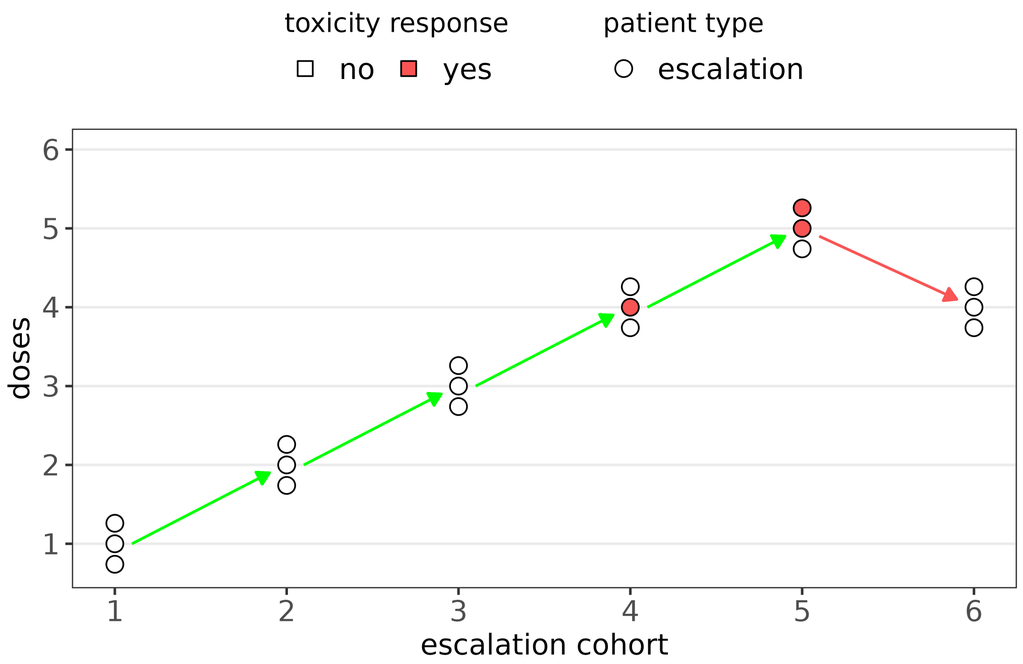
The impact of backfilling on early phase dose optimisation trials in oncology

James Willard[1], Thomas Jaki[1,2], Burak Kürsad Günhan[3], Christina Habermehl[3], Anja Victor[3], and Pavel Mozgunov[1]

[1] MRC Biostatistics Unit, University of Cambridge, United Kingdom
[2] Faculty of Informatics and Data Science, University of Regensburg, Germany
[3] Merck Healthcare KGaA, Darmstadt, Germany

2026-06-15

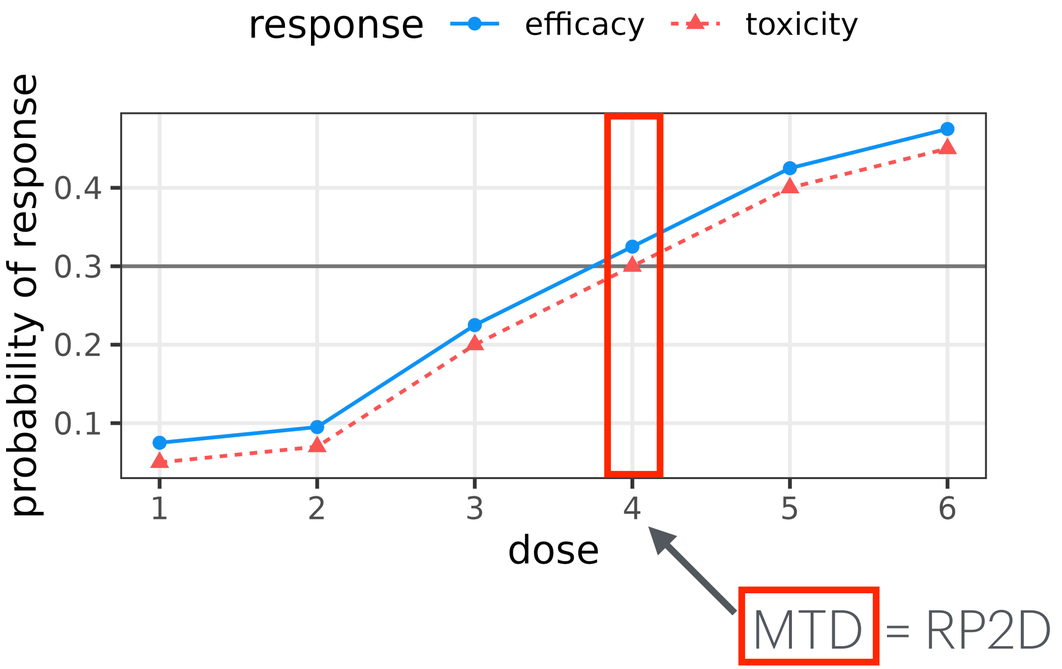
Phase 1 Dose escalation



- Phase 1 trials in oncology evaluate first-in-human doses of experimental cancer therapy
- Dose escalation trials typically enroll only a small number of patients (e.g., 20-40)
- Due to unknown toxicity of drug, must sequentially escalate from smaller to larger doses - so-called “dose escalation”
- Decisions (arrows) are guided by a statistical model using accruing trial data

