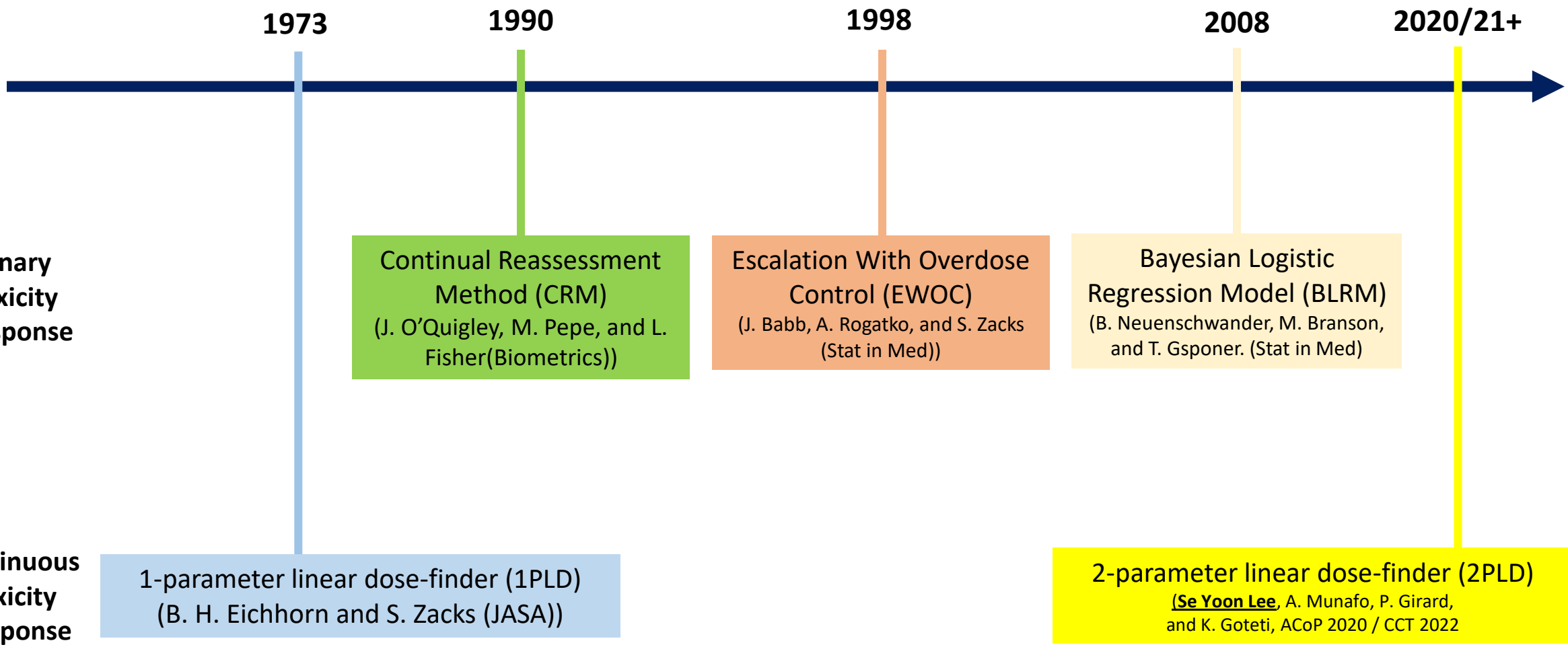
Fully sequential design setting

- We assume that patients are introduced to the trials sequentially one by one, and each patient is assumed to be assigned with an optimal estimate of MTD based on the accumulated patients' information at interim

- To protect patients from being overdosed, a drug is initiated at a very low dose and slowly escalated toward the targeted MTD as the trials go on.



Development of Bayesian adaptive clinical trial model for phase I cancer clinical trials



Review paper: O. Sverdlov et al. Adaptive clinical trial designs for phase I cancer studies, (2014), *Statistics Survey*

Extensions of Bayesian adaptive clinical trial models

Extensions	Authors (Year, Journal)
Incorporation of covariates	Babb et al. (2001, Stat in Med), Rogatko et al. (2008, Pharmaceutical Medicine), Tighiouart et al. (2012, Journal of Prob and Stat), Bailey et al. (2009, Journal of Biopharmaceutical Stat.)
Accounting for late-onset toxicities	Braun et al. (2006, Stat in Med), Moller et al. (1995, Stat in Med)
Choice of cohort size	Tighiouart et al. (2012, Journal of Prob and Stat)
Incorporation of multiple Adverse Events from CTCAE in determining MTD	Yuan et al. (2007, Biometrics) Chen et al. (2012, Contemporary Clinical Trial) ❖ Key Idea: 1. Introduce the concept of 'Toxicity Score' belong to interval [0,1] 2. Use 'pseudo-Bernoulli likelihood' instead of 'exact Bernoulli likelihood' ❖ Drawback: Loss of probabilistic characteristics of Bayesian model

Common Toxicity Criteria for Adverse Event (CTCAE) Version 5.0

- In the current version of CTCAE v.5.0, according to their severities and types of the Adverse Event (AE), toxicities are classified into five toxicity grades plus no toxicity grade as follows:

CTCAE Grade	Description	Note
Grade 0	**No toxicity**	
Grade 1	**Mild**	Asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
Grade 2	**Moderate**	Minimal, local or noninvasive intervention indicated
Grade 3	**Severe**	Medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling
Grade 4	**Life-threatening**	Urgent intervention indicated
Grade 5	**Death**	

A format of toxicity report system for CTCAE v.5.0

- Suppose that clinical scientists are particularly interested in J number of Adverse Events (AEs).
- After introducing a patient with a dose x, clinicians can record an Adjusted Toxicity Grade $G_j \in \{0,1,2,3,4\}$ for the j-th AE $(j = 1,2,\cdots,J)$ as the Right Table.
- Clinicians can specify a weight $w_j \in [0,1]$ corresponding to the j-th AE based on prior experiences such as early dose escalation studies, preclinical finding, etc. $(\sum_{j=1}^{J} w_j = 1)$
- A value w_j closer to 1 indicates a relatively higher importance of the j-th AE in determining the MTD than other AEs.
- When the j-th AE expects 'No toxicity(CTCAE Grade 0)', then we can set $w_j = 0$.

