Dose-Finding Oncology Trials Through the Lens of FDA's Draft Guidance on Bayesian Methods

Bayesian Biostatistics Conference 2026, Leiden

Burak Kürşad Günhan[1], Pavel Mozgunov[2], Anja Victor[1]

22 October 2026

[1]Merck Healthcare KGaA, Darmstadt, Germany
[2]MRC Biostatistics Unit, Cambridge, UK

FDA draft guidance (U.S. FDA, 2026) and dose finding

- "Dose-Finding Trials in Oncology" including the Bayesian Logistic Regression Model (BLRM) (Neuenschwander et al., 2008) is listed as a situation where Bayesian methods have been used.[1]
- Success criterion (Section IV.A): $\Pr(d > a) > c$

  d: population-level summary, a: threshold, c: probability level
- No universal values for a and c, unlike the one-sided $\alpha = 0.025$.
- **This talk: what are d, a and c in dose escalation, and how do we justify them?**

[1] *Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products: Draft Guidance for Industry*, January 2026. FDA. `https://www.fda.gov/media/190505/download`

FDA draft guidance (U.S. FDA, 2026) and dose finding

- "Dose-Finding Trials in Oncology" including the Bayesian Logistic Regression Model (BLRM) (Neuenschwander et al., 2008) is listed as a situation where Bayesian methods have been used.[1]
- Success criterion (Section IV.A): $\Pr(d > a) > c$

  d: population-level summary, a: threshold, c: probability level
- No universal values for a and c, unlike the one-sided $\alpha = 0.025$.
- **This talk: what are d, a and c in dose escalation, and how do we justify them?**

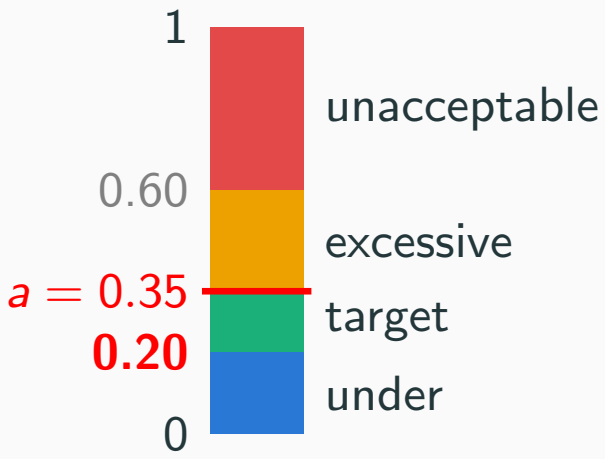
[1] *Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products: Draft Guidance for Industry*, January 2026. FDA. `https://www.fda.gov/media/190505/download`

The DLT estimand: what is π_x? (Mercier et al., 2025)

Estimand attributes

Objective	Characterize the dose–toxicity relationship and determine the MTD
Endpoint	DLTs observed within the DLT period (yes/no)
Intercurrent events (composite strategy)	• $< 70\%$ of the planned dose due to an AE • Growth factors, or transfusions during the DLT period
Summary measure	Probability of a DLT during the DLT period at dose x: $\boldsymbol{\pi_x}$

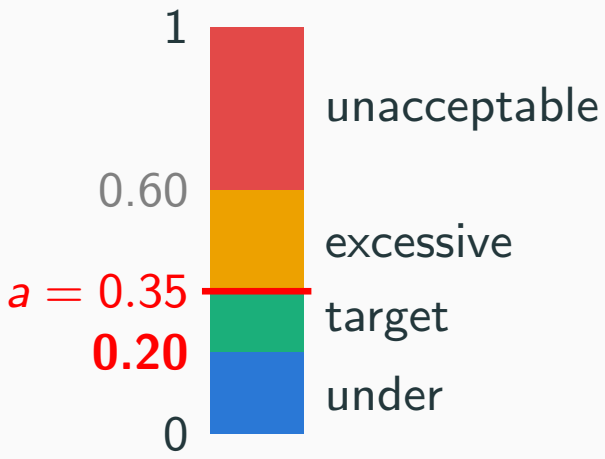
Toxicity intervals and the threshold a (Neuenschwander et al., 2008)



- 0.20 and $a = 0.35$ bound the targeted toxicity interval and are used directly as a in $\Pr(\pi_x < a) > c$ (e.g. for the EWOC approach).
- 0.60 is used for the loss function approach (weight w_U).

Toxicity intervals are informed by clinical knowledge.