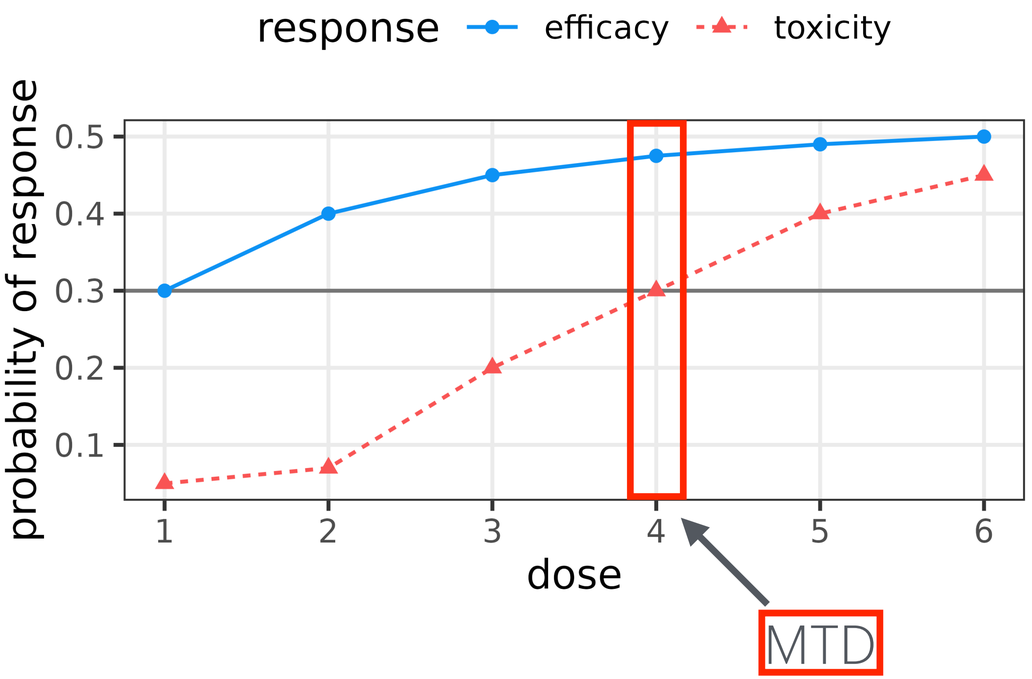
Traditional chemotherapies

- Paradigm: “more is better”
  - greater toxicity implies greater efficacy
  - Binary responses, $P(\text{toxicity}) = \theta$ and $P(\text{efficacy}) = \phi$
  - assumed that ϕ closely follows θ
  - Thus, largest “safe” dose was commonly RP2D
- Goal: identify *maximum tolerated dose* (MTD)
  - i.e., largest dose j where $\theta_j \leq 0.3$
  - Example: MTD is dose 4



Newer therapies

- Challenge the old “more is better” paradigm
    - e.g., immunotherapies, targeted therapies
- Sufficient efficacy may be observed at doses lower than MTD (Le Tourneau et. al. 2009)
- Possible that lower doses yield similar efficacy *but with less toxicity*
- So RP2D may be $\leq$ MTD

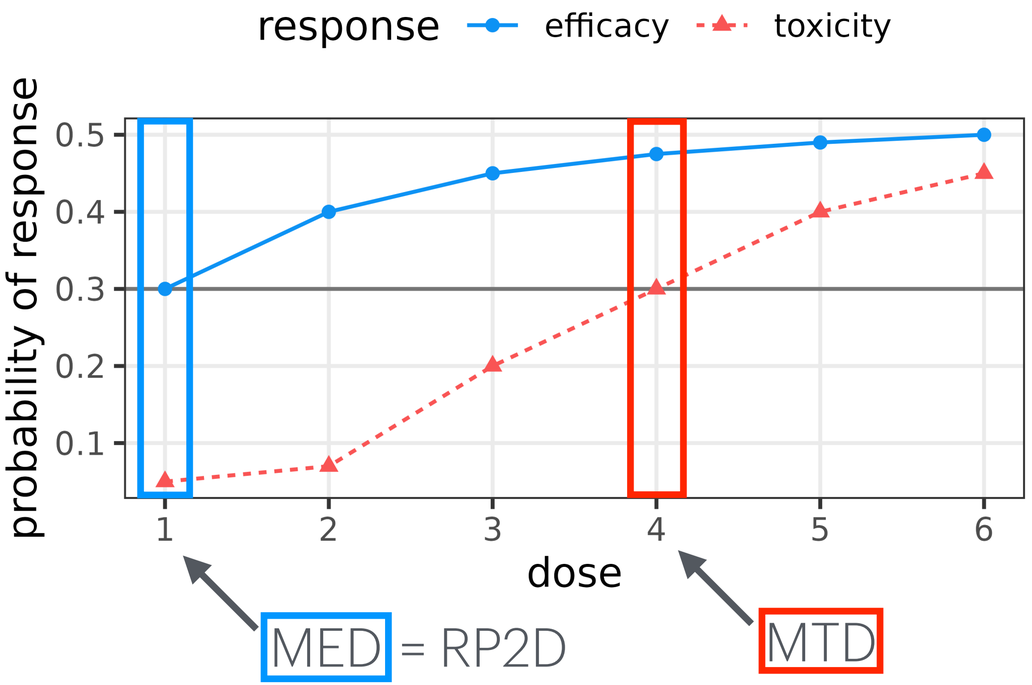


Project Optimus

- Recognizing this, FDA's Project Optimus argues for a revised paradigm toward early phase dose finding in oncology (FDA 2024a)
- Specifically, a greater emphasis is place on *dose optimisation* in Phase 1 trials (FDA 2024b)
- *Goal*: find RP2D which *balances tradeoff* between efficacy and toxicity
- Example RP2D:
    - Smallest dose with sufficient efficacy - "minimum effective dose" (Groenland et. al. 2021)
    - Safe dose at which efficacy curve begins to plateau (Dehbi et. al. 2021)
    - Dose which optimizes utility function that explicitly defines risk-benefit tradeoff (Yuan et. al. 2024)

Newer therapies

- Interest now in identifying different RP2D with a better risk/benefit tradeoff than the MTD
- Define *minimum effective dose* (MED) as RP2D
  - smallest safe dose j s.t. $\phi_j \geq 0.3$
- May still want to identify MTD along with MED, however

