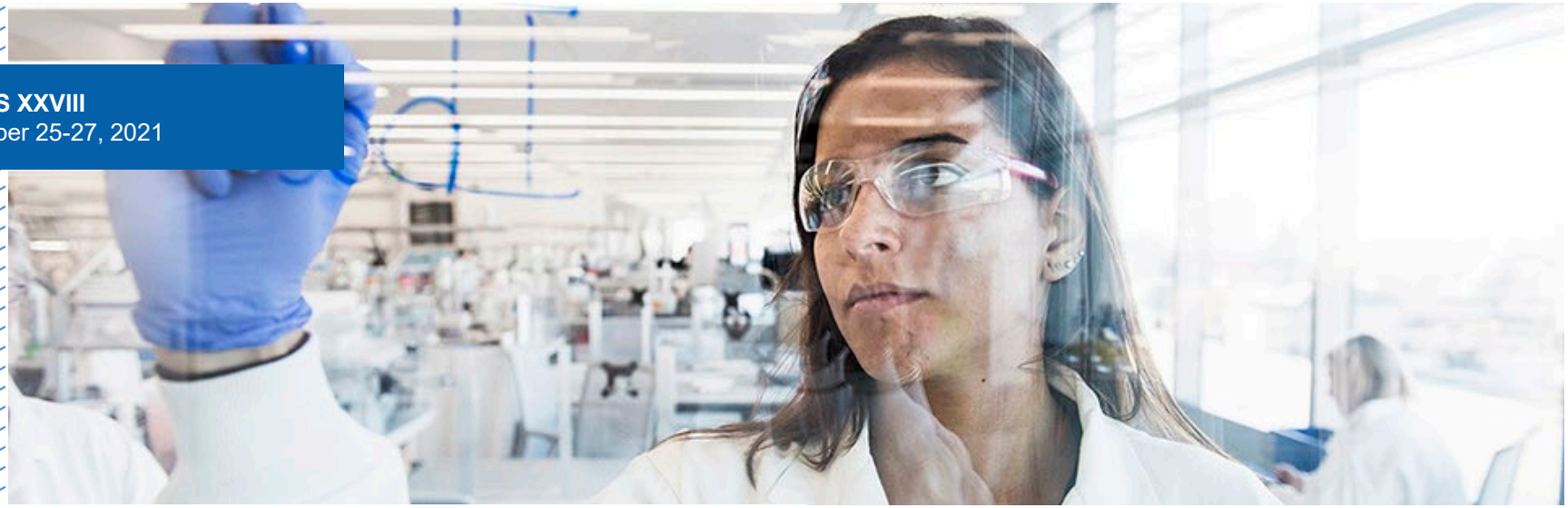
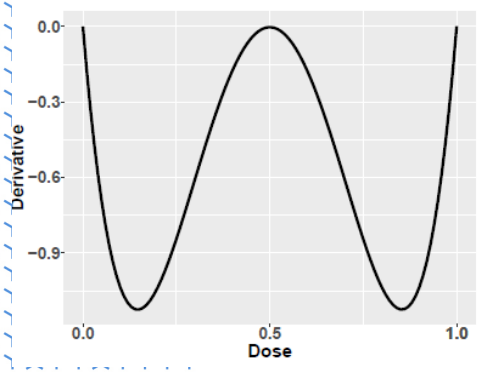
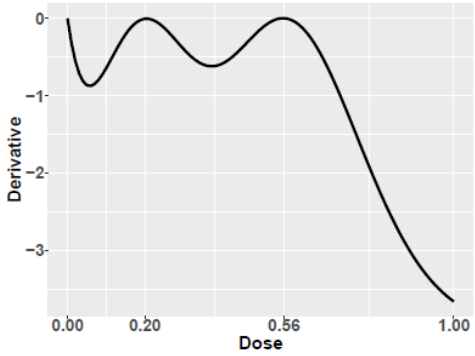
BASS XXVIII
October 25-27, 2021







Optimal Design in Clinical Trials

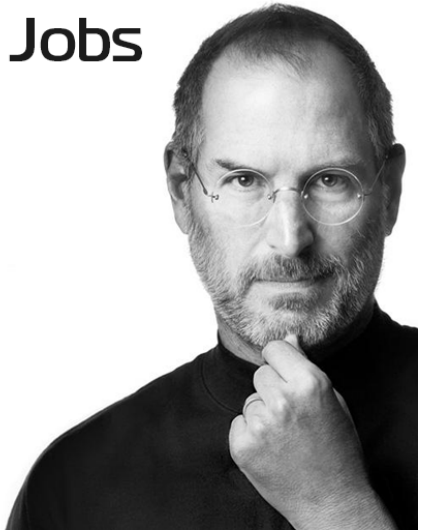
Alex Sverdlov, PhD

NOVARTIS | Reimagining Medicine

Quoting the great...

Design is not just what it looks like and feels like.
Design is how it works.

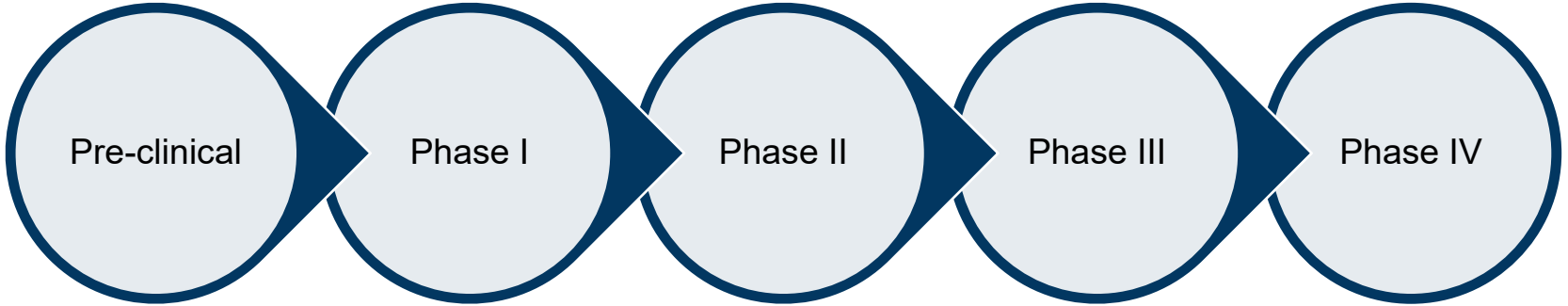
Steve Jobs



Outline

1. Background on clinical trials and optimal designs
2. Phase I dose–toxicity studies
3. Phase I/II efficacy–toxicity studies
4. Phase II dose-ranging studies
5. Phase III randomized controlled trials
6. Population PK/PD experiments
7. Summary

Drug development programs include multiple studies of increasing complexity



- Animal studies to calibrate dose range for phase I studies in humans
- Utilize safety, PK, and PD data to project dose and regimen for phase II
- Update PK/PD models to optimize dose and regimen in phase III
- Assess probability of success
- Test clinical hypotheses in RCTs
- Refine dose and regimen for special populations
- Optimization of drug use
- Population PK/PD models

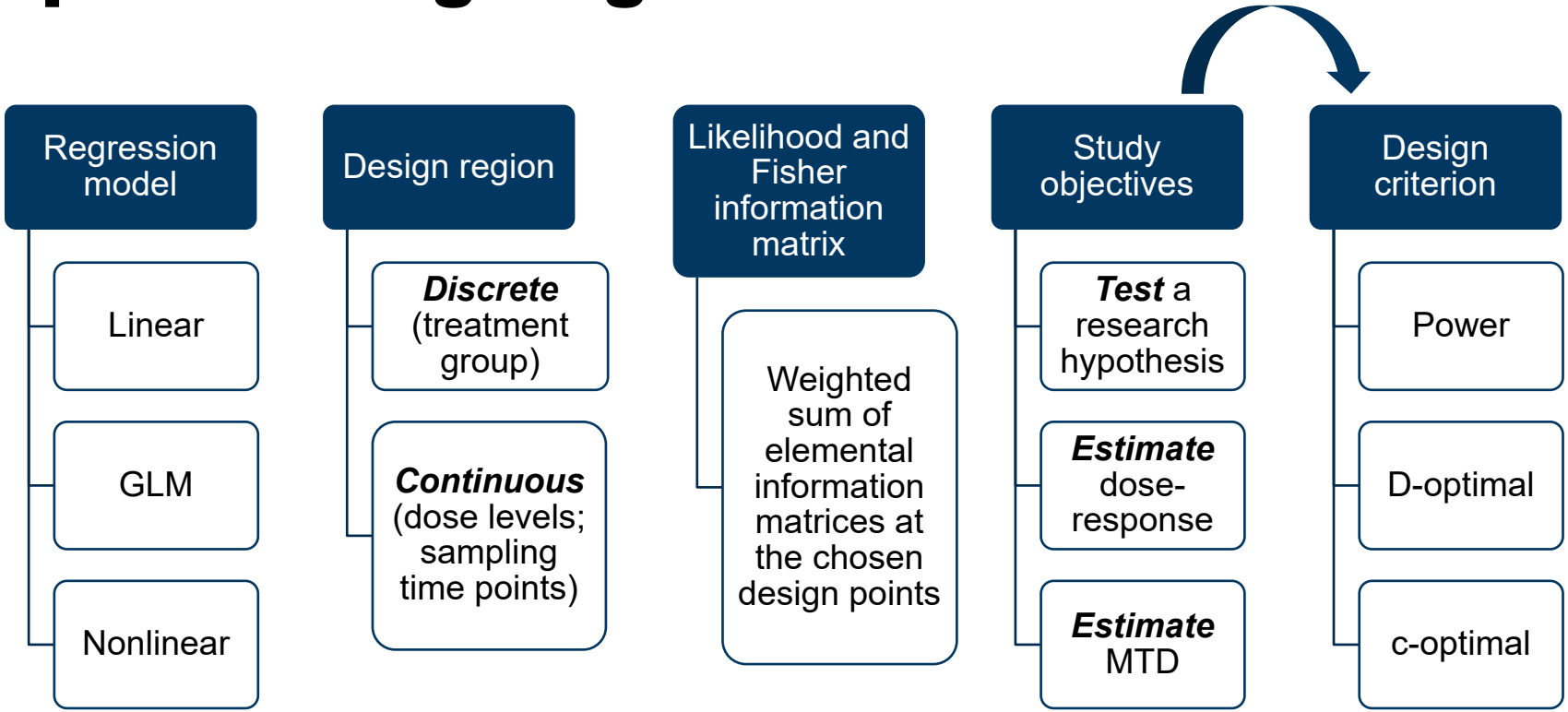
Why is it important to consider optimal designs for clinical trials?

