The Effect of Dosage Errors and Step Selection Method on the Performance of the Up-and-Down Method

Ninth US Army Conference on Applied Statistics
29 to 31 October 2003
Napa Valley Marriott, Napa Valley, CA

Douglas R. Sommerville, PE

Modeling Simulation & Analysis Team

DISCLAIMER: The findings presented in this briefing are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

Edgewood Chemical Biological Center
5183 Blackhawk Road, ATTN: AMSRD-ECB-RT-IM, Bldg. E5951
Aberdeen Proving Ground, Maryland, USA 21010-5424

Email: Douglas.Sommerville@us.army.mil
Phone: (410) 436-4253
FAX: (410) 436-2742

Background

- Up-and-Down (UaD) Method is a common approach for estimating median effective stress
  - Originally developed by the Explosives Research Laboratory, Bruceton, PA in early 1940's, and the Bruceton Method was basis for work of Dixon and Mood (1948)
  - Extensive documentation shows that the method provides an accurate estimate while keeping number of trials to a minimum
  - Numerous testing applications
    - Explosives
    - Metal fatigue
    - Development of Medical Procedures
    - Toxicology and Pharmacology
      - Median effective stress ==> median effective dose/dosage (ED_{50})

Background--Conducting an UaD Bioassay Experiment

- Technique is simple to execute but has inspired extensive discussions on how to analyze the data
  - Trials are conducted one at a time until stopping criteria is met
  - For trial *i*, a subject is given some administered log dose (d_i) and the binary response is recorded (R_i)
  - The log dose for the next trial (d_{i+1}) is dependent on value for R_i
    - If a response occurs (R_i = X or 1): $d_{i+1} = d_i - \Delta d$
    - If a response does not occur (R_i = O or 0): $d_{i+1} = d_i + \Delta d$
    - Δd is the step size
      - In original method, Δd kept fixed in value throughout experiment
- Several modifications of the basic method have been developed
  - Example: various schemes for changing Δd during course of experiment