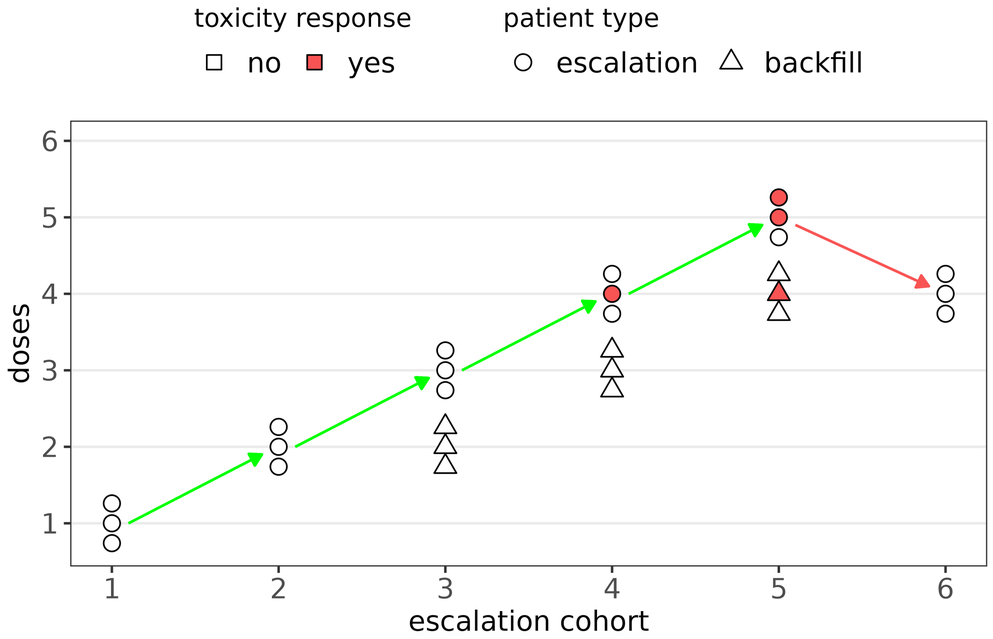


Challenges of dose optimisation

- Ideal world:
    - Collect toxicity and efficacy at same time and search for MED directly!
- Rarely possible in practice: efficacy responses often come *long after* toxicity in oncology
    - e.g., obtain toxicity response at 1 month vs 3 months for efficacy
- Important: *cannot delay dose escalation*
    - Bring new therapies to patients quickly, 11 month trial vs 33 month trial
- Thus, *only toxicity responses used for dose escalation*
- BUT: we still want more efficacy information before moving to next phase - backfill and dose expansion!

Dose optimisation: escalation w/ backfill + dose expansion

Dose escalation w/ backfill

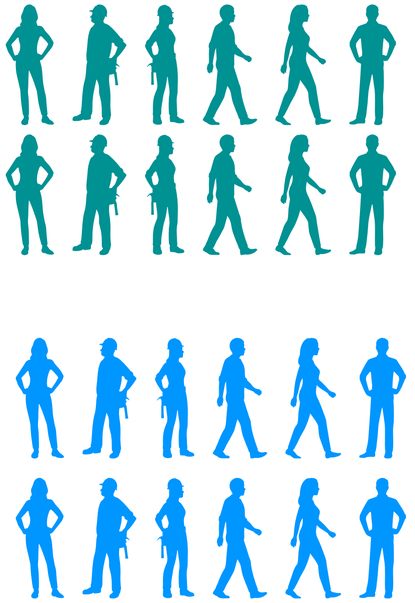


2 doses

Dose expansion

d_a

d_b



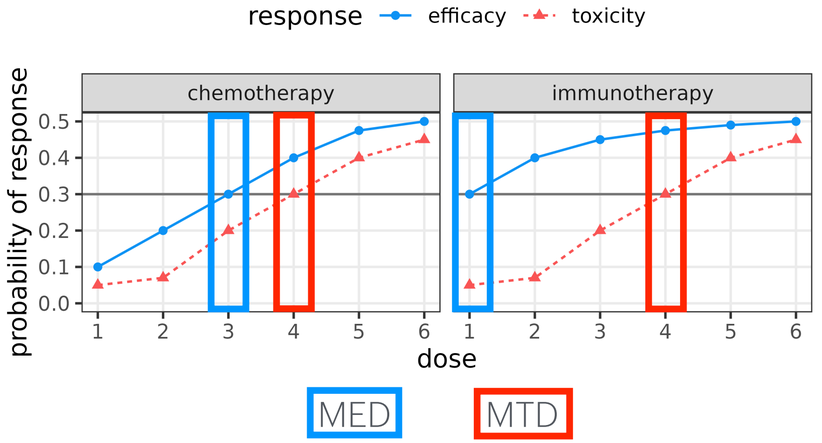
RP2D

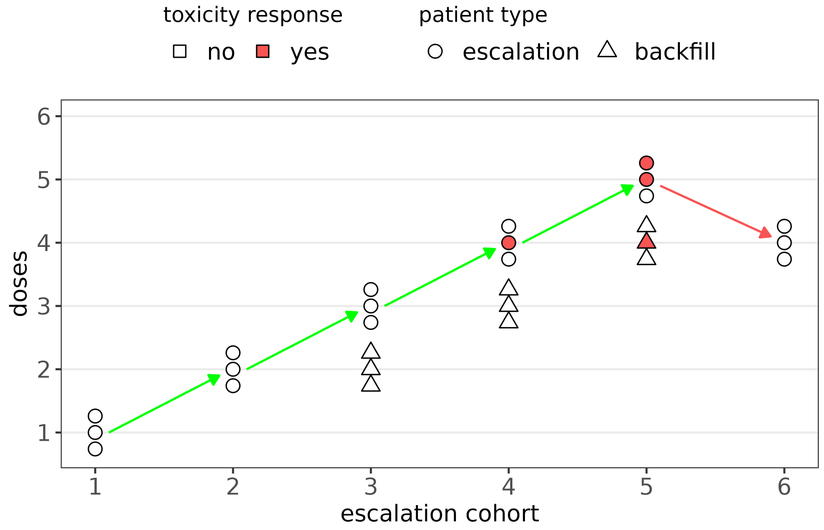
Dose escalation w/ backfill

- Dose escalation targets MTD only
- When MTD is high dose and distance b/w MED and MTD is large
  - less information near MED

Collect more via backfill

- Backfill: place *additional* patients on lower (safe) doses
  - ER/BRAR (Dehbi et. al. 2021, Pin et. al. 2024)
  - One below current (Barnett et. al. 2023)
- After escalation, select two doses for expansion:
  - Estimated *MTD and MED*