Higher quality results (increased statistical power, more accurate estimates of treatment effects) for the given resource constraints

Lower sample size and/or decreased study cost (and potentially faster study completion) for the given data quality objectives

Study participant benefit: maximize information from the trial while minimizing exposure of study subjects to suboptimal treatments

Optimal design ingredients



Phase I dose-toxicity study

- Goals of phase I: characterize safety, tolerability, and PK of a compound
  - In oncology: to determine the maximum tolerated dose (MTD)
- For ethical reasons, phase I studies are cast as dose-escalation designs
  - Only when the previous dose is deemed ‘safe’ would the next cohort of subjects be assigned to the next dose
- Various methodologies to determine MTD are available

Statistics Surveys
Vol. 8 (2014) 2–44
ISSN: 1935-7516
DOI: 10.1214/14-SS106

Adaptive clinical trial designs for phase I cancer studies

Oleksandr Sverdlov

Biometrics and Statistical Sciences
Integrated Information Sciences
Novartis Institutes for Biomedical Research
One Health Plaza, East Hanover, NJ 07936
USA
e-mail: alex.sverdlov@novartis.com

Weng Kee Wong

Department of Biostatistics
Fielding School of Public Health
University of California, Los Angeles
Los Angeles, CA 90095
USA
e-mail: wkwong@ucla.edu

and

Yevgen Ryeznik

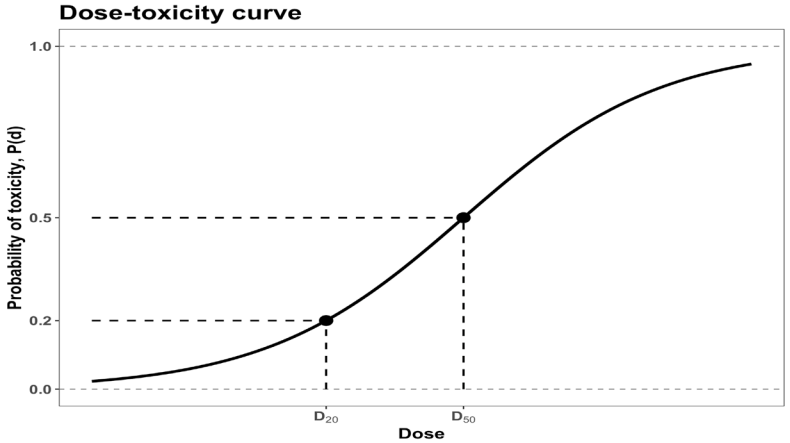
CS Ltd., 42A Tobolskaya Street
Kharkov 61072
Ukraine
e-mail: yevgan@csltd.com.ua

Phase I dose-toxicity study

- $d_1 < d_2 < \cdots < d_K$ - study doses
- Outcome: Toxicity (Yes/No)
- Probability of toxicity is modeled using a 2-parameter logistic curve:

$$P(d) = \Pr(Y = 1|d) = \frac{1}{1 + e^{-(\alpha+\beta d)}}$$

- α and $\beta > 0$ are unknown parameters; monotone increasing dose-toxicity relationship



- Estimands of interest:
  - $P(d)$ for a given $d > 0$
  - 50th percentile of the dose-tox curve
  - MTD - say, 20th percentile of the dose-tox curve: $D_{20} = (\log\left(\frac{0.2}{1-0.2}\right) - \alpha)/\beta$

Phase I dose-toxicity study
How do we estimate dose-toxicity curve?

- Data structure: $\{(d_i, n_i, x_i), i = 1, \ldots, K\}$

$$x_i \sim Bin(n_i, P_i), \text{ where } P_i = \frac{1}{1+e^{-(\alpha+\beta d_i)}}$$

- Likelihood: $\mathcal{L}(\alpha, \beta) = \prod_{i=1}^{m} P_i^{x_i}(1-P_i)^{n_i - x_i}$
- MLE $(\hat{\alpha}, \hat{\beta})$ is obtained as a solution to the system of score equations:

$$\frac{\partial}{\partial \alpha} \log \mathcal{L}(\alpha, \beta) = \sum_{i=1}^{m}(x_i - n_i P_i) = 0$$
$$\frac{\partial}{\partial \beta} \log \mathcal{L}(\alpha, \beta) = \sum_{i=1}^{m}(x_i - n_i P_i) d_i = 0$$

- By the invariance property of MLE, other parameters can be readily estimated, e.g., $\widehat{D}_{20} = (\log\left(\frac{0.2}{1-0.2}\right) - \hat{\alpha})/\hat{\beta}$

Phase I dose-toxicity study
How do we quantify uncertainty?

- Uncertainty quantification:

$$\widehat{\mathrm{Var}}(\hat{\alpha}, \hat{\beta}) \approx \boldsymbol{M}^{-1}(\hat{\alpha}, \hat{\beta}) \text{ (inverse of the observed Fisher information)}$$

- Once we have $(\hat{\alpha}, \hat{\beta})$ and $\boldsymbol{M}^{-1}(\hat{\alpha}, \hat{\beta})$, we can construct (asymptotic) 95% CI's for α and β:

$$\hat{\alpha} \pm 1.96\sqrt{\mathrm{var}(\hat{\alpha})} \text{ and } \hat{\beta} \pm 1.96\sqrt{\mathrm{var}(\hat{\beta})}$$

- Using delta-method, we can obtain 95% CI's for other parameters, such as MTD:

$$\widehat{\mathrm{Var}}(\hat{D}_{20}) = \frac{1}{\hat{\beta}^2}\left(\mathrm{var}(\hat{\alpha}) + \hat{D}_{20}^2\mathrm{var}(\hat{\beta}) - 2\hat{D}_{20}\mathrm{cov}(\hat{\alpha}, \hat{\beta})\right)$$

$$\hat{D}_{20} \pm 1.96\sqrt{\widehat{\mathrm{Var}}(\hat{D}_{20})}$$

Phase I dose-toxicity study
Example

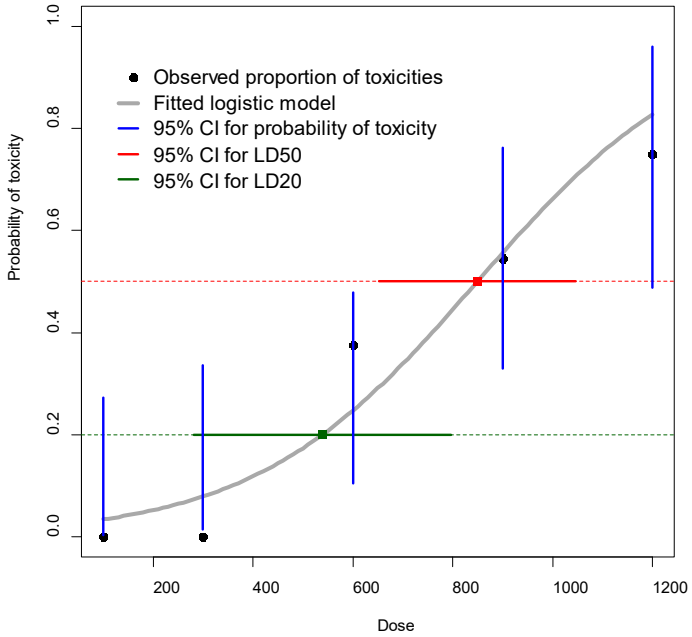
- Phase I study of the ChemoTx agent R115777 conducted at the University of Maryland School of Medicine (Karp et al., 2001)
- n=34 patients with acute leukemia, treated at 5 different doses

Dose	100mg	300mg	600mg	900mg	1200mg
Assigned	6	5	8	11	4
Number of toxicities	0	0	3	6	3
Proportion of toxicities	0	0	0.375	0.545	0.750

- 2-parameter logistic model was fitted:

	Estimate	95% CI
α	-3.7958	(-7.1276, -1.59015)
β	0.004468	(0.0016986, 0.0084355)

Phase I dose-toxicity study
Example contd.



