Some Issues in Response-Adaptive Designs for Dose-Finding Studies

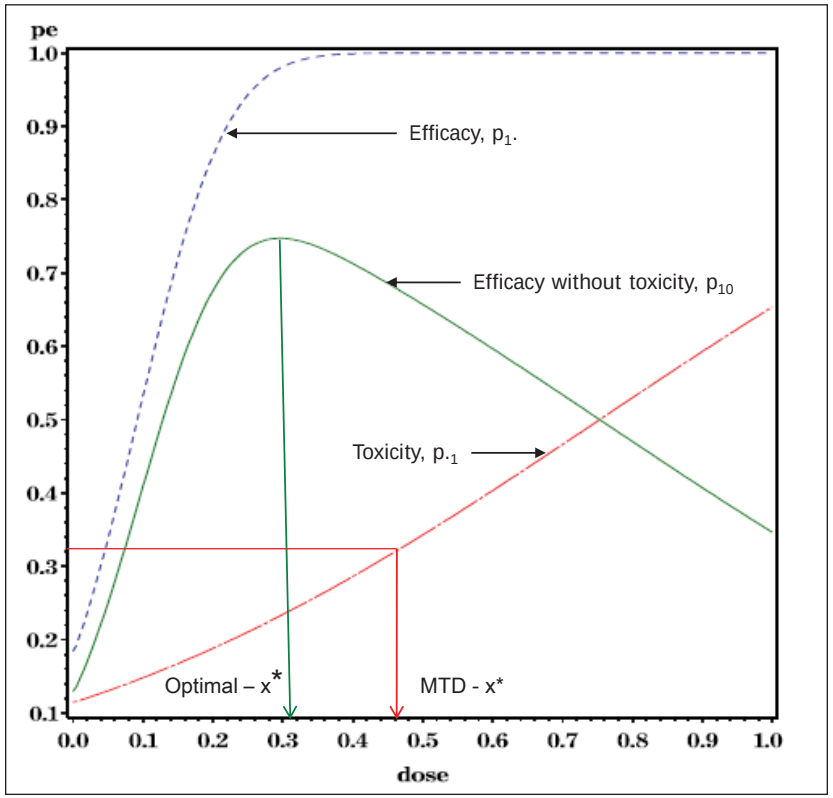
Nancy Flournoy

Department of Statistics
University of Missouri, Columbia, MO 65211
flournoyn@missouri.edu

Design of Experiments in Healthcare
Isaac Newton Institute for Mathematical Sciences
19 August 2011

Basic Response Function Shape

Typical dose-response curves



curtesy of V. Fedorov

Outline

1. Up and Down $\neq 3+3$
2. Treatment Allocations by Up & Down and CRM
3. Best Intention Designs for Toxicity, e.g., CRM
4. Best Intention Designs
   in the context of Efficacy and Toxicity,
   compared with Optimal Designs

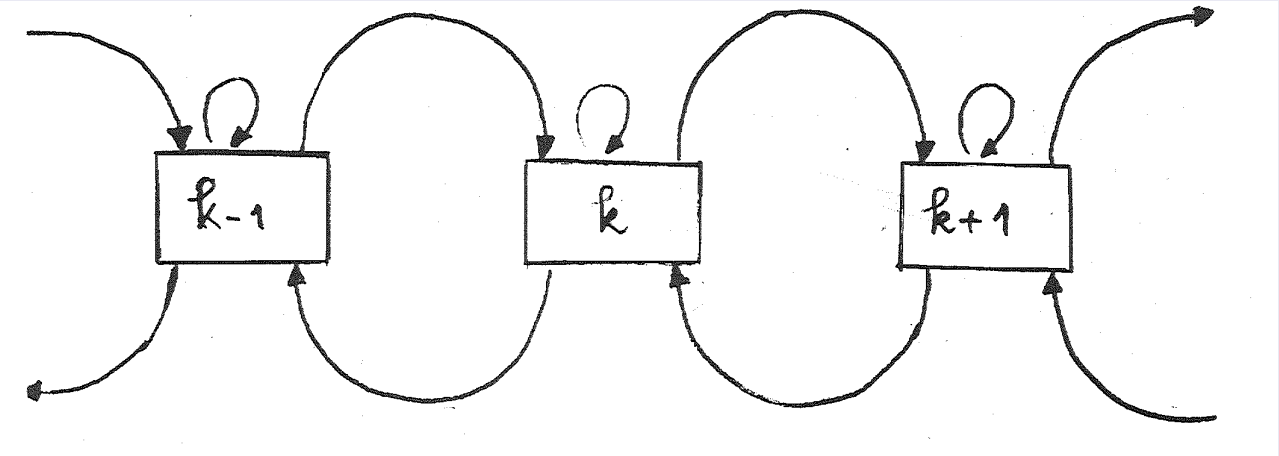
Outline

Up and Down $\neq 3 + 3$

$3 + 3$

$3 + 3$ is a dose escalation procedure, period.
There is no going down.