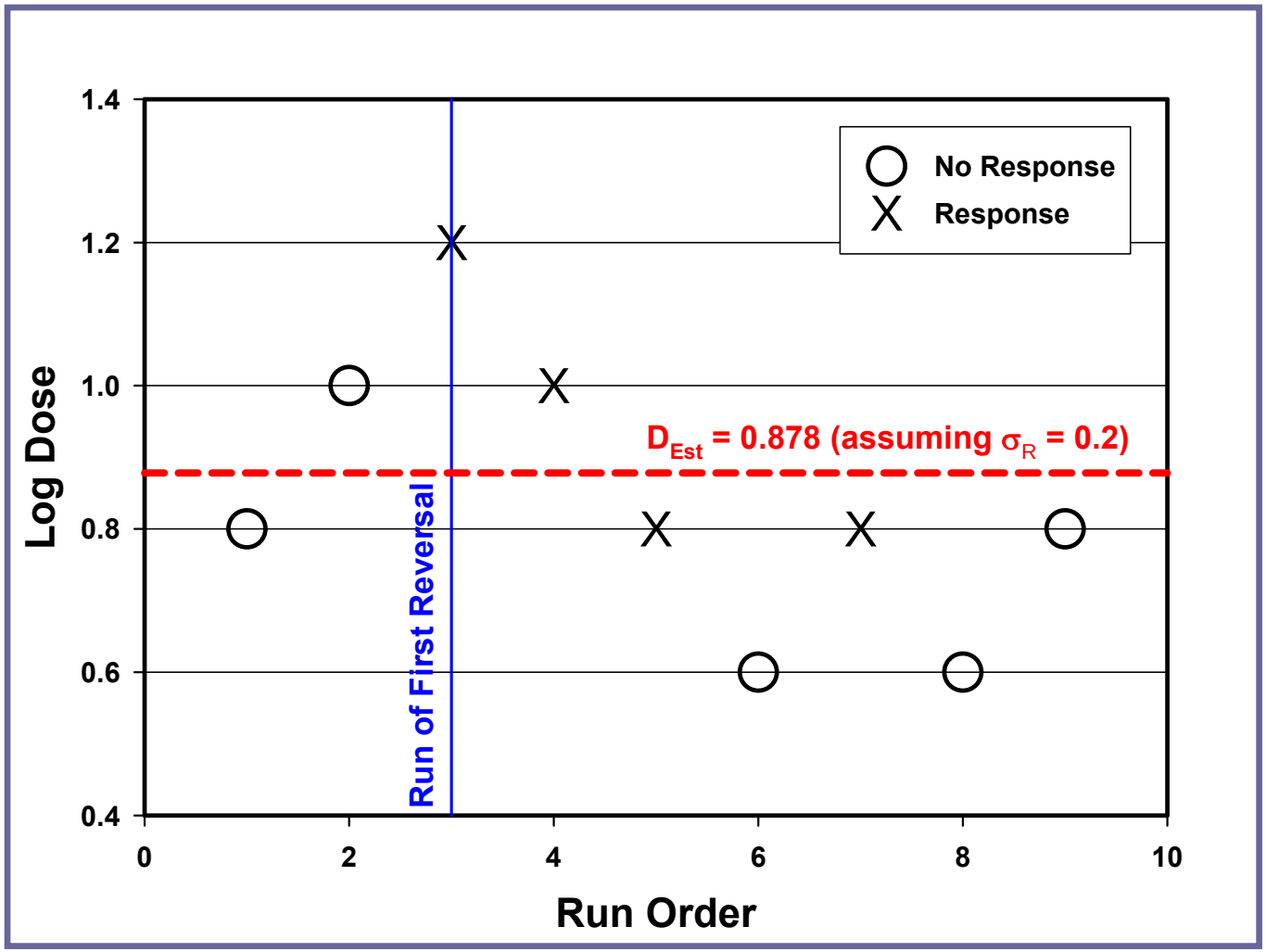
Obtaining an Estimate of the ED_{50}

- Maximum Likelihood Estimation (MLE) used to find estimate, D_{Est}, of the true $\log(ED_{50})$, D_{Tr}
  - Normal distribution is favorite though others used (ex. Little (1974))
    - For bioassays, normal distribution obtained by working with the logarithms of the doses
  - With small sample sizes, D_{Est} is solved for while σ_R is set equal to Δd (Dixon (1965))
- Precision of D_{Est} increases as Δd decreases
- Speed of convergence towards region of D_{Tr} decreases as Δd decreases

Characteristics of UaD Bioassay Experiment

- Optimum Δd range:
  - $(\sigma_R / 2) < \Delta d < (2\sigma_R)$
    - σ_R -- standard deviation of distribution of effective dosages, which needs to be estimated prior to start of experiment
  - Above range represents a trade-off between precision and efficiency
- Standard error (σ_{ED}) of D_{Est} from UaD experiment
  - σ_{ED} approximately equals $\{(\sigma_R)(2 / N)^{(1/2)}\}$ (from Dixon (1965))
    - It is assumed that σ_R equals Δd
    - N is the nominal sample size (versus N_T the total number of trials conducted)
      - Common practice is to base N on N_{FR} (run of first reversal {O to X, or X to O})
      - $N = N_T - N_{FR} + 2$

Example of an Up and Down Experiment



$\Delta d = 0.2$

$N_T = 9$

$N_{FR} = 3$

$N = 8$

UaD Method and Errors in Dosage Administration and Measurement

- What happens when irregular spacing of dosages occurs--found with UaD in inhalation toxicology?
  - Dosage errors (assumed to be normally distributed)
    - Administration error: σ_A
      - Difficulty in precisely generating the target dosage
    - Measurement error: σ_M
      - Error involved in measuring the actual dosage produced by generation device
    - Total error: $\sigma_{Total} = (\sigma_A^2 + \sigma_M^2)^{(1/2)}$
- Vast majority of work/theory on UaD method assumes that both types of errors essentially equal zero.
- How are the accuracy and efficiency of UaD method affected by nonzero σ_A and/or σ_M?

