

Adaptive Dose Ranging Studies: Flexible, Adaptive Dose-Finding Designs

Frank Bretz and José Pinheiro

Novartis Pharmaceuticals

Tokyo University of Science, July 28, 2006

Outline

- Background and motivation
- Adaptive Dose Ranging Studies PhRMA initiative: goals and scope
- Evaluating DF methods: simulation study
- Simulation results
- Conclusions
- Preliminary recommendations

Background

- Pharmaceutical industry **pipeline problem**: decreasing number of approved drugs, despite advances in basic science
- FDA's **Critical Path Initiative** — “Innovation vs. Stagnation” White Paper
- Pharmaceutical industry (PhRMA) reaction: different working groups (WGs) addressing **key drivers** of poor performance
- Adaptive Dose Ranging Studies (ADRS) group formed to address problems with inefficient **dose finding** performance

Adaptive Dose Ranging Studies core WG members

- Alex Dmitrienko, Eli Lilly
- Amit Roy, BMS
- Brenda Gaydos, Eli Lilly
- Frank Bretz, Novartis
- Frank Shen, BMS
- Greg Enas, Eli Lilly
- **José Pinheiro**, Novartis
- Michael Krams, Pfizer
- Qing Liu, J & J
- **Rick Sax**, AstraZeneca
- Tom Parke, Tessella

ADRS additional WG members

- Björn Bornkamp, University of Dortmund
- Beat Neuenschwander, Novartis
- Chyi-Hung Hsu, Pfizer
- Franz König, Med. Univ. Vienna

ADRS initiative – Motivation

- Poor understanding of dose response (DR) for both **efficacy** and **safety** is pervasive in drug development
- Indicated by both FDA and industry as one of **root causes** of late phase attrition and post-marketing problems with approved drugs
- Current dose finding designs and methods focus on selection of target dose (e.g., minimum effective dose) out of fixed, generally small number of dose levels, via pairwise hypothesis testing \implies **inefficient**

ADRS initiative – Goals

- Investigate and develop designs and methods for efficiently **learning** about safety and efficacy DR profile \implies benefit/risk profile
- More accurate and faster **decision making** on dose selection and improved labeling
- Evaluate statistical operational characteristics of alternative designs and methods to make recommendations on their use in practice
- Increase awareness about this class of designs, promoting their use, when advantageous

ADRS – Definition and Scope

- Adaptive dose-ranging designs allowing **dynamic** allocation of patients and possibly variable number of dose levels based on accumulating information
- Intended to strike **balance** between need for additional DR information and increased costs and time-lines
- Emphasis on modeling/estimation (**learning**) as opposed to hypothesis testing (**confirming**)
- Investigate existing and new ADRS methods via simulation
- Evaluate potential benefits over traditional dose-ranging designs over variety of scenarios to make recommendations on practical usefulness of ADRS methods

Dose Finding Methods – Fixed Doses

- Traditional **ANOVA** based on pairwise comparisons and multiplicity adjustment (Dunnett); common approach used in dose finding studies
- **MCP-Mod** combination of multiple comparison procedure (MCP) to identify presence of DR and modeling, to estimate target dose(s) and DR profile (Bretz, Pinheiro and Branson, 2005)
- **MTT**: novel method based on Multiple Trend Tests
- Bayesian Model Averaging: **BMA**
- Nonparametric local regression fitting: **LOCFIT**