up-and-down



von Bétsky, 1947; Dixon & Mood, 1948

Selecting the MTD

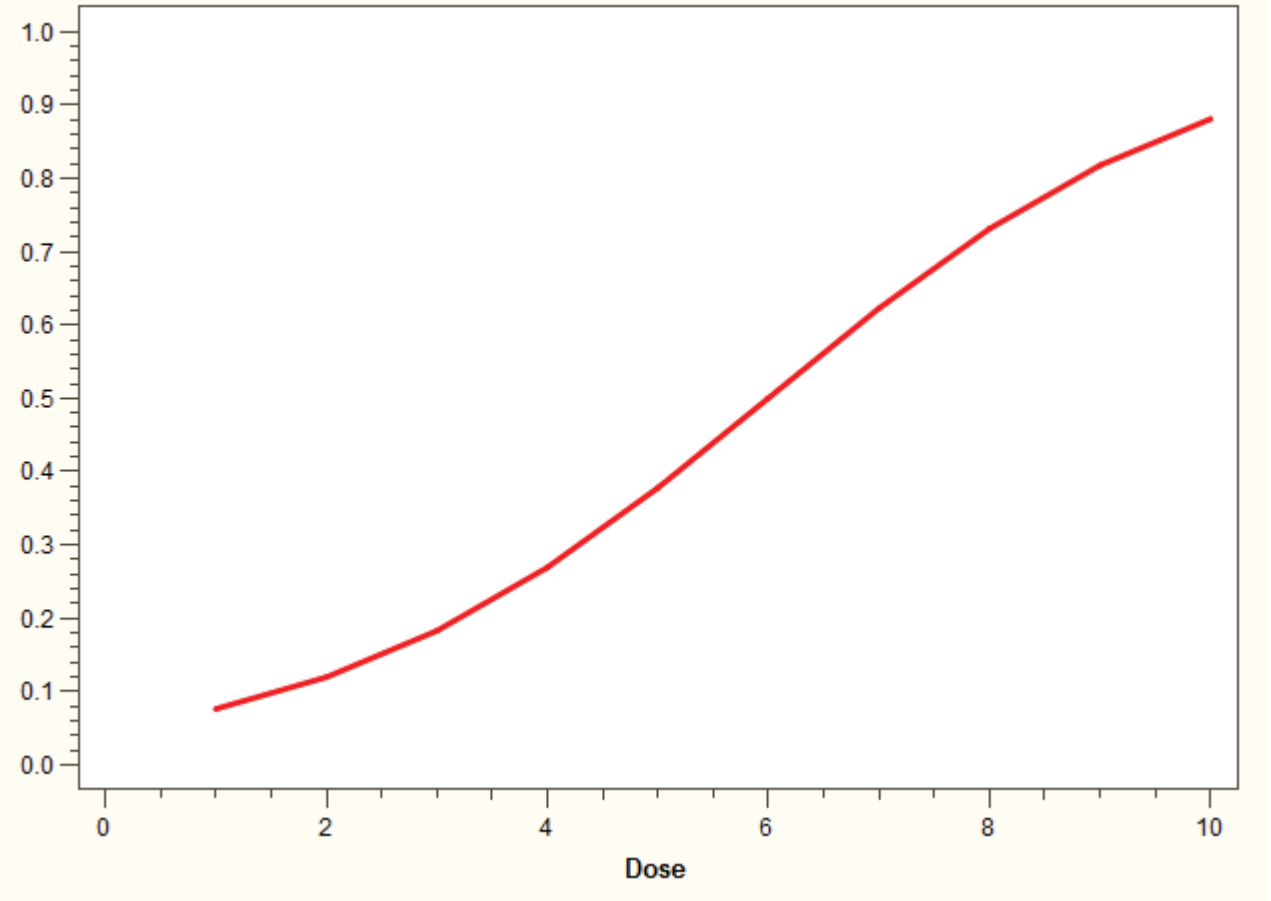
Proportion of Correct Selection of MTD
by CRM and
by U&D with groups of size 2 after 8 cohorts

Senario	MTD	CRM	U&D
Uniform	1	50.2	54.0
Gamma	2	36.6	40.8
Normal	3	57.8	56.4
Lognormal	4	46.7	33.0
Weibull	5	39.0	38.1
Logistic	6	26.0	32.2

From Oron & Hoff (2011)

