THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center

Making Cancer History®

Bayesian Data Augmentation Continual Reassessment Method (DA-CRM) for Phase I Trials with Delayed Toxicities

Ying Yuan

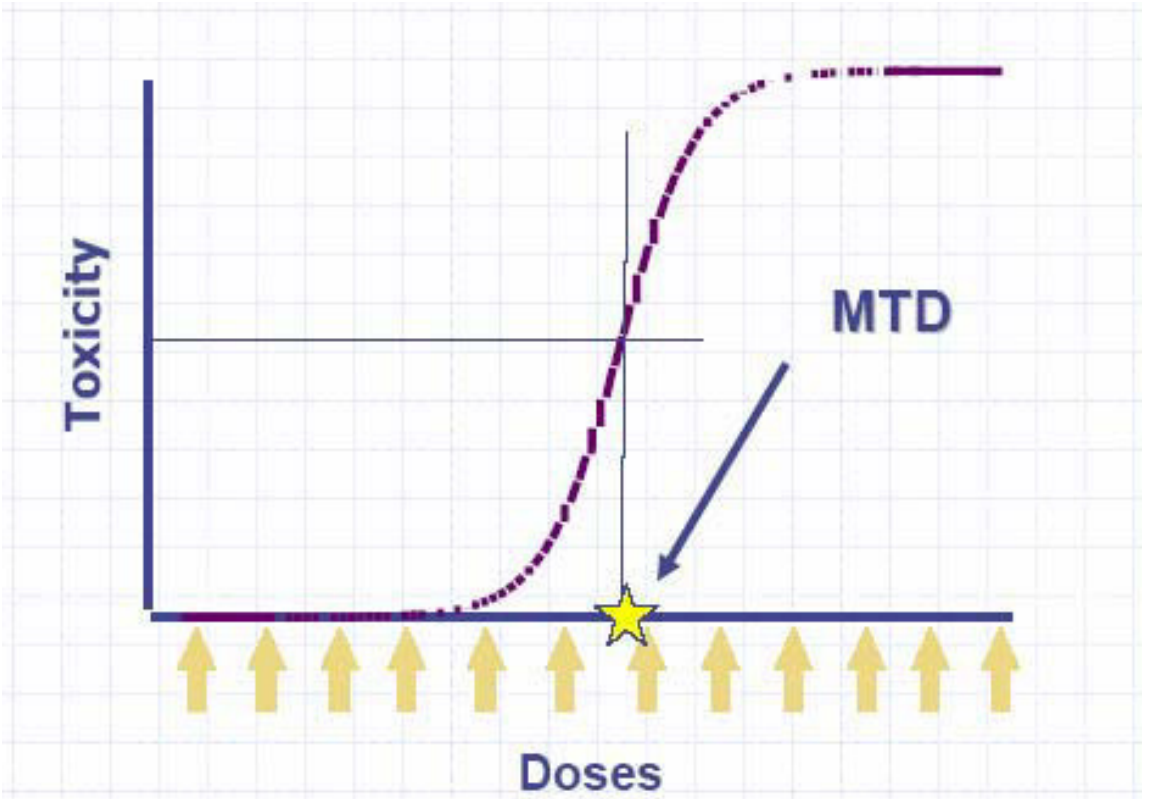
Department of Biostatistics
The University of Texas, MD Anderson Cancer Center

Outline

- Background
- Late-onset toxicity and missing data
- Bayesian data augmentation for late-onset toxicity
- Simulations
- Concluding remarks

Phase I trials

- In phase I oncology trials, the primary objective is to find the maximum tolerated dose (MTD).



Phase I trial designs

- "3+3" design (Storer, 1989)
- Continual reassessment method (CRM; O'Quigley et al., 1990)
- Decision theoretic approach (Whitehead and Brunier, 1995)
- Bayesian optimal interval (BOIN) design (Liu and Yuan, 2015)

.....