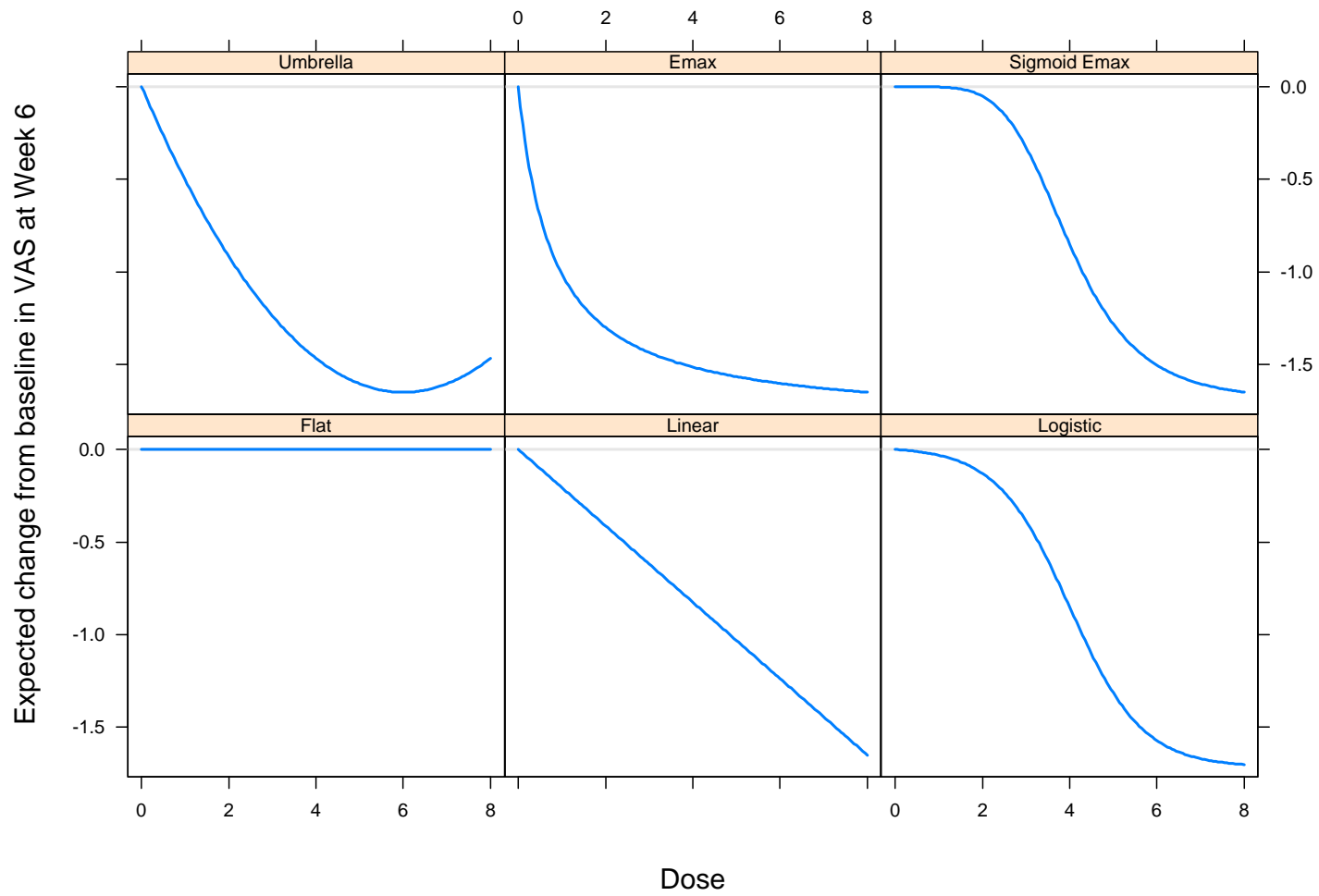
Dose Finding Methods – ADRS

- **GADA**: Dynamic dose allocation based on Bayesian normal dynamic linear model (Krams, Lees and Berry, 2005); allocation of patients to dose adaptively changed according to model-based optimization criteria (e.g., variance of target dose estimate)
- **D-opt**: novel adaptive dose allocation based on D-optimality criterion used with sigmoid- E_{\max} model; model parameters re-estimated at interim analysis and corresponding D-optimal allocation determined for next interval

Simulation study: design and assumptions

- Proof-of-concept + dose finding trial, motivated by neuropathic pain indication
- Key questions: whether there is evidence of dose response and, if so, which dose level to bring to confirmatory phase and how well dose response (DR) curve is estimated
- Primary endpoint: change from baseline in VAS at Week 6
- Dose design scenarios:
 - 5 equally spaced doses levels 0, 2, 4, 6, 8
 - 7 unequally spaced dose levels: 0, 2, 3, 4, 5, 6, 8
 - 9 equally spaced dose levels: 0, 1, ..., 8
- Significance level: one-sided FWER $\alpha = 0.05$
- Sample sizes: 150 and 250 patients (total)

Dose response profiles



Measuring performance

- Probability of identifying dose response: $Pr(DR)$
- Probability of identifying clinical relevance and selecting a dose for confirmatory phase: $Pr(dose)$
- Dose selection
 - Distribution of selected doses (rounded to nearest integer, if continuous estimate possible)

Dose selection performance (cont.)

- Target dose interval – doses that produce effect within $\pm 10\%$ of target effect Δ

Model	Target dose		Target interval	
	actual	rounded	actual	rounded
Linear	6.30	6	(5.67, 6.93)	{6,7}
Logistic	4.96	5	(4.65, 5.35)	{5}
Umbrella	3.24	3	(2.76, 3.81)	{3,4}
E _{max}	2.00	2	(1.44, 2.95)	{2,3}
Sig-E _{max}	5.06	5	(4.68, 5.58)	{5}

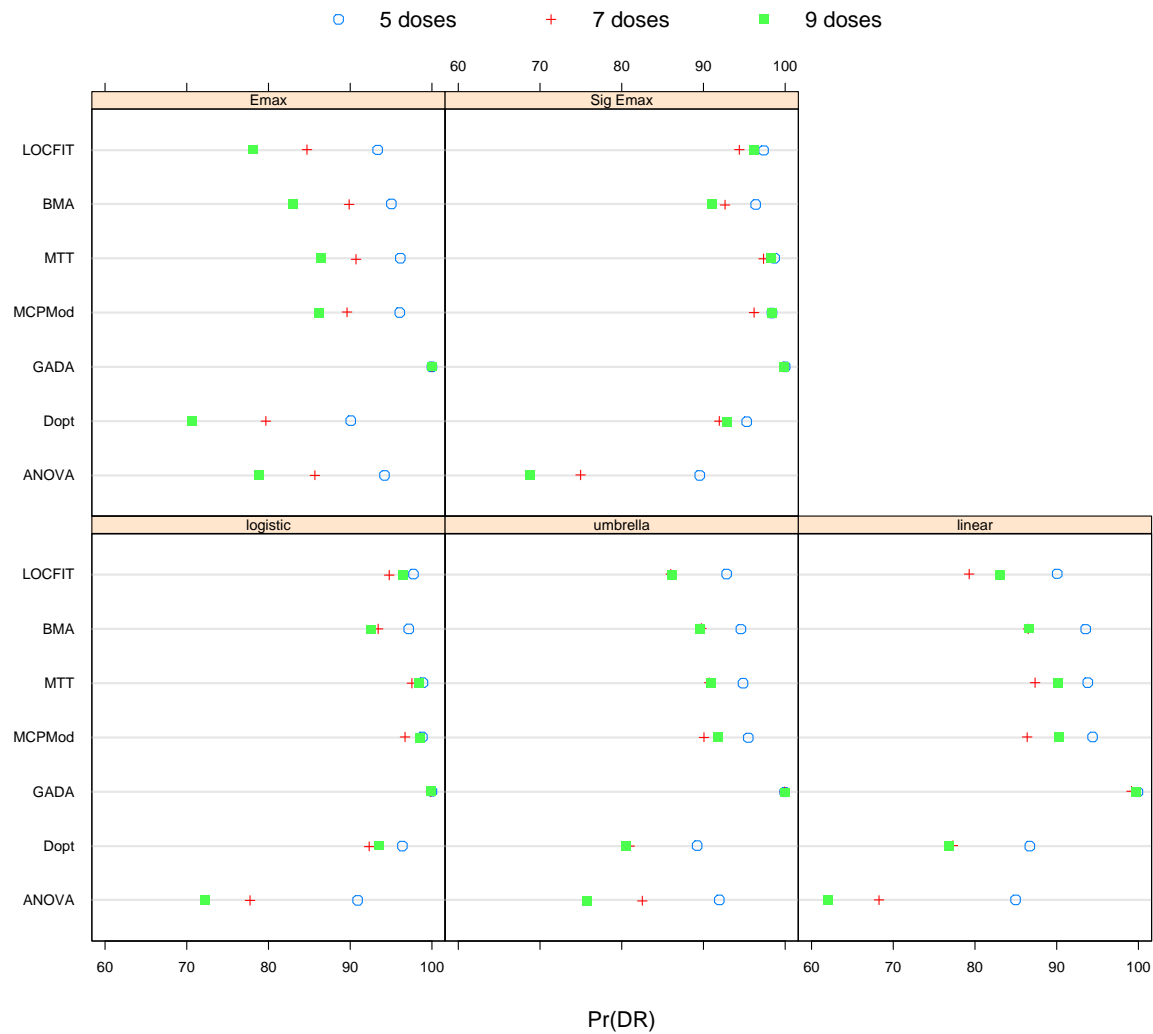
- Probabilities of under-, over-, and correct interval estimation:

$$P^- = P(\hat{d}_{\text{targ}} < d_{\text{min}}), \quad P^+ = P(\hat{d}_{\text{targ}} > d_{\text{min}}),$$