- Estimates of toxicity probabilities at study doses were obtained
- Estimates of D_{20} and D_{50}:

	Estimate	95% CI
D_{20}	539	(282, 797)
D_{50}	850	(653, 1046)

- The dataset is small (n=34) ⇒ estimation uncertainty is high

Phase I dose-toxicity study
Can we optimize the design of the next study?

- Design: $\boldsymbol{\xi} = \{(d_i, \rho_i), i = 1, \ldots, K\}$
  - d_i's – dose levels
  - $\rho_i = n_i/n$ – allocation proportion for d_i
- Fisher Information Matrix (FIM) for design $\boldsymbol{\xi}$:

  $\boldsymbol{M}(\boldsymbol{\xi}, \alpha, \beta) = n \sum_{i=1}^{K} \rho_i \boldsymbol{M}_i(\alpha, \beta)$, where $\boldsymbol{M}_i(\alpha, \beta)$ = information at dose d_i
- Optimal design problem:

  minimize $\Phi\left(\boldsymbol{M}^{-1}(\boldsymbol{\xi}, \alpha, \beta)\right)$ w.r.t. $\boldsymbol{\xi}$

  $\Phi = \det \Rightarrow$ D-optimality $\Rightarrow$ min(volume of the confidence ellipsoid for $\boldsymbol{\theta}$)
- Beautiful convex design theory, algorithms and numerical techniques, all beyond the scope of presentation; see Fedorov and Leonov (2014) and references therein

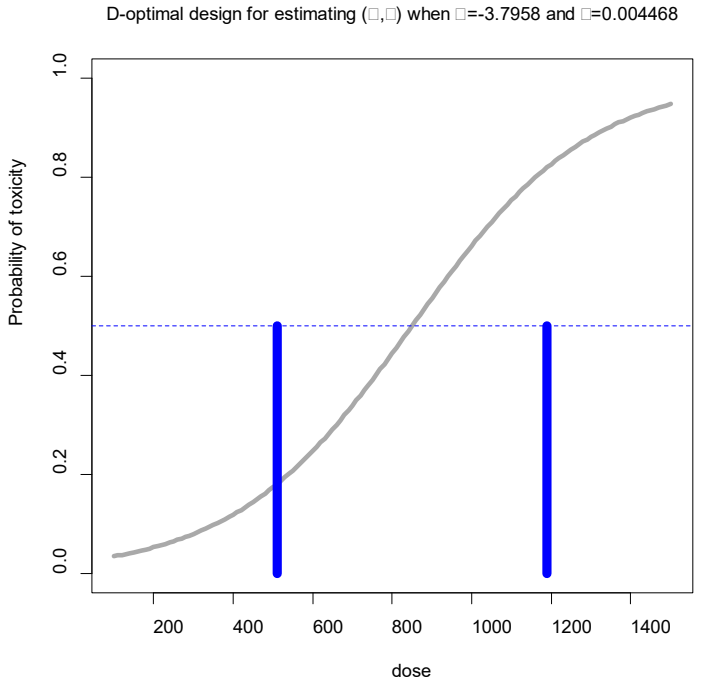
Phase I dose-toxicity study
D-optimal design

- The D-optimal design minimizing $|\boldsymbol{M}^{-1}(\alpha, \beta)|$ is a 2-point design, symmetric around D_{50}, equally supported at the 18th and 82nd percentiles of the dose–toxicity curve

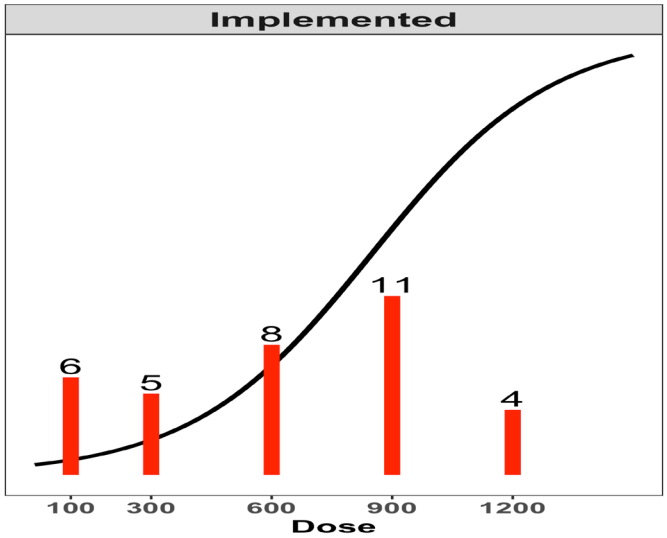
$$\boldsymbol{\xi}^{*}_{D-opt} = \left\{(D_{18}, \frac{1}{2}), (D_{82}, \frac{1}{2})\right\}$$

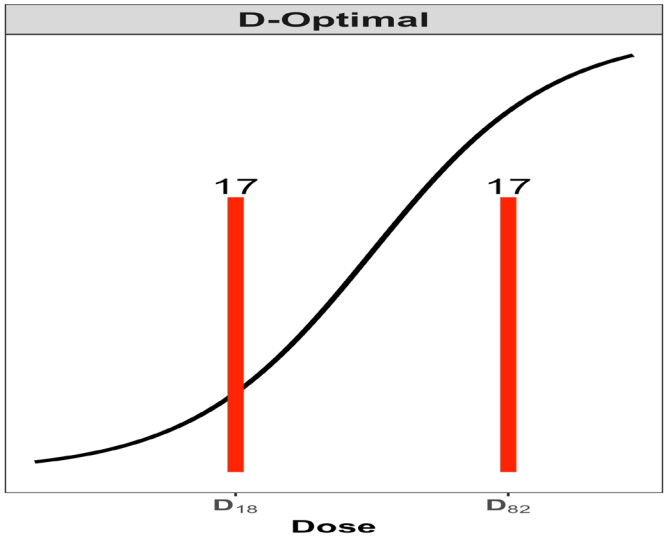
$$D_{18} = \frac{\log(0.18/0.82) - \alpha}{\beta}, \text{ and}$$

$$D_{82} = \frac{\log(0.82/0.18) - \alpha}{\beta}$$



Phase I dose-toxicity study
How to facilitate a comparison among designs?





- Efficiency of the implemented design ξ (Karp et al., 2001) relative to the D-optimal design ξ^*(for the same sample size) is $D_{eff} = \frac{|M^{-1}(\xi^*,\theta)|}{|M^{-1}(\xi,\theta)|} = \left\{\frac{3.45\cdot10^{-7}}{5.16\cdot10^{-7}}\right\}^{1/2} = 0.82$

Phase I dose-toxicity study
What are merits and limitations of optimal designs?

Merits

- ODs provide important theoretical benchmarks to compare various designs w.r.t. selected optimality criteria
- If properly implemented, ODs can help achieve study goals with a reduced sample size/study cost
- D-optimal design maximizes information for estimating the entire dose-toxicity curve

Limitations

- ODs depend on the choice of statistical model
- ODs frequently depend on the true parameter values (local optimality)
- D-optimal design allocates 50% of subjects to the doses with toxicity probabilities 18% and 82% - may not be ‘clinically optimal’
- Frequently require advanced numerical optimization

Phase I dose-toxicity study
Can more elaborate models be considered?

- Instead of a 2-parameter logistic dose–toxicity model, one can consider 3- or 4-parameter logistic models (Li and Majumdar, 2008)
  - Since phase I trials are small, striking the right balance between model parsimony and rigor is important
- Optimal designs for other types of outcomes (e.g., ordinal toxicity grades with a proportional odds model) are available (Perevozskaya et al., 2003)
- Design space may be more complex
  - Optimal designs for drug combination studies (Fedorov and Leonov, 2014)

Phase I dose-toxicity study
How to overcome the issue of local optimality?

- Bayesian optimal design – minimizes average value of the criterion given the prior distribution of the parameters (Chaloner and Larntz, 1989)
- Minimax design – minimizes worst value of the criterion over the range of model parameter values (King and Wong, 2000)
- Adaptive design – sequentially updates model parameters and directs future dose assignments to the targeted optimal design
  - Any phase I dose-escalation study is adaptive; however, adaptations are performed based on individual patient safety considerations, not based on statistical precision
  - Adaptive designs that formally combine “treatment” and “learning” goals merit consideration (Haines et al., 2003; Bartroff and Lai, 2011)

Phase I dose-toxicity study
Can a design combine 'treatment' and 'learning' goals?