Open questions: backfill

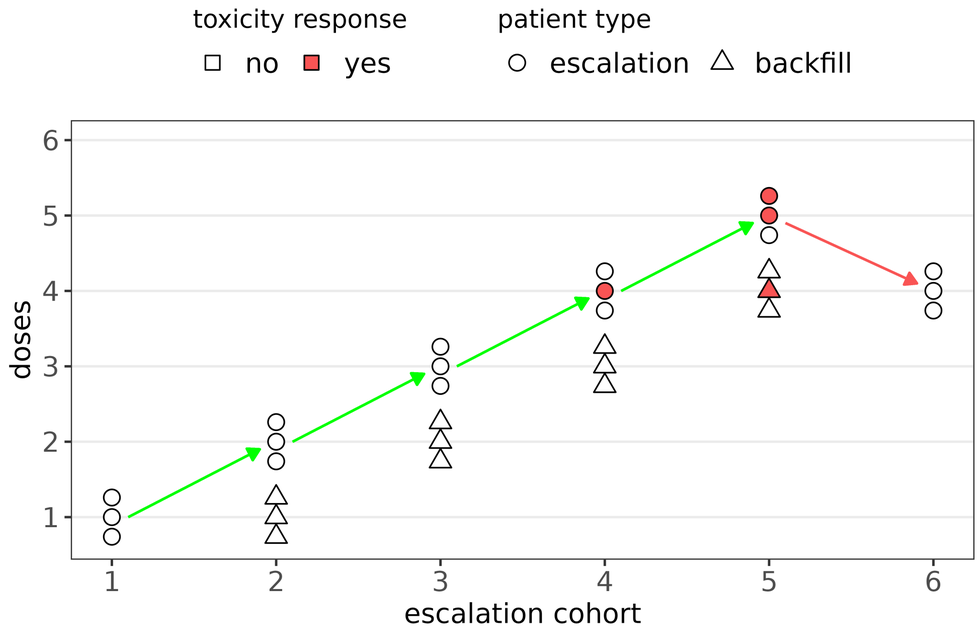
1. When to start backfill?
2. How to select backfill doses?
3. How many patients to use for backfill?
4. How does backfilling impact selection of the expansion doses?

Investigate these via simulation study

Simulation study: backfill

When to start/how many patients for backfill?

