### Time-to-event Bayesian Optimal Interval Design to Accelerate Phase I Trials

Ying Yuan

Professor, Chief of Section of Adaptive Clinical Trials MD Anderson Cancer Center

Joint work with Ruitao Lin (MDA), Daniel Li (Juno), Lei Nie (FDA) and Kathy Warren (NCI)

#### **Outline**

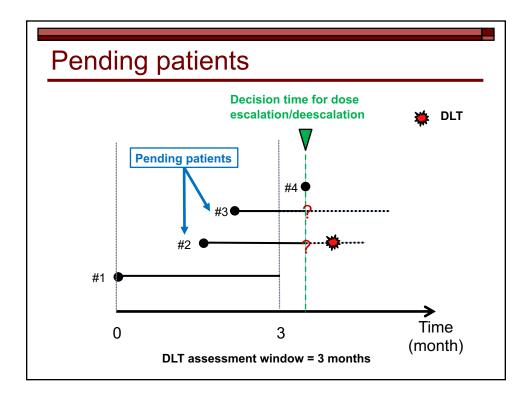
- Logistic difficulty associated with late-onset toxicity and fast accrual
- Time-to-event Bayesian Optimal Interval (TITE-BOIN) design
- Numerical Study
- Software

#### Late-onset toxicity

- Late-onset toxicity is common in the era of immunotherapy and targeted therapy.
  - In 36 clinical trials involving molecularly targeted agents, more than half of the 445 patients developed their high grade toxicity after the first cycle (Postel-Vinay et al., 2011, JCO).
  - Immuno-toxicity is often late-onset (June et al., 2017, Nat Med; Weber et al., 2015, JCO).
- Late-onset toxicity is also common in conventional radiochemotherapy.

#### Logistic difficulty with late-onset toxicity

- Late-onset toxicity causes a major logistic difficulty for conducting phase I trials.
- For example, if the DLT takes up to 8 weeks to evaluate and the accrual rate is 1 patient/week, on average, 5 new patients will be accrued while waiting to evaluate the previous 3 patients' outcomes.
- Question: How can new patients receive timely treatment when the previous patients' outcomes are pending?



#### Logistic difficulty with fast accrual

- The same logistic difficulty arises when the accrual is fast.
- Suppose that the DLT of a new agent can be assessed in the first 28-day cycle.
- If the accrual rate is 8 patients/28 days, then on average, 5 new patients will accrue while waiting to evaluate the previous 3 patients' outcomes.
- Question: How can new patients receive timely treatment when the previous patients' outcomes are pending?

#### Methods for late-onset toxicity

- Model-based approach: Time-to-event CRM (TITE-CRM; Cheung and Chappell, 2000), data argumentation CRM (DA-CRM; Liu et al., 2013)
  - Perform well, but are complicated to implement and subject to the influence of model misspecification
- Algorithm-based approach: Rolling 6 design (Skolnik et al., 2008), Rapid enrollment design (Ivanova, et al., 2016)

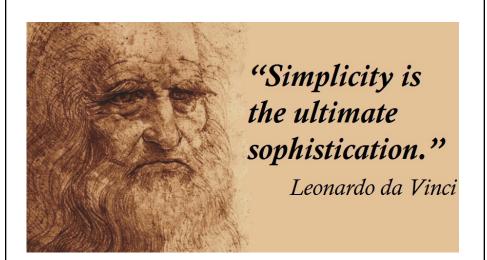
DLT Data   DLT Data									
No. Enrolled         No. DLTs         No. Without DLT         No. With Data Pending         3 + 3         Rolling Six         3 + 3         Rolling           2         0, 1         Any         Any         n         n         n         n         1         2         2         0         0         n - 1         n - 1         n - 1         n - 1         n         1         3         0         n - 1         <						Enrolling Dose Level*			
2 0,1 Any Any n n n 2 2 0 0 0 n-1 n-1 3 0 0,1,2 3,2,1 Suspend n 3 0 3 0 n+1 n+1 3 1 0,1 2,1 Suspend n 3 1 2 0 n n			DLT Data		MTD No	t Exceeded	MTD Exceeded		
2 2 0 0 n-1 n-1 3 0 0,1,2 3,2,1 Suspend n 3 0 3 0 n+1 n+1 3 1 0,1 2,1 Suspend n 3 1 2 0 n n	No. Enrolled	No. DLTs	No. Without DLT	No. With Data Pending	3 + 3	Rolling Six	3 + 3	Rolling Six	
3 0 0,1,2 3,2,1 Suspend n 3 0 3 0 n+1 n+1 3 1 0,1 2,1 Suspend n 3 1 2 0 n	2	0, 1	Any	Any	n	n			
3 0 3 0 n+1 n+1 3 1 0,1 2,1 Suspend n 3 1 2 0 n n	2	2	0	0	n – 1	n – 1			
3 1 0,1 2,1 Suspend n 3 1 2 0 n n	3	0	0, 1, 2	3, 2, 1	Suspend	n			
3 1 2 0 n n	3	0	3	0	n + 1	n + 1			
	3	1	0, 1	2, 1	Suspend	n			
3 ≥ 2 Any Any n - 1 n - 1	3	1	2	0	n	n			
	3	≥ 2	Any	Any	n – 1	n – 1			