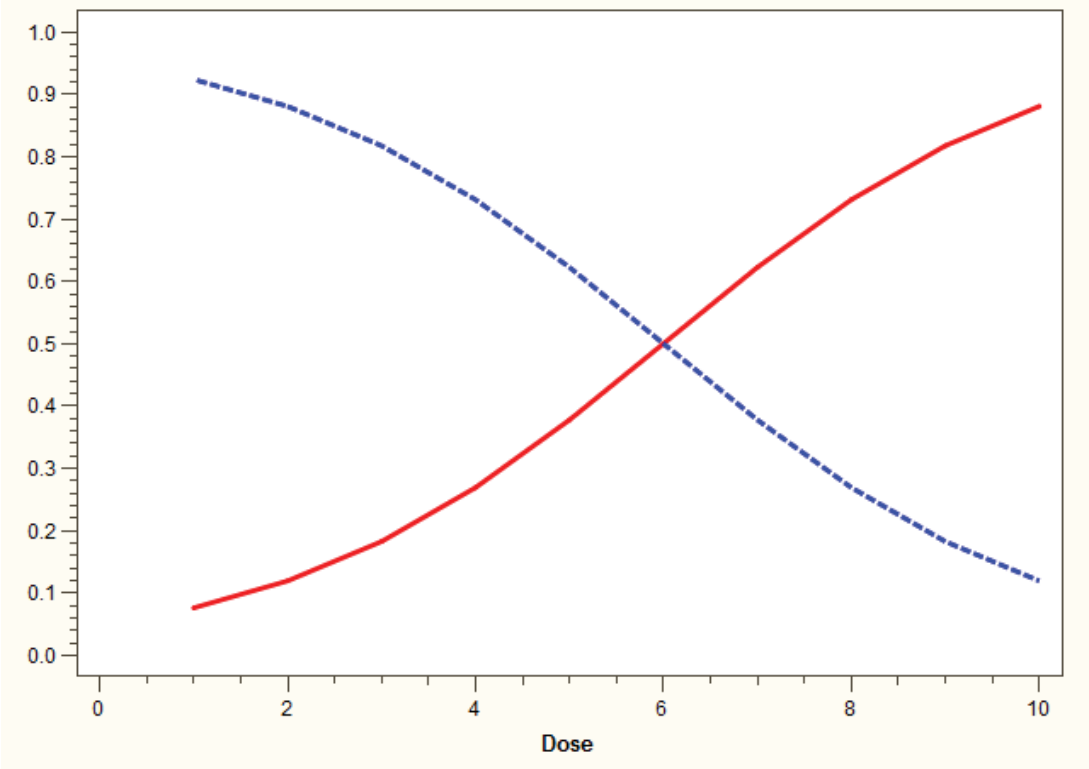
Oron & Hoff give similar results for 16 cohorts;
similar estimation performance also found by Durham, Flournoy and Rosenberger (1997)

P(Toxicity)



Suppose the dose is decreased if a toxicity is observed

P(Toxicity) and P(No Toxicity)

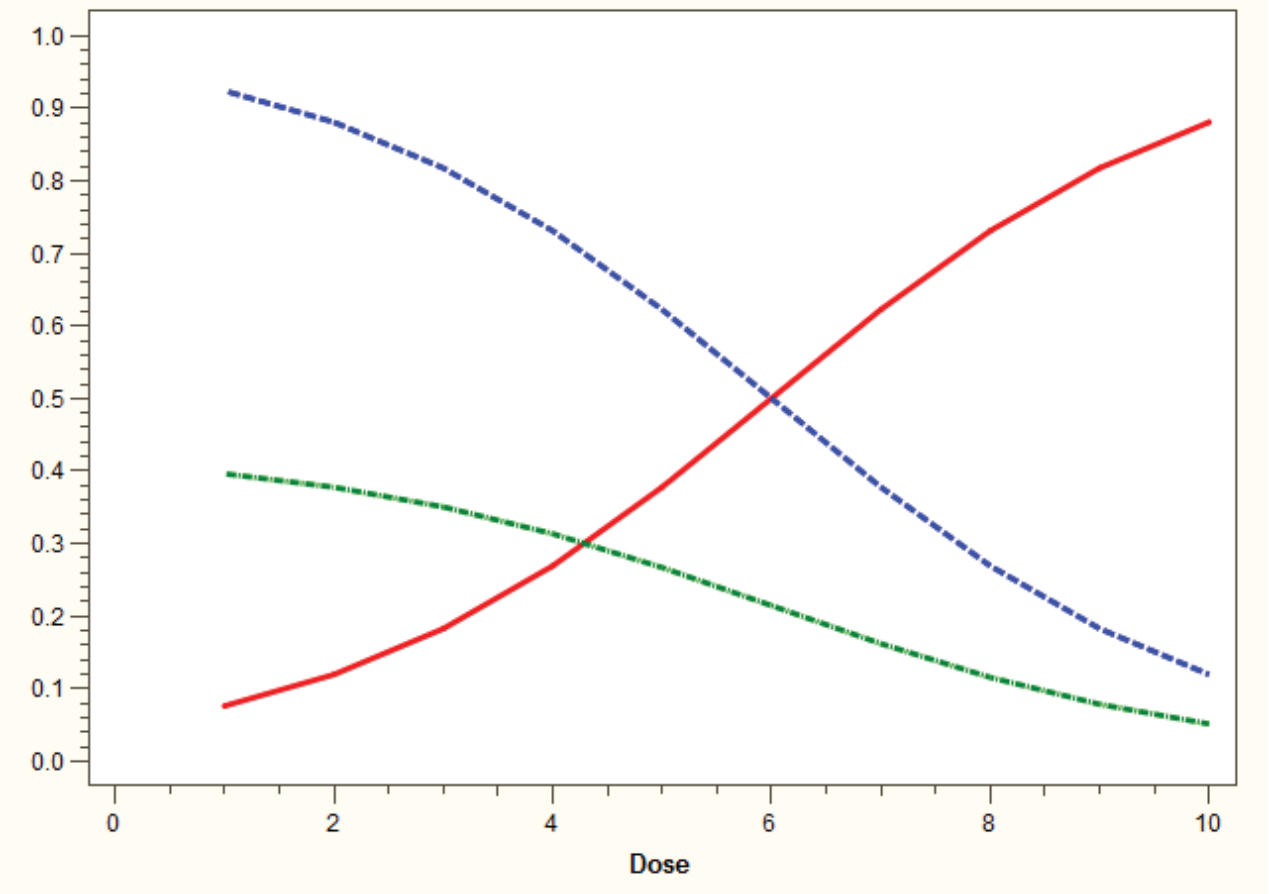


Cluster allocations around LD50

1. decrease dose if toxicity is observed
2. increase dose if no toxicity is observed