Experimental Method

- Monte Carlo simulations were performed using MINITAB® v. 13
- Three different dosage parameters were tracked
  - d_T -- The target value for the log dosage to be administered
  - d_A -- The actual log dosage administered
  - d_M -- The measured/observed value of the administered log dosage
- Binary responses generated from the following underlying normal distribution of effective dosages
  - D_{Tr} set equal to $\log_{10}(80)$
  - σ_R set equal to (1/15) (or 0.0667)
- Each UaD experiment consisted of 10 trials
  - Wish to examine situations where number of available runs at premium

Binary Response Simulation Procedure

- $d_{T,i}$ is calculated from result of previous trial or chosen as initial guess
  - Dosage step can be taken from either d_T or d_M of previous step
    - Method T (previous target): $d_{T,i+1} = d_{T,i} \pm \Delta d$
    - Method M (previous measured): $d_{T,i+1} = d_{M,i} \pm \Delta d$
- $d_{A,i+1}$ is randomly generated from $d_{T,i+1}$
  - Distribution: $N(d_{T,i+1}, \sigma_A^2)$
- $d_{M,i+1}$ is randomly generated from $d_{A,i+1}$
  - Distribution: $N(d_{A,i+1}, \sigma_M^2)$
- R_{i+1} generated from binomial distribution
  - Probability of response ($R_{i+1} = X$) equals the area under $N(D_{Tr}, \sigma_R^2)$ between the limits of $-\infty$ to $d_{A,i+1}$

Calculation of ED_{50} for Individual Experiment

- A MLE solution for D_{Est} was obtained numerically using the Newton-Raphson method and a probit link function
  - A MINITAB® macro was written to perform the calculations
  - d_M values were used in the calculations for D_{Est}
  - As default for MLE calculations, σ_R was set equal to the Δd value used for the experiment
    - However, sometimes large values of σ_A and/or σ_M will produce situations where no solution was possible (using MINITAB®) unless σ_R was increased
    - Calculation artifacts would occur due to arbitrary cutoff of normit values by MINITAB®
      - Values below -7 were rounded up to -7, and values above 7 were rounded down to 7.
- Several other MLE related parameters were calculated
  - Wald Test statistic
  - Score Test statistic
  - Likelihood-Ratio Test statistic
  - Log Likelihood of final MLE fit

Factors Investigated

Factor	Description	Definition
A	Step Size	$(\Delta d / \sigma_R)$
B_A	Administration Error	(σ_A / σ_R)
B_M	Measurement Error	(σ_M / σ_R)
C	Location Initial Log Dosage	$([d_{T,1} \pm D_{Tr}] / \sigma_R)$
D	Basis for Calculation of Next Dosage	Method T: $d_{T,i+1} = d_{T,i} \pm \Delta d$ Method M: $d_{T,i+1} = d_{M,i} \pm \Delta d$

Factor Level Values

Factor	Level Values								
	1	2	3	4	5	6	7	8	9
A	0.500	0.707	1.000	1.414	2.000	2.828	4.000	5.657	
B_A	0.000	0.125	0.250	0.500	1.000	2.000			
B_M	0.000	0.125	0.250	0.500	1.000	2.000			
C	-3.000	-2.250	-1.500	-0.750	0.000	0.750	1.500	2.250	3.000
D	T	M							