#### Objective

Can we have a design that combines the good performance of the TITE-CRM with the simplicity of rolling 6 design?



A model-assisted design!



#### **Notation**

- Let T denote the pre-specified DLT assessment window
  - T should be long enough to cover all DLTs that are relevant to defining the MTD
- $y_i$  is the DLT indicator, such that  $y_i$  = 1 if patient experiences DLT in (0, T], otherwise  $y_i$  =0
- Suppose that at a moment of decision making, n patients are enrolled at the current dose, r patients have completed the DLT assessment (i.e., their DLT data y<sub>i</sub> are observed), denoted as 0.

#### **Notation**

- c = n r patients have not completed the DLT assessment (i.e., their DLT data  $y_i$  are pending/missing).
- Denote these pending patients as M.
- $t_i(< T)$  denotes the follow-up time for the patient whose DLT data are pending, i.e.,  $i \in M$ .

# Bayesian Optimal Interval (BOIN) Design $\frac{\leq \lambda_e}{\text{Compute the DLT rate at the current dose}}$ Within $(\lambda_e, \lambda_d)$ Retain the current dose $\frac{\text{DLT rate at the current dose}}{\text{Compute the DLT rate at the current dose}} = \frac{\text{No. of patients experienced DLT at the current dose (ntox)}}{\text{No. of patients treated at the current dose (n)}}$

#### Escalation/de-escalation boundaries

Table 1. Dose esca	Table 1. Dose escalation and de-escalation boundaries										
		Target toxicity rate for the MTD									
Boundary	0.1	0.15	0.2	0.25	0.3	0.35	0.4				
$\lambda_{\rm e}$ (escalation)	0.078	0.118	0.157	0.197	0.236	0.276	0.316				
$\lambda_{d}$ (de-escalation)	0.119	0.179	0.238	0.298	0.358	0.419	0.479				

• Escalation and de-escalation boundaries  $\lambda_e$  and  $\lambda_d$  are derived to minimize the probability of making incorrect decisions of dose escalation and de-escalation.

Yuan Y, Hess K, Hilsenbeck, S and Gilbert M (2016), *Clinical Cancer Research*, 22, 4291-4301.

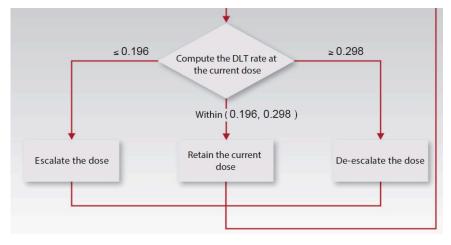
#### Escalation/de-escalation boundaries

Table 1. Dose escalation and de-escalation boundaries										
	Target toxicity rate for the MTD									
Boundary	0.1	0.15	0.2	0.25	0.3	0.35	0.4			
$\lambda_{ m e}$ (escalation)	0.078	0.118	0.157	0.197	0.236	0.276	0.316			
$\lambda_{d}$ (de-escalation)	0.119	0.179	0.238	0.298	0.358	0.419	0.479			

• Escalation and de-escalation boundaries  $\lambda_e$  and  $\lambda_d$  are derived to minimize the probability of making incorrect decisions of dose escalation and de-escalation.