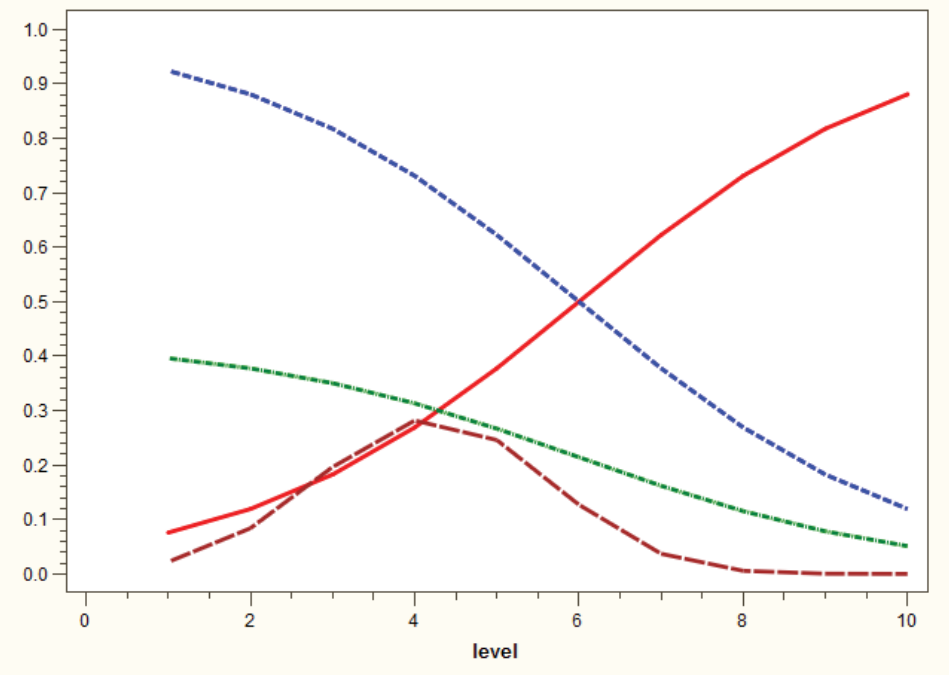
Derman, 1957; Durham & Flournoy, 1994

$P(\text{Going Up}) = (.3/.7)P(\text{No Toxicity})$ targets 30% toxicity rate



Derman, 1957; Durham & Flournoy, 1994; Durham & Flournoy, 1995; Durham, Flournoy and Haghighi, 1995; Giovagnoli & Pintacuda, 1998, Bortot, Giovagnoli, 2005; Baldi Antognini, Bortot, & Giovagnoli, 2008

Treatment Allocations Targeting .3



Change P(UP) to cluster allocations around LD30

1. go up if no toxicity and heads wp .3/.7
2. go up if no toxicity in group of size 2

(1) Durham & Flournoy, 1994, (2) Gezmu & Flournoy, 2006; Stylianou and Flournoy (2002) recommend estimating target by isotonic regression.

Up if no Efficacy and and Down if Toxicity

Sounds good, but wrong target

1. Go down if toxicity
2. Go up if no efficacy

Allocations cluster around dose where
P('no toxicity')=P('efficacy')
which is not the
dose that maximizes P('no toxicity'and 'efficacy').

(Corollary of Durham & Flournoy, 1994).

simulated in Gooley et al, 1994; corollary discussed in Kpamegan, 2001, further analyzed in Ivanova, 2003

Up and Down: Discretized Kiefer-Wolfowitz

Right target; but clinically unappealing

If the *nth* pair of subjects has been treated at $(n) - \frac{\Delta}{2}$ and $X(n) + \frac{\Delta}{2}$, the midpoint of the $(n+1)th$ pair is

$$X(n+1) = X(n) + \Delta V(n),$$

$$\mathbf{V(n)} = \begin{cases} -\mathbf{1} & \text{if treatment at } X(n) - \frac{\Delta}{2} \text{results in success} \\ & \text{and treatment at } X(n) + \frac{\Delta}{2} \text{ in failure.} \\ \mathbf{0} & \text{if the nth pair of treatments result in 2 successes} \\ & \text{or 2 failures.} \\ \mathbf{1} & \text{if the treatment at } \mathbf{X(n)} - \frac{\Delta}{2} \text{ results in failure} \\ & \text{and the treatment } \mathbf{X(n)} + \frac{\Delta}{2} \text{ in success.} \end{cases}$$