Toxicity intervals and the threshold a (Neuenschwander et al., 2008)



- 0.20 and $a = 0.35$ bound the targeted toxicity interval and are used directly as a in $\Pr(\pi_x < a) > c$ (e.g. for the EWOC approach).
- 0.60 is used for the loss function approach (weight w_U).

Toxicity intervals are informed by clinical knowledge.

Objective: determining the MTD

> *"... c is a minimum probability level that would support a conclusion of effectiveness."*

	Guidance: efficacy	**Dose escalation**
Claim	the treatment is effective	dose x is tolerable and/or MTD
Posterior probability	$\Pr(d > a) > c$	$\Pr(\pi_x < 0.35) > c_{\text{EWOC}}$
c guards against	a false claim of effectiveness	a false claim of tolerability
Decision-theoretic	*"success criteria that minimize the expected loss"*	loss function (Neuenschwander et al., 2008)

Like a success criterion, the dose-escalation criterion needs to be pre-specified.

Objective: determining the MTD

> *"... c is a minimum probability level that would support a conclusion of effectiveness."*

	Guidance: efficacy	**Dose escalation**
Claim	the treatment is effective	dose x is tolerable and/or MTD
Posterior probability	$\Pr(d > a) > c$	$\Pr(\pi_x < 0.35) > c_{\text{EWOC}}$
c guards against	a false claim of effectiveness	a false claim of tolerability
Decision-theoretic	*"success criteria that minimize the expected loss"*	loss function (Neuenschwander et al., 2008)

Like a success criterion, the dose-escalation criterion needs to be pre-specified.

Decision criteria after each cohort

(i) Recommendation for the next cohort

Next dose	**EWOC** (Babb et al., 1998)	*or*	**Loss** (Neuenschwander et al., 2008)
	highest dose with $\Pr(\pi_x < 0.35) > c_{\text{EWOC}}$		lowest $\sum_k w_k \Pr(\pi_x \in I_k)$

(ii) Is the trial completed?

Declare MTD	$n_x \geq 6$ and $\Pr(\pi_x < 0.35) > c_{\text{MTD}}$
Stop	$\Pr(\pi_{x_1} \geq 0.35) > c_{\text{stop}}$ (x_1: lowest dose)

Loss weights	Under-dosing	Target	Excessive	Unacceptable
I_k	$[0, 0.20)$	$[0.20, 0.35)$	$[0.35, 0.60)$	$[0.60, 1]$
w_k	1	0	1	2

Example (Neuenschwander et al., 2008)