Continual reassessment method (CRM)

- CRM specifies a working dose-toxicity model, such as

$$\pi_j(\alpha) = p_j^{\exp(\alpha)}, \quad j = 1, \ldots, J, \tag{1}$$

  where π_j is the toxicity probability of dose level j, p_j is the prior estimate of the toxicity probability of dose level j (i.e., skeleton), and α is an unknown parameter.
- Based on observed data, the CRM continuously updates the estimate of the dose-toxicity model $\pi_j(\hat{\alpha})$ to direct the dose escalation and de-escalation.

CRM dose-finding rule

- A new cohort of patients is assigned to dose level j^* such that

$$j^* = \text{argmin}_{j\in(1,\ldots,J)}|\hat{\pi}_j - \phi|.$$

  where ϕ is the target toxicity probability.
- The trial continues until the exhaustion of the total sample size, and then the dose with a posterior toxicity probability closest to ϕ is selected as the MTD.

An important assumption

- **The toxicity outcome is observed immediately**, such that by the time of the next dose assignment, the currently treated patients must have complete information on toxicity.
- Delayed or late-onset toxicities, however, are common in practice.

Late-onset toxicity

- In radiotherapy trials, dose-limiting toxicities often occur long after the treatment is finished.
- Late-onset toxicity is common for novel molecularly targeted agents.
  - A recent review paper in the *Journal of Clinical Oncology* found that among a total of 445 patients included in 36 trials, 57% of the grade 3 and 4 toxicities were late-onset; the authors called for particular attention to the issue of late-onset toxicity (Postel-Vinay et al., 2011).

Late-onset toxicity

- Patients who have not experienced toxicity at the moment of dose escalation may experience toxicity later during the remaining follow-up.
- Conventional dose-finding methods, e.g., CRM, often underestimates the toxicity probabilities and leads to an undesirably large number of patients treated at overly toxic doses.

Available methods for late-onset toxicity

- Cheung and Chappell (2000) proposed the time-to-event CRM (TITE-CRM) by weighting the likelihood with the followup time. This often results in pseudo-likelihood.
- we propose a likelihood-based approach built upon the missing data methodology.

Setup

- Patients enter the study sequentially, and are followed for a fixed period of time $(0, T)$ to assess the toxicity of the drug.
- During this evaluation window $(0, T)$, we measure a binary toxicity outcome for subject i,

$$Y_i = \begin{cases} 1 & \text{if toxicity observed in } (0, T) \\ 0 & \text{if no toxicity observed in } (0, T). \end{cases}$$

- The length of the assessment period T is chosen so that if a drug-related toxicity occurs, it would occur within $(0, T)$.
- T can be weeks, months or longer.

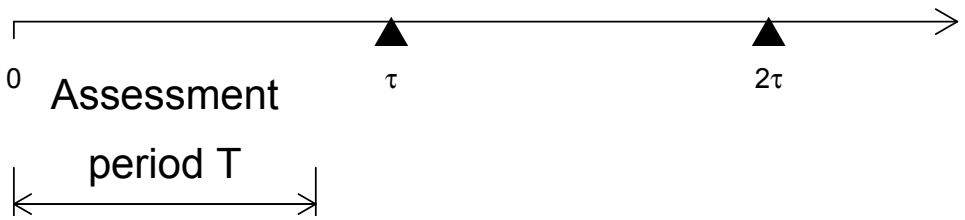
Late-onset toxicity

- Whether or not the toxicity is of late onset depends on the relative length of the assessment period and the patient's inter-arrival time, that is,