CTCAE Grade	Description	Adjusted Grade
Grade 0	**No toxicity**	$G_j = NA$
Grade 1	**Mild**	$G_j = 0$
Grade 2	**Moderate**	$G_j = 1$
Grade 3	**Severe**	$G_j = 2$
Grade 4	**Life-threatening**	$G_j = 3$
Grade 5	**Death**	$G_j = 4$

A format of toxicity report system

Adverse Events	AE 1	AE 2	⋯	AE j	⋯	AE J
Adjusted Grade	G_1	G_2	⋯	G_j	⋯	G_J
Weight	w_1	w_2	⋯	w_j	⋯	w_J

Continuous Toxicity Score: $Y(x) = \sum_{j=1}^{J} w_j G_j \in [0,4]$

Definition of Maximum Tolerated Dose (MTD) (Eichhorn and Zacks, JASA, 1973)

- Assume the monotonic toxicity-dose relationship: "Continuous Toxicity Score ($Y(x) = \sum_{j=1}^{J} w_j G_j$) increases when dose x increases".

[Definition] Given prespecified values $\eta \in (0,4), \gamma \in (0.5,1)$, and dose range (x_{min}, x_{max}), the MTD is defined to be the dose x that satisfies:

$$\text{MTD } \xi = \underset{x \in (x_{min}, x_{max})}{arg\max} \{ \Pr[Y(X) \geq \eta \mid X = x] \leq 1 - \gamma \}$$

Event of DLT at dose x

Upper bound

- Note that the definition of MTD ξ involves medically interpretable variables:
  - **Maximum toxicity score**: $\eta \in (0,4)$: A threshold dividing toxicity-dose plain into DLT and non-DLT region.
  - **Dose range**: (x_{min}, x_{max}) : Interval to contain the MTD ξ
  - **Homogeneity constant**: $\gamma \in (0.5,1)$: Scientist's belief about the degree of homogeneity of patients.

    Higher value for the γ leads to a stronger homogeneity. Default value is γ=0.99.

Definition of Maximum Tolerated Dose (MTD) (Eichhorn and Zacks, JASA, 1973)

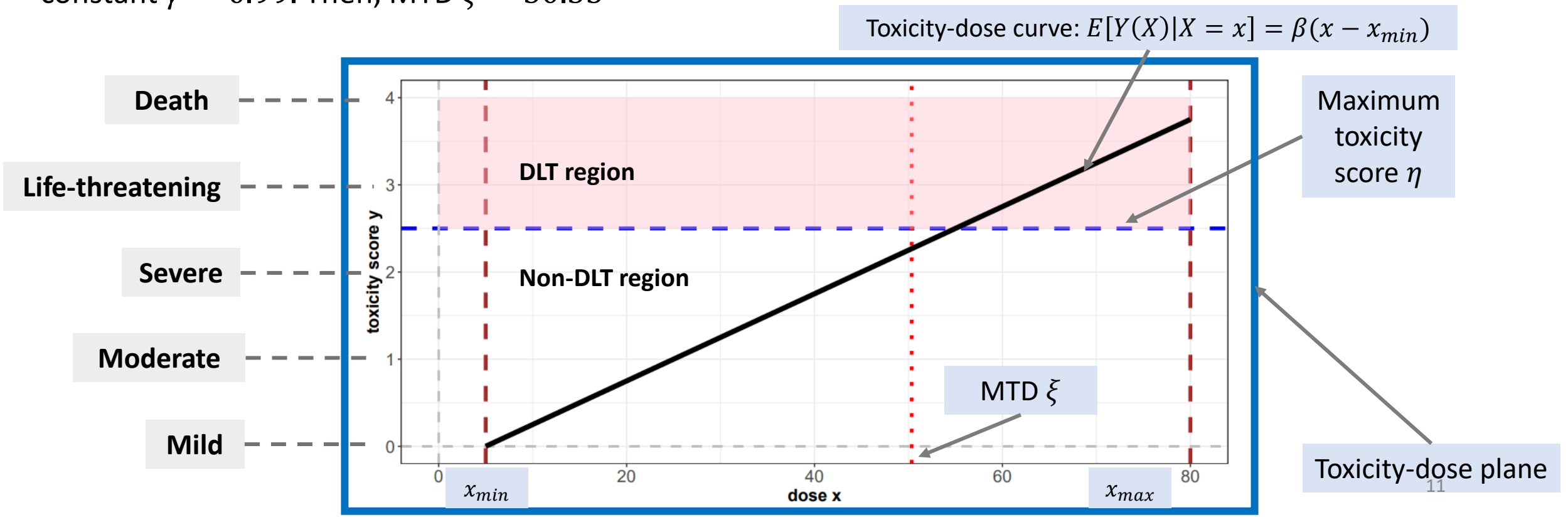
[Lemma] Assume the linear relationship between the toxicity score $y = Y(x)$ and dose x:

$$y = \beta(x - x_{min}) + \epsilon, \quad \epsilon \sim N(0, \sigma^2).$$

Then, the MTD ξ is given by

$$\xi = \xi(\beta, \sigma) = x_{min} + \frac{\eta - \sigma\,\Phi^{-1}(\gamma)}{\beta}.$$