Constrained Bayesian optimal designs (Haines et al., 2003):

- 2-parameter logistic dose–toxicity model, with a prior distribution for $\boldsymbol{\theta} = (\alpha, \beta)$
- Constrained optimization problem:

$$E(\log|\boldsymbol{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})|) \rightarrow \min \text{ (w.r.t. } \boldsymbol{\xi})$$

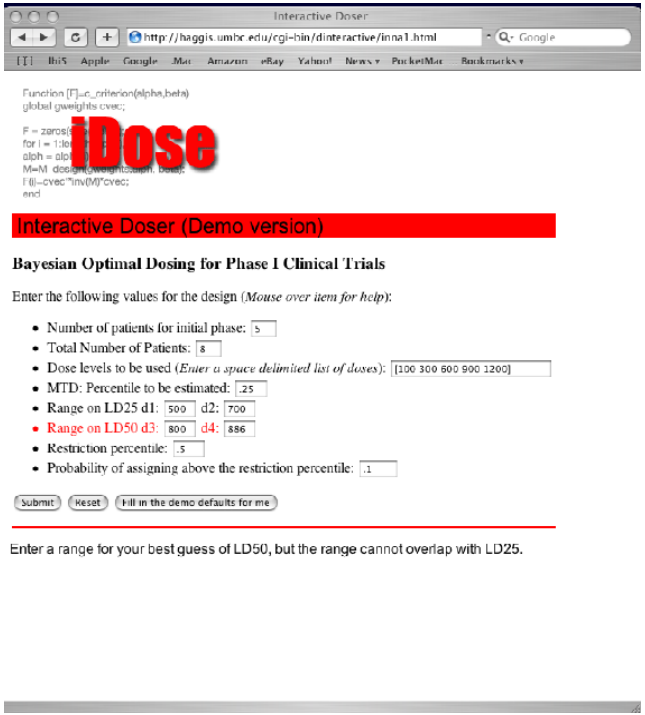
subject to to an "overdose" constraint: $\sum_{i=1}^{K} \rho_i \Pr(\mu_R \leq d_i) \leq \varepsilon$

(μ_R=maximum dose that cannot be exceeded; $\varepsilon > 0$ small, investigator-specified constant)

- Implementation in practice:
  - 2-stage: n_0 subjects are assigned to doses using some pilot design + n_1 subjects are assigned to doses according to updated optimal design
  - Sequential: Small pilot design + subsequent sequential assignments to maximize incremental gain in information while protecting patient safety

Phase I dose-toxicity study
Why consider constrained Bayesian optimal designs?

- The methodology (Haines et al., 2003) incorporates clinicians' prior knowledge on the dose-toxicity curve and utilizes Bayesian OD theory to efficiently estimate target quantities with as few patients as possible
- The sequential design has established asymptotic properties (convergence to the targeted design) (Roy et al., 2009)
- Statistical software (iDose) is available (Rosenberger et al., 2005); yet, need to check if it is still supported



Phase I/II efficacy-toxicity study

- Development of a targeted therapy in oncology is different from that of a cytotoxic drug
  - Lower risk of toxicity
  - Efficacy may plateau at doses below MTD
- Seamless phase I/II designs incorporate toxicity and efficacy (response) in dose-finding objectives
  - Joint modeling of a dose–toxicity–efficacy relationship
  - Phase I/II trial is typically larger that a single phase I trial
  - Avoids administrative wait between phase I and II protocol activation

Phase I/II efficacy-toxicity study
Bivariate binary outcomes

- $\Omega = \{d_1 < \cdots < d_K\}$ - study doses
- Dose-toxicity and dose-efficacy probability curves

$$p_T(d) = \Pr(Y_T = 1|d) \text{ (Tox)}$$

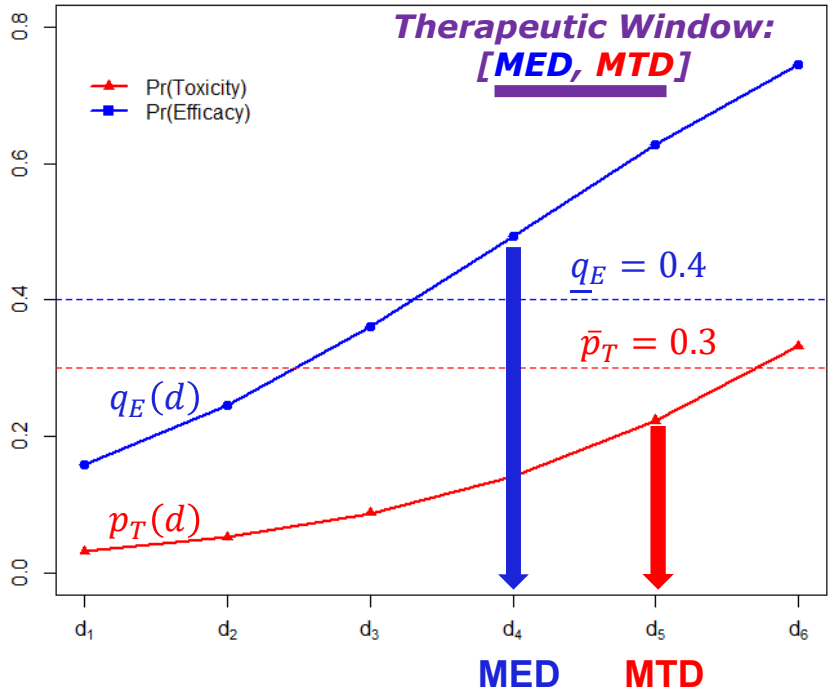
$$q_E(d) = \Pr(Y_E = 1|d) \text{ (Eff)}$$

- Maximum tolerated dose (MTD):

$$\max\{d \in \Omega: p_T(d) \leq \bar{p}_T\}$$

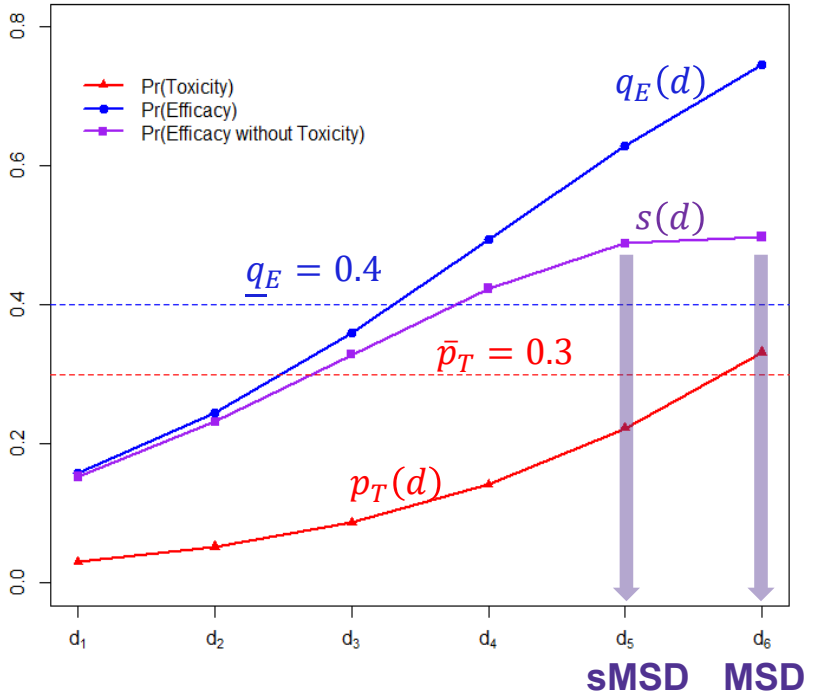
- Minimum efficacious dose (MED):

$$\min\{d \in \Omega: q_E(d) \geq \underline{q}_E\}$$



Phase I/II efficacy-toxicity study
Joint outcomes

- "Success" = Efficacy without toxicity: $(Y_E, Y_T) = (1,0)$
- Probability of Success:
  $s(d) = \Pr(Y_E = 1 | Y_T = 0, d) \times \Pr(Y_T = 0 | d)$
- Most Successful Dose (MSD):
  A dose $d^* \in \Omega$ that maximizes $s(d)$
- Safe MSD (sMSD):
  A dose $d^* \in \Omega$ that maximizes $s(d)$ while satisfying the MTD constraint:
  $\Pr(Y_T = 1 | d^*) \leq \bar{p}_T$

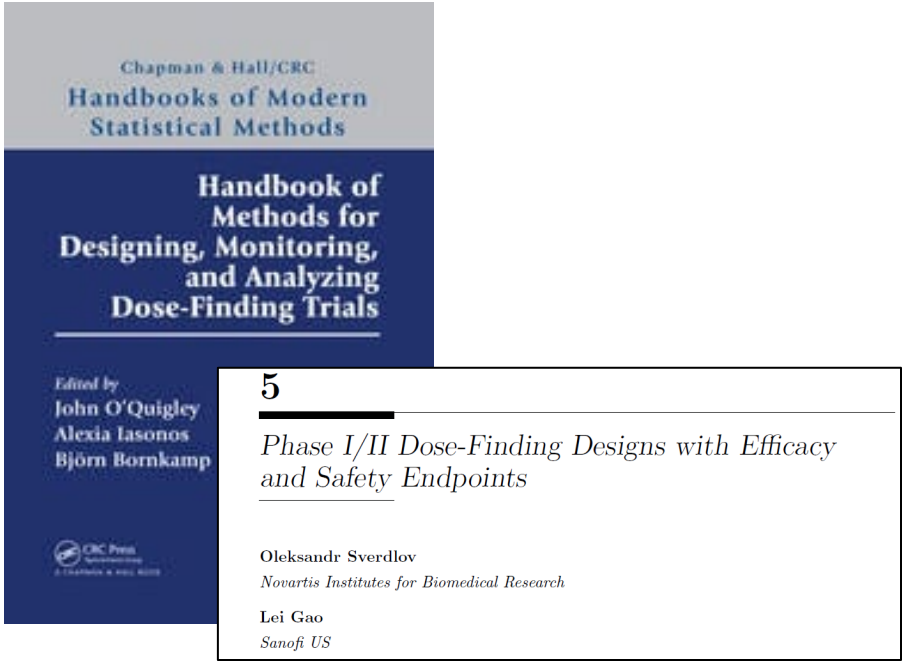


Phase I/II efficacy-toxicity study
Objectives

- ‘Treatment’ goal: To cluster dose assignments at and around safe MSD
- ‘Learning’ goals:
  - To identify safe MSD (or stop the trial early if no dose satisfies safety & efficacy requirements)
  - To estimate safe MSD, MTD, MED, and possibly other parameters at the end of the study
- Seemingly similar goals of identification/estimation of safe MSD may require different design considerations
- These objectives must be reliably achieved with as few patients as possible