$$\text{A/I ratio} = \frac{\text{assessment period}}{\text{interarrival time}}$$

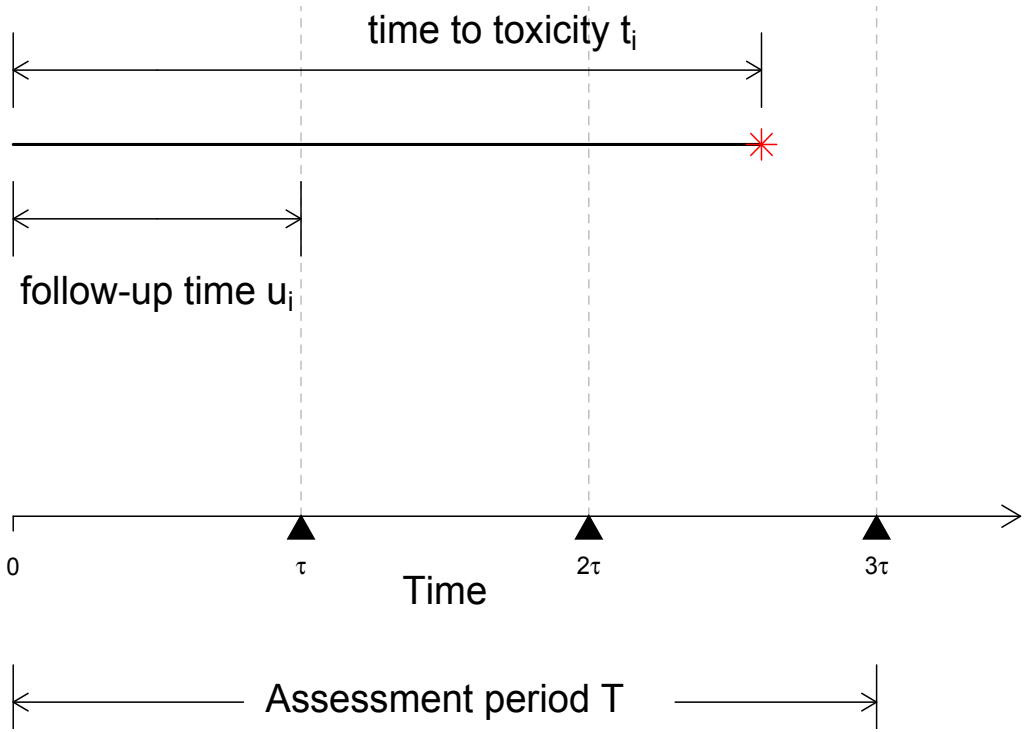
.

- If the assessment period is not longer than the patient's inter-arrival time (i.e., A/I ratio ≤ 1), the toxicity is not of late onset.



Late-onset toxicity

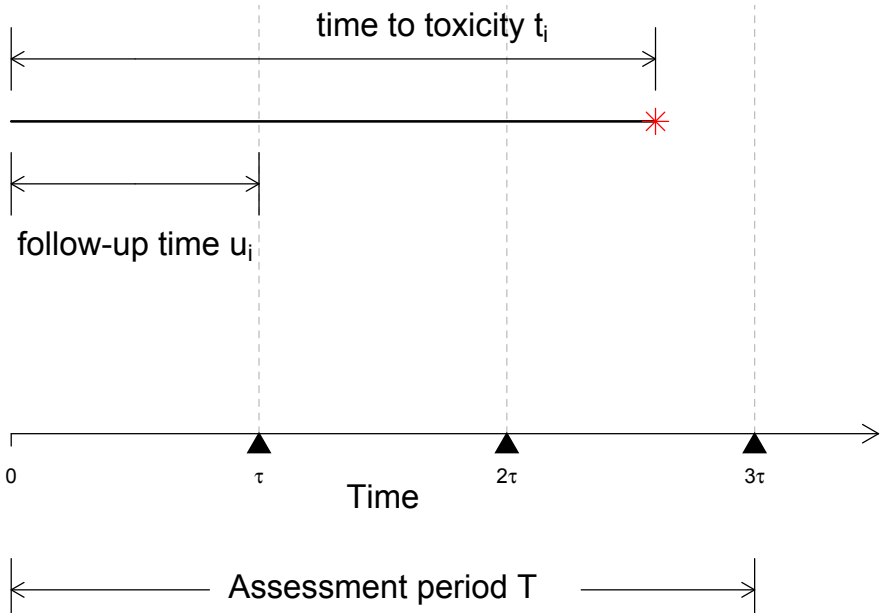
- However, if the assessment period is longer than the patient's inter-arrival time (i.e., A/I ratio > 1), e.g., under a fast accrual, the toxicity may be of late onset.