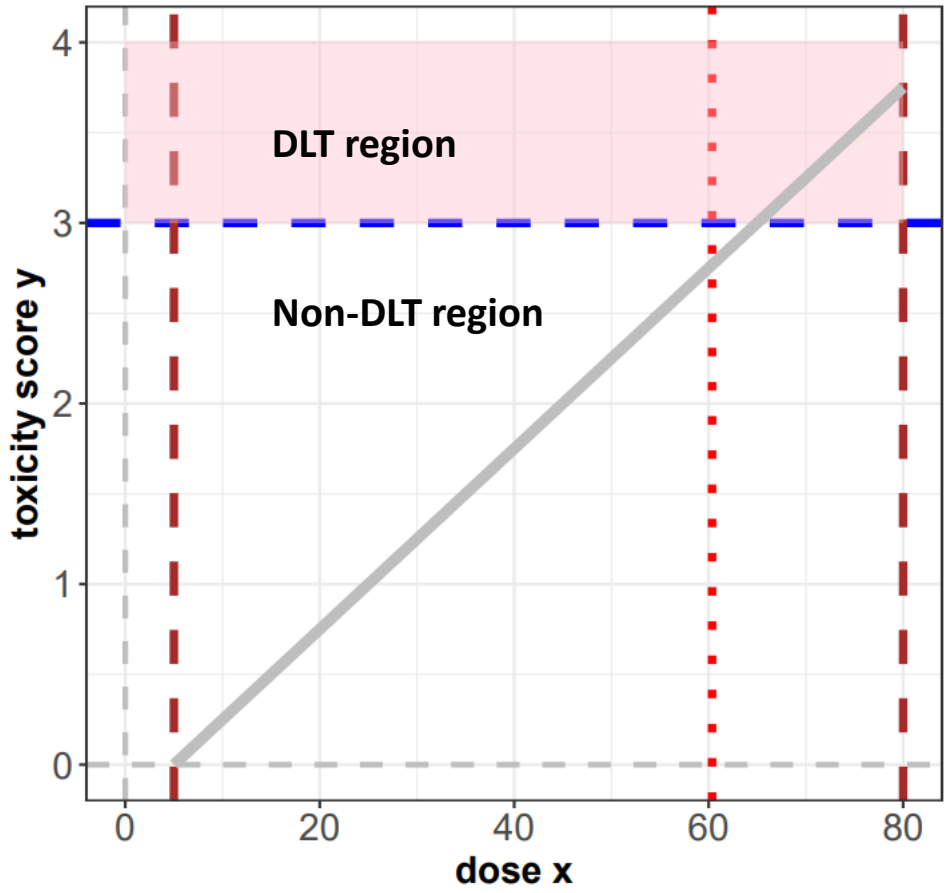
- Example: Slope $\beta = 0.05$, $\sigma = 0.1$, $x_{min} = 5$, $x_{max} = 80$, maximum toxicity score $\eta = 2.5$, homogeneity constant $\gamma = 0.99$. Then, MTD $\xi = 50.35$



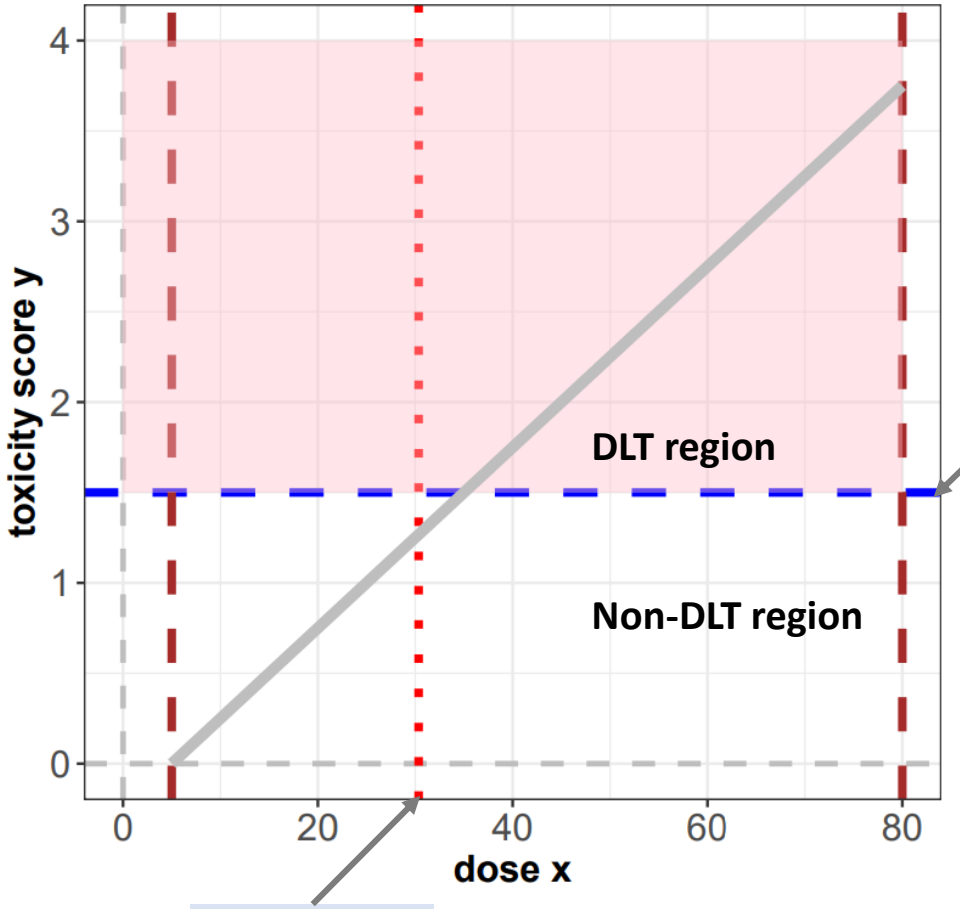
Clinical meaning of maximum toxicity score η

- Example: Slope $\beta = 0.05$, $\sigma = 0.1$, $x_{min} = 5$, $x_{max} = 80$, homogeneity constant $\gamma = 0.99$.