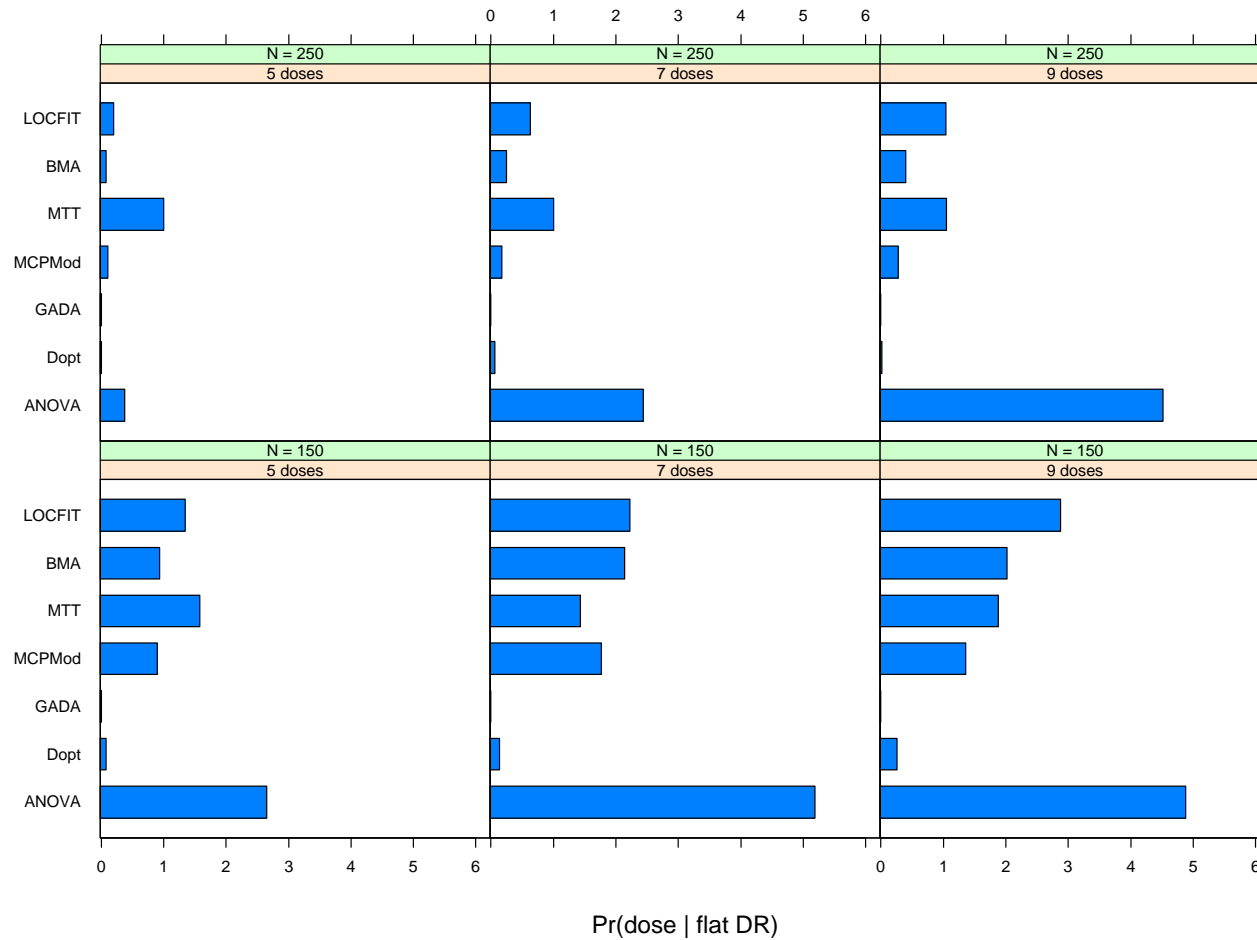
$$P^\circ = 1 - (P^- + P^+)$$

Sample of Simulation Results

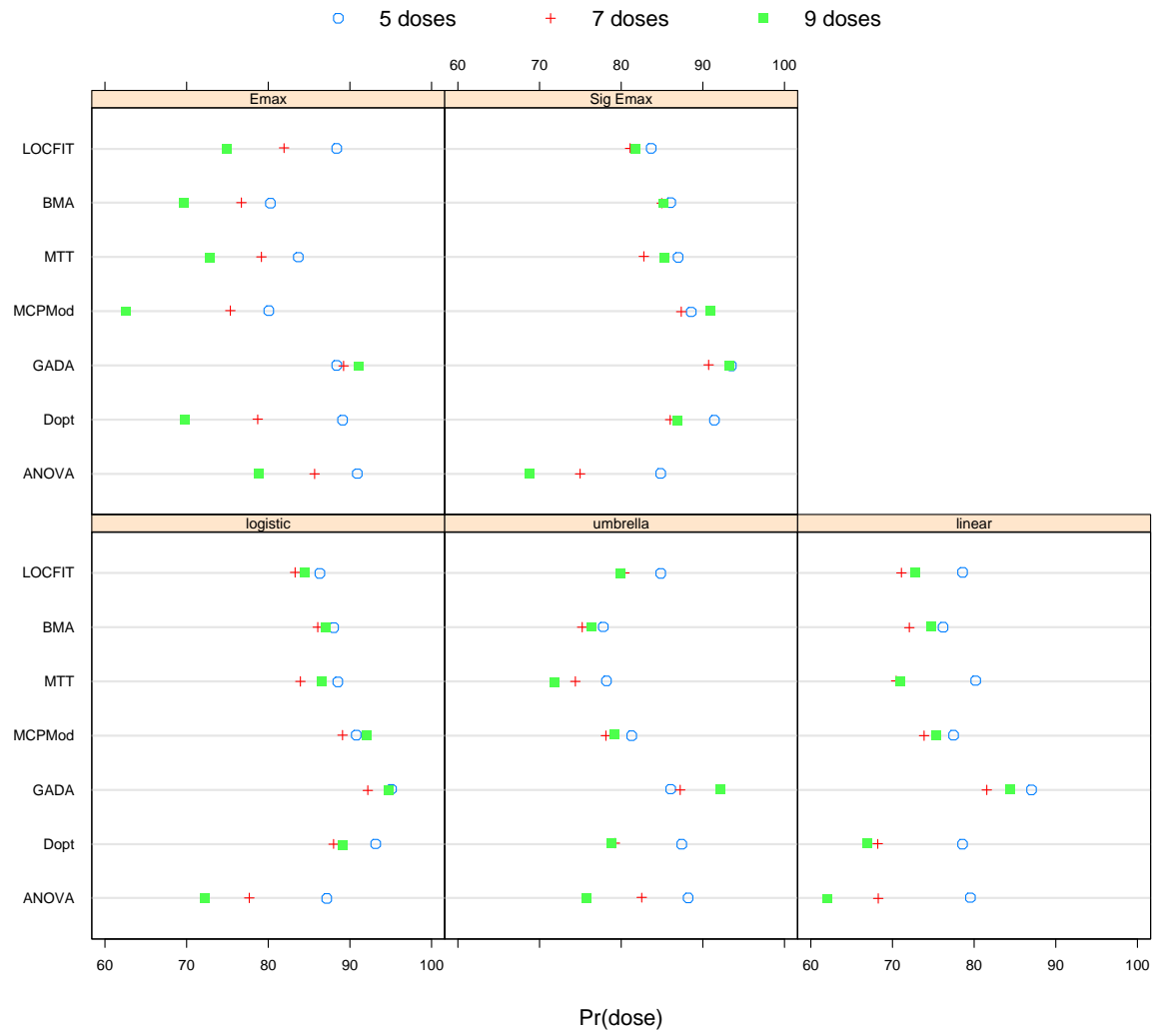
