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

Se Yoon Lee[1,3], Shankar Lanke[1], Alain Munafo[2], Pascal Girard[2], Kosalaram Goteti[1]

[1]EMD Serono, Inc., Billerica, MA, USA, a business of Merck KGaA, Darmstadt, Germany;
[2]Merck Institute of Pharmacometrics, Lausanne, Switzerland, an affiliate of Merck KGaA, Darmstadt, Germany;
[3]Department of Statistics, Texas A&M University, College Station, TX, USA

INTRODUCTION

- One of the main objectives of a phase I cancer clinical trial (CCT) is to estimate maximum tolerated dose (MTD) of a new drug associated with dose limiting toxicity (DLT) status.
- National Cancer Institute (NCI) provides the Common Toxicity Criteria for Adverse Events (CTCAE) for Adverse Event (AE) reporting. CTCAE v5.0 specifies a list of 26 AEs, and each AE is recorded with 5 level of grade according to the severity of a toxic reaction.
- Two pioneering Bayesian adaptive designs for the phase I CCT are **continual reassessment method** (CRM) and **escalation with overdose control** (EWOC) [1]. Two methods have two drawbacks: firstly, AE is recorded with a binary response, DLT or non-DLT; and secondly, they cannot accommodate grade information for multiple AEs.
- Recently, to accommodate grade information from multiple AEs, EWOC using **normalized equivalent toxicity score** (EWOC-NETS) had been developed, however, the results cannot be reproduced by simulation due to the theoretical limits inherent to the pseudo-Bernoulli likelihood used in the EWOC-NETS estimation.
- We aim to develop a fully Bayesian statistical model to estimate MTD for anti-cancer drugs in Phase I CCT by accommodating the grade information from multiple AEs.

METHODS

CTCAE Report System

- After introducing a patient with a dose x, a toxicity grade G_j for the j^{th} AE ($j = 1, \cdots, 26$) can be measured based on CTCAE with weight w_j prespecified by clinician.

Table 1: A format of CTCAE report system

AE	Nervous	Endocrine	Cardiac	⋯	Vascular
Grade	G_1	G_2	G_3	⋯	G_{26}
Weight	w_1	w_2	w_3	⋯	w_{26}

Grade G_j: 0 (mild), 1 (moderate), 2 (severe), 3 (life-threatening), and 4 (death)

Weight w_j: Importance of the j^{th} AE, $0 \le w_j \le 1$, $\sum_{j=1}^{26} w_j = 1$.

- Table 1 describes a format of CTCAE report system that can be recorded from an individual patient. Toxicity score for the patient is a weighted average obtained from 26 AEs:

toxicity score: $y = Y(x) = \sum_{j=1}^{26} w_j G_j \in [0,4]$.

- A higher/lower value of $Y(x)$ indicates a severer/lesser toxicity.
- In phase I CCT, to give the patients an optimal chance of a favorable treatment effect, one may be willing to accept a nonnegligible probability of severe toxic reaction.
- Clinician can specify weights w_j based on prior experiences such as early dose escalation cohorts, preclinical findings, etc.

METHODS

Two Parameter Linear Dose-finder (2PLD)

- 2PLD is a fully Bayesian adaptive clinical trial design model that assumes a linear relationship between the toxicity score $y = Y(x)$ and dose x.
- Hierarchical formulation of the 2PLD is given by

$$y|\beta, \sigma \sim N\big(\beta(x - x_{min}), \sigma^2\big) \qquad (1)$$

$$\beta|\sigma \sim U(l(\sigma), u(\sigma)),\ \sigma \sim C^+(0,1) I_{\left(0, \eta/\Phi^{-1}(\gamma)\right)} \qquad (2)$$

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

(Notations:

- $N(\mu, \sigma^2)$: normal distribution with mean μ and standard deviation σ;
- $U(l, u)$: uniform distribution on interval (l, u);
- $\Phi(z)$: cumulative distribution of the standard normal distribution;
- $C^+(0,1)I_{(0,d)}$: truncated unit-scaled half-Cauchy distribution supported on interval $[0, d]$.)

Maximum Tolerated Dose (MTD) [2]

Definition. Given pre-specified values for $\eta \in (0,4)$, $\gamma \in (0.5,1)$, and an interval (x_{min}, x_{max}), the maximum tolerated dose ξ is defined to be the largest value of x that satisfies the inequalities:

$$\Pr[Y(x) \le \eta | x] \ge \gamma \ \text{ and } x_{min} \le x \le x_{max}.$$

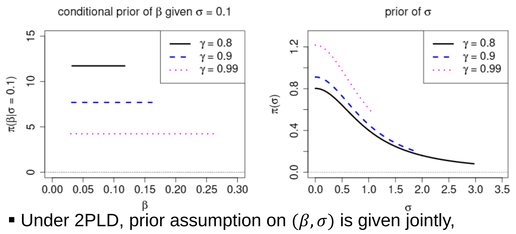
- Under 2PLD (1) – (2), MTD is given by

$$\xi = \xi(\beta, \sigma) = x_{min} + \frac{\eta - \sigma\Phi^{-1}(\gamma)}{\beta} \qquad (3)$$

- Note that MTD ξ (3) is a function of random variables β and σ, hence, MTD itself is also a random variable. That means that we can quantify the uncertainty of MTD in a fully Bayesian way *a posteriori*.