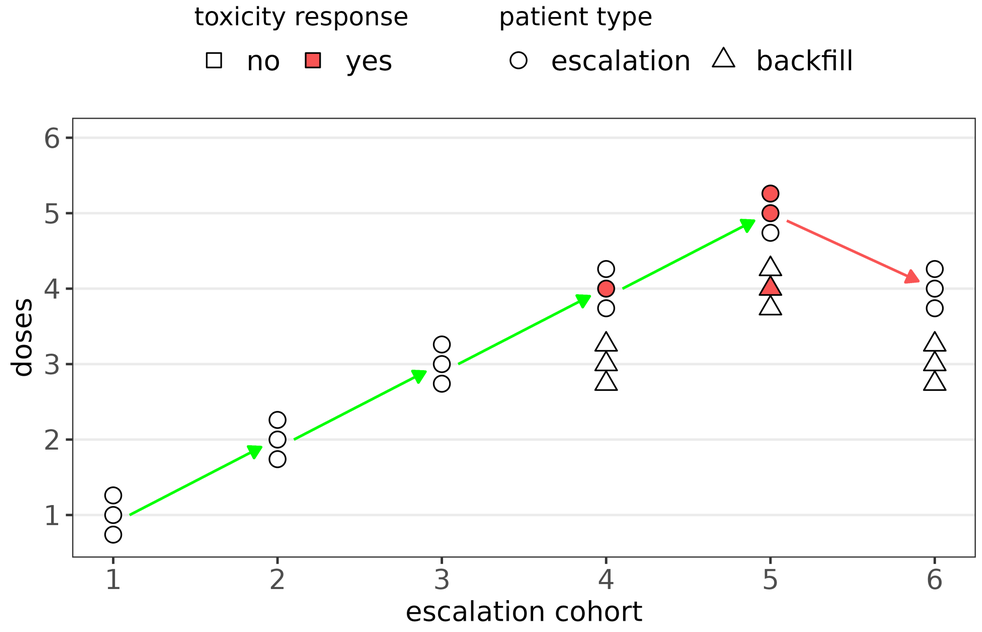
Start sooner/permit more BF patients



Min n escalation patients before bf begins: 3

Max n bf patients: 12

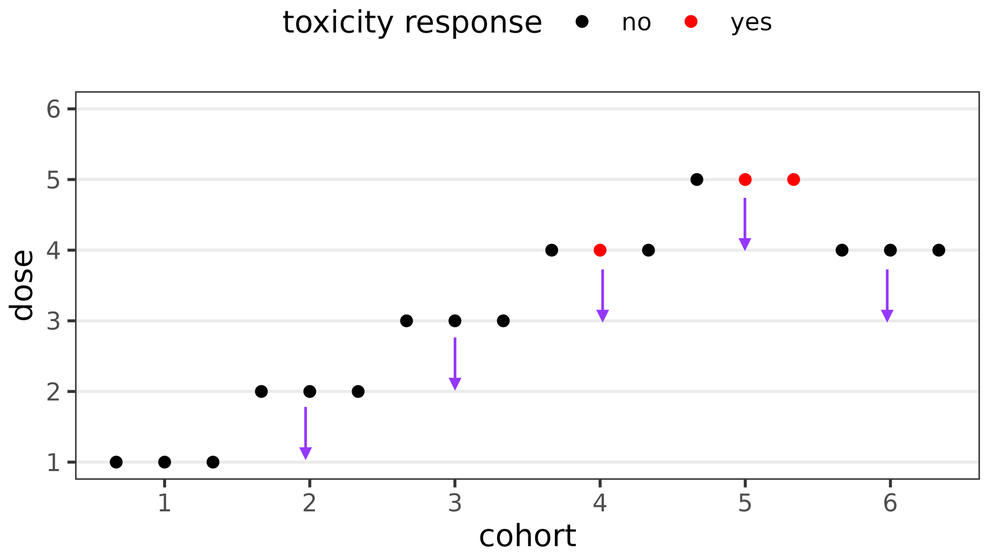
Start later/permit fewer BF patients



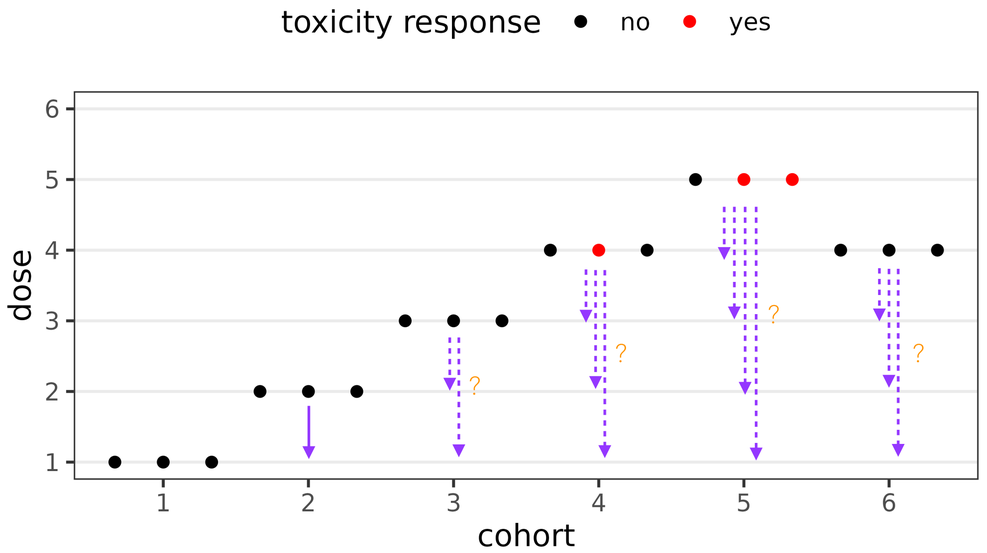
Min n escalation patients before bf begins: 9

Max n bf patients: 9

How to select backfill doses?



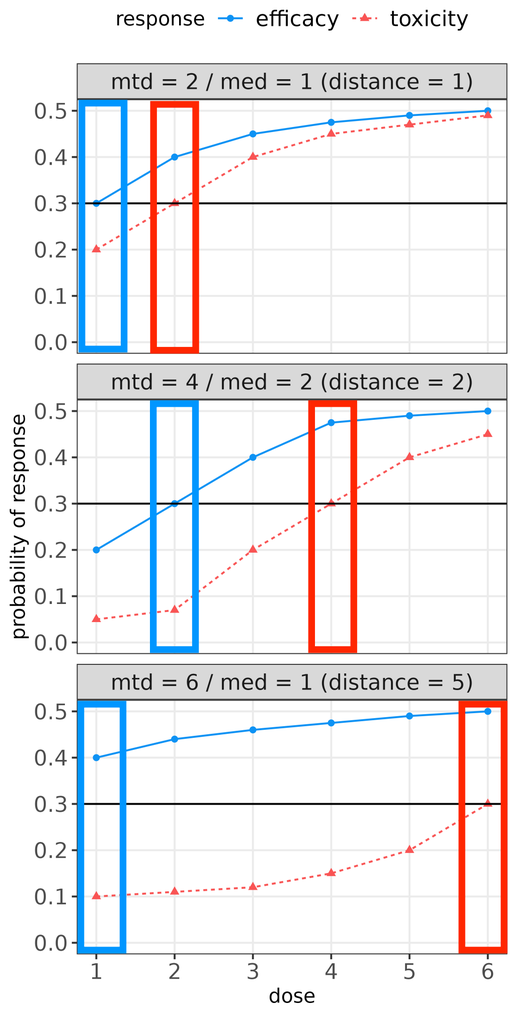
- *One below*: bf dose one below highest dose used in testing
- Uses only toxicity information



- *Efficacy*: bf dose smallest (safe) dose with some evidence of efficacy
- $\pi(\phi_j > 0.3 \mid Data) > \delta$ (where $\delta = 0.2$ here)
- Uses efficacy responses *which have accrued up to that point*

Simulation setup

- Maximum sample size escalation + expansion = 90
- Two month delay between tox/eff responses
- Three scenarios on right
  - low/medium/large MTD
  - small/medium/large distance between MTD and MED
- Dose escalation
  - Bayesian logistic regression model
  - Escalation decision: $d_{t+1}^{ESC} = \text{argmax}_{\mathcal{A}_t} \pi(\theta_j \in [0.2, 0.3] \mid Data)$
  - Ends when stopping rule is triggered