Probability of identifying DR, N = 150



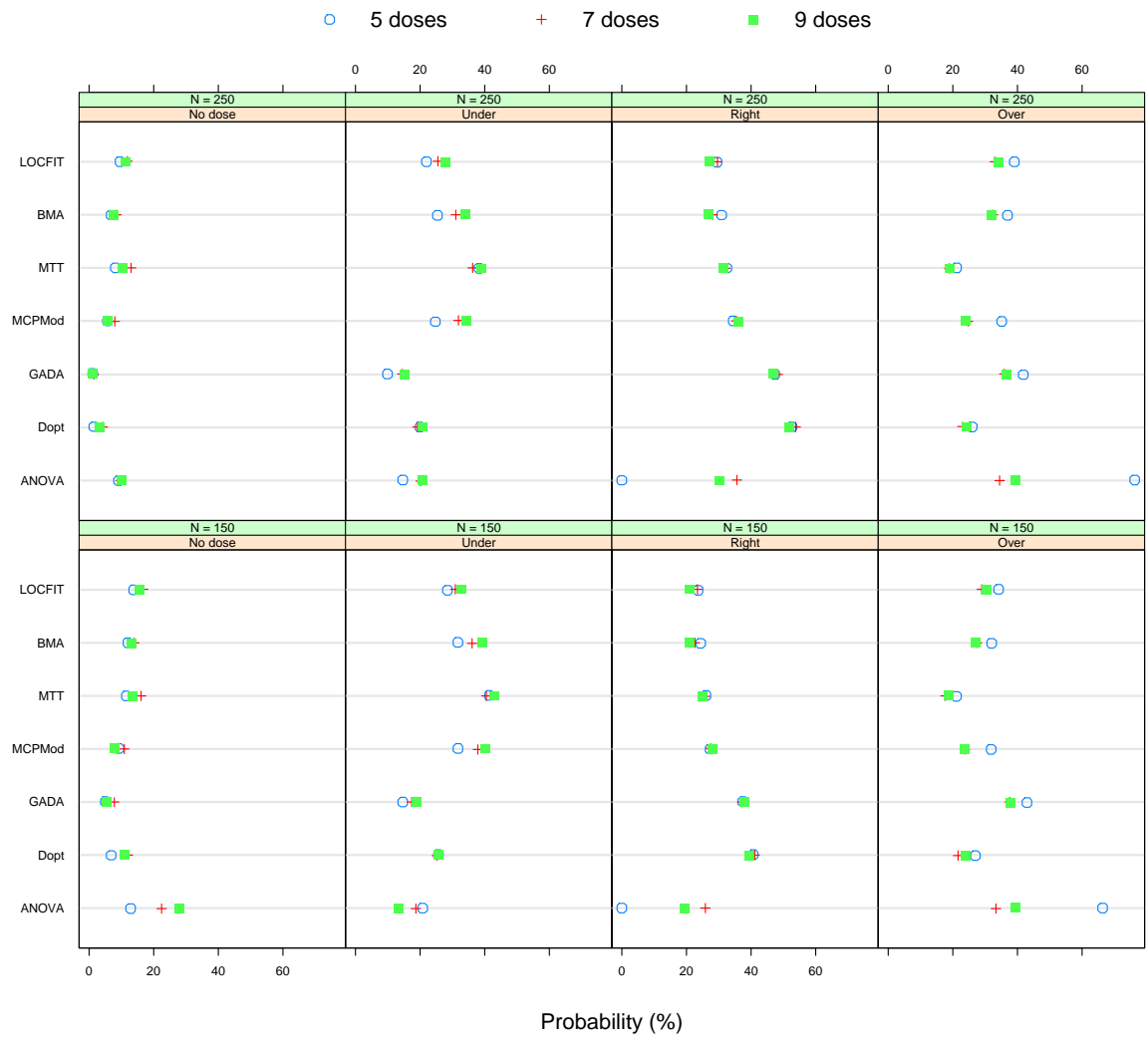
Probability dose selection – flat DR, N = 150