Missing data induced by late-onset toxicity

- Let t_i denote the time to toxicity, and u_i ($0 \leq u_i \leq T$) denote the actual follow-up time at the moment of interim decision.
- $M_i(u_i)$ denote the missing data indicator, then

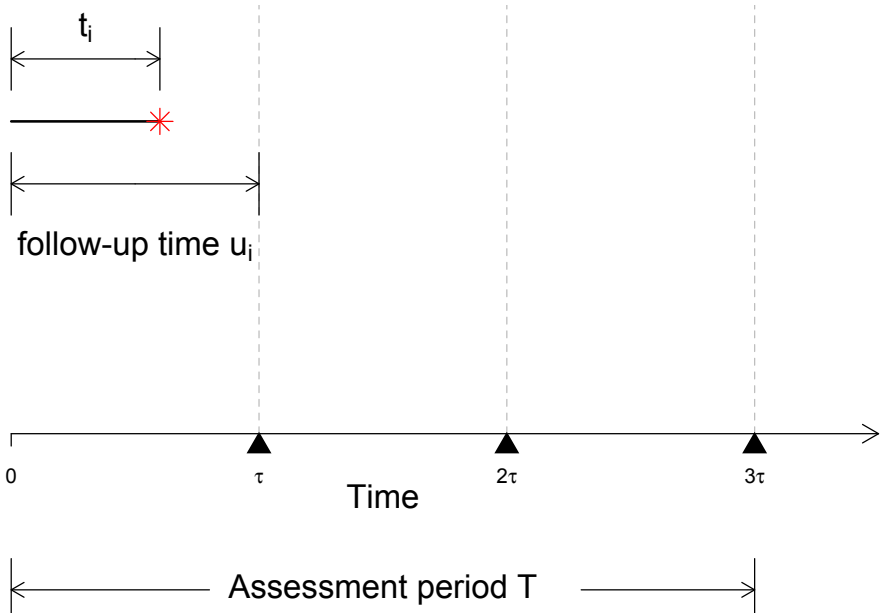
$$M_i(u_i) = \begin{cases} 1 & \text{if} \quad t_i > u_i \text{ and } u_i < T, \end{cases}$$



Missing data induced by late-onset toxicity

- Let t_i denote the time to toxicity, and u_i ($0 \leq u_i \leq T$) denote the actual follow-up time at the moment of interim decision.
- $M_i(u_i)$ denote the missing data indicator, then

$$M_i(u_i) = \begin{cases} 1 & \text{if} \quad t_i > u_i \text{ and } u_i < T, \\ 0 & \text{if} \quad t_i < u_i \text{ or } t_i > u_i = T. \end{cases}$$



Nonignorable Missing Data

- The missing data are nonignorable or informative! because

$$\text{pr}(M_i = 1 | Y_i = 0) > \text{pr}(M_i = 1 | Y_i = 1).$$

- Patients who will not experience toxicity ($Y_i = 0$) in the assessment period are more likely to be missing than patients who will experience toxicity ($Y_i = 1$).

Two implications

- Because the missing data are nonignorable, the simple way by discarding the missing data and making inference solely based on the observed data is problematic.
- When modeling toxicity, we need to account for the missing data mechanism, which is known:

$$M_i = \begin{cases} 1 & \text{if} \quad t_i > u_i \text{ and } u_i < T, \\ 0 & \text{if} \quad t_i < u_i \text{ or } t_i > u_i = T. \end{cases}$$

Impute missing data

- An intuitive approach to dealing with the unobserved toxicity outcomes is to impute the missing data, so that the standard CRM methodology can be applied.
- Under the Bayesian paradigm, this can be achieved using data augmentation.
  1. Imputation (I) step, in which the missing data are imputed,
  2. posterior (P) step, in which the posterior samples of unknown parameters are simulated based on imputed data.