Phase I/II efficacy-toxicity study
Various designs have been developed for this purpose

- Nonparametric up-and-down design (Ivanova, 2003)
- Bayesian 'best intention' designs
  - Bivariate CRM (Braun, 2002)
  - Eff-tox method (Thall and Cook, 2004)
- Adaptive penalized optimal designs (Dragalin and Fedorov, 2006)
- It is difficult to recommend any particular design as "best"



NOVARTIS | Reimagining Medicine

Phase I/II efficacy-toxicity study
Adaptive penalized ODs (Dragalin and Fedorov, 2006)

- Postulate a statistical model for dose–response: $\pi_{T,E}(d, \boldsymbol{\theta})$, $d \in \Omega$
- Fisher information matrix: $\mathbf{M}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \sum_{k=1}^{K} \rho_k \pi_{T,E}(d_k, \boldsymbol{\theta})$
- Cost function penalizing doses with low success and high toxicity: $\phi(d, \boldsymbol{\theta}, C_E, C_T) > 0$, where $C_E \geq 0$ and $C_T \geq 0$ are user-specified constants
- Penalized optimal design problem:

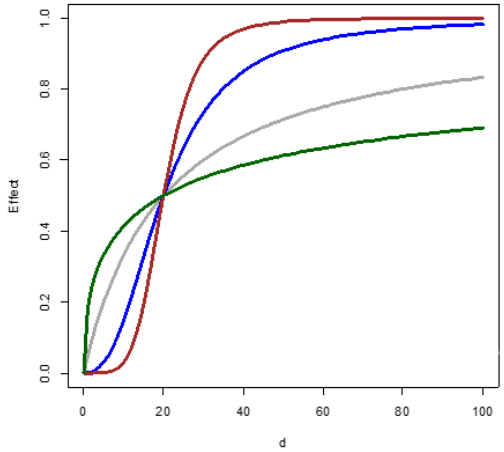
$$\log \frac{|\mathbf{M}^{-1}(\boldsymbol{\xi}, \boldsymbol{\theta})|}{\sum_{k=1}^{K} \rho_k \phi(d_k, \boldsymbol{\theta}, C_E, C_T)} \rightarrow \min \text{ (w.r.t. } \boldsymbol{\xi})$$

- Implementation: some 'start-up' dose-escalation design to ascertain initial data for estimating $\boldsymbol{\theta}$, then sequential dose assignments to maximize incremental increase of information per cost unit

Phase I/II efficacy-toxicity study
Why consider adaptive penalized optimal designs?

- Substantial improvement in accuracy of dose–response estimation compared to ‘best intention’ designs (Dragalin and Fedorov, 2006)
- Good balance between ‘treatment’ and ‘learning’ goals in small-to-moderate experiments; known asymptotic properties (Pronzato, 2010)
- Bayesian adaptive penalized D-optimal design has competitive performance to Thall and Cook’s Eff-Tox method (Gao and Rosenberger, 2013)
- While in practice it may be difficult to gain IRB clinical approval for these designs, they may be more readily applicable in animal studies where ethical issues are not as high as in human experiments

Phase II dose-ranging study



- Randomized, placebo- and/or active-controlled trial with several doses of an investigational drug, with sample sizes up to several hundred patients
- Research questions in phase II:
  - Is there any evidence of a drug effect (proof-of-concept)?
  - Which dose(s) exhibit a response different from the control?
  - What is the dose–response relationship?
  - What is the “optimal” dose for taking into phase III?
- Design considerations:
  - Sample size
  - Dose levels and allocation proportions for the chosen doses

Phase II dose-ranging study
Statistical analysis approaches

- Multiple comparison procedures
  - Dose is regarded as a classification factor
  - ANOVA test + stepwise procedures to identify dose(s) that are significantly different from placebo
- Modeling techniques
  - Dose is regarded as a continuous predictor for the mean response
  - Regression modeling allows borrowing strength across dose levels and extrapolate the results beyond the study doses
- MCP-Mod (Bretz et al., 2005):
  - A combination of multiple comparisons (MCP) with modeling techniques (Mod)

Phase II dose-ranging study
Design optimization challenges

- In practice, the true model is unknown at the trial design stage ⇒ uncertainty in the planning of the experiment
- There are several (seemingly similar) research questions which require different choice of study design; e.g.:
  a) Is there a dose-related effect at all?
  b) What is the smallest dose that achieve a clinically relevant effect Δ over placebo?
  c) Which dose achieves a 50% of the maximum effect (ED_{50})?
  d) Where does the dose–response curve start to plateau?
- Question a) involves hypothesis testing; questions b) – d) involve estimation

Phase II dose-ranging study
Example: D-optimal design for time-to-event outcomes

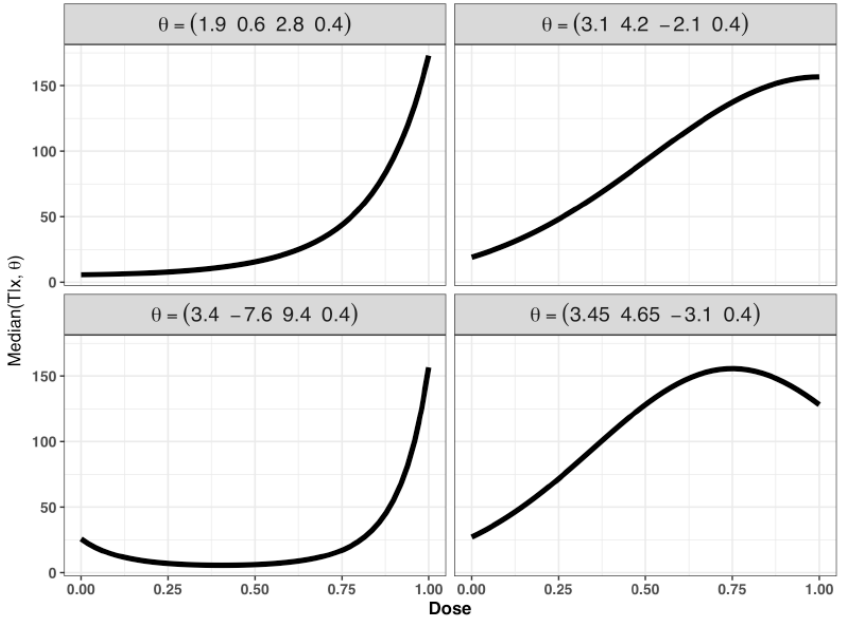
- Quadratic dose-response model for event times: $T \sim Weibull$ with

$$Median(T) = \exp(\beta_0 + \beta_1 d + \beta_2 d^2)\,(\ln 2)^b$$

Doses: $d \in [0,1]$; 0 =placebo; 1=MTD
Parameters: $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, b)$, $b > 0$

- Observations may be right-censored
- Objective: estimate dose–response as accurately as possible by allocating n subjects to 'most informative' dose levels