Non-fatal drug disease:
Maximum toxicity score $\eta = 3$
MTD $\xi = 60.34$



Life-threatening drug disease:
Maximum toxicity score $\eta = 1.5$
MTD $\xi = 30.34$



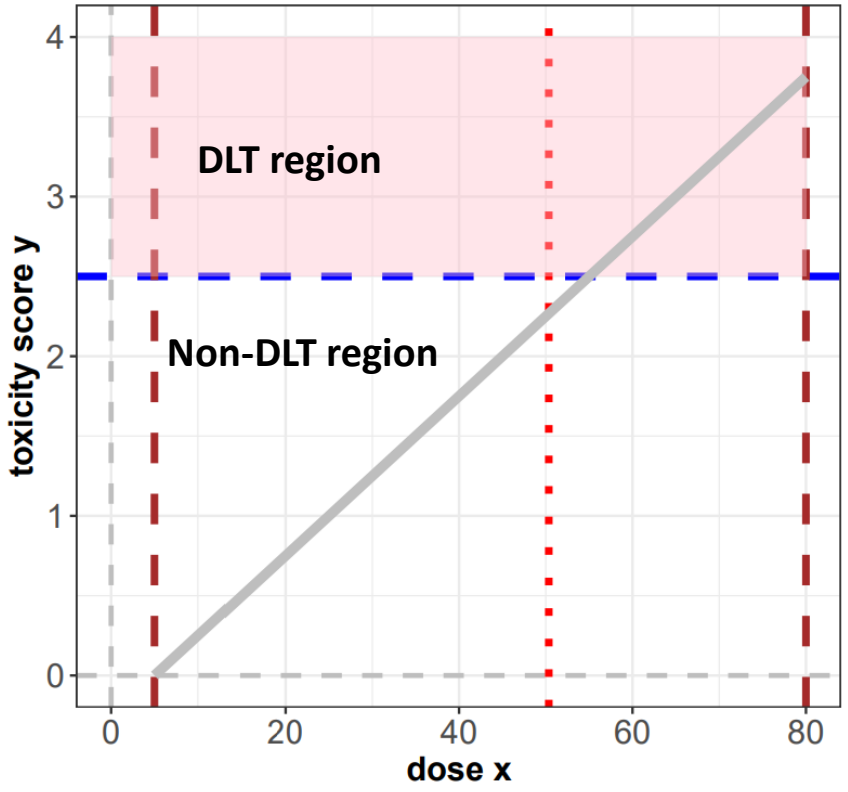
Maximum toxicity score η

MTD ξ

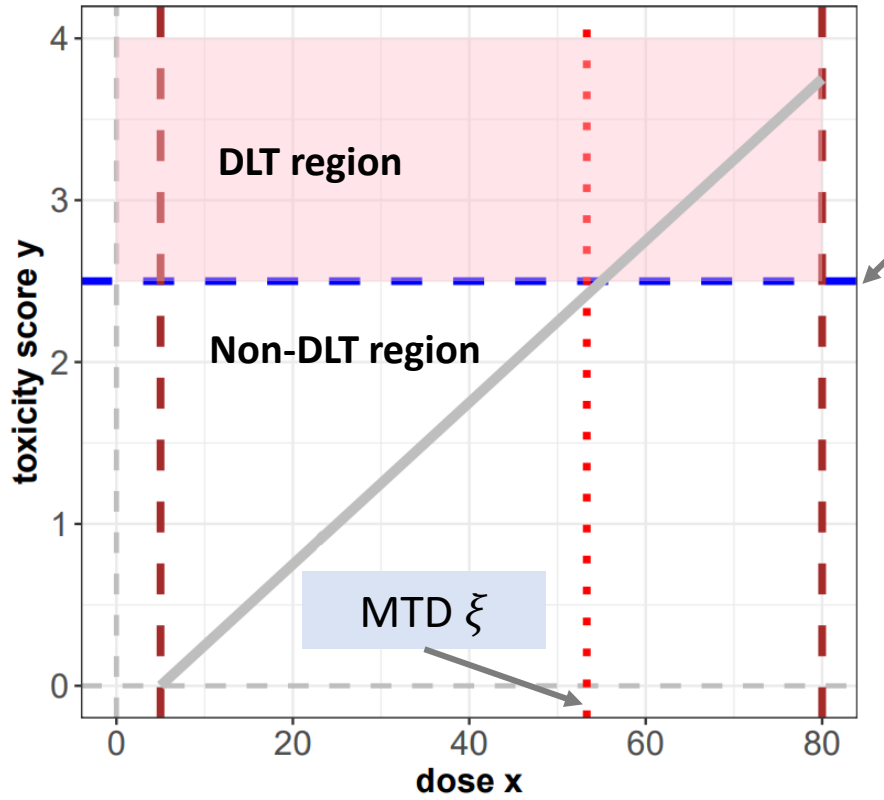
Clinical meaning of homogeneity constant γ

- Example: Slope $\beta = 0.05$, $\sigma = 0.1$, $x_{min} = 5$, $x_{max} = 80$, Maximum toxicity score $\eta = 2.5$

homogeneity constant $\gamma = 0.99$
$(\Pr[Y(X) \geq \eta \mid X = x] \leq 1 - \gamma = 0.01)$
MTD $\xi = 50.34$



homogeneity constant $\gamma = 0.8$
$(\Pr[Y(X) \geq \eta \mid X = x] \leq 1 - \gamma = 0.2)$
MTD $\xi = 53.31$



Maximum toxicity score η

We recommend to use $\gamma = 0.99$ for the default value for the safety concern.