Published data

Dose (mg)	1	2.5	5	10	25
Patients	3	4	5	4	2
DLTs	0	0	0	0	2

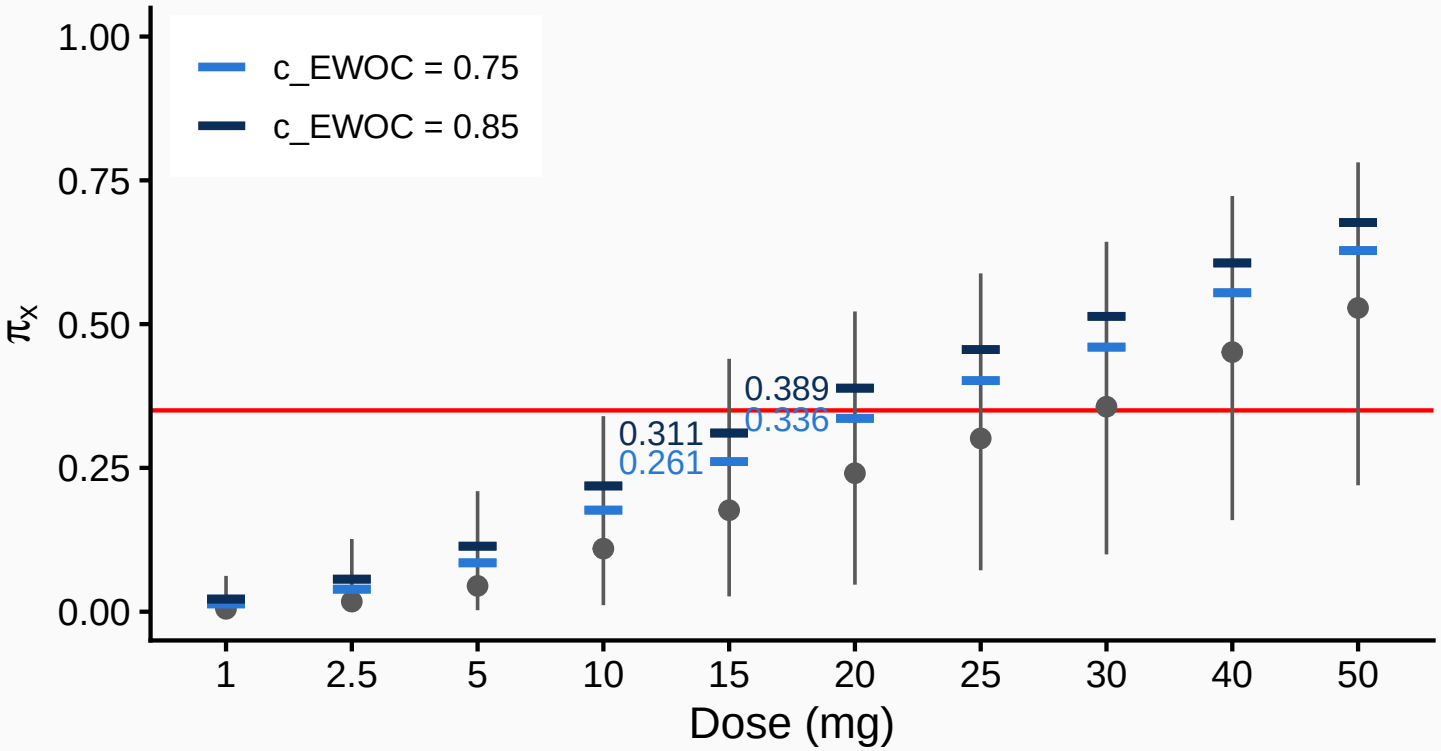
Decision rules for the next dose

EWOC	$\Pr(\pi_x < 0.35) > c_{\text{EWOC}}$
Loss function	lowest $E[\text{loss}] = \sum_k w_k \Pr(\pi_x \in I_k)$

I_k: toxicity intervals (under-dosing, target, excessive, unacceptable)

Same data and model throughout: only c_{EWOC} and w change.

Next dose depends on c_{EWOC}

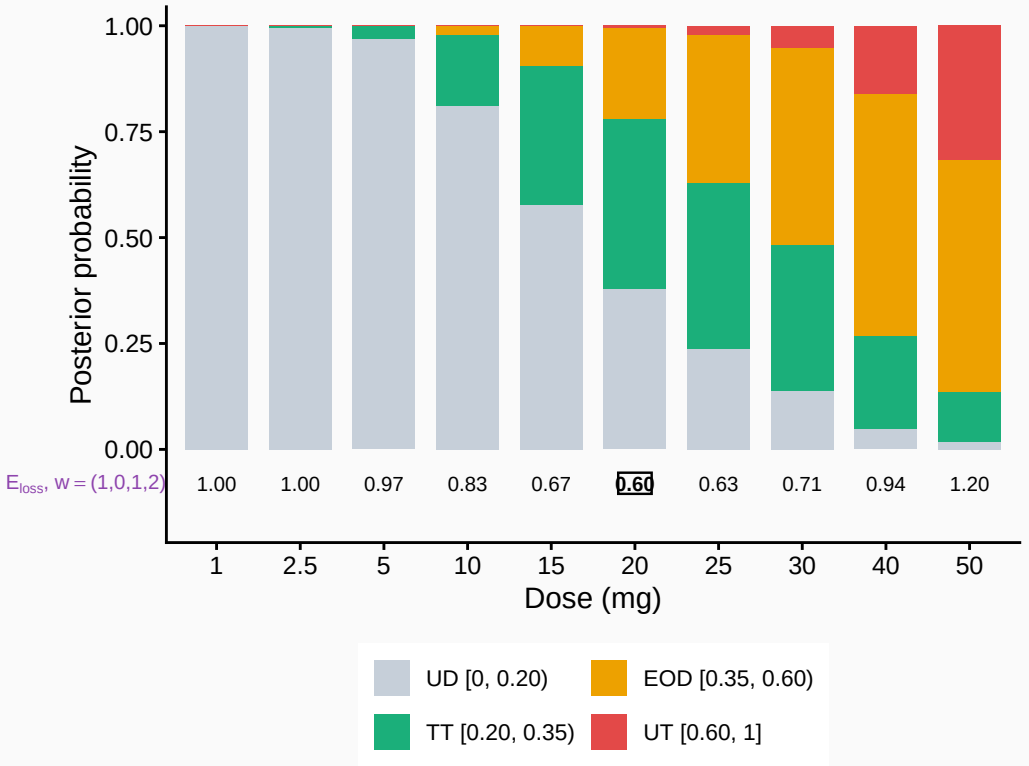


A dose is allowed if it satisfies the EWOC criterion.