MED MTD

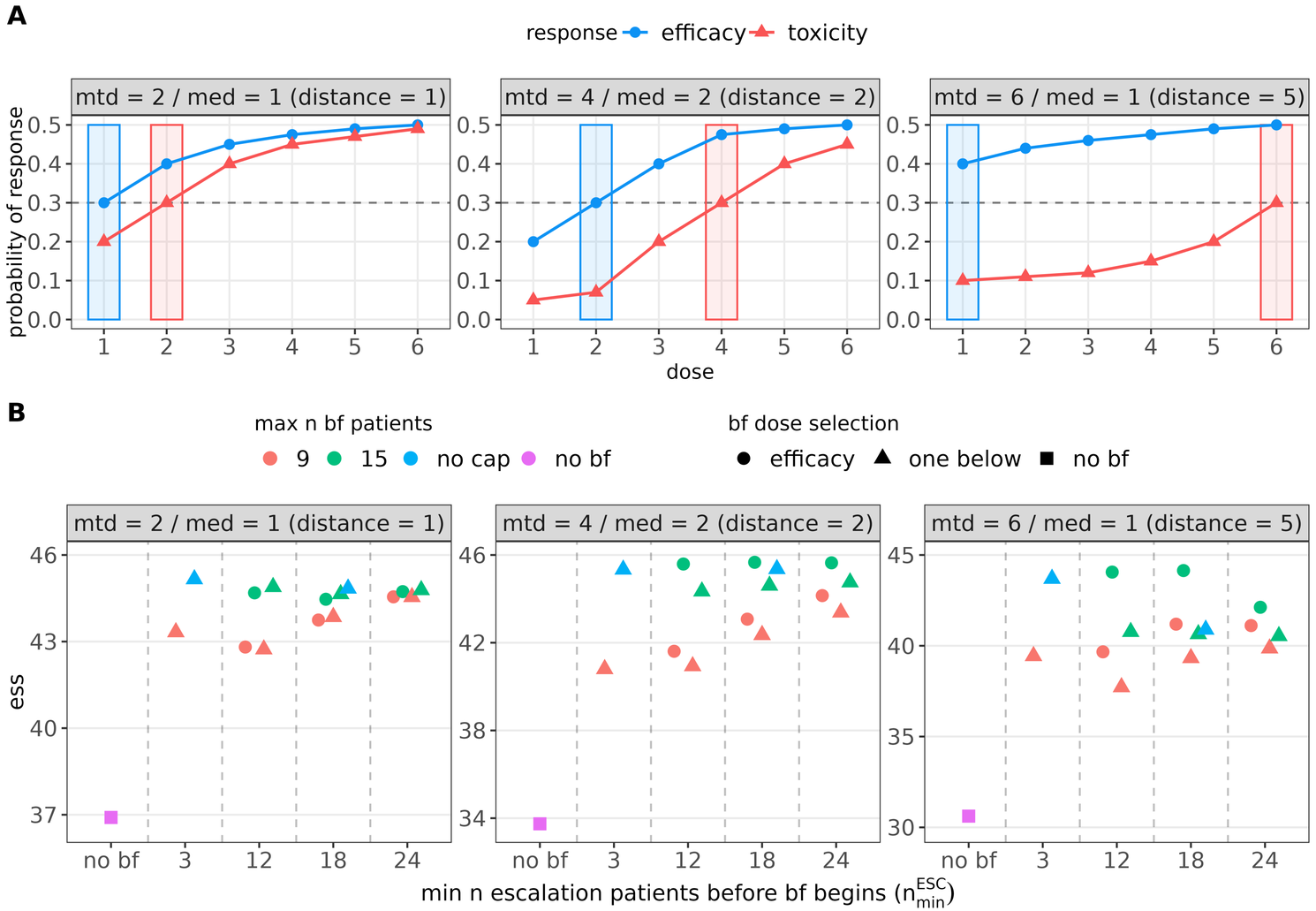
Simulation setup

- At end of escalation, all efficacy responses collected/two expansion doses selected
  - Dose 1: estimated MTD
  - Dose 2: estimated MED, smallest dose j where $\pi(\phi_j > 0.3 \mid Data) > 0.2$
- *Min n escalation patients before bf begins*: 3, 12, 18, and 24 responses
- *Max n bf patients*: 9, 15, and no cap on n bf patients
- *Backfill selection rules*: no backfill, one below rule, efficacy rule
- Performance metrics
  - ESS, PCSMTD, PCSMED, and PEDIR

Performance Metrics

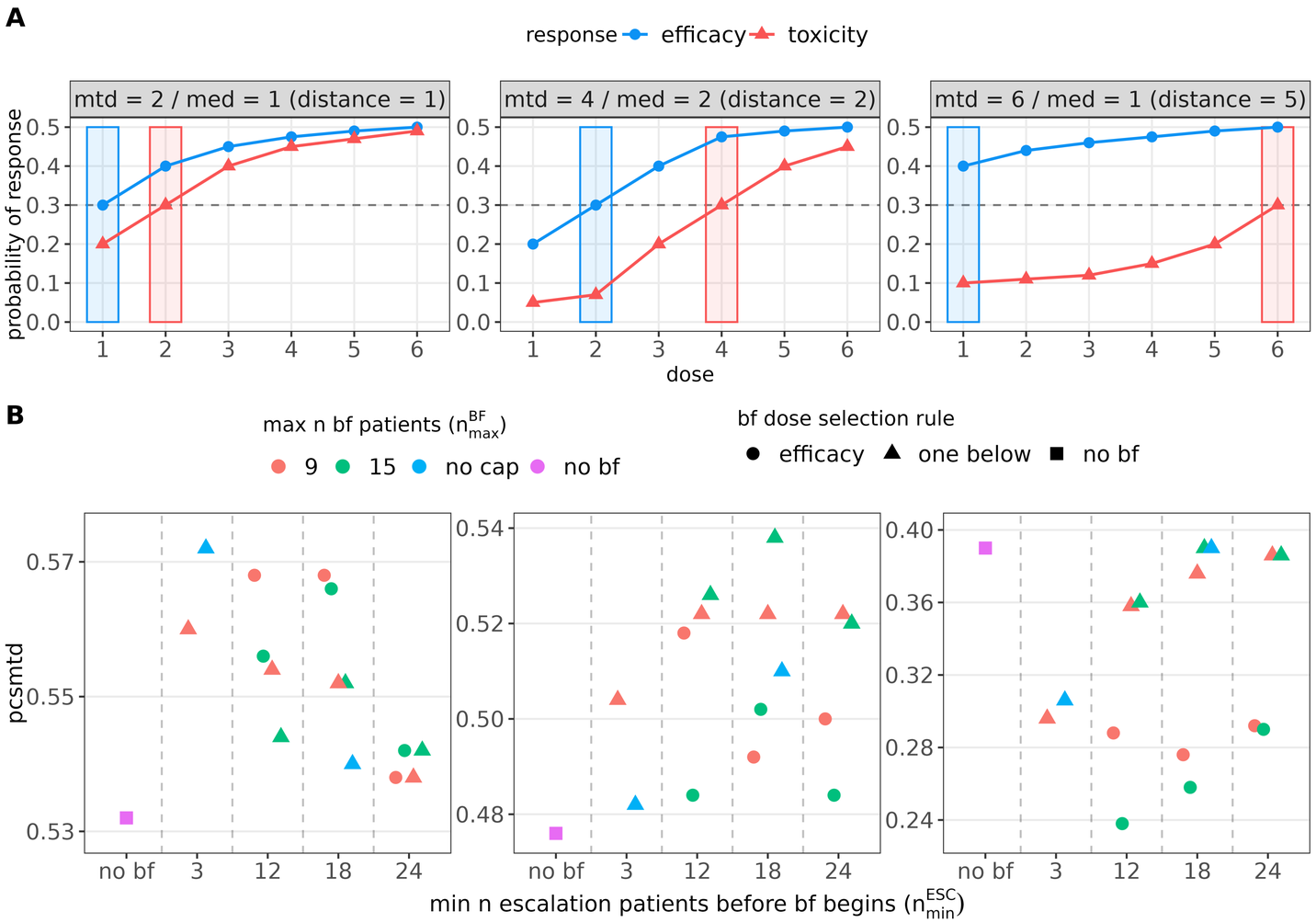
- *Monte Carlo estimates* of the following are of interest:
  - Expected sample size of dose escalation (ESS)
  - Probability of correctly selecting the MTD (PCSMTD)
  - Probability of correctly selecting the MED (PCSMED)
  - Probability that at least one expansion dose is in range [MED, MTD] (PEDIR), i.e., at least one dose is safe and effective