Impute missing data

- The missing data we consider here is a special case of nonignorable missing data with a known missing data mechanism as defined previously.

Model for time to toxicity

- We use a piecewise exponential model for the time to toxicity for patients who will experience DLTs.
- Partition the follow-up period $[0, T]$ into a finite number K of disjoint intervals $[0, h_1), [h_1, h_2), \ldots, [h_{K-1}, h_K = T]$ and assume a constant hazard λ_k in the kth interval.

Model for time to toxicity

- Define the observed time $x_i = \min(u_i, t_i)$ and $\delta_{ik} = 1$ if the ith subject experiences toxicity in the kth interval.
- Letting $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_K\}$, when $\{y_i\}$ are completely observed, the likelihood function of $\boldsymbol{\lambda}$ for n subjects is given by

$$L(\boldsymbol{\lambda}) = \prod_{i=1}^{n} \prod_{k=1}^{K} (\lambda_k)^{\delta_{ik}} \exp\{-y_i \lambda_k e_{ik}\},$$

where $e_{ik} = h_k - h_{k-1}$ if $x_i > h_k$; $e_{ik} = x_i - h_{k-1}$ if $x_i \in [h_{k-1}, h_k)$; otherwise $e_{ik} = 0$.

Prior for λ_k

- We assume that *a priori* toxicity occurs uniformly throughout the assessment period $(0, T)$.
- The hazard in the middle of the kth partition is $\tilde{\lambda}_k = K/\{T(K - k + 0.5)\}$. Thus, we assign λ_k a gamma prior,
$$f(\lambda_k) = Ga(\tilde{\lambda}_k/C, 1/C).$$
- This prior has the mean of $\tilde{\lambda}_k$ and the variance of $C\tilde{\lambda}_k$, where C is a constant determining the size of the variance with respect to the mean.

Data Augmentation

- At the I step of the DA, we "impute" the missing data by drawing samples from their full conditional distribution.
  - Conditional on $\mathcal{D}_{\text{obs}} = (\mathbf{y}_{\text{obs}}, \mathbf{M})$ and model parameter θ, the full conditional distribution of $y_i \in \mathbf{y}_{\text{mis}}$ is

$$f(y_i|\mathcal{D}_{\text{obs}}, \theta) = \text{Bernoulli}\left(\frac{p_{d_i}^{\exp(\alpha)}\exp(-\sum_{k=1}^{K}\lambda_k e_{ik})}{1 - p_{d_i}^{\exp(\alpha)} + p_{d_i}^{\exp(\alpha)}\exp(-\sum_{k=1}^{K}\lambda_k e_{ik})}\right).$$

Posterior Step

- At the P step, given the imputed data $\mathbf{y}$, we sequentially sample the unknown model parameters from their full conditional distributions.
  - sample α from $f(\alpha|\boldsymbol{\theta}, \mathbf{y})$;
  - sample $\lambda_k, k = 1, \ldots, K$, from $f(\lambda_k|\boldsymbol{\theta}, \mathbf{y})$
- The DA iteratively draws samples of the missing data and model parameters through the imputation (I) step and posterior (P) step until convergence.

Dose-finding algorithm

- Patients in the first cohort are treated at the lowest dose d_1, or the physician-specified dose.
- At the current dose level j^{curr}, conditional on the cumulated data, we obtain estimates of the toxicity probabilities, $\hat{\pi}_j$ $(j = 1, \ldots, J)$, using DA.
- We then find dose level j^* that has a toxicity probability closest to ϕ,

$$j^* = \text{argmin}_{j \in (1,\ldots,J)} |\bar{\pi}_j - \phi|.$$

- If $j^{\text{curr}} > j^*$, we de-escalate the dose level to $j^{\text{curr}} - 1$;
- if $j^{\text{curr}} < j^*$, we escalate the dose level to $j^{\text{curr}} + 1$;
- otherwise, the dose stays at the same level as j^{curr} for the next cohort of patients.