$$\Pr(\pi_x < 0.35) > c_{EWOC} \iff q_c(\pi_x) < 0.35$$

$c_{EWOC} = 0.75 \rightarrow 20$ mg $\qquad$ $c_{EWOC} = 0.85 \rightarrow 15$ mg

Next dose depends on weights w

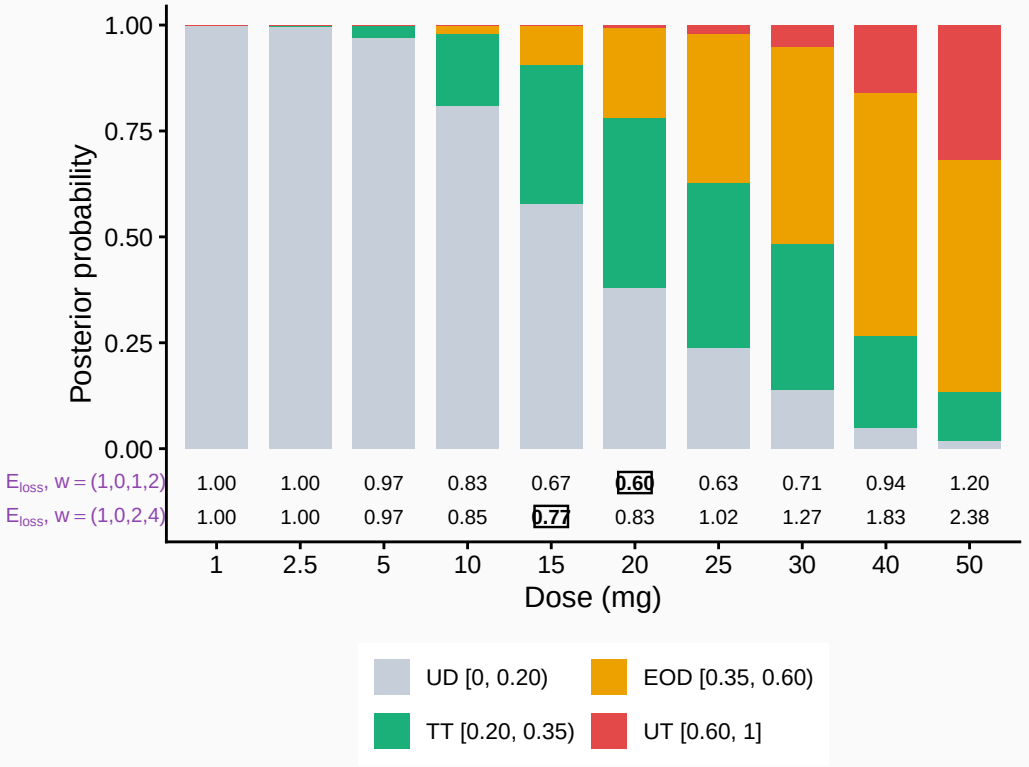


next dose = lowest expected loss

$$E[\text{loss}] = \sum_k w_k \Pr(\pi_x \in I_k)$$

$$w = (1, 0, 1, 2) \rightarrow 20 \text{ mg}$$

Next dose depends on weights w



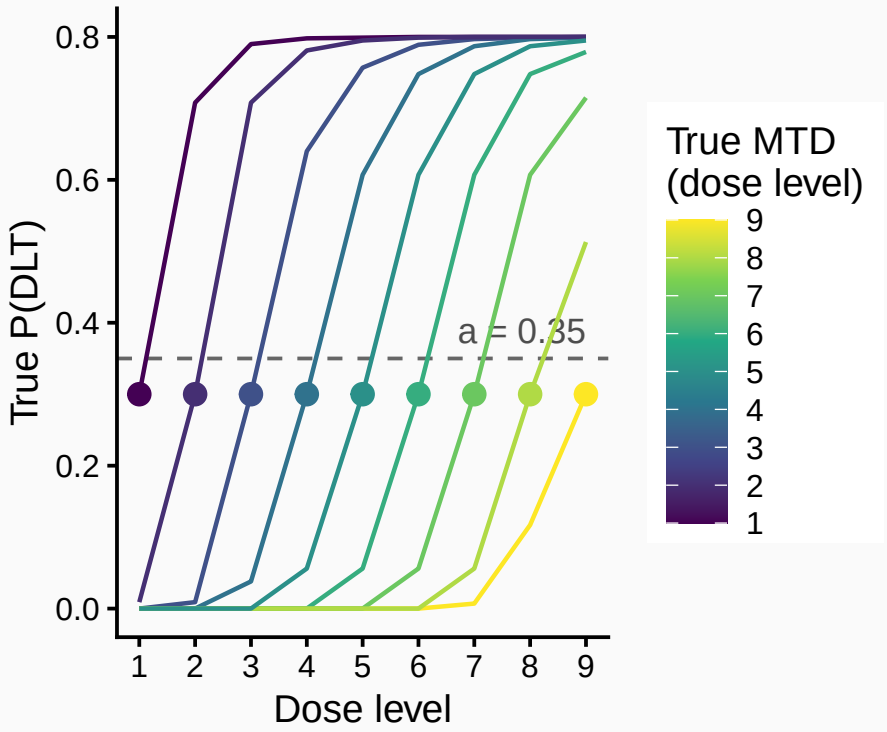