Results: backfill (ESS)



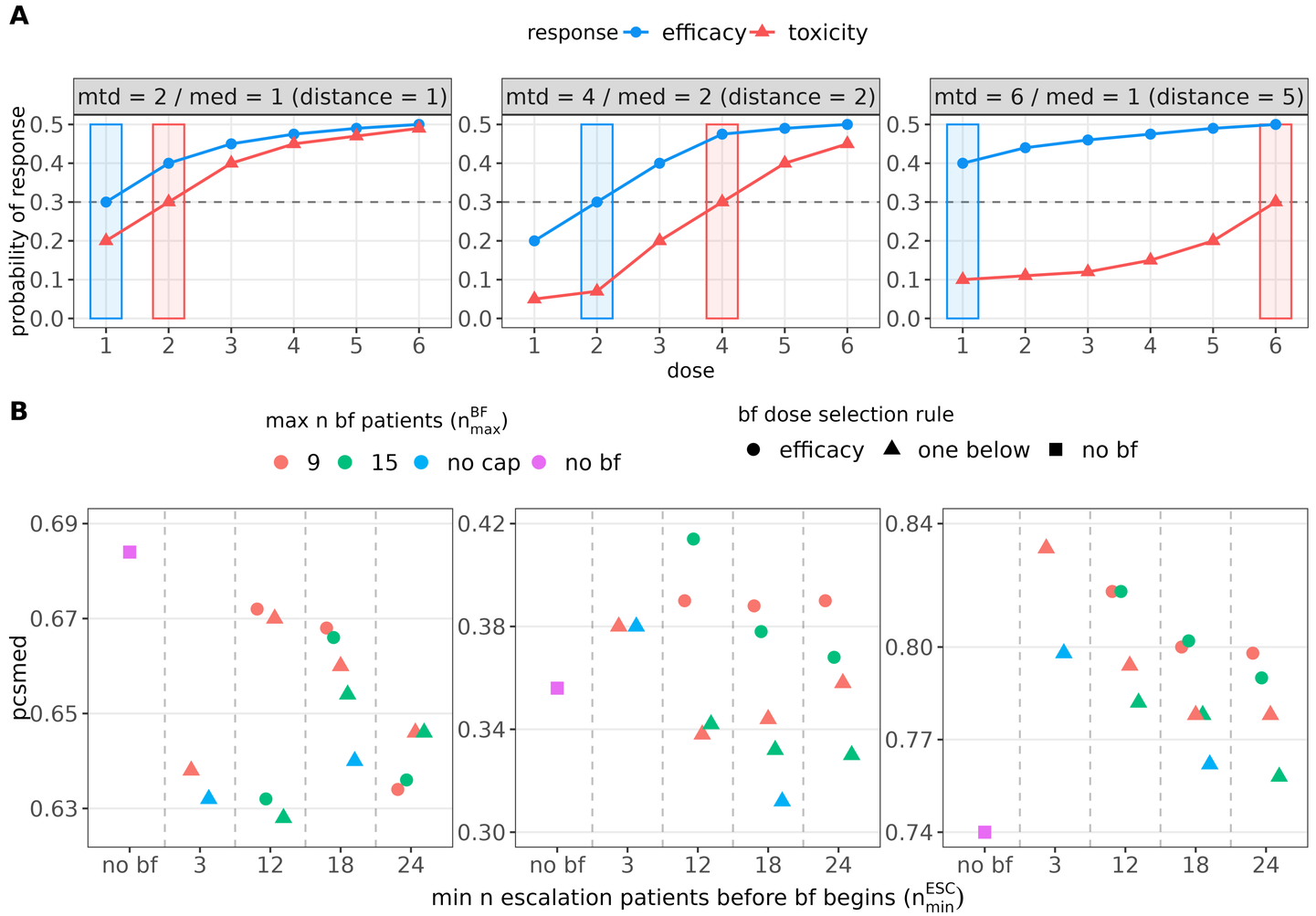
- ESS = expected sample size for dose escalation
- In all plots that follow
  - *Magenta square* = *no backfilling* - our comparator
  - Y-axis scales are different
- Backfill leads to 7-12 more patients in escalation on average
- Total number depends on exact scenario

Results: backfill (PCSMTD)