Dose-finding algorithm

- Once the maximum sample size is reached, the dose that has the toxicity probability closest to ϕ is selected as the MTD.
- We require an early termination of a trial if the lowest dose is too toxic,

$$\text{pr}(\pi_1|D) > c.$$

where c is a constant, such as 96%.

Simulation

- Six dose levels and a maximum number of 12 cohorts in size of three
- The assessment period was $T = 3$ weeks
- The interarrival time between every two consecutive cohorts was $\tau = 0.5$
- Time to toxicity generated from Weibull distribution with approximately 70% of toxicities would occur in $(T/2, T)$.

Simulation

- We compared DA-CRM with CRM and TITE-CRM (with adaptive weight) based on 5000 simulated trials.
- In the CRM, we suspended the accrual until all of the toxicity outcomes in the trial were completely observed prior to the next dose assignment, i.e., no missing data.
- Such a complete-data CRM provide a benchmark for comparison.

Results: scenario 1

Table: Simulation with the CRM, TITE-CRM and LOT-CRM.

Design	Recommendation percentage at dose level						$N_{\text{MTD}+}$	Dur
	1	2	3	4	5	6		
Scenario 1	0.1	0.15	0.3	0.45	0.6	0.7		
CRM	0.6	13.8	61.9	22.9	0.6	0.0	9.0	36.4
# patients	4.8	7.2	14.9	7.6	1.3	0.1		
TITE-CRM	3.4	23.1	55.9	16.5	0.6	0.0	15.5	9.0
# patients	5.1	6.3	9.0	8.1	4.9	2.5		
DA-CRM	0.9	14.7	56.4	25.1	1.5	0.0	10.4	8.9
# patients	9.4	7.6	8.3	6.0	3.0	1.3		

Results: scenario 2

Design	Recommendation percentage at dose level						$N_{\text{MTD}+}$	Dur
	1	2	3	4	5	6		
Scenario 2	0.08	0.1	0.2	0.3	0.45	0.6		
CRM	0.0	1.4	23	55.9	18.8	0.8	6.6	36.5
# patients	4.1	4.1	9.0	12.2	5.5	1.0		
TITE-CRM	0.1	2.6	29.2	52.4	14.9	0.8	11.2	9.0
# patients	4.2	4.7	7.1	8.8	6.8	4.4		
DA-CRM	0.0	1.5	24.4	54.0	17.7	1.4	7.3	8.9
# patients	8.4	6.3	7.0	6.7	4.5	2.8		

Results: scenario 3

Design	Recommendation percentage at dose level							
	1	2	3	4	5	6	$N_{\text{MTD}+}$	Dur
Scenario 4	0.15	0.3	0.45	0.6	0.7	0.8		
CRM	17.8	56.9	22.7	0.8	0.0	0.0	10.6	36.0
# patients	10.2	14.8	8.9	1.5	0.1	0.0		
TITE-CRM	32.6	42.9	18.3	0.9	0.0	0.0	17.7	8.8
# patients	9.0	8.8	8.6	5.5	2.6	1.0		
DA-CRM	19.3	47.4	27.0	1.6	0.0	0.0	12.2	8.7
# patients	13.4	9.4	7.3	3.4	1.2	0.4		

What is next?

Software, Software, **Software!**

Conclusion

- We naturally casted the late-onset toxicity as a missing problem.
- We proposed the DA-CRM based on Bayesian data augmentation to address late-onset toxicities.
- Simulation study shows that the DA-CRM has good operating characteristics and outperforms available methods.

Reference

- Liu, S., Yin, G. and Yuan, Y. (2013) Bayesian Data Augmentation Dose Finding with Continual Reassessment Method and Delayed Toxicity. *Annals of Applied Statistics*, **4**, 2138-2156.
- The software is available for download from `http://odin.mdacc.tmc.edu/~yyuan/index_code.html`,
  or MD Anderson Biostatistics software download website `https://biostatistics.mdanderson.org/SoftwareDownload/`

Thank you !