Yuan Y, Hess K, Hilsenbeck, S and Gilbert M (2016), Clinical Cancer Research, 22, 4291-4301

#### BOIN design for target = 25%



DLT rate at the current dose  $=\frac{\text{No. of patients experienced DLT at the current dose (ntox)}}{\text{No. of patients treated at the current dose (n)}}$ 

#### **BOIN** under late-onset toxicity

 BOIN makes decision based on the empirical (maximum likelihood) estimate of the toxicity rate at the current dose

$$\hat{p} = \frac{\sum_{i \in O} y_i + \sum_{i \in M} y_i}{n}$$

- **Problem**:  $y_i$  is not observed for pending patients (i.e.,  $i \in M$ )
- **Strategy**: to replace unobserved  $y_i$  with its predicted value  $\hat{y}_i$

$$\hat{p} = \frac{\sum_{i \in O} y_i + \sum_{i \in M} \widehat{y}_i}{n}$$

#### Impute missing/pending data

Assuming that the time to DLT  $X_i$  follows a uniform distribution over [0,T], the expected value of  $y_i$ ,  $i \in M$ , for a pending patient with follow-up time  $t_i$  is

$$\hat{y}_i = E(y_i | X_i > t_i) = \Pr(y_i = 1 | X_i > t_i)$$

$$= \frac{p\left(1 - \frac{t_i}{T}\right)}{p\left(1 - \frac{t_i}{T}\right) + (1 - p)} \approx \frac{p\left(1 - \frac{t_i}{T}\right)}{(1 - p)}$$

#### Impute missing/pending data

Thus

$$\hat{p} = \frac{\sum_{i \in O} y_i + \sum_{i \in M} \hat{y}_i}{n}$$

$$= \frac{s + \frac{p}{1 - p}(c - \text{STFT})}{n}$$

where **STFT** = $\sum_{i \in M} t_i / T$  is the standardized total follow-up time (STFT) for pending patients at the current dose, and s is the number of patients who experienced DLT at the current dose

#### TITE-BOIN decision table (target=0.3)

**Table S1.** Dose escalation and de-escalation boundaries for TITE-BOIN with a target DLT rate of 0.3 and cohort size of 3.

No.	No.	No.		STFT		No.	No.	No.		STFT	
treated	DLTs	data pending	Escalate	Stay	De- escalate	treated	DLTs	data pending	Escalate	Stay	De- escalate
3	0	≤1	Y			12	2	5	≥2.72	<2.72	
3	0	≥2	Sus	pend acci	rual	12	2	6	≥4.11	<4.11	
3	1	0		Y		12	2	≥7	Sus	pend acc	rual
3	1	1		>0.88	≤0.88	12	3	≤6		Y	
3	1	≥2	Sus	pend acci	rual	12	3	≥7	Sus	pend acc	rual
3	2	≤1			Y	12	4	0		Y	
3	3	0			Y&Elim	12	4	1		>0.43	≤0.43
6	0	≤3	Y			12	4	2		>1.50	≤1.50
6	0	≥4		pend acci	rual	12	4	3		>2.57	≤2.57
6	1	≤1	Y			12	4	4		>3.65	≤3.65
6	1	2	≥0.60	< 0.60		12	4	5		>4.72	≤4.72
6	1	3	≥1.96	<1.96		12	4	6		>5.79	≤5.79
6	1	≥4	Sus	pend acci	rual	12	4	≥7	Sus	pend acc	rual
6	2	0		Y		12	5, 6	≤7			Y
6	2	1		>0.73	≤0.73	12	≥7	≤5			Y&Elim
6	2	2		>1.80	≤1.80	15	0	≤7	Y		
6	2	3		>2.87	≤2.87	15	0	≥8		pend acc	rual
6	2	≥4	Sus	pend acci		15	1	≤7	Y		
6	3	≤3			Y	15	1	≥8	Sus	pend acc	rual
6	≥4	≤2			Y&Elim	15	2	≤5	Y		
9	0	≤4	Y			15	2	6	≥0.35	< 0.35	
9	0	≥5	Sus	pend acci	rual	15	2	7	≥2.07	< 2.07	



4					
	Ma	N.	No.	STFT	
	No. treated	No. DLTs	data pending	Escalate Stay De- escalate	
	3	0	≤1	Y	
	3	0	≥2	Suspend accrual	
- 1	2	- 1	^	3.7	

STFT (Standardized Total Follow-up Time) =

Sum of the follow up time for pending patients at the current dose

The length of DLT assessment window

-	_		Suspense accium
6	1	≤1	Y
6	1	2	≥0.60 <0.60
6	1	3	≥1.96 <1.96
6	1	≥4	Suspend accrual
6	2	0	Y
6	2	1	>0.73 ≤0.73
6	2	2	>1.80 ≤1.80
6	2	3	>2.87 ≤2.87

#### Incorporate prior information

- Partition the assessment window [0, T] into three parts: the initial part [0, T/3], the middle part (T/3, 2T/3] and the final part (2T/3, T]
- Let  $(\pi_1, \pi_2, \pi_3)$  be the prior probability that the DLT would occur at the three parts of the assessment window
- Weighted STFT (WSTFT) weights follow-up time using  $(\pi_1, \pi_2, \pi_3)$
- Remarkably, using an informative prior for the time to DLT does not alter the decision table!