Values in shaded boxes were not used with Method M

50 experiments per coordinate point

Method	T	M	Total
# Coordinates	2592	1440	4032
# Experiments	129600	72000	201600

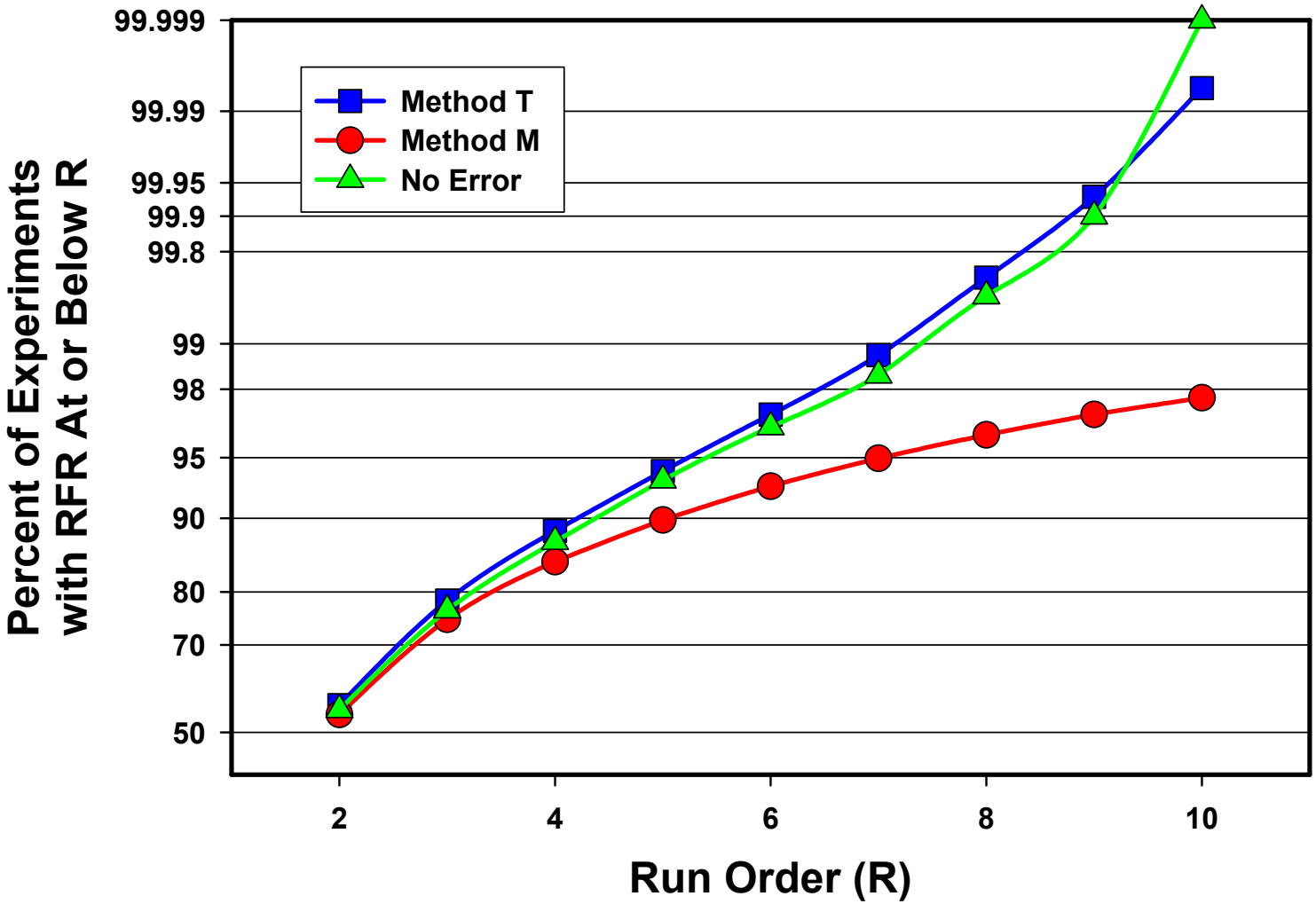
Results

- Efficiency
  - Number of experiments that "failed"
    - Experiments that consists of either all responses or all non-responses
  - Run of first reversal (RFR)
    - Indication of how quickly region of the median has been reached
  - Number of responses in an experiment
    - Indication of how well region of the median has been covered
- Precision
  - Standard error (SE) of D_{est}
    - Expressed as multiples of σ_R
  - Mean square error (MSE) of D_{est}
    - Expressed as multiples of σ_R^2