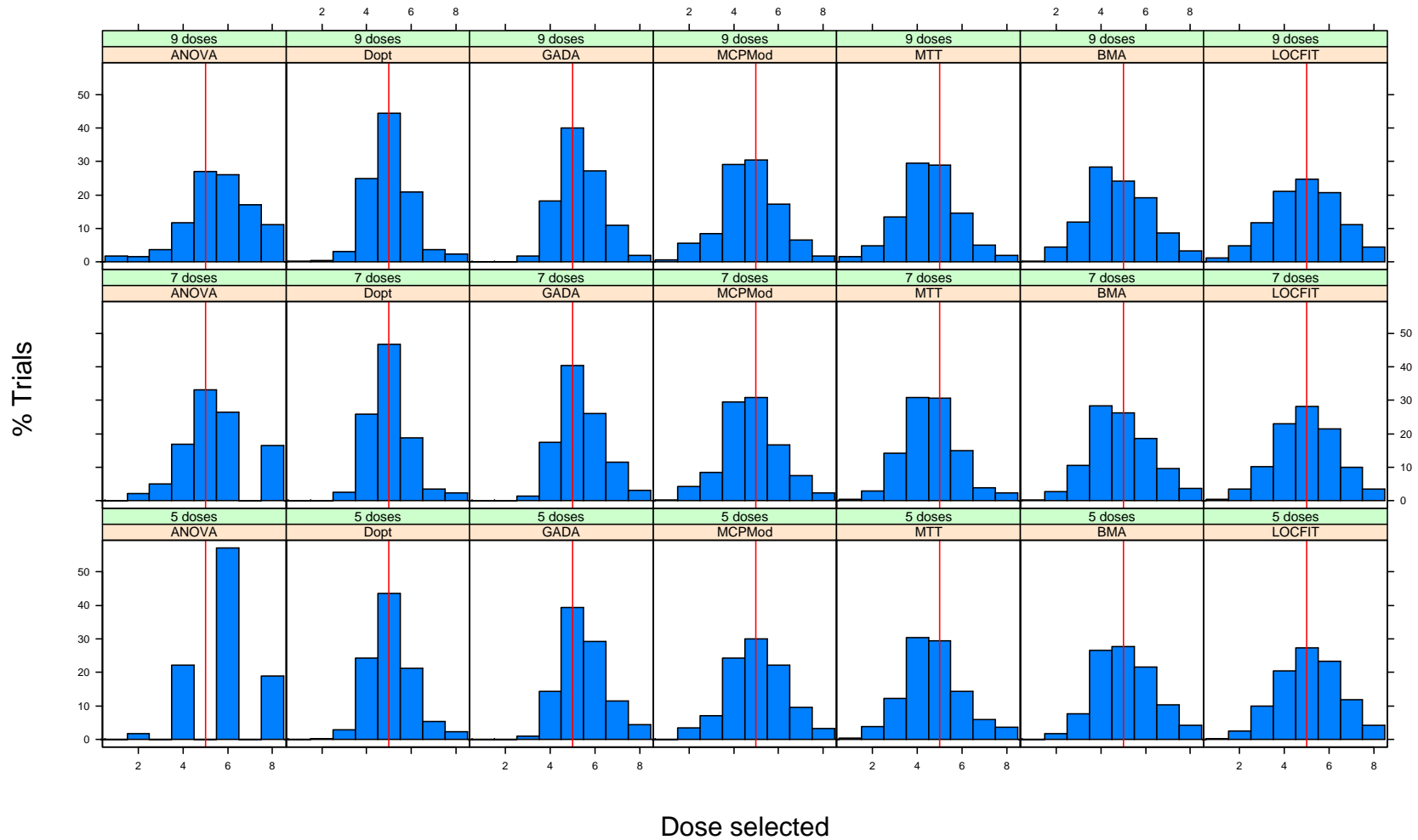
Probability dose selection, N = 150



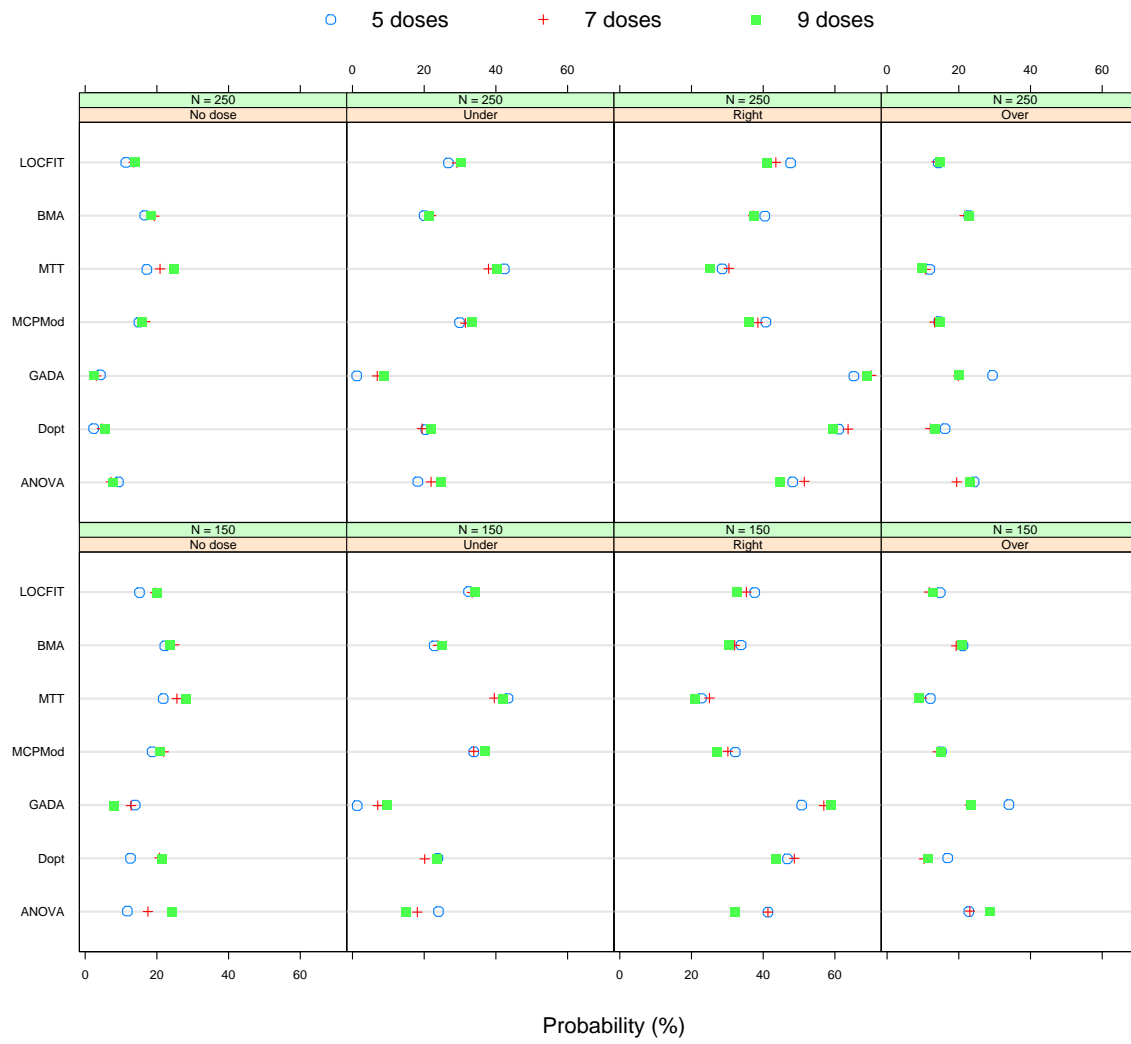
Prob. of interval dose selection, Logistic model