Prior analysis of 2PLD

Figure 1: Priors of 2PLD.
Conditional prior of β given $\sigma = 0.1$ (left) and prior of σ (right)

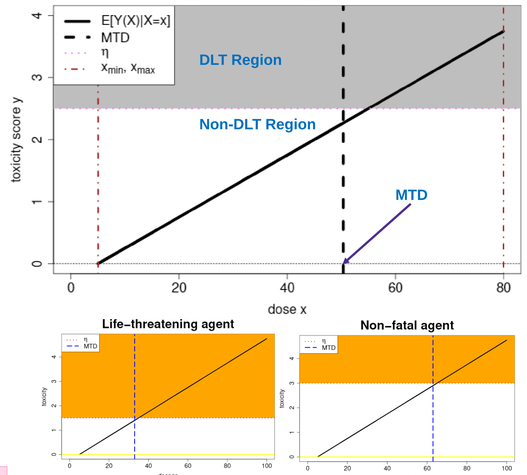


- Under 2PLD, prior assumption on (β, σ) is given jointly, contrasting to EWOC [2].
- Setting a higher/lower value for the γ allows a lower/higher variability of patients' toxicity score responses, respectively, given a dose, a *priori*: see the right panel in Figure 1.
- The γ determines the degree of homogeneity of the patients, where a higher value for the γ leads to a stronger homogeneity.

METHODS

Clinical Interpretation of 2PLD

Figure 2: Mode description of 2PLD.
($\beta = 0.05, \sigma = 0.1, \eta = 2.5, \gamma = 0.99$, $x_{min} = 5$, and $x_{max} = 80$)

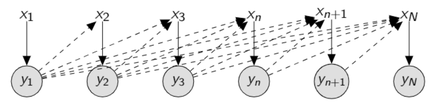


- *Maximum toxicity score* $\eta \in (0,4)$: A threshold dividing toxicity-dose plain into DLT region and non-DLT region: refer to Fig 2. For a life-threatening/non-fatal agent, it is set to low/high, respectively.
- *Homogeneity constant* $\gamma \in (0.5,1)$: Scientist's belief about the degree of homogeneity of patients. Default value for the γ is 0.99 [2].
- *Dose range* (x_{min}, x_{max}): Interval to contain the MTD ξ (3). It should be set to wide enough to include the MTD. Typically, x_{min} and x_{max} are inferred from pre-clinical studies.

Adaptive clinical trial design in fully sequential design setting

- 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.

Figure 3: Fully sequential design setting using N patients.
x_i: dose assigned to the i^{th} patient;
$y_i = Y(x_i)$: toxicity score for the i^{th} patient



METHODS

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 $\Pi_n(x) = \Pr[\xi \le x | \mathcal{F}_n] = \int_{x_{min}}^{x} \pi(\xi|\mathcal{F}_n) d\xi$ is the posterior cumulative distribution function of MTD ξ (3).
- Given specified feasibility bound $\alpha \in (0,1)$ [1,2], the selected dose for the $(n+1)^{\text{th}}$ patient is the posterior α-quartile for the MTD ξ satisfying

$$\Pi_n(x_{n+1}) = \Pr[\xi \le x_{n+1} | \mathcal{F}_n] \le \alpha \qquad (4).$$

- The feasibility bound α is an upper bound of the posterior probability of the event of overdose $\{\xi \le x_{n+1}\}$.
- EWOC also utilized a similar idea except that EWOC is based on binary toxicity information, whereas the 2PLD is based on continuous toxicity information.
- Default value for the feasibility bound is $\alpha = 0.05$. [2]

RESULTS

Simulation set-up

- To experiment with a simulated phase I CCT, the experimenter needs to specify the values: (a) the clinical hyper-parameters γ, η, and (x_{min}, x_{max}); (b) the model parameters β and σ; (c) the initial dose x_1 for the first patient; (d) the feasibility bound α; and (e) the total number of patients for the phase I CCT N.
- Once variables in (a) and (b) are set, then the targeted MTD ξ (3) is automatically determined.
- To simulate a toxicity score $y = Y(x)$ corresponding to a dose x, we generate y from the likelihood of 2PLD (1)

$$y = \beta_0(x - x_{min}) + \sigma_0\epsilon,\ \ \epsilon \sim N(0,1) \qquad (5),$$

with some fixed values for the slope β_0 and measurement error σ_0, that leads to true MTD ξ_0 based on the formula (3).