#### Safety rules

- If >50% patient's DLT data are pending at the current dose, we suspend the accrual.
- During trial conduct, we impose the following overdose control / safety stopping rule:

If  $\Pr(p > \phi | y, n) > 0.95$  and  $n \ge 3$ , eliminate the current and higher doses from the trial; if the lowest dose is eliminated, terminate the trial early for safety.

where  $\phi$  is the target DLT rate, and  $\Pr(p_j \ge \phi | n_j, y_j)$  can be evaluated based on a beta-binomial model.

#### Selection of the MTD

When the maximum sample size is reached, stop the trial and select the dose whose isotonic estimate of the toxicity rate is closest to the target φ as the MTD.

#### Simulation

- A phase I trial with 7 dose levels.
- The DLT assessment window is 3 months, the accrual rate is 2 patients/month.
- The time to DLT is sampled from a Weibull distribution, with 50% of DLTs occurring in the second half of the assessment window.
- The maximum sample size is 36 patients, treated in cohorts of 3.
- The target DLT rate = 0.2 or 0.3, with 8 representative scenarios for each rate, resulting in 16 scenarios

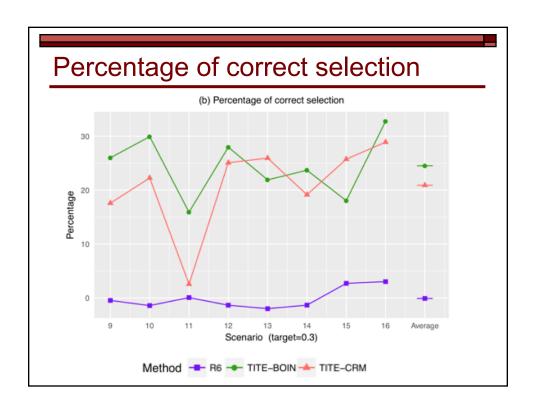
Scenarios									
Scena	rio			Dose	level				
Scena	1110	1	2	3	4	5	6	7	
				Target	DLT rate	e is 0.2			
1		0.05	0.20	0.46	0.50	0.60	0.70	0.80	
2		0.02	0.05	0.20	0.28	0.34	0.40	0.44	
3		0.01	0.05	0.10	0.20	0.32	0.50	0.70	
4		0.01	0.04	0.07	0.10	0.50	0.70	0.90	
5		0.01	0.05	0.10	0.14	0.20	0.26	0.34	
6		0.01	0.02	0.03	0.05	0.20	0.40	0.50	
7		0.01	0.04	0.07	0.10	0.15	0.20	0.25	
8		0.01	0.02	0.03	0.04	0.05	0.20	0.45	
				Target	DLT rate	e is 0.3			
9		0.30	0.40	0.50	0.60	0.70	0.80	0.90	
10		0.14	0.30	0.39	0.48	0.56	0.64	0.70	
11		0.07	0.23	0.41	0.49	0.62	0.68	0.73	
12		0.05	0.15	0.30	0.40	0.50	0.60	0.70	
13		0.05	0.12	0.20	0.30	0.38	0.49	0.56	
14		0.01	0.04	0.08	0.15	0.30	0.36	0.43	
15		0.02	0.04	0.08	0.10	0.20	0.30	0.40	
16		0.01	0.03	0.05	0.07	0.09	0.30	0.50	