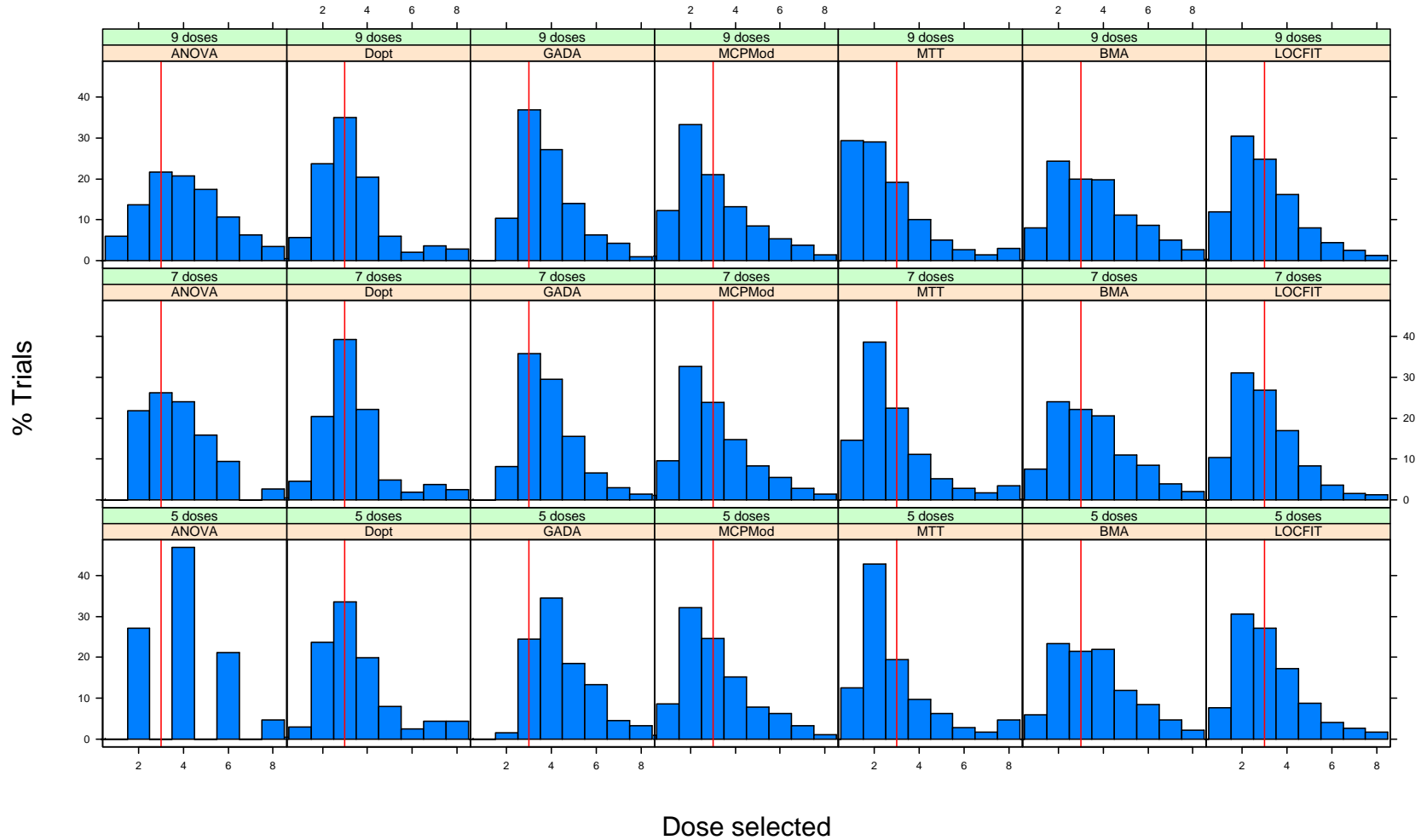
Estimated dose distrib., Logistic model and N = 150