Allocations cluster around dose with maximum

$$\underline{P\{\text{'no toxicity' and 'efficacy'}\}}$$

Kpamegan & Flournoy, 2001, 2008

Outline

Number of Cohorts Allocated to MTD: CRM left; U&D right; Oron, et al., 2011

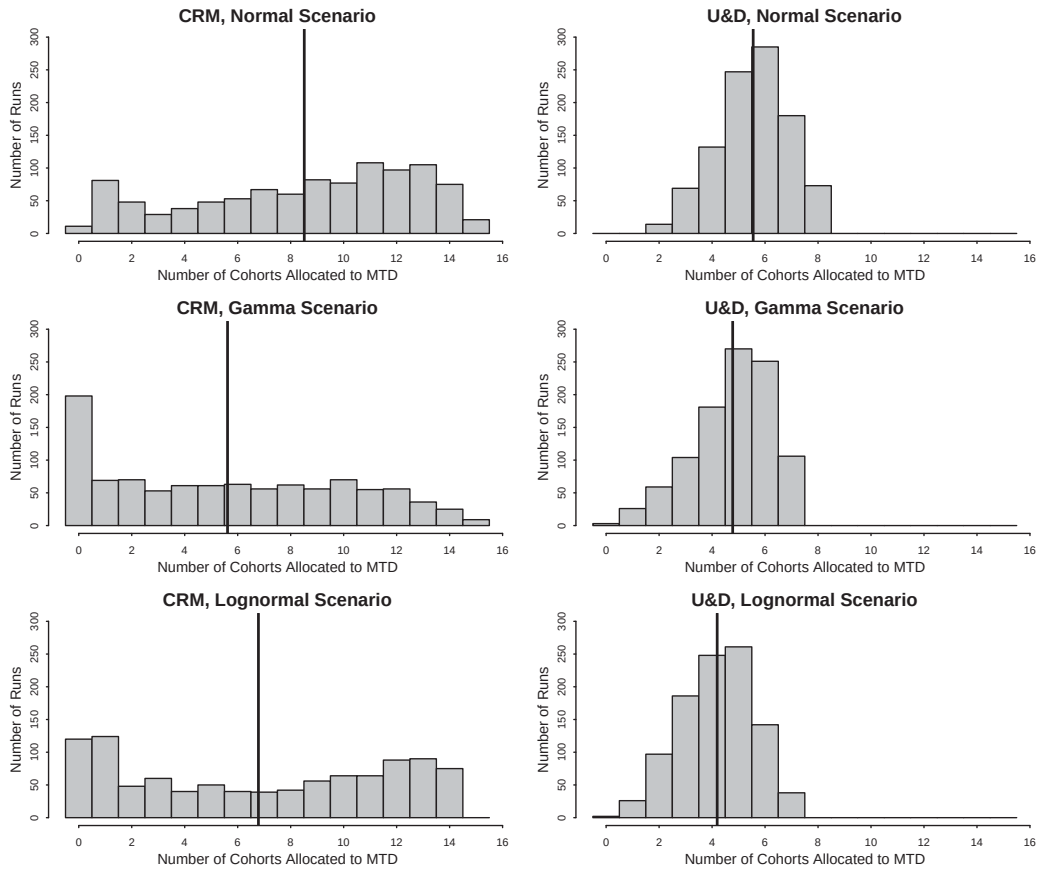


Figure 5: Between-run and between-scenario variability. The histograms depict the distribution of the number of cohorts (excluding the first one) that have been allocated the true MTD during a single specific run. The ensemble size is 1000 runs. Scenarios are Normal (top), Gamma (middle) and Lognormal (bottom); designs are CRM one-parameter 'power' (left) and GU&D (right), both with cohort size 2.

Outline

Impossibility Theorem: $\mathbf{d}_j = MTD$

$\mathbf{x}_n \in \mathcal{F}_{n-1},\ y_n \mid \mathcal{F}_{n-1} \sim \mathbf{Bernoulli}\left(f\left(x_n\right)\right)$

$\{x_n\}_{n=1}^{\infty}$ is a design

Sequence of estimators $\left\{\widehat{MTD}_n\right\}_{n=1}^{\infty}$ is said to be strongly consistent with respect to a given design

if $\widehat{MTD}_n \to d_j$ as for all increasing functions f,

or equivalently, by discreteness,

$P(\exists N \text{ such that for all } n \geq N, \widehat{MTD}_n = d_j) = 1$

Under this framework, there exists no design that satisfies for all increasing functions f,

$$P(\exists N \text{ such that for all } n \geq N, x_n = d_j) = 1) = 1,$$

or equivalently, that

$$P(x_n \neq d_j i.o.) = 0$$

Azriel, Mandel & Rinott, 2011

Azriel (2011) proves Cheung & Chappell (2002) conditions for CRM consistency

Oron et al. (2011) sampled dose-response curves and found conditions met 26% of the time.

Azriel sampled curves that satisfy the commonly used power function prior and found conditions satisfied 43% of the time.

How to use this information?

Cheung reports (personal communication here) that he provides a method for constructing priors that will satisfy his conditions. Is this consistent with Bayesian philosophy?

Designs that are consistent for the MTD

from Azriel, et al., 2011

Using $\widehat{MTD}$ based on isotonic estimators, consistent designs have been proposed by

1. Ivanova, et al. (2003??)
2. Ivanova and Kim (2009)
3. Azriel, et al., (2011)

Tradeoffs

- For increasing response functions, the dose that minimizes the variance of the estimated target is the target itself, except if target quantile is extreme.
- This suggests a good design will converge quickly to the target dose.
- Converging quickly to the target dose costs in terms of information about slope of toxicity function at the target.
- Converging quickly can cause convergence to the wrong dose; it may converge before reaching the target. (Chang & Ying, 2009)

Tradeoffs

- For long memory designs, first dose is very influential on converge rate.
  This has been long known for stochastic approximation and is periodically rediscovered, as by Resche-Rigon, Zohar and Chevret (2008), concerning the CRM.
- Long memory designs, based on maximum likelihood or Bayesian estimates, may not converge or converge to the wrong dose with substantial probability.
- Up & down designs with Markovian rules for changing doses are short memory designs, with well-known analytical finite operating characteristics, that converge exponentially fast to a distribution clustered around the target.