Plausible median time-to-event dose-response

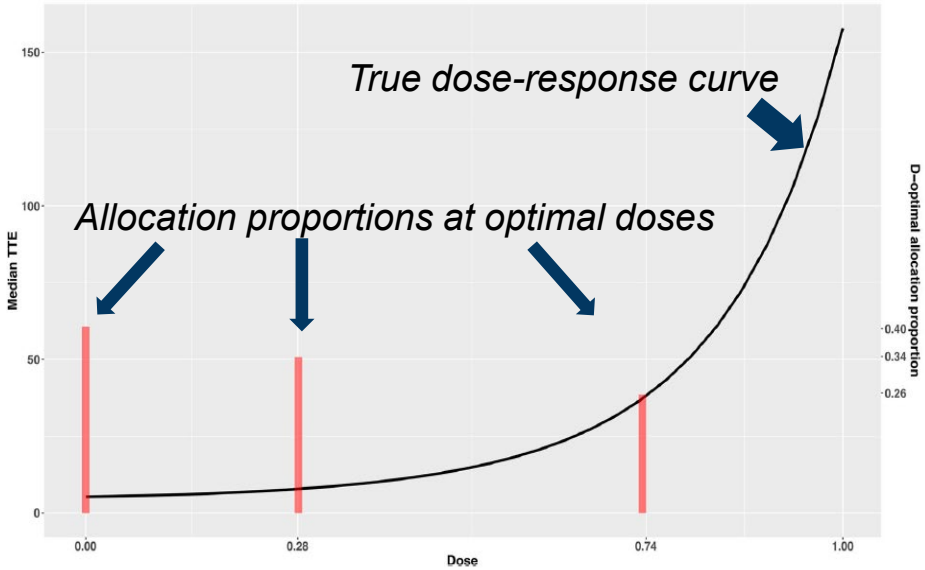


NOVARTIS | Reimagining Medicine

Phase II dose-ranging study
Example: D-optimal design for time-to-event outcomes

- Without censoring, D-optimal design is a 3-point, uniform design
  - Equal allocation proportions (1/3) to $d = 0$ (placebo); $d = 0.5$ (middle dose); and $d = 1$ (highest dose)
- In the presence of censoring, D-optimal design still has 3 points, but the doses are different and allocation proportions are unequal
  - Higher amount of censoring ⇒ greater degree of skewness from the uniform design

Weibull model with $\beta_0 = 1.9$, $\beta_1 = 0.6$, $\beta_2 = 2.8$, $b = 0.65$, and average probability of event = 50%



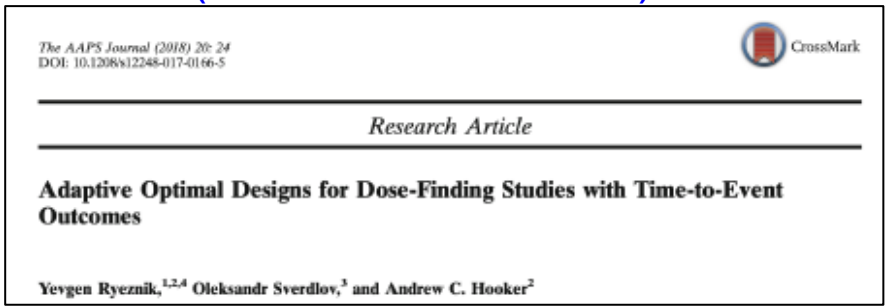
Phase II dose-ranging study
Adaptive D-optimal design for time-to-event outcomes

- Stage 1: n_1 subjects are randomized equally among the doses 0, 0.5, and 1
- Interim analysis ($k = 2, \ldots, v$):
  - Update model estimates $\widehat{\boldsymbol{\theta}}$ based on accrued data from stages $1, \ldots, k-1$
  - Check the pre-specified stopping rule: Has the desired estimation precision been achieved?
    - YES ⇒ STOP
    - NO ⇒ Compute D-optimal design $\xi^*(\widehat{\boldsymbol{\theta}})$
- Stage k: Randomize n_k patients to doses according to $\xi^*(\widehat{\boldsymbol{\theta}})$
  - Thus, the design applies response-adaptive randomization to cohorts of subjects, to concentrate assignments at 'most informative' dose levels

Phase II dose-ranging study
Why consider adaptive D-optimal design?

- D-optimal design depends on the underlying model and the amount of censored data
- Equal allocation (uniform) design can be highly inefficient
- Adaptive D-optimal design with early stopping facilitates learning about the model and can potentially reduce study size with better estimation accuracy than the uniform design

More details in the two papers (2018, AAPS Journal):



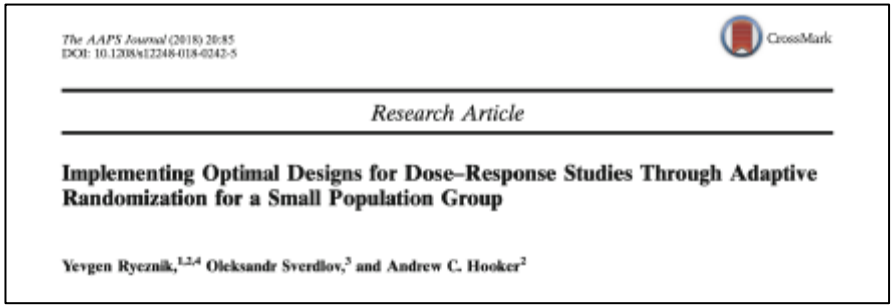
The AAPS Journal (2018) 20: 24
DOI: 10.1208/s12248-017-0166-5

CrossMark

Research Article

Adaptive Optimal Designs for Dose-Finding Studies with Time-to-Event Outcomes

Yevgen Ryeznik,[1,2,4] Oleksandr Sverdlov,[3] and Andrew C. Hooker[2]



The AAPS Journal (2018) 20:85
DOI: 10.1208/s12248-018-0242-5

CrossMark

Research Article

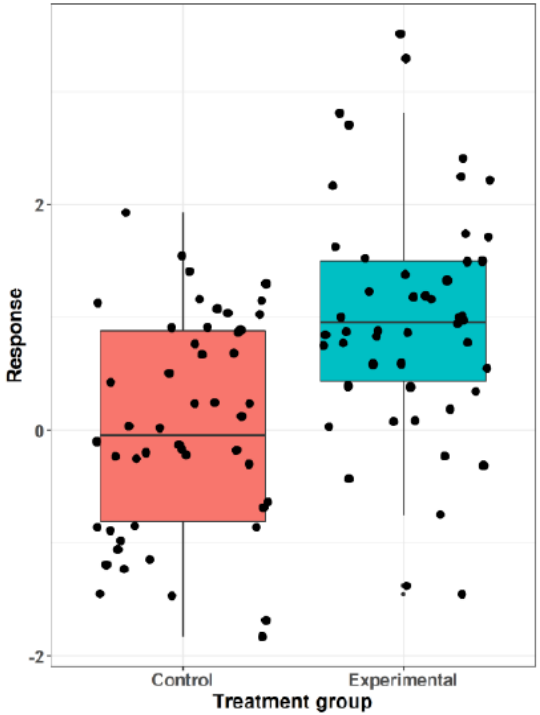
Implementing Optimal Designs for Dose–Response Studies Through Adaptive Randomization for a Small Population Group

Yevgen Ryeznik,[1,2,4] Oleksandr Sverdlov,[3] and Andrew C. Hooker[2]

NOVARTIS | Reimagining Medicine

Phase III randomized controlled trial (RCT)

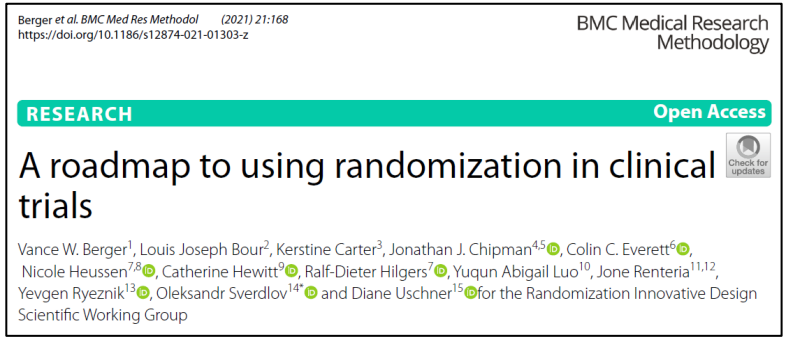
- RCT is regarded as the ‘gold standard’ clinical research design
  - Evidence from RCTs is widely used as a basis for submissions in request of marketing authorization of new drugs, biologics and medical devices
- Randomization is used to ensure group comparability w.r.t. known and unknown confounders
- Sample size is chosen to have sufficient statistical power
- Data analysis involves estimation/test of treatment group difference (with possible adjustment for important covariates)



Phase III randomized controlled trial
Design considerations

- Primary outcome
- Number of treatment arms
- Choice of control group
- Sample size
- Interim analyses
- Allocation ratio
- Randomization procedure

- Allocation ratio
  - Equal (1:1) allocation is frequently (but not always) optimal
- Randomization procedure
  - Tradeoff between treatment balance and allocation randomness


Berger et al. BMC Med Res Methodol (2021) 21:168
https://doi.org/10.1186/s12874-021-01303-z

BMC Medical Research Methodology

RESEARCH Open Access

A roadmap to using randomization in clinical trials

Vance W. Berger[1], Louis Joseph Bour[2], Kerstine Carter[3], Jonathan J. Chipman[4,5], Colin C. Everett[6], Nicole Heussen[7,8], Catherine Hewitt[9], Ralf-Dieter Hilgers[7], Yuqun Abigail Luo[10], Jone Renteria[11,12], Yevgen Ryeznik[13], Oleksandr Sverdlov[14*] and Diane Uschner[15] for the Randomization Innovative Design Scientific Working Group

NOVARTIS | Reimagining Medicine

Phase III randomized controlled trial
Unequal allocation ratio may sometimes be preferred