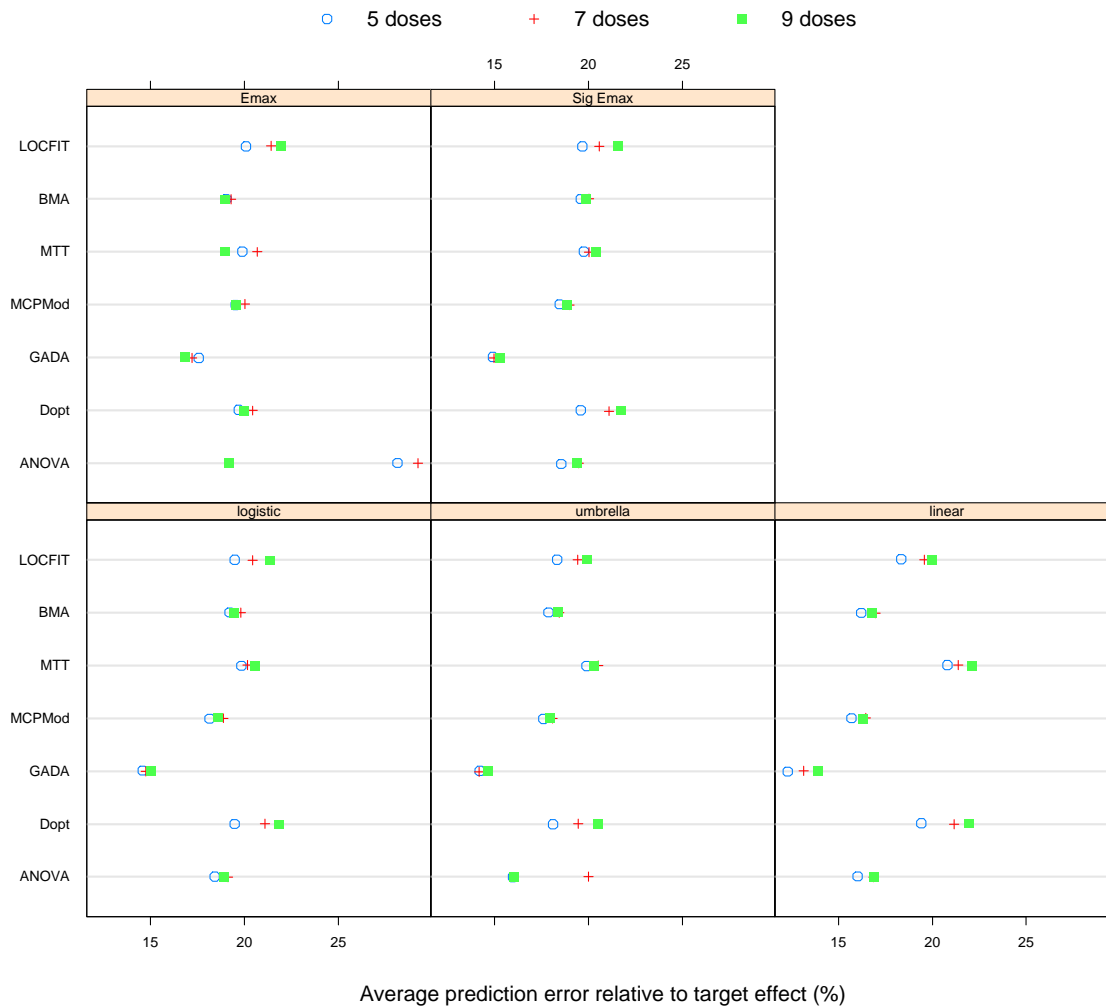
Prob. of interval dose selection, Umbrella model



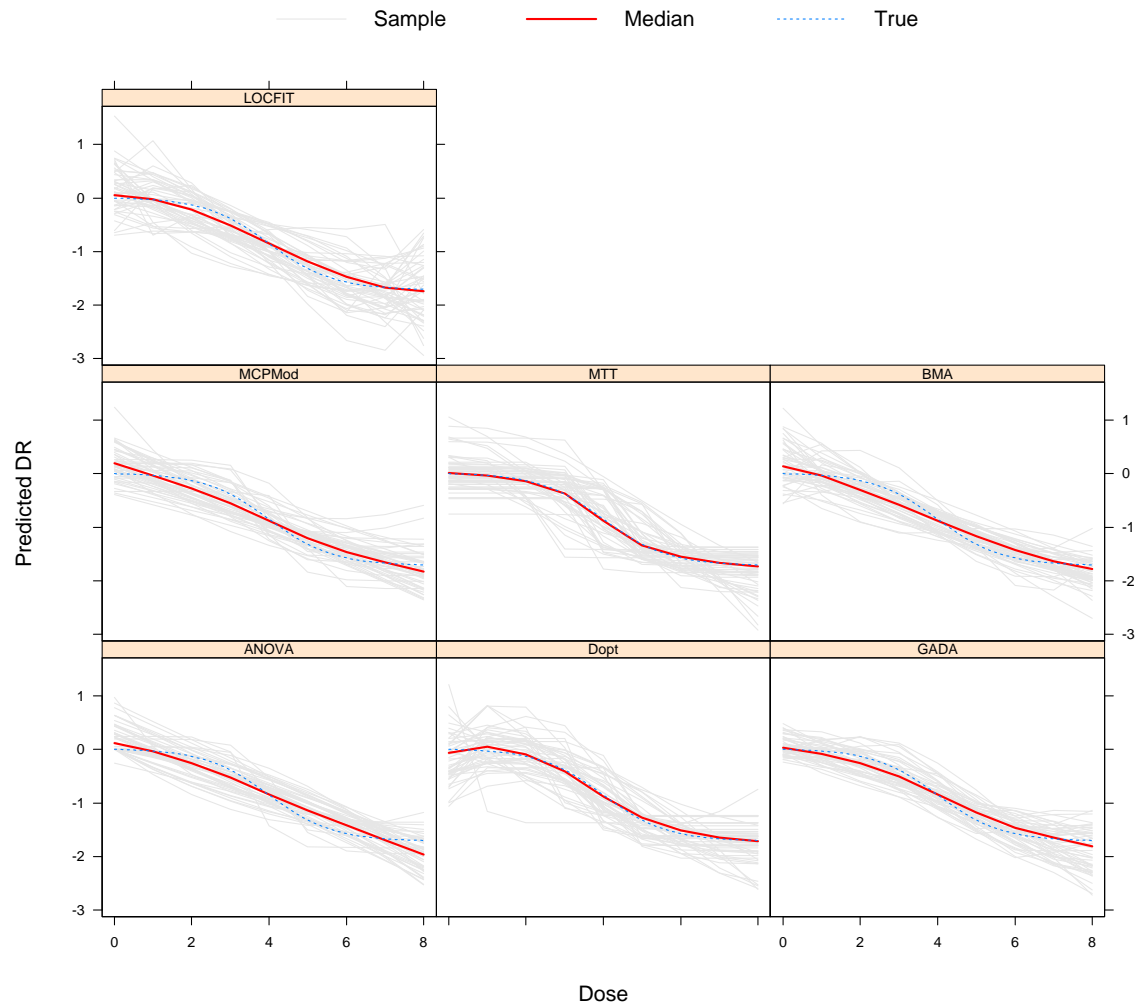
Estimated dose distrib., Umbrella model and N = 150