Outline

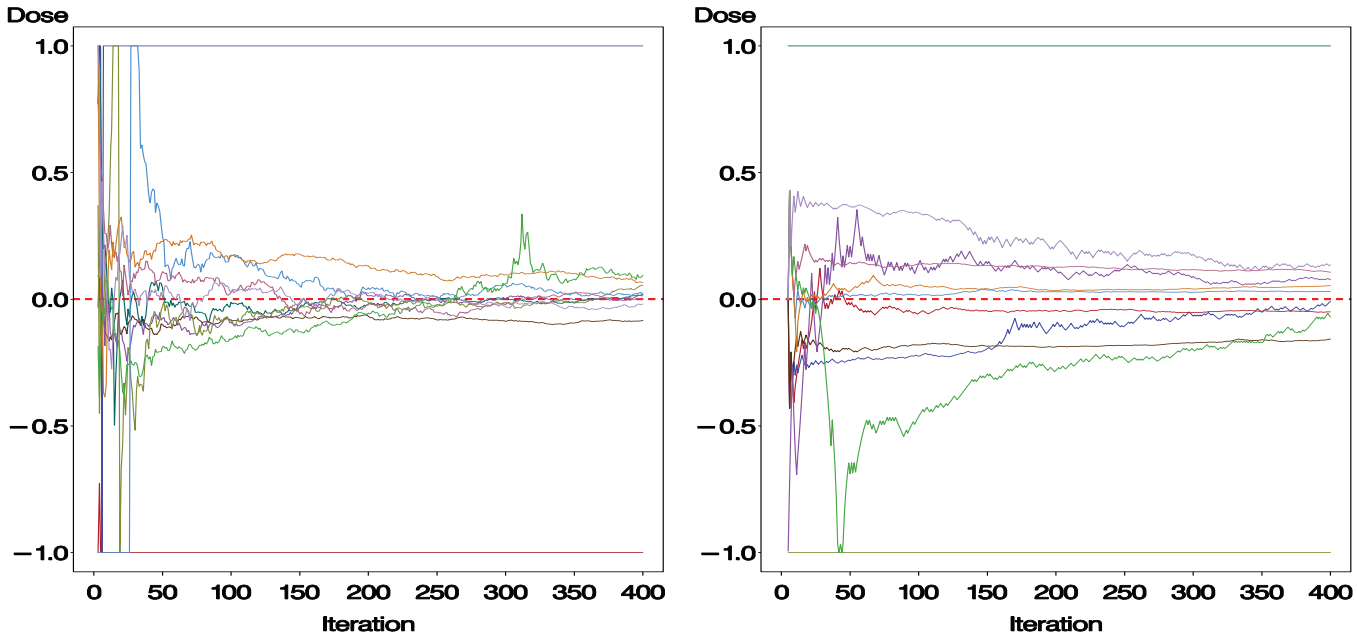


Figure 1: Some dose allocation vs. sample size for 10 trials under first order model, model(6)(left), dichotomized responses (9)(right).