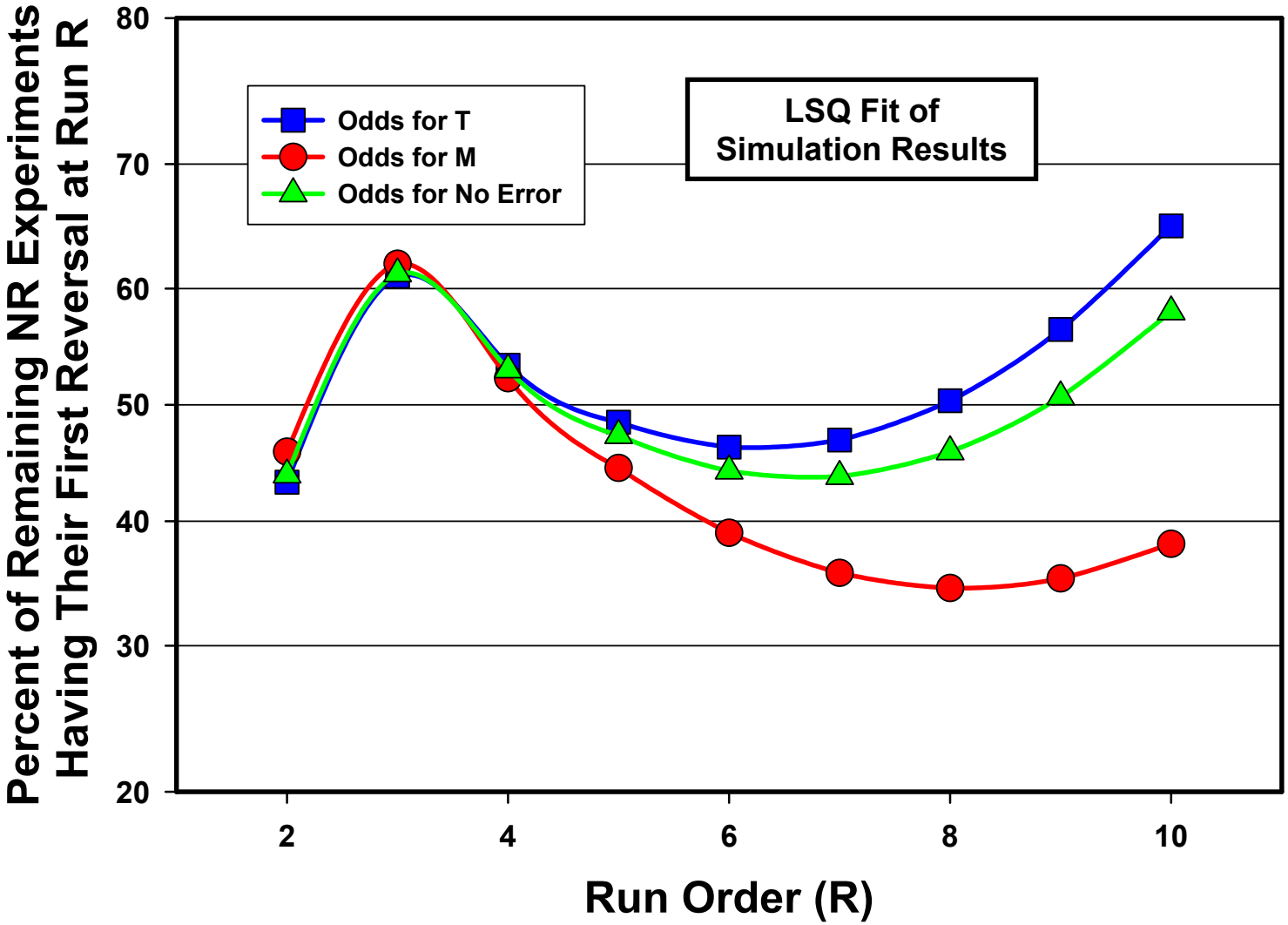
Efficiency Results

- Obtaining Early Run of First Reversal (RFR)
  - Factors in order of decreasing importance: A, D, CC, B_A, B_M and AA
  - Many statistically significant interactions also exist
  - Target dosage basis (D) more important than dosage errors (B_A and B_M)
- Success rate (T vs. M): Method T slightly better
  - Success rate of Method T closely parallels that of experiments executed