$$\text{next dose} = \text{lowest expected loss}$$

$$E[\text{loss}] = \sum_k w_k \Pr(\pi_x \in I_k)$$

$w = (1, 0, 1, 2) \rightarrow 20$ mg $\qquad$ $w = (1, 0, 2, 4) \rightarrow 15$ mg

Same data, heavier overdose weights, lower next dose.

Simulation settings



9 simulated designs

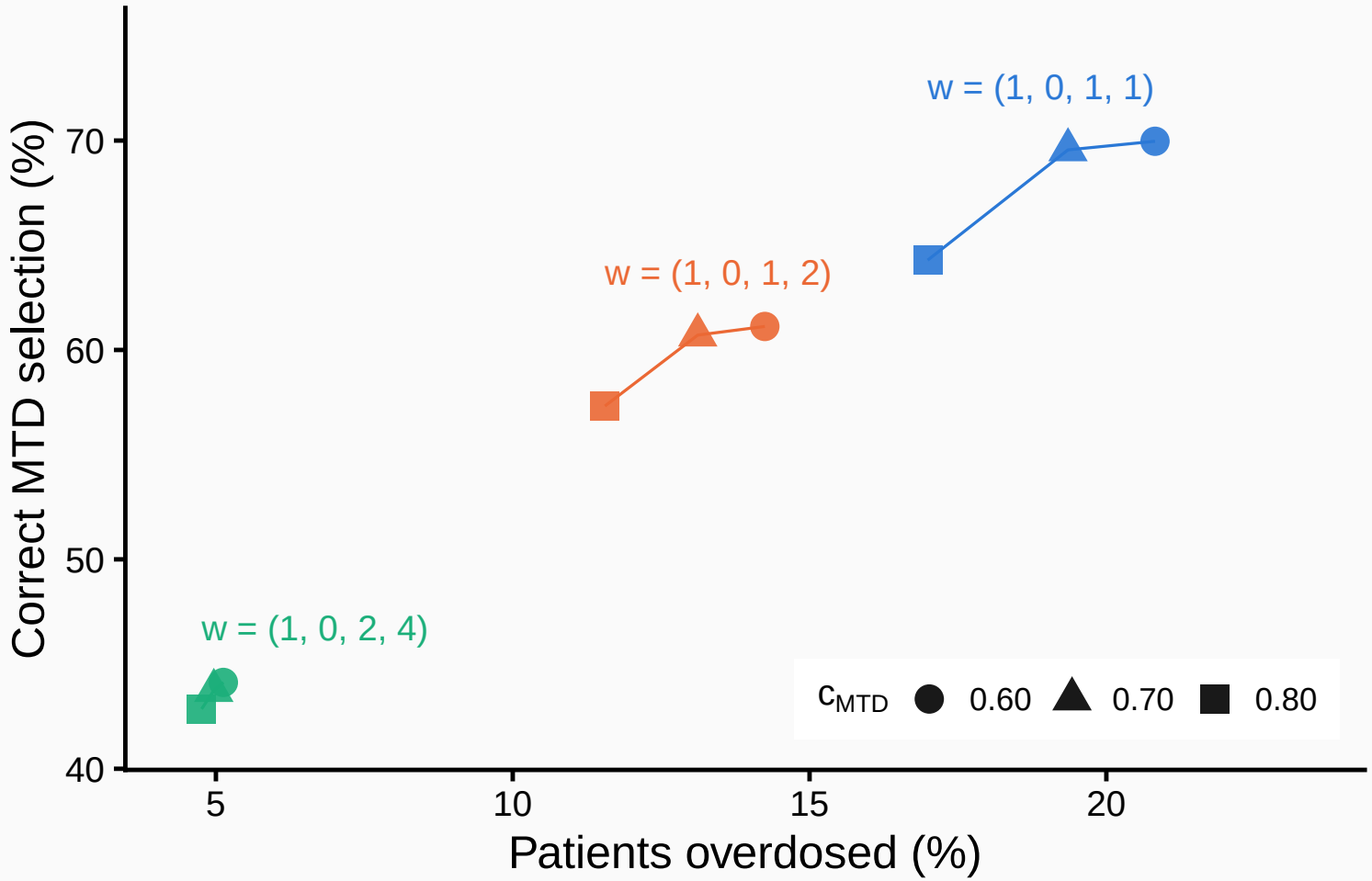
	c_{MTD}		
w	0.60	0.70	0.80
(1, 0, 1, 1)	•	•	•
(1, 0, 1, 2)	•	•	•
(1, 0, 2, 4)	•	•	•