1-Parameter Linear Dose-Finder (1PLD) (Eichhorn and Zacks, JASA, 1973)

Hierarchy of the 1PLD is given by:

- Likelihood:

$$y|\beta \sim N(\beta(x - x_{min}), \sigma^2)$$

- Prior:

$$\beta \sim \pi(\beta) = N(a, b^2)$$

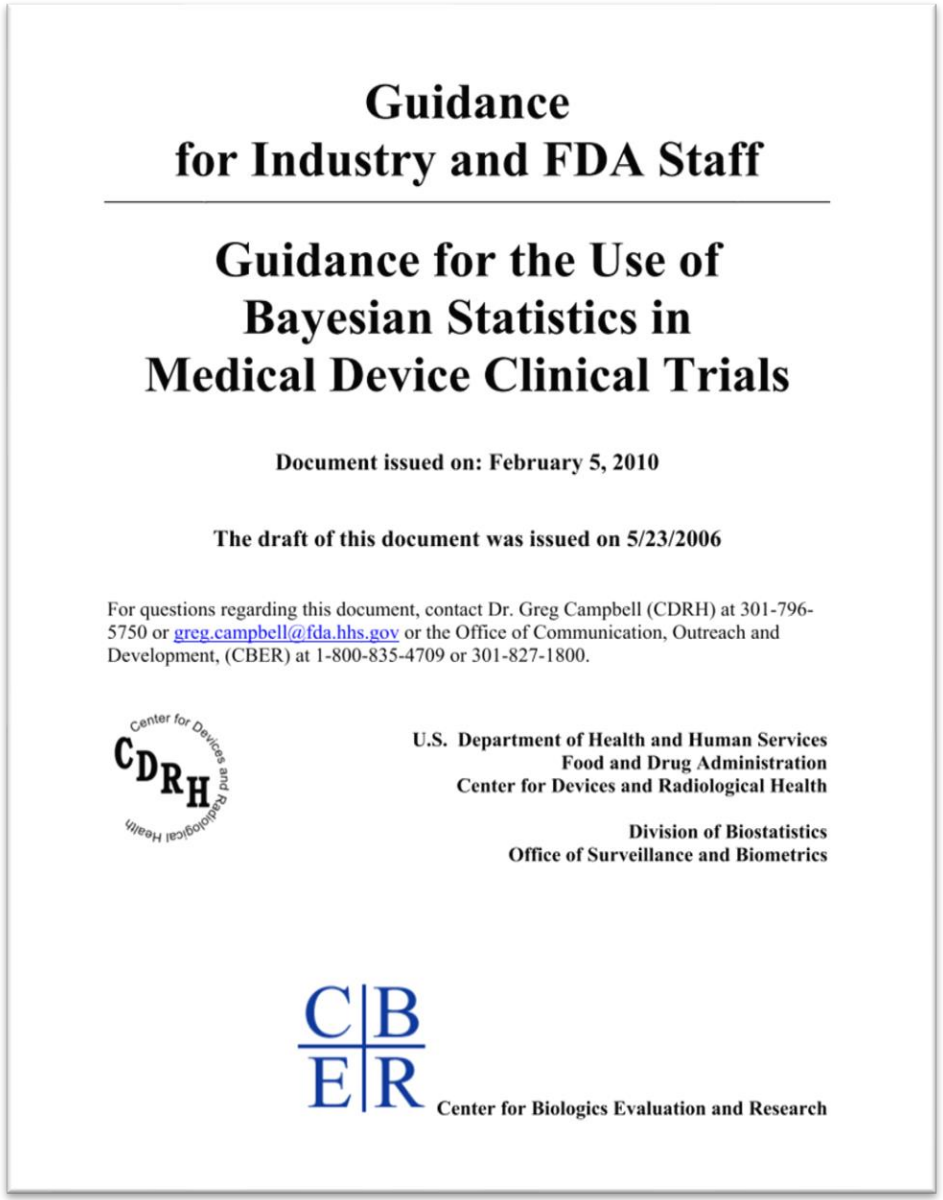
- Hyperparameters: x_{min}, σ, a, and b

- Interpretation of parameters:
  - Slope β : rate of the increment of the toxicity score per unit of dose
  - Standard deviation σ : measurement error of the toxicity score of patients at given dose x

- Drawbacks of 1PLD:
  1. β can take a negative value: violating monotonicity assumption of toxicity-dose curve
  2. STD σ is not estimated: ignoring heterogeneity of patients.
  3. Lack of practical consideration of phase I CCT: Maximum dose x_{max} is not incorporated in the model

- Contributions of Eichhorn and Zacks to phase I CCT :
  1. Shed light on how Bayesian paradigm can be used phase I CCT
  2. Introduce the notion of '(Bayesian) feasibility bound' to control the overdose: later, adopted to EWOC.

Guidance for the use of Bayesian statistics in medical device clinical trials



Guidance
for Industry and FDA Staff

Guidance for the Use of
Bayesian Statistics in
Medical Device Clinical Trials

Document issued on: February 5, 2010

The draft of this document was issued on 5/23/2006

For questions regarding this document, contact Dr. Greg Campbell (CDRH) at 301-796-5750 or greg.campbell@fda.hhs.gov or the Office of Communication, Outreach and Development, (CBER) at 1-800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Division of Biostatistics
Office of Surveillance and Biometrics

Center for Biologics Evaluation and Research

"Prior information should allow a Bayesian model to be flexible and efficient in identifying any pattern during trials."

❖ Bottlenecks in selecting an appropriate prior

- Based on the Guidance, we should prevent a suggested Bayesian model from being highly dominated by the prior information, and to that end, **imposing a lesser (weaker) prior information seems reasonable.**

- **However, allowing too weak prior information may lead to unstable parameter estimations for the first few patients in adaptive designs**, and this may contradict ethics that every patient has the right to be treated with an optimal dose.