- Heteroscedastic outcomes: If standard deviation is different between study arms ⇒ Neyman allocation $\sigma_E : \sigma_C$ (experimental to control) maximizes power
- Unequal treatment cost: $\frac{\sigma_E}{\sqrt{w_E}} : \frac{\sigma_C}{\sqrt{w_C}}$ minimizes study cost for a given power
  - If experimental is 4 times as expensive as control ⇒ 1:2 allocation experimental to control is most cost-efficient
- Vaccine RCTs frequently utilize 2:1 allocation experimental to control
- Ethical considerations: If disease is severe and/or rare ⇒ it is desirable to minimize number of treatment failures while maintaining study power
- Platform trials: In a multi-arm trial with a shared control group, allocation ratio to control may be gradually decreased

Phase III randomized controlled trial
Optimal response-adaptive randomization (RAR)

- Hu and Rosenberger (2003) developed 'optimal RAR framework' for a K-arm RCT:
  1. Derive optimal allocation proportions to satisfy chosen experimental objectives:
  $$\rho_j^* = \rho_j(\boldsymbol{\theta}), j = 1, \dots, K, \sum \rho_j = 1.$$
  2. Construct a RAR procedure with low variability and fast convergence to the chosen optimal allocation:
  $$\pi_{ij} = \Pr(ith \text{ subject is randomized to treatment } j) = \pi_j\left(\frac{N_j(i)}{i}, \rho_j(\widehat{\boldsymbol{\theta}})\right), i \geq n_0, j = 1, \dots, K$$
  3. Evaluate properties of the RAR procedure (type I error rate, power) through simulation under standard to worst-case experimental scenarios
- Fixed total sample size (no early stopping) is assumed
- Estimators and tests have known asymptotic properties under widely satisfied conditions on $\rho_j(\boldsymbol{\theta})$ and the allocation function π_j
- Conceptually different from 'Thompson's sampling' (Bayesian RAR)

Phase III randomized controlled trial
Optimal response-adaptive randomization (RAR)

- This approach (Hu and Rosenberger, 2003) was applied to develop optimal RAR designs in various settings
  - Binary, continuous, time-to-event outcomes
  - Two-arm and multi-arm RCTs
  - Longitudinal RCTs
  - Sample size reassessment
- The approach relies on certain assumptions (common to RAR trials), such as:
  - 'Better' treatment is not more toxic
  - Outcomes are observed relatively quickly to enable design adaptations
  - There is no drift in patient characteristics over time
  - Measures to protect study integrity (blinding of sponsors/investigators, etc.) are in place

Phase III randomized controlled trial
Optimal RAR for multi-arm survival trials

- Several papers developed optimal RAR designs for $K \geq 2$-arm survival RCTs (Sverdlov et al., 2011, 2014; Frieri and Zagoraiou, 2021)
- Value:
  - Increased allocation to more variable treatment arms exhibiting longer survival ⇒ dual goals of statistical efficiency and individual patient benefit are addressed
  - Simulation evidence shows that type I error rate is maintained at the nominal level, and power is the same or higher compared to equal allocation with the same sample size
  - Robustness to model misspecification (distribution of event times)
- Challenge:
  - Events are naturally delayed ⇒ RAR is meaningful only when ~60% or more outcomes are observed during the recruitment phase of the trial

Phase III randomized controlled trial
Regulatory view on RAR has evolved over time…

- FDA finalized guidance 'Adaptive Designs for Clinical Trials of Drugs and Biologics' in November 2019
- RAR is mentioned in Section V (Adaptive Designs Based on Comparative Data), subsection E (Adaptations to Patient Allocation)

The second type is response-adaptive randomization, an adaptive feature in which the chance of a newly-enrolled subject being assigned to a treatment arm varies over the course of the trial based on accumulating outcome data for subjects previously enrolled. There are a variety of response-adaptive randomization techniques, some of which go by names such as *play the winner* designs. Statistical, ethical, and pragmatic rationales are all sometimes given for using response-adaptive randomization. In statistical terms, response-adaptive techniques can in some circumstances minimize the variance of the test statistics, leading to shorter trials, smaller sample sizes, and/or greater statistical power. The ethical argument for response-adaptive randomization is that this design feature can lead to more trial subjects being assigned to the more promising of the treatment arms. Finally, a pragmatic argument is that clinical trials with this design feature can be appealing to potential participants, thereby increasing speed and ease of accrual. Note that the arguments for response-adaptive randomization are controversial, and some researchers feel that inconclusive interim results should not be used to alter randomization in an ongoing trial and/or that statistical efficiency is not substantially improved in two-arm trials to justify adjusting randomization ratios (Hey and Kimmelman 2015, and accompanying commentaries).

Response-adaptive randomization alone does not generally increase the Type I error probability of a trial when used with appropriate statistical analysis techniques. It is important to ensure that the analysis methods appropriately take the design of the trial into account. Finally, as with many other adaptive techniques based on outcome data, response-adaptive randomization works best in trials with relatively short-term ascertainment of outcomes.

Response-adaptive randomization
Further challenges and opportunities

Platform trials:

- How to define ‘optimal allocation’ given that the number of experimental treatment arms is not known upfront?
- How to modify allocation to the shared control over time given that experimental arms may be added/dropped during the study? (Kaizer et al., 2018)
- Incorporating stratification factors (genetic signatures and other predictive biomarkers)
- Strong control of the type I error rate

One recent manuscript (Robertson et al., 2021) provides a fresh outlook at methodological and practical aspects of RAR in clinical trials

Population PK/PD studies

- PK: description of the plasma concentration of a drug as a function of time (what the body does to the drug)
- PD: description of the drug effects (what the drug does to the body)
- PK/PD model links the effect of dose on trug concentration and drug response over time
  - Mechanistic modeling of individual subject profiles with an assessment of corresponding uncertainty

Nonlinear mixed effects model (NLMEM):

$$\boldsymbol{Y}_i = \boldsymbol{f}(\boldsymbol{t}_i, \boldsymbol{d}_i, \boldsymbol{\theta}, \boldsymbol{\eta}_i) + \varepsilon_i,\ i = 1, \ldots, n$$

- $\boldsymbol{Y}_i$ = vector of responses
- $\boldsymbol{f}$ = nonlinear (vector) function
- $\boldsymbol{d}_i$ = vector of administered doses
- $\boldsymbol{t}_i$ = vector of sampling time points
- $\boldsymbol{\theta}$ = vector of typical parameter values
- $\boldsymbol{\eta}_i \sim MVN(\mathbf{0}, \boldsymbol{\Omega})$ = inter-individual variabilities
- $\boldsymbol{\varepsilon}_i \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$ = measurement errors

Population PK/PD studies
Parameter Estimation

- MLE of parameters $(\boldsymbol{\theta}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$ is found by maximizing the integrated likelihood

$$(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Omega}}, \widehat{\boldsymbol{\Sigma}}) = \arg\max_{\boldsymbol{\theta},\boldsymbol{\Omega},\boldsymbol{\Sigma}} \sum_{i=1}^{n} \log\left(\iint \ell_i(\boldsymbol{y}_i, \boldsymbol{\eta}|\boldsymbol{\theta}, \boldsymbol{\Sigma}) \cdot p(\boldsymbol{\eta}|\boldsymbol{\Omega}) d\boldsymbol{\eta}\right)$$

- No closed-form solution; advanced numerical optimization is required
  - First-order (FO) method (Beal and Shiner, 1982), originally implemented in NONMEM
  - First-order conditional estimation (FOCE) (Lindstrom and Bates, 1990), implemented in NONMEM
  - Laplace integration
  - Stochastic and Monte Carlo methods
- Software: SAS PROC NLMIXED, Perl-speaks-NONMEM (PsN), R packages (nlme, nlmer, saemix, brms, nlmixr), etc.

Population PK/PD studies
Optimal Designs

- Optimal design problem in this context involves maximization of some criterion of the population Fisher information matrix (which is not a closed-form expression) (Mentre et al., 1997)
- Elements to be optimized: dose; sampling times; sampling frequency, etc.
- Value:
  - ODs can help characterize a typical pattern of PK over time and uncertainty in the observations (very important in small studies)
  - Number of sampling times may be reduced ⇒ savings in the study cost
  - Population ODs may help improve existing therapies or diagnostics
  - Population ODs may help bridge different populations (e.g., adult to children)
- Software: a head-to-head comparison of 5 different tools (Nyberg et al., 2015)

Population PK/PD studies
Model-based adaptive optimal designs (MBAOD)

- MBAOD attempts to overcome potential non-robustness to changes in the parameter values of locally optimal designs
- Examples:
  - PK bridging study from adults to children (Strömberg, 2016) – MBAOD requires fewer children to fulfill the FDA precision criteria compared to traditional estimation methodologies
  - Robust optimality criterion in MBAOD (Strömberg and Hooker, 2017) – reduced sensitivity to model misspecification and improved practicality of experimental design
- Software: R package MBAOD (https://github.com/andrewhooker/MBAOD)

Summary
The Punchline…

Optimal designs are applicable in all stages of drug development, and they serve at least two important purposes:

1. ODs provide important theoretical benchmarks for judging alternatives
   - If a simple heuristic procedure is shown to be robust and nearly as efficient as the optimal one, its use may be well justified in a given trial
   - If a simple procedure exhibits high loss in efficiency, then alternatives should be considered
2. Adaptive ODs (stage-wise or sequential) can be constructed and implemented in practice
   - Study goals can be potentially achieved with a reduced sample size
   - Careful calibration is required, multi-stakeholder collaboration