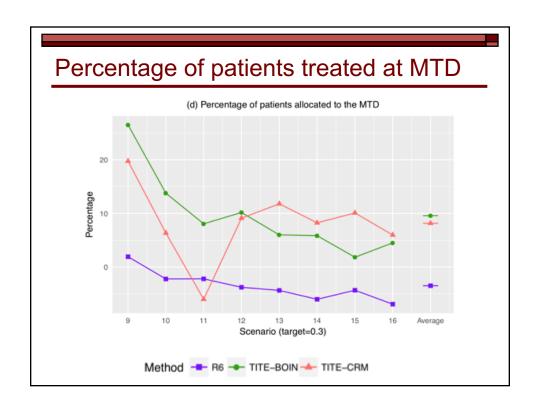
#### **Simulation**

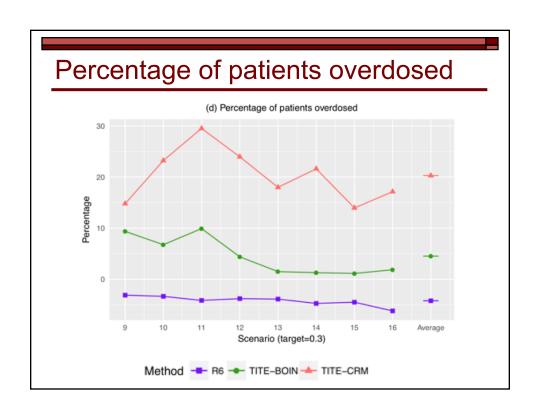
- Compare TITE-BOIN, 3+3 design, R6 design, and TITE-CRM
- For the 3+3 design, a new cohort is enrolled only when the previous cohort's DLT data are cleared
- Because the 3+3 and R6 designs often stopped the trial early (e.g., when 2 of 3 patients experienced DLT) before reaching 36 patients, in these cases, the remaining patients are treated at the selected "MTD" as the cohort expansion, such that the four designs have comparable sample sizes

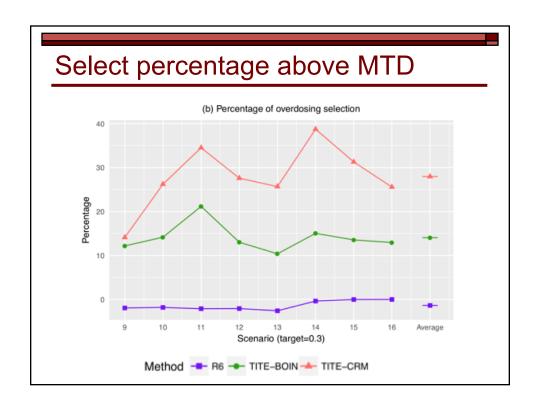
#### Performance metrics

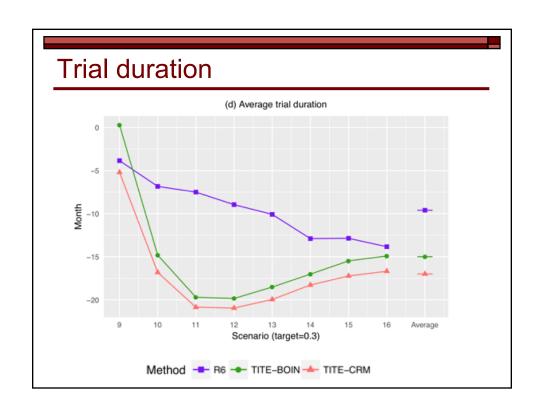
- Percentage of correct selection of the MTD
- Percentage of patients allocated to the MTD
- Percentage of overdosing selection (i.e., selecting a dose above the MTD)
- Percentage of patients overdosed (i.e., treated at doses above the MTD)
- Average trial duration











#### Summary

- By leveraging the follow-up time data from pending patients, TITE-BOIN is more efficient than rolling 6 design, and yields comparable accuracy to identify the MTD as TITE-CRM.
- TITE-BOIN is safer than TITE-CRM, and can be implemented in a simple and transparent way as rolling 6 design.
- TITE-BOIN has great potential to shorten the trial duration and accelerate drug development.

#### Application (protocol 2018-0899)

- Phase I study of a PARP inhibitor in combination with radiation therapy for recurrent gynecologic cancers
- DLT assessment window = 5 months
- 4 doses
- Target = 0.3
- Elicited prior toxicity probability in the trimesters of the assessment window  $(\pi_1, \pi_2, \pi_3)$ =(0.43, 0.46, 0.11)
- FDA protocol, IRB approved, ready to accrue

ecision table								
0010.0								
		ision rules for large-field	cohort	D 11 D				
# of Treated		# of Patients with	E1-4-	Decision Ru				
Patients	with DLTs	Pending Information	Escalate	Stay	De-escalate			
_	0	≤1	Y					
2	0	2		Suspend accru				
	≥1	≤1			Y			
4	0	≤2	Y					
	0	≥3		Suspend accru	ıal			
	1	0	Y					
	1	1	≥0.76	< 0.76				
	1	2	≥1.84	<1.84				
	1	≥3		Suspend accru	ıal			
	2	≤2		•	Y			
	≥3	≤1			Y&Elim			
	0	≤3	Y					
	0	≥4		Suspend accru	ıal			
	1	≤2	Y	•				
6	1	3	≥1.4	<1.4				
	1	≥4		Suspend accru	ıal			
	2	_ ≤4			Y			
	≥3	≤3			Y&Elim			