➢ **A good model should retain a reasonable balance between these two competing practical requirements.**

2-Parameter Linear Dose-Finder (2PLD) (Se Yoon Lee, A. Munafo, P. Girard, and K. Goteti, ACoP 2020 / CCT 2022)

Hierarchy of the 2PLD is given by:

- Likelihood:

$$y|\beta,\sigma \sim N(\beta(x - x_{min}), \sigma^2)$$

- Prior:

$$\beta|\sigma \sim \pi(\beta|\sigma) = \mathcal{U}\big(l(\sigma), u(\sigma)\big) \text{ and } \sigma \sim \pi(\sigma) = C^{+}(0,1)I_{(0,\eta/\Phi^{-1}(\gamma))}$$

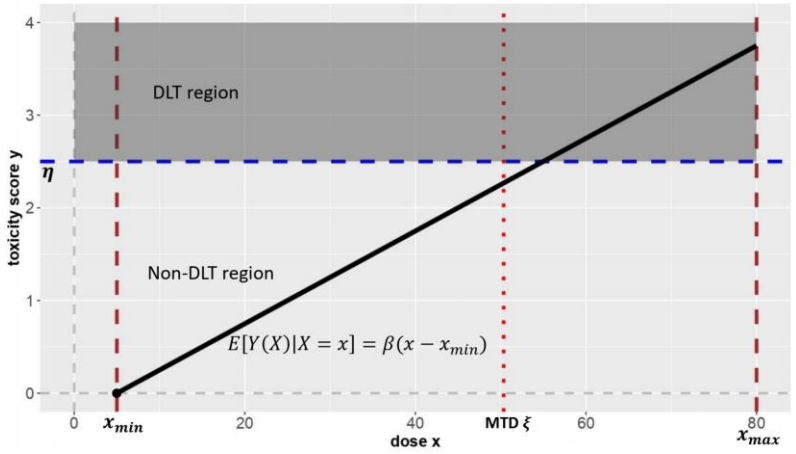
Uniform distribution — Truncated Half-Cauchy distribution

where

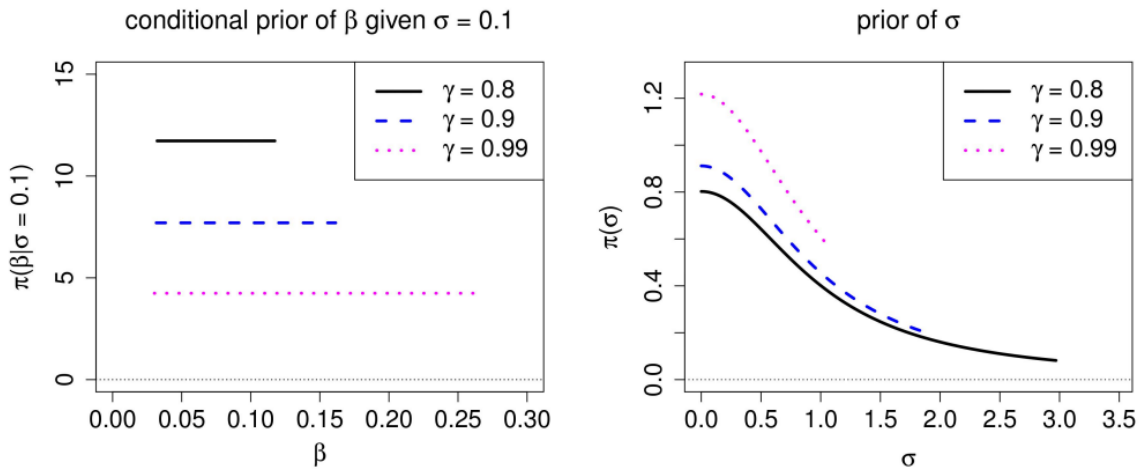
$$l(\sigma) = \frac{\{\eta - \sigma\,\Phi^{-1}(\gamma)\}}{x_{max} - x_{min}} \text{ and } u(\sigma) = \frac{\eta}{x_{max} - x_{min}} + \sigma\Phi^{-1}(\gamma).$$

- Hyperparameters: Dose range (x_{min}, x_{max}), maximum toxicity score η, and homogeneity constant γ

Likelihood



Prior



Good Operating Characteristics of 2PLD in determining MTD

[Theorem 1] Under the 2PLD, the dose-toxicity curve is monotonically increasing.

(Outline of proof) Marginal prior of the slope, $\pi(\beta)$, is supported on the positive real line $(0, \infty)$

[Theorem 2] Under the 2PLD, the prior of MTD ξ is compactly supported on the dose range (x_{min}, x_{max}).

(Outline of proof) Marginal prior of the MTD ξ, $\pi(\xi)$, is supported on the open interval (x_{min}, x_{max}).

[Corollary] Under the 2PLD, the posterior distribution of MTD ξ is consistent at any value of ξ_0 within the dose range (x_{min}, x_{max}).

(Outline of proof) Use Doob's theorem.

- **[Corollary]** guarantees that as long as the dose range (x_{min}, x_{max}) contains the true MTD ξ_0, the posterior distribution of the MTD ξ is consistent at MTD ξ_0.

- **This means that 2PLD is theoretically guaranteed to find the truth MTD as the sample size grows.**

  - Reference 'A review of consistency and convergence of posterior distribution' By Subhashis Ghosal.

A Dose Search Procedure Based on 2PLD: Overdosing control via feasibility bound

Let $\mathcal{F}_n = \{(x_i, y_i)\}_{i=1}^n$ be accrued information up to n patients where x_i is an assigned dose and $y_i = Y(x_i)$ is a toxicity score of the i-th patient.

Let $\Pi_n(x) = \Pr[\xi \le x \mid \mathcal{F}_n] = \int_{x_{min}}^{x} \pi(\xi \mid \mathcal{F}_n)\, d\xi$ is the posterior cumulative distribution of MTD ξ.

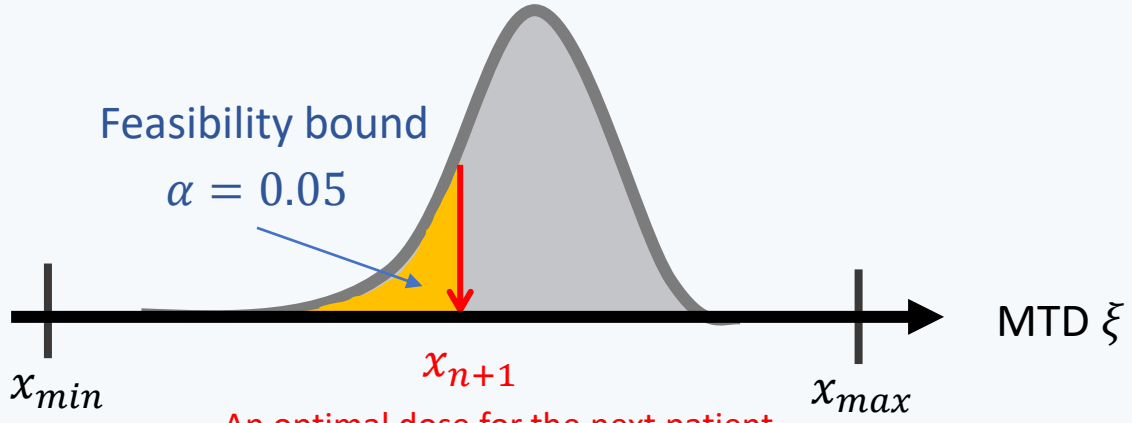
Given specified feasibility bound $\alpha \in (0,1)$, the selected dose for the $(n+1)$-th patient is the posterior lower α-quartile for the MTD ξ:

$$\Pi_n(x_{n+1}) = \Pr[\xi \le x_{n+1} \mid \mathcal{F}_n] \le \alpha \qquad \text{(default value is } \alpha = 0.05\text{)}$$

or, equivalently, we can denote with:

$$x_{n+1} = \Pi_n^{-1}(\alpha) \stackrel{\text{def}}{=} \mathcal{D}_\alpha(\mathcal{F}_n)$$

$\pi(\xi|\mathcal{F}_n)$: Posterior distribution of MTD given information from n patients



A Dose Search Procedure Based on 2PLD: Overdosing control via feasibility bound

Let $\mathcal{F}_n = \{(x_i, y_i)\}_{i=1}^{n}$ be accrued information up to n patients where x_i is an assigned dose and $y_i = Y(x_i)$ is a toxicity score of the i-th patient.

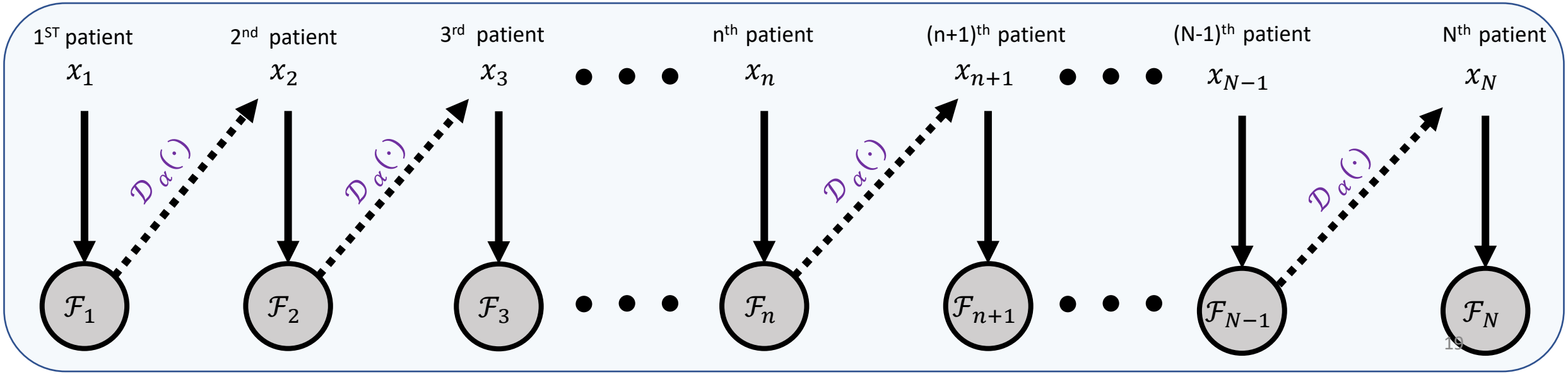
Let $\Pi_n(x) = \Pr[\xi \le x \mid \mathcal{F}_n] = \int_{x_{min}}^{x} \pi(\xi \mid \mathcal{F}_n)\, d\xi$ is the posterior cumulative distribution of MTD ξ.

Given specified feasibility bound $\alpha \in (0,1)$, the selected dose for the $(n+1)$-th patient is the posterior lower α-quartile for the MTD ξ:

$$\Pi_n(x_{n+1}) = \Pr[\xi \le x_{n+1} \mid \mathcal{F}_n] \le \alpha \quad \text{(default value is } \alpha = 0.05\text{)}$$

or, equivalently, we can denote with:

$$x_{n+1} = \Pi_n^{-1}(\alpha) \stackrel{\text{def}}{=} \mathcal{D}_\alpha(\mathcal{F}_n) \text{ : a mapping from information space to dose range}$$



Ancillary Options to Dose Find Rule

- **Discrete Dosage**: Let $\Omega = \{d_1 < d_2 < \cdots < d_K\}$ denote a set of prespecified dose level, based on the data from pre-clinical studies. Then one may use $x_{n+1} = argmin_{d\in\Omega}|d - \Pi_n^{-1}(\alpha)|$ on top of the vanilla algorithm.

- **Stopping Rule:** To directly protect patients, clinician can terminate the trials whenever any patient experiences the CTCAE Grade 3 or 4 for any Adverse Event.

- **Monotonic-increasing dosage:** To induce strictly escalating at each trial, one may use $x_{n+1} = \max\{x_n, \Pi_n^{-1}(\alpha)\}$ on top of the vanilla algorithm.

- **Upper bounding dose increment:** To protect patient from overdose due to large increment between doses, one may use $x_{n+1} = \min\{x_n + M, \Pi_n^{-1}(\alpha)\}$ for a fixed constant $M > 0$ on top of the vanilla algorithm.

Simulation Studies: Simulation Setup

- To experiment with a simulated phase I CCT, the experimenter needs to specify the following values:

  (1) Clinical hyper-parameters:
  maximum toxicity score $\eta \in (0,4)$;
  homogeneity constant $\gamma \in (0.5,1)$;
  and dose range (x_{min}, x_{max}).
  (2) Model parameters:
  Slope β and standard deviation σ;
  (3) The initial dose for the 1st patient $x_1 \in (x_{min}, x_{max})$.
  (4) Feasibility bound $\alpha \in (0,1)$
  (5) Total number of patients N.