More details in our two JSTP papers:

Journal of Statistical Theory and Practice, 7:753–773, 2013

ISSN: 1559-8608 print / 1559-8616 online
DOI: 10.1080/15598608.2013.783726

On Recent Advances in Optimal Allocation Designs in Clinical Trials

OLEKSANDR SVERDLOV[1] AND
WILLIAM F. ROSENBERGER[2]

[1]Translational Sciences, Novartis Pharmaceuticals Corporation, East Hanover, New Jersey, USA
[2]Department of Statistics, George Mason University, Fairfax, Virginia, USA

Journal of Statistical Theory and Practice (2020) 14:10
https://doi.org/10.1007/s42519-019-0073-4

ORIGINAL ARTICLE

Check for updates

On Optimal Designs for Clinical Trials: An Updated Review

Oleksandr Sverdlov[1] · Yevgen Ryeznik[2,3] · Weng Kee Wong[4]

References (1 of 5)

1. Bartroff J, Lai TL. Incorporating individual and collective ethics into phase I cancer trial designs. *Biometrics*, 67(2), 596-603, 2011. https://doi.org/10.1111/j.1541-0420.2010.01471.x
2. Beal SL, Sheiner LB. Estimating population kinetics. *Critical Reviews in Biomedical Engineering*, 8(3), 195-222, 1982. https://europepmc.org/article/MED/6754254
3. Berger V, Bour L, Carter K, et al. A roadmap to using randomization in clinical trials. *BMC Med Res Methodol* 21:68, 2021. https://doi.org/10.1186/s12874-021-01303-z
4. Braun TM. The bivariate continual reassessment method: Extending the CRM to phase I trials of two competing outcomes. *Controlled Clinical Trials*, 23(3), 240-256, 2002. https://doi.org/10.1016/s0197-2456(01)00205-7
5. Bretz F, Pinheiro J, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*, 61, 738-748, 2005. https://doi.org/10.1111/j.1541-0420.2005.00344.x
6. Chaloner K, Larntz K. Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, 21, 191-208, 1989. https://doi.org/10.1016/0378-3758(89)90004-9
7. Dragalin V, Fedorov V. Adaptive designs for dose-finding based on efficacy–toxicity response. *Journal of Statistical Planning and Inference*, 136(6), 1800-1823. https://doi.org/10.1016/j.jspi.2005.08.005
8. Fedorov V Leonov S. *Optimal Design for Nonlinear Response Models*. CRC Press, 2014. ISBN: 9780367379803.

References (2 of 5)

9. Frieri R, Zagoraiou M. Optimal and ethical designs for hypothesis testing in multi-arm exponential trials. *Statistics in Medicine*, 40, 2578-2603, 2021. https://doi.org/10.1002/sim.8919

10. Gao L, Rosenberger WF. Adaptive Bayesian design with penalty based on toxicity-efficacy response. In D. Ucinski, A. C. Atkinson, M. Patan (eds.), *mODa 10—Advances in Model-Oriented Design and Analysis*, pp. 91–98, Springer International Publishing, Basel, Switzerland, 2013.

11. Haines L, Perevozskaya I, Rosenberger WF. Bayesian optimal designs for phase I clinical trials. *Biometrics*, 59, 591-600, 2003. https://doi.org/10.1111/1541-0420.00069

12. Hu F, Rosenberger WF. Optimality, variability, power: Evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association* 98, 671-678, 2003. https://doi.org/10.1198/016214503000000576

13. Ivanova A. A new dose-finding design for bivariate outcomes. *Biometrics*, 59(4), 1001-1007, 2003. https://doi.org/10.1111/j.0006-341X.2003.00115.x

14. Kaizer AM, Hobbs BP, Koopmeiners JS. A multi-source adaptive platform design for testing sequential combinatorial treatment strategies. *Biometrics*, 74(3), 1082-1094, 2018. https://doi.org/10.1111/biom.12841

15. Karp JE, Lancet JE, Kaufmann SH, et al. Clinical and biologic activity of the farnesyltransferase inhibitor R115777 in adults with refractory and relapsed acute leukemias: a phase 1 clinical-laboratory correlative trial. Blood, 97(11), 3361-9, 2001. https://doi.org/10.1182/blood.v97.11.3361

16. King J, Wong WK. Minimax D-optimal designs for the logistic model. *Biometrics*, 56, 1263-1267, 2000. https://doi.org/10.1111/j.0006-341X.2000.01263.x

References (3 of 5)

17. Li G, Majumdar D. D-optimal designs for logistic models with three and four parameters. *Journal of Statistical Planning and Inference*, 138, 1950-1959, 2008. https://doi.org/10.1016/j.jspi.2007.07.010

18. Lindstrom MJ, Bates DM. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3), 673-687, 1990. https://pubmed.ncbi.nlm.nih.gov/2242409/

19. Mentré F, Mallet A, Baccar D. Optimal design in random-effects regression models. *Biometrika*, 84(2), 429-442, 1997. https://doi.org/10.1093/biomet/84.2.429

20. Nyberg J, Bazzoli C, Ogungbenro K, Aliev A, Leonov S, Duffull S, Hooker AC, Mentré F. Methods and software tools for design evaluation in population pharmacokinetics-pharmacodynamics studies. *British Journal of Clinical Pharmacology*, 79(1), 6-17, 2015. https://doi.org/10.1111/bcp.12352

21. Perevozskaya I, Rosenberger WF, Haines LM. Optimal design for the proportional odds model. Canadian Journal of Statistics, 31(2), 225-235, 2003. https://doi.org/10.2307/3316068

22. Pronzato L. Penalized optimal adaptive designs for dose finding. *Journal of Statistical Planning and Inference*, 140, 283-296, 2010. https://doi.org/10.1016/j.jspi.2009.07.012

23. Robertson DS, Lee KM, Lopez-Kolkovska BC, Villar SS. Response-adaptive randomization in clinical trials: from myths to practical considerations. https://arxiv.org/abs/2005.00564

References (4 of 5)

24. Roy A, Ghosal S, Rosenberger WF. Convergence properties of sequential Bayesian D-optimal designs. *Journal of Statistical Planning and Inference*, 139, 425-440, 2009. https://doi.org/10.1016/j.jspi.2008.04.025

25. Rosenberger WF, Canfield GC, Perevozskaya I, Haines LM, Hausner P. Development of interactive software for Bayesian optimal phase 1 clinical trial design. *Drug Information Journal*, 39, 89-98, 2005. https://doi.org/10.1177/009286150503900112

26. Ryeznik Y, Sverdlov O, Hooker A. Adaptive optimal designs for dose-finding studies with time-to-event outcomes. *The AAPS Journal*, 20(1):24, 2018. https://doi.org/10.1208/s12248-017-0166-5

27. Ryeznik Y, Sverdlov O, Hooker A. Implementing optimal designs for dose–response studies through adaptive randomization for a small population group. *The AAPS Journal*, 20(5):85, 2018. https://doi.org/10.1208/s12248-018-0242-5

28. Sverdlov O, Gao L. Phase I/II dose-finding designs with efficacy and safety endpoints. In: J. O'Quigley, A. Iasonos, and B. Bornkamp, editors, pp.81–107, *Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials*. Boca Raton, FL, CRC Press, 2017.

29. Sverdlov O, Rosenberger WF. On recent advances in optimal allocation designs in clinical trials. *Journal of Statistical Theory and Practice*, 7(4), 753-773, 2013. https://doi.org/10.1080/15598608.2013.783726

References (5 of 5)

30. Sverdlov O, Ryeznik Y, Wong WK. Efficient and ethical response-adaptive randomization designs for multi-arm clinical trials with Weibull time-to-event outcomes. *Journal of Biopharmaceutical Statistics*, 24(4), 732-754, 2014. https://doi.org/10.1080/10543406.2014.903261

31. Sverdlov O, Ryeznik Y, Wong WK. On optimal designs for clinical trials: An updated review. *Journal of Statistical Theory and Practice*, 14(1), 1-29, 2020. https://doi.org/10.1007/s42519-019-0073-4

32. Sverdlov O, Tymofyeyev Y, Wong WK. Optimal response-adaptive randomized designs for multi-armed survival trials. *Statistics in Medicine*, 30, 2890-2910, 2011. https://doi.org/10.1002/sim.4331

33. Sverdlov O, Wong WK, Ryeznik Y. Adaptive clinical trial designs for phase I cancer studies. *Statistics Surveys*, 8, 2-44, 2014. https://doi.org/10.1214/14-SS106

34. Strömberg EA. *Simulated Model Based Adaptive Optimal Design Using FDA Stopping Criteria.* https://www.page-meeting.org/pdf_assets/7437-MBAOD_simulation_ES_2.pdf

35. Strömberg EA, Hooker AC. The effect of using a robust optimality criterion in model based adaptive optimization. *Journal of Pharmacokinetics and Pharmacodynamics*, 44, 317-324, 2017. https://doi.org/10.1007/s10928-017-9521-5

36. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*, 60(3), 684-693, 2004. https://doi.org/10.1111/j.0006-341x.2004.00218.x