#### Application (protocol 2018-1129)

- Phase I study of BMS-986301 in advanced solid cancers
- 8 dose levels
- DLT assessment window = 28 days, but accrual is expected to be fast
- Accelerated titration + TITE-BOIN

#### **Decision table**

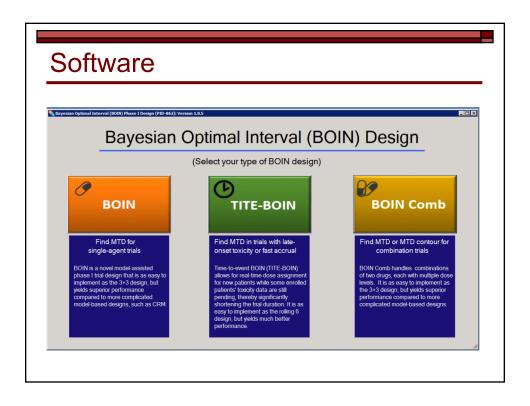
Table 1. Dose escalation/de-escalation rule for the TITE-BOIN design, pg. 1 of

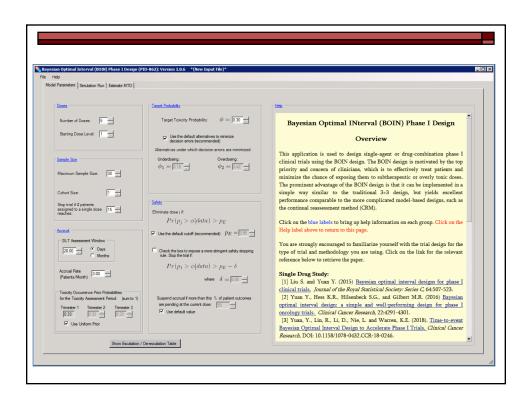
No.	No.	No.	STFT					
treated	DLTs	data pending	Escalate	Stay	De-escalate			
3	0	≤1	Υ					
3	0	≥2		Suspend accrua	ı			
3	1	0		Υ				
3	1	1		>0.88	≤0.88			
3	1	≥2		Suspend accrua				
3	2	≤1			Υ			
3	3	0			Y&Elim			
6	0	≤3	Υ					
6	0	≥4		Suspend accrua				
6	1	≤1	Υ					
6	1	2	≥0.6	<0.6				
6	1	3	≥1.96	<1.96				
6	1	≥4		Suspend accrua				

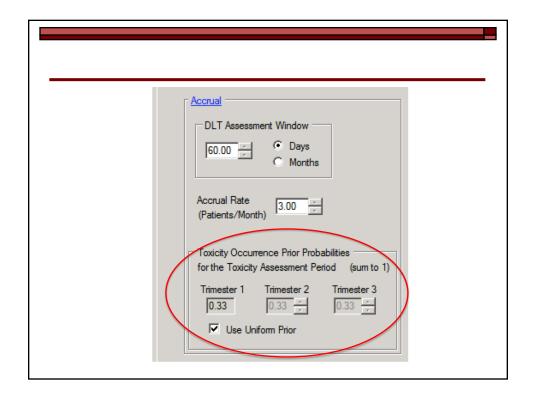
#### Software

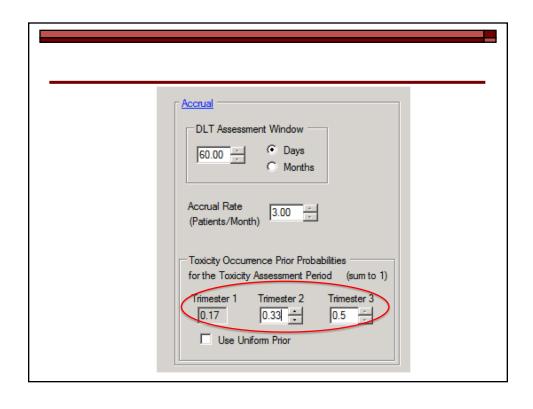
- Windows desktop program for TITE-BOIN is freely available at the MD Anderson Software Download Website
  - https://biostatistics.mdanderson.org/softwaredownload/ SingleSoftware.aspx?Software\_ld=81.
- Web applications for TITE-BOIN is freely available at http://www.trialdesign.org.