Simulation 1: Varied slope β

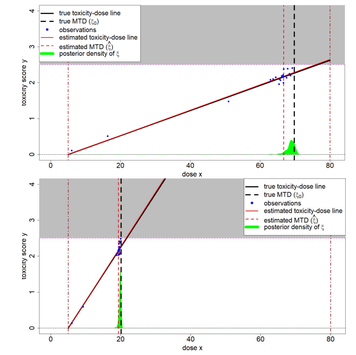
- Slope β is set to be 0.035 and 0.15 with $\sigma = 0.1$.
- Corresponding MTD ξ are 69.78 and 20.12.
- Other variables are set with $\eta = 2.5$, $\gamma = 0.99$, $(x_{min}, x_{max}) = (5,80)$, $x_1 = 6, \alpha = 0.05$, and $N = 20$.

Table 2: Sequential trials with different slope β

$(\beta_0, \sigma_0, \xi_0) = (0.035, 0.1, 69.78)$						$(\beta_0, \sigma_0, \xi_0) = (0.15, 0.1, 20.12)$					
i	x_i	y_i	i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	6.00	0.02	11	68.66	2.26	1	6.00	0.15	11	19.64	2.17
2	12.89	0.39	12	68.97	2.21	2	9.50	0.62	12	19.76	2.26
3	51.30	1.49	13	69.42	2.42	3	19.71	2.19	13	19.82	2.38
4	66.67	2.21	14	68.07	2.40	4	19.63	2.26	14	19.66	2.26
5	63.97	2.12	15	66.82	2.00	5	19.85	2.24	15	19.74	2.09
6	65.22	2.18	16	66.67	2.22	6	20.02	2.49	16	19.68	2.05
7	66.03	1.98	17	66.90	2.38	7	18.88	2.03	17	19.63	2.14
8	65.98	2.12	18	65.89	2.08	8	19.05	2.08	18	19.69	2.41
9	67.14	2.18	19	66.26	2.19	9	19.29	2.16	19	19.46	2.34
10	67.89	2.15	20	66.54	2.16	10	19.49	2.20	20	19.36	2.04

RESULTS

Figure 4: Sequential trials of 20 patients: $(\beta_0, \sigma_0, \xi_0) = (0.035, 0.1, 69.78)$ (top) and $(\beta_0, \sigma_0, \xi_0) = (0.15, 0.1, 20.12)$ (bottom)



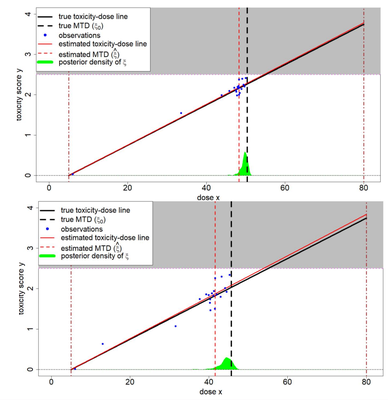
Simulation 2: Varied measurement error σ

- Measurement error σ is set to be 0.1 and 0.2 with $\beta = 0.05$.
- Corresponding MTD ξ are 50.35 and 45.69.
- Other variables are set with the same as **Simulation 1**.

Table 3: Sequential trials with different measurement error σ

$(\beta_0, \sigma_0, \xi_0) = (0.05, 0.1, 50.35)$						$(\beta_0, \sigma_0, \xi_0) = (0.05, 0.2, 45.69)$					
i	x_i	y_i	i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	6.00	0.03	11	49.53	2.26	1	6.00	0.01	11	44.12	2.01
2	23.02	1.02	12	49.76	2.21	2	13.03	0.63	12	44.53	1.92
3	46.48	1.95	13	50.07	2.42	3	31.52	1.07	13	45.28	2.34
4	45.16	2.06	14	49.15	2.40	4	39.98	1.84	14	43.24	2.30
5	46.13	2.11	15	48.28	2.00	5	37.71	1.75	15	41.53	1.50
6	47.15	2.18	16	48.17	2.22	6	39.32	1.86	16	41.29	1.94
7	47.75	1.98	17	48.33	2.38	7	40.42	1.46	17	41.61	2.26
8	47.64	2.12	18	47.63	2.08	8	40.33	1.74	18	40.24	1.65
9	48.46	2.18	19	47.89	2.19	9	41.95	1.86	19	40.75	1.88
10	49.00	2.15	20	48.08	2.16	10	43.00	1.80	20	41.13	1.82

Figure 5: Sequential trials of 20 patients: $(\beta_0, \sigma_0, \xi_0) = (0.05, 0.1, 50.35)$ (top) and $(\beta_0, \sigma_0, \xi_0) = (0.05, 0.2, 45.69)$ (bottom)



CONCLUSION

- This novel 2PLD can be an attractive tool for clinical scientists because of its parsimonious description of a toxicity-dose curve, medically interpretable hyperparameters, and an automated posterior computation.

REFERENCES

[1] Babb, J., Rogatko, A. and Zacks, S., 1998. Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in medicine*, *17*(10), pp.1103-1120.

[2] Eichhorn, B.H. and Zacks, S., 1973. Sequential search of an optimal dosage, I. *Journal of the American Statistical Association*, *68*(343), pp.594-598.