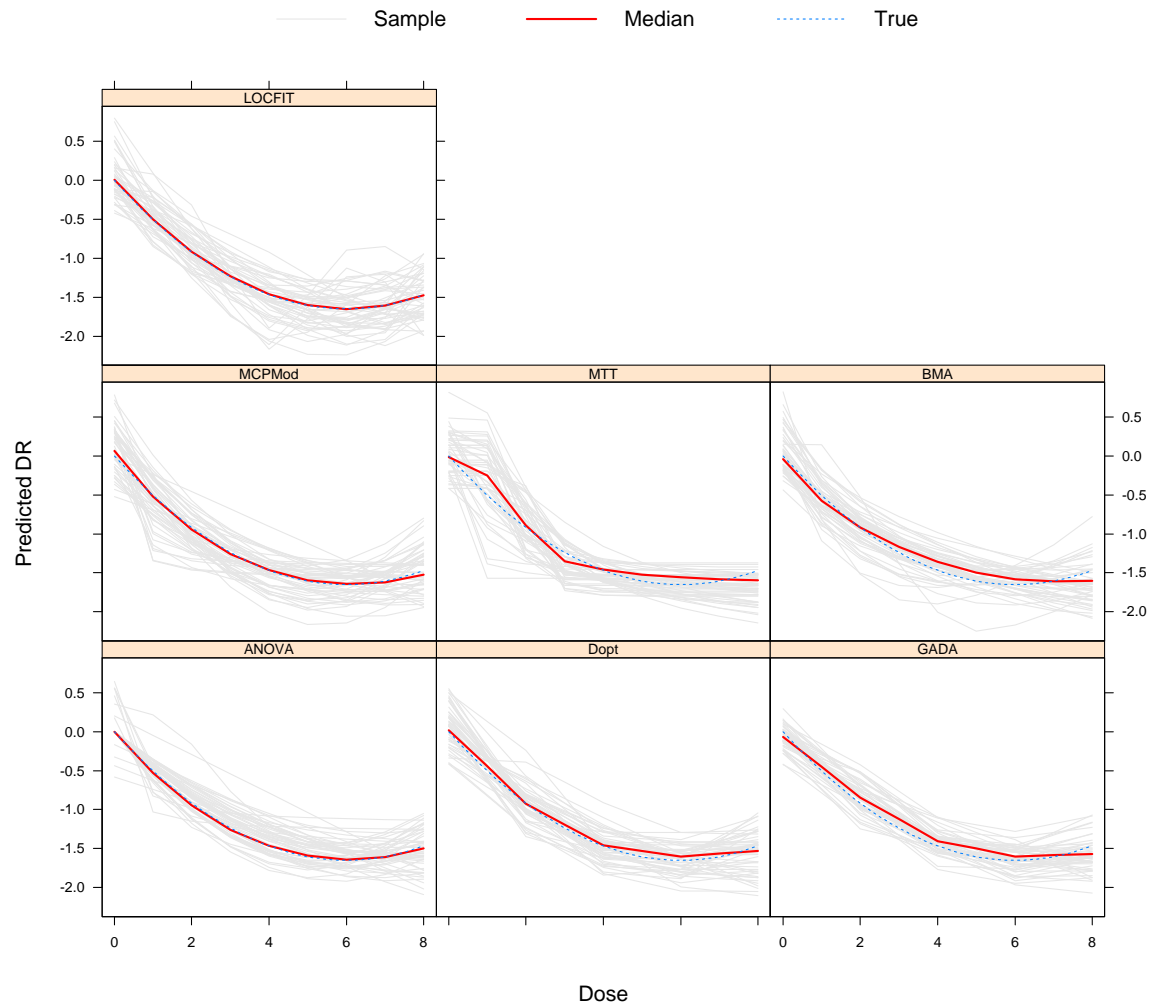
Average prediction error per dose, N = 150



Sample predicted curves: Logistic, 9 doses and N = 150



Sample predicted curves: Umbrella, 5 doses and N = 250



Conclusions

- Detecting DR is considerably easier than estimating it
- Current sample sizes for DF studies, based on power to detect DR, are inappropriate for dose selection and DR estimation
- None of methods had good performance in estimating dose in the correct target interval: maximum observed percentage of correct interval selection – 60% \implies larger N needed
- Adaptive dose-ranging methods (i.e., ADRS) lead to gains in power to detect DR, precision to select target dose, and to estimate DR – greatest potential in the latter two
- GADA had best overall performance, especially on DR estimation

Conclusions (cont.)

- Model-based methods have superior performance compared to methods based on hypothesis testing
- Number of doses larger than 5 does not seem to produce significant gains (provided overall N is fixed) \implies trade-off between more detail about DR and less precision at each dose
- In practice, need to balance gains associated with adaptive dose ranging designs approach against greater methodological and operational complexity

Preliminary Recommendations

- Adaptive, model-based dose-ranging designs should be used routinely in drug development, as they can lead to substantial gains in performance over traditional DF methods
- Sample size calculations for Phase II studies should take into account desired precision of estimated target dose and possibly also estimated DR (current methods are not appropriate)
- When resulting sample size is not feasible, should consider selecting two or three doses for confirmatory phase to increase likelihood of including “correct” dose – adaptive designs could be used in confirmatory phase for greater efficiency (e.g., dropping less efficient doses earlier)

Preliminary Recommendations (cont.)

- Proof-of-concept (PoC) and dose selection should be combined, when feasible, into one seamless trial
- Early stopping rules, for both efficacy and futility, should be used when feasible to allow greater efficiency in adaptive designs – Bayesian methods are particularly well-suited for this purpose
- Trial simulations should be used to determine appropriate sample sizes, as well as for estimating operational characteristics of designs/methods under consideration

References

Bretz, F., Pinheiro, J. and Branson, M. (2005). Combining multiple comparisons and modeling techniques in dose-response studies, *Biometrics* **61**(3): 738–748.

Krams, M., Lees, K. R. and Berry, D. A. (2005). The past is the future: Innovative designs in acute stroke therapy trials, *Stroke* **36**(6): 1341–7.