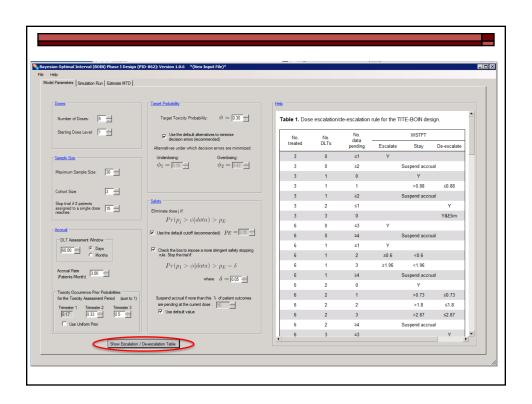


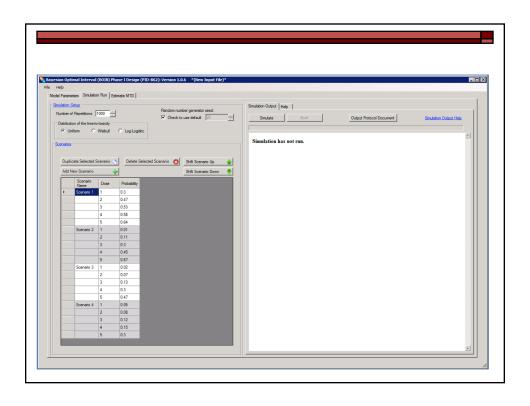


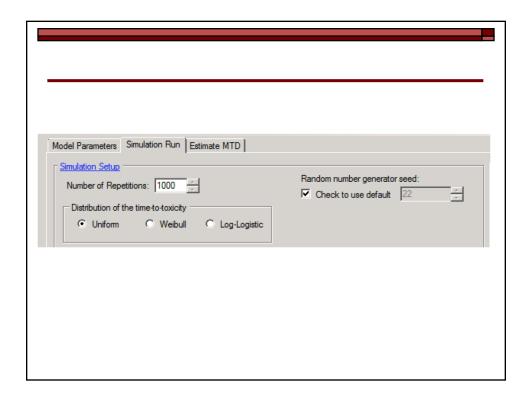


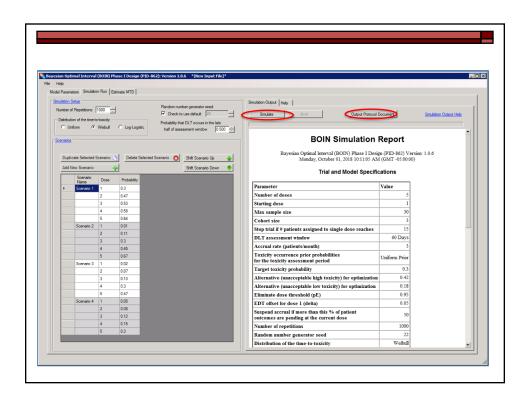


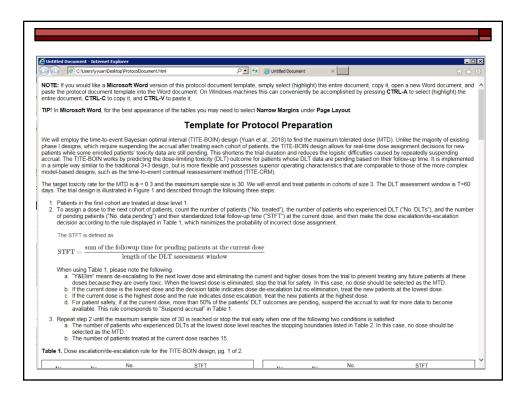


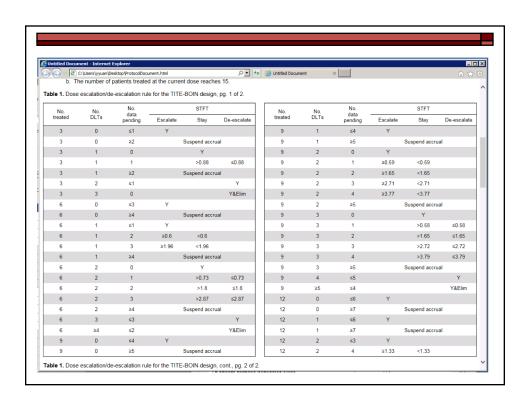


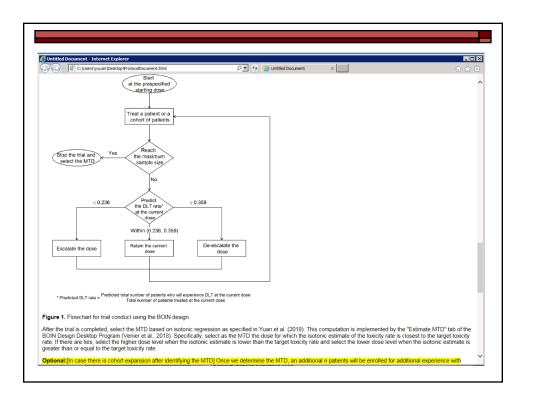


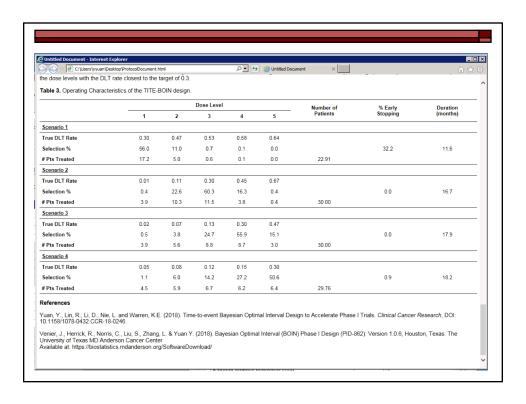








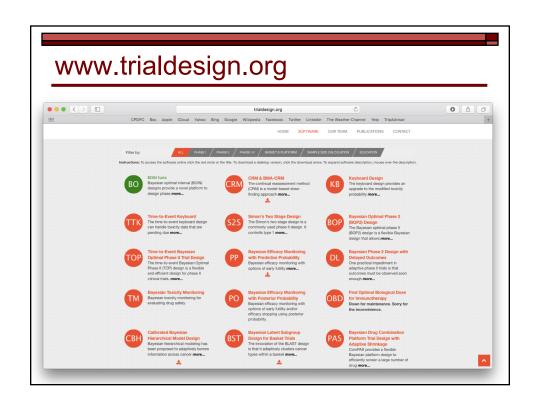


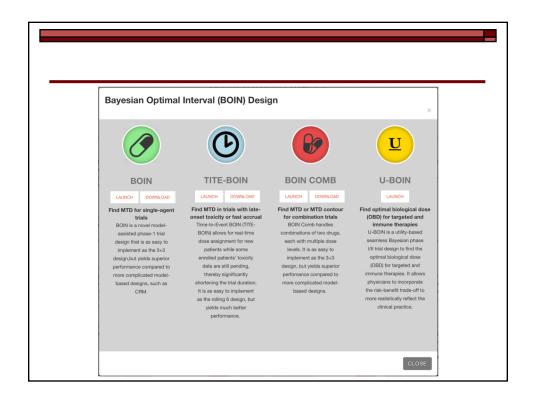


## Anything better?

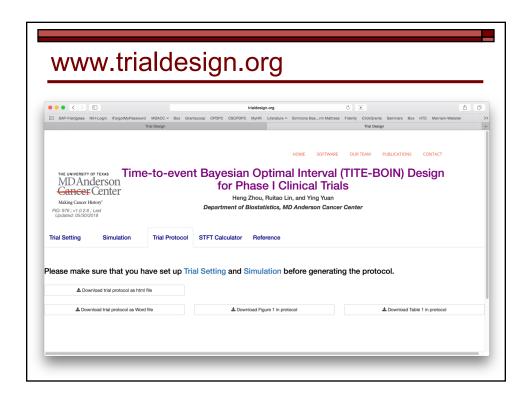












#### References

- Zhou H, Yuan Y and Nie L (2018) Accuracy, safety and reliability of novel Bayesian phase I trial designs. *Clinical Cancer Research*, 24(18):4357-4364
- Yuan Y, Lin R, Li D, Nie L and Warren KE (2018) Time-to-event Bayesian optimal interval design to accelerate phase I trials, Clinical Cancer Research, 24, 4921-4930.
- Lin R and Yuan Y (2019) Time-to-event model-assisted designs for dose-finding trials with delayed toxicity, *Biostatistics*, in press (discussed TITE-keyboard design and TITE-mTPI design)