curtesy of V. Fedorov, 2011

Estimated target doses, ARM left; D-optimal right

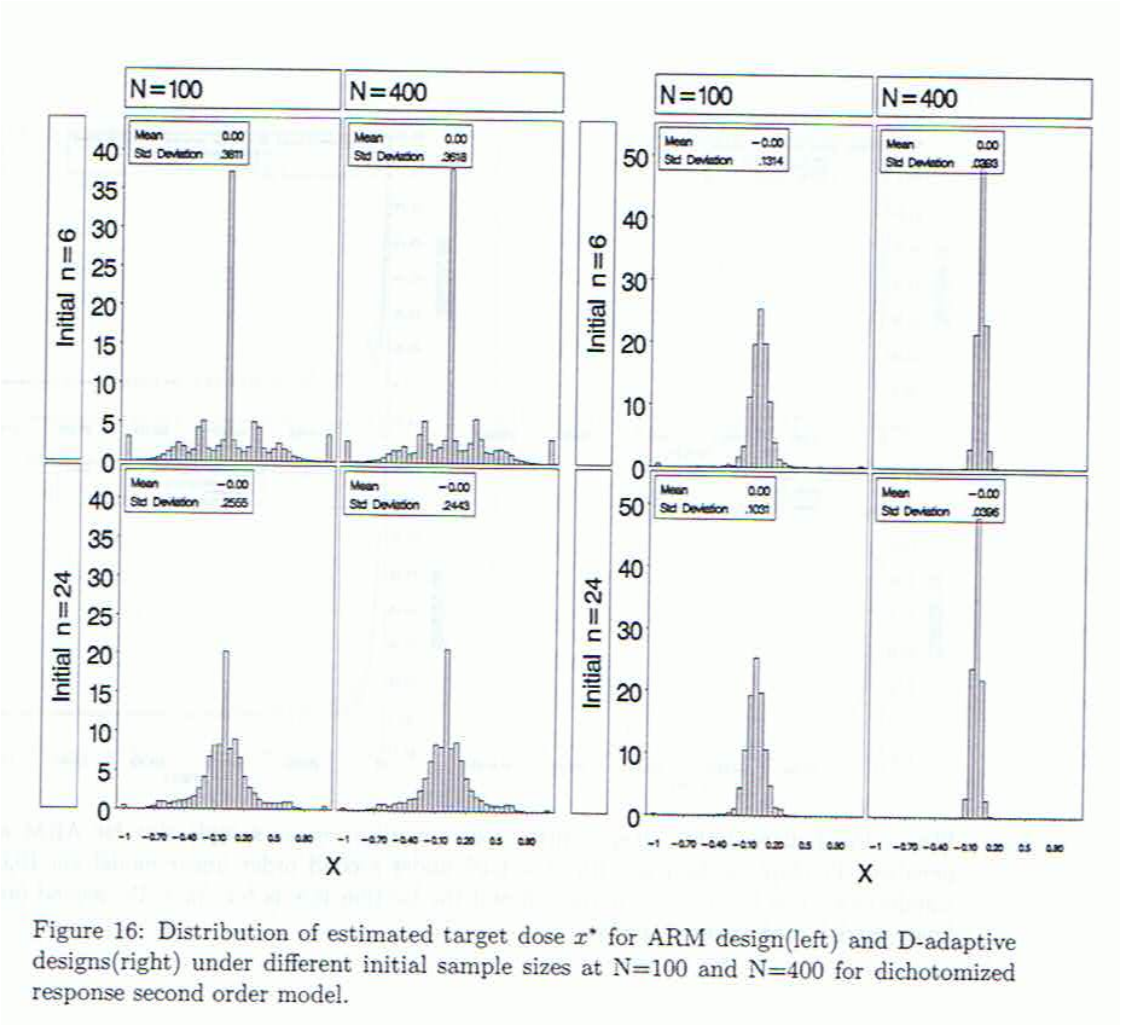


Figure 16: Distribution of estimated target dose x^* for ARM design(left) and D-adaptive designs(right) under different initial sample sizes at N=100 and N=400 for dichotomized response second order model.

curtesy of V. Fedorov, 2011

It is Important to Understand Interactions between Objective Criteria

In typical phase one dose-finding studies,

- inference and patient gain objectives *compliment* each other.

In typical phase two dose-finding studies,

- inference and patient gain objectives *contradict* each other.

Conflict Between Objectives with Unimodal Success Probability Functions

- With Best Intentions Designs maximizing patient gain, goal is a One Point Design at $argmax_x\{P(Success|x)\}$
- Designs that minimizes the variance of MLE of $argmax\{P(Success|x)\}$ typically are at least two point designs
  - This has been proven for several parametric models of toxicity and efficacy and believed to hold very generally when marginal $P\{toxicity|dose\}$ and $P\{efficacy|dose\}$ are common two parameter models.
    Fan and Chaloner (2003),
    Rabie and Flournoy (2004),
    Dragalin, Fedorov and Wu (2006).

Ways to Deal with the Conflict between Criteria

- Choose between the Goals of Estimation and Patient Gain
- Use Compound Criteria
- Penalize Undesirable Outcomes (Dragalin, Fedorov and Wu (2006)
- Pronzato (2010) suggests starting with estimation objective and then converting to patient gain objective after enough information about the response function has been obtained to give reliable estimates.

Thank you!

Outline

References

1. Azriel, D. (2010) Dissertation and Technical Report. Israel
2. Aziel, D. (2011) A note on the robustness of the continual reassessment method. Preprint.
3. Azriel, D., Mandel, M., Rinott, Y., 2011. The treatment versus experimentation dilemma in dose-finding studies. J. Statist. Plann. Inference 141 (8), 2759–2768.
4. Baldi Antognini, Alessandro · Paola Bortot · Alessandra GiovagnoliAISM (2008) Randomized group up and down experiments AISM 60:4559 DOI 10.1007/s10463-006-0081-5 Randomized group up and down experiments
5. Bortot, P., Giovagnoli, A. (2005) Up-and-down experiments of first and second order. Journal of Statistical Planning Inference 134(1), 236-253

References continued

1. Chang, H-H., Ying, Z. (2009) Nonliner sequential designs for logistic item response theory models with applications to comperterized adaptive tests. Annals of Statistics 37 1466-1488.
2. Cheung, Y., 2002. Letter to the editor. Biometrics 58 (1), 237–240.
3. Cheung, Y., Chappell, R., 2002. A simple technique to evaluate model sensitivity in the continual reassessment method. Biometrics 58 (3), 671–674.
4. Derman, C., 1957. Non-parametric up-and-down experimentation. Ann. Math. Statist. 28, 795–798.
5. Dixon, W. J., Mood, A. M., 1948. A method for obtaining and analyzing sensitivity data. J. Am. Statist. Assoc. 43, 109–126.

References continued

1. Dragalin, V., Fedorov, V., and Wu, Y. (2006) Adaptive designs for selecting drug combinations based on efficacy-toxicity response. JSPI
2. Durham, SD, Flournoy, N (1995). Up-and-down designs I: Stationary treatment distributions. Adaptive Designs. Lecture-Notes Monograph Series 25 139-157. Flournoy, N, Rosenberger, WF (eds.). Hayward, CA: Institute of Mathematical Statistics.
3. Durham, SD, Flournoy, N, Haghighi, AA (1995). Up-and-down designs II: Exact treatment moments. Adaptive Designs. Lecture-Notes Monograph Series 25 158-178. Flournoy, N, Rosenberger, WF (eds.). Hayward, CA: Institute of Mathematical Statistics.
4. Durham, SD, Flournoy, N, Rosenberger, WF (1997). A random walk rule for phase I clinical trials. Biometrics 53 745-760.