[Lemma] Assume the linear relationship between the toxicity score $y = Y(x)$ and dose x:

$$y = \beta(x - x_{min}) + \epsilon, \quad \epsilon \sim N(0, \sigma^2).$$

Then, the MTD ξ is given by

$$\xi = \xi(\beta, \sigma) = x_{min} + \frac{\eta - \sigma\,\Phi^{-1}(\gamma)}{\beta}.$$

- Once above variables are set, then the truth MTD ξ_0 is automatically derived by the [Lemma].

- To simulate a toxicity score $y = Y(x)$ corresponding to the dose x, we generate y from the likelihood of 2PLD:

$$y = \beta_0(x - x_{min}) + \epsilon, \qquad \epsilon \sim N(0, \sigma_0^2)$$

for some fixed values for the β_0 and σ_0.

Simulation Studies: Experiment 1

- Set the variables: $\eta = 2.5, \gamma = 0.99, (x_{min}, x_{max}) = (5,80), x_1 = 6, \alpha = 0.05,\ N = 20$
- Model parameter $\beta_0 = 0.035$ and $\sigma_0 = 0.1$

➢ True MTD $\xi_0 = 69.78$

Patient i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Dose x_i	6.00	12.89	51.30	66.67	63.97	65.22	66.03	65.98	67.14	67.89	68.66	68.97	69.42	68.07	66.82	66.67	66.90	65.89	66.26	**66.54**
Response y_i	0.02	0.39	1.49	2.21	2.12	2.18	1.98	2.12	2.18	2.15	2.26	2.21	2.42	2.40	2.00	2.22	2.38	2.08	2.19	2.16
DLT ($y_i > \eta$)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Overdose ($x_i >$ MTD ξ_0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

❖ Results:

1. Final estimate of MTD $\hat{\xi} =$ **66.54**
2. # of DLT patients = 0
3. # of Overdosed patients = 0

Simulation Studies: Experiment 2

- Set the variables: $\eta = 2.5, \gamma = 0.99, (x_{min}, x_{max}) = (5,80), x_1 = 6, \alpha = 0.05,\ N = 20$
- Model parameter $\beta_0 = 0.05$ and $\sigma_0 = 0.2$

➢ True MTD $\xi_0 = 45.69$

Patient i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Dose x_i	6.00	13.03	31.52	39.98	37.71	39.32	40.42	40.33	41.95	43.00	44.12	44.53	45.28	43.24	41.53	41.29	41.61	40.24	40.75	**41.13**
Response y_i	0.01	0.63	1.07	1.84	1.75	1.86	1.46	1.74	1.86	1.80	2.01	1.92	2.34	2.30	1.50	1.94	2.26	1.65	1.88	1.82
DLT ($y_i > \eta$)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Overdose ($x_i >$ MTD ξ_0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

❖ Results:

1. Final estimate of MTD $\hat{\xi} =$ **41.13**
2. # of DLT patients = 0
3. # of Overdosed patients = 0

Simulation Studies: Experiment 3

- Set the variables: $\eta = 1.5, \gamma = 0.99, (x_{min}, x_{max}) = (5,80), x_1 = 6, \alpha = 0.005,\ N = 20$
- Model parameter $\beta_0 = 0.05$ and $\sigma_0 = 0.1$

➢ True MTD $\xi_0 = 30.35$

Patient i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Dose x_i	8.00	18.86	24.39	29.85	31.10	32.15	28.15	29.41	29.24	28.53	27.92	28.47	28.94	28.01	27.85	28.20	28.51	28.83	28.96	**29.24**
Response y_i	0.13	0.71	1.00	1.23	1.31	1.21	1.09	1.27	1.30	1.29	1.10	1.12	0.98	0.96	1.01	1.16	1.16	1.22	1.17	1.26
DLT ($y_i > \eta$)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Overdose ($x_i >$ MTD ξ_0)	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

❖ Results:

1. Final estimate of MTD $\hat{\xi} =$ **29.24**
2. # of DLT patients = 0
3. # of Overdosed patients = 2

- This overdose occurred due to huge jump: need to use **Ancillary Option: Upper bounding dose increment** (Supplemental Material)

Summary

- We proposed a Bayesian model-based design for Phase I cancer clinical trials, called 2PLD, which we demonstrated to be very safe.
- 2PLD can incorporate multiple AEs from CTCAE.
- 2PLD is also theoretically guaranteed to estimate MTD.
  (cf. Posterior Consistency)
- Ancillary Options can be incorporated into the 2PLD to reflect more practical requirement for Phase I CCT.

References

1. **Se Yoon Lee**, Shankar Lanke, Alain Munafo, Pascal Girard, and Kosalaram Goteti. (2020) “Optimization of dose selection using multiple surrogates of toxicity as continuous variable in Phase I cancer trial,” *ACoP11, ISSN:2688-3953, Vol 2*

2. **Se Yoon Lee**, Alain Munafo, Pascal Girard, and Kosalaram Goteti. (2022) “Optimization of dose selection using multiple surrogates of toxicity as continuous variable in Phase I cancer trial,” *Contemporary Clinical Trials*

3. John O’Quigley, Margaret Pepe, and Lloyd Fisher. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics, pages 33–48, 1990.*

4. Benjamin H Eichhorn and S Zacks. Sequential search of an optimal dosage, i. *Journal of the American Statistical Association, 68(343):594–598, 1973.*

5. James Babb, Andre Rogatko, and Shelemyahu Zacks. Cancer phase i clinical trials: efficient dose escalation with overdose control. *Statistics in medicine, 17(10):1103–1120, 1998.*

6. National Cancer Institute. Common toxicity criteria for adverse events v5.0. 2017.

7. US FDA. Guidance for the use of Bayesian statistics in medical device clinical trials. Guidance for industry and FDA staff. US FDA Docket, (2006D-0191):50, 2010.