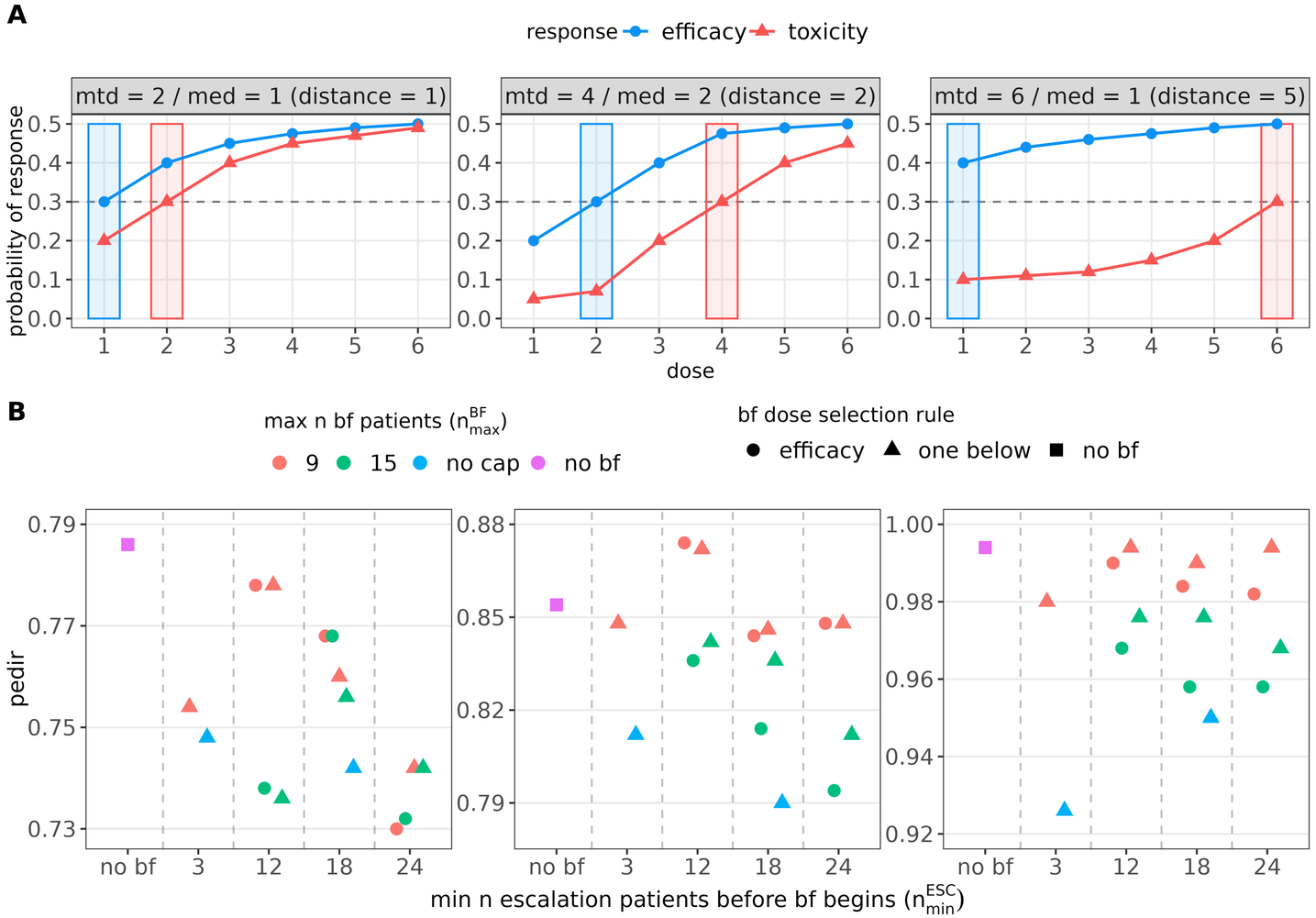
- PCSMTD = probability of correctly selecting the MTD
- When MTD is *low* dose or distance between MTD/MED is *small* —> backfill may *increase* PCS
- When MTD is *high* dose and distance between MTD and MED is *large* —> backfill may *reduce* PCS

Results: backfill (PCSMED)



- PCSMED = probability of correctly selecting the MED
- When MTD is *low* dose, backfill can increase PCSMTD but *also increase chance of declaring neither dose safe* (not shown), so can *decrease* PCSMED
- As MTD gets *larger* and distance b/w MTD and MED *grows*, backfill becomes *more beneficial*
  - Efficacy selection rule improves PCSMED the most since it is able to explore lower doses

Results: backfill (PEDIR)



- PEDIR = probability at least one expansion dose in [MED, MTD], i.e., safe and effective
- Surprisingly, backfill *decreases* the probability of taking at least one safe and effective dose into expansion
- If backfill patients experience toxic events, pushes up estimates of θ_j making doses seem more toxic
  - Increases the chance that no dose is declared safe and decreases PEDIR

Conclusions - backfill

1. *When to start backfill?* — starting backfill too early interferes with dose escalation (results not shown) and may decrease probability of selecting true MTD

2. *How to select backfill doses?* — using efficacy responses can help better explore lower doses, but δ must be tuned

3. *How many patients to use for backfill?* — using too many backfill patients may decrease performance selecting MTD and MED - depends on the scenario

4. *How does backfilling impact selection of the expansion doses?*

    1. Backfill most useful to identify MED when MTD is *high dose* and distance b/w MED and MTD is *large*

    2. *However*, under the considered approaches/scenarios, backfilling appears to *decrease* the probability that at least one expansion dose is safe and effective - concerning!

        1. Future work should investigate alternative backfill strategies

References

- Barnett H, Boix O, Kontos D, Jaki T. Backfilling cohorts in phase I dose-escalation studies. Clinical Trials. 2023 Jun;20(3):261-8.
- Dehbi HM, O'Quigley J, Iasonos A. Controlled backfill in oncology dose-finding trials. Contemporary clinical trials. 2021 Dec 1;111:106605.
- FDA. Project Optimus: Reforming the Dose Optimization and Dose Selection Paradigm in Oncology. 2024 June. https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus. Accessed: 2025-12-31.
- FDA. Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases. 2024 August. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/optimizing-dosage-human-prescription-drugs-and-biological-products-treatment-oncologic-diseases. Accessed: 2025-11-13.
- Groenland SL, Ratain MJ, Chen LS, Gandhi V. The Right Dose: From Phase I to Clinical Practice. In American Society of Clinical Oncology Educational book. American Society of Clinical Oncology. Annual Meeting 2021 Mar 1 (Vol. 41, pp. 92-106).
- Iasonos A, O'Quigley J. Integrating the escalation and dose expansion studies into a unified Phase I clinical trial. Contemporary clinical trials. 2016 Sep 1;50:124-34.
- Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. JNCI: Journal of the National Cancer Institute. 2009 May 20;101(10):708-20.
- Neuenschwander B, Branson M, Gsponer T. Critical aspects of the Bayesian approach to phase I cancer trials. Statistics in medicine. 2008 Jun 15;27(13):2420-39.
- Pin L, Villar SS, Dehbi HM. Implementing and assessing Bayesian response-adaptive randomisation for backfilling in dose-finding trials. Contemporary Clinical Trials. 2024 Jul 1;142:107567.
- Yuan Y, Zhou H, Liu S. Statistical and practical considerations in planning and conduct of dose-optimization trials. Clinical Trials. 2024 Jun;21(3):273-86.