References continued

1. Durham, S. D., Flournoy, N., 1994. Random walks for quantile estimation. In: Statistical decision theory and related topics V (West Lafayette, IN, 1992). Springer, New York, pp. 467–476.
2. Fan, SK and Chaloner, K (2003) Optimal designs and limiting optimal designs for a trinomial response. JSPI
3. Gezmu, M., Flournoy, N., 2006. Group up-and-down designs for dose-¯nding. J. Statist. Plann. Inference 136 (6), 1749–1764.
4. Giovagnoli, A., Pintacuda, N., 1998. Properties of frequency distributions induced by general up-and-down methods for estimating quantiles. J. Statist. Plann. Inference 74 (1), 51–63.

References continued

1. Gooley, T. A., Martin, P. J., Lloyd, D. F., Pettinger, M. Simulation as a design tool for Phase I/II clinical trials: An example from bone marrow transplantation. *Controlled Clinical Trials* **15,** 450-460.
2. Goodman, S., Zahurak, M., Piantadosi, S., 1995. Some practical improvements in the continual reassessment method for phase I studies. Stat. Med. 14, 1149–1161.
3. Hardwick, J., Meyer, M.C. and Stout, Q.F. (2003) Directed walk designs for dose response problems with competing failure modes. *Biometrics*
4. Kiefer, J. and Wolfowitz, J. (1952). Stochastic approximation of the maximum of a regression function. *Annals of Mathematical Statistics* **25,** 529-532.

References continued

1. Kpamegan, EE, Flournoy, N (2001). An optimizing up and down design. Optimum Design 2000 211-224. Atkinson, A.C., Bogacka, B., Zhigljavsky, A. (eds.). The Netherlands: Kluwer Academic Publishers.
2. Kpamegan, EE, Flournoy, N (2008). Up-and-down designs for selecting the dose with maximum success probability. Sequential Analysis 27 78–96.
3. Ivanova, A., Flournoy, N., Chung, Y., 2007. Cumulative cohort design for dose–finding. J. Statist. Plann. Inference 137, 2316–2327.

References continued

1. Ivanova, A., Kim, S.H., 2009.Dosefinding for continuous and ordinal outcomes with a monotone objective function: a unified approach.Biometrics 65, 307315.
2. Ivanova, A., Haghighi, A., Mohanty, S., Durham,S.,2003.Improvedup-and-down designs for phasel trials. Statistics in Medicine 22, 6982.
3. Oron, A. P., 2009. Small-sample behavior of long-memory Phase I cancer designs. preprint and In: DAE 2009, Columbia, MO. Invited Conference Talk.
4. Oron, A.P., Hoff, P.D. (2011). Small–sample behavior of novel phase I designs. preprint.
5. Oron, AP, Azriel, D, and Hoff, PD (2011) Dose-finding designs: The role of convergence properties. International Journal of Biostatistics. In Press.

References continued

1. O'Quigley, J, Shen, LZ. (1996) Continual reassessment method: a likelihood approach. Biometrics 52 673-84.
2. Pronzato, L (2010). Penalized optimal designs for dose-finding. J. Statist. Plann. Inference 140 283–296.
3. Rabie, H., Flournoy, N. (2004). Optimal designs for contingent response models. MoDa 7 Advances in Model-Oriented Design and Analysis 133-142. A.D. Bucchianico, H. Lauter and H.P. Wynn, editors. Heidelberg: Physica-Verlag.
4. Resche-Rigon, M., Zohar, S., Cheveret, S. (2008) Adaptive designs or dose-finding in non-cancer phase II trials: influence of early unexpected outcomes.
5. Stylianou, M, Flournoy, N (2002). Dose finding using isotonic regression estimates in an up-and-down biased coin design. Biometrics 58 171-177.

Other Best Intention References

Best Intention (Treatment above Everything)

- Adaptive R – M (ARM)

Plug-in estimated parameters and search for the needed dose, made the next observation at this dose

G. Wetherill (1963) Sequential Estimation of Quantal Response Curves, JRSS (B), 25, 1-48

T. Lai, H. Robbins (1982), Adaptive Design in Regression and Control, Proc. Natl. Acad. Sci. USA, 75, 586-587

C.F.J. Wu (1985), Efficient Sequential Designs with Binary Data, JASA, 80, 974-984

- Continual Reassessment Method (CRM)

CRM ~ ARM + Discrete doses + Dose escalation + Bayesian blending

J. O'Quigley, M. Pepe, L. Fisher (1990), Continual Reassessment Method, Biometrics, 46, 33-48

J. O'Quigley, L. Shen (1996), Continual Reassessment method: Likelihood Approach, Biometrics, 52, 673

- Adaptive Response (Utility) Optimization

Plug-in estimated parameters and search for the optimal dose, made an observation at this dose

Zh. Li, S. Durham, N. Flournoy (1995), An Adaptive Design for Maximization of a Contingent Binary Response, In "Adaptive design", IMS lecture Notes, Volume 25.

- Desirability Maximization

Plug-in estimated parameters and search for the dose that maximizes a desirability function + Dose escalation + Bayesian blending

P. Thall, J. Cook (2004), Dose-Finding Based on Efficacy-Toxicity Trade-offs, Biometrics, 60, 684-693

curtesy of V. Fedorov