MTD criterion: $n_x \geq 6$ and c_{MTD} (varied)

Stopping rule: $c_{\text{stop}} = 0.65$ (fixed)

Anonymised design of a trial (different from the example). Only loss function approaches (no EWOC). 1000 trials per scenario.

Operating characteristics: loss weights w and c_{MTD}



FDA draft guidance applied to dose escalation

Guidance	Dose escalation
$\Pr(d > a) > c$	π_x, toxicity intervals, c_{EWOC}, c_{MTD}, c_{stop}
decision-theoretic	loss function for the next dose, weights w
each decision point	after every cohort: next dose, declare MTD, stop
operating characteristics	**c and w should also be evaluated through OCs at the design stage**

Conclusions

- The DLT estimand defines π_x; clinical knowledge informs the toxicity-interval thresholds.
- Given the same data and prior, different c or w can lead to different decisions.
- **There are no universal values for c and w, evaluate them by simulations at the design stage.**

References

Babb, J., Rogatko, A., & Zacks, S. (1998). Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine*, 17(10), 1103–1120.

Mercier, F., Homer, V., Geng, J., Zhang, H., Englert, S., Kan-Dobrosky, N., & Victor, A. (2025). Estimands in oncology early clinical development: assessing the impact of intercurrent events on the dose-toxicity relationship. *Statistics in Biopharmaceutical Research*, 17(1), 78–86.

Neuenschwander, B., Branson, M., & Gsponer, T. (2008). Critical aspects of the Bayesian approach to phase I cancer trials. *Statistics in Medicine*, 27(13), 2420–2439.

U.S. FDA (2026). *Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products: Guidance for Industry (Draft Guidance)*. Technical report, U.S. Food and Drug Administration.

Thanks!

- Email: burak-kuersad.guenhan@merckgroup.com
- GitHub: https://github.com/gunhanb

Backup: operating characteristics, all 9 designs

w	c_{MTD}	Correct (%)	Over-toxic (%)	Under-toxic (%)	No MTD (%)	Overdosed (%)
w = (1, 0, 1, 1)	0.60	70.0	8.7	13.4	0.5	20.8
w = (1, 0, 1, 1)	0.70	69.6	5.3	15.8	1.6	19.4
w = (1, 0, 1, 1)	0.80	64.3	2.7	18.7	6.3	17.0
w = (1, 0, 1, 2)	0.60	61.1	6.8	25.4	0.1	14.2
w = (1, 0, 1, 2)	0.70	60.7	3.9	27.7	0.4	13.1
w = (1, 0, 1, 2)	0.80	57.3	2.6	29.8	3.0	11.6
w = (1, 0, 2, 4)	0.60	44.1	1.8	49.0	0.0	5.1
w = (1, 0, 2, 4)	0.70	43.7	1.9	49.2	0.0	5.0
w = (1, 0, 2, 4)	0.80	42.9	1.6	49.9	0.2	4.8

Mean over 9 MTD scenarios, 1000 trials each; $c_{\text{stop}} = 0.65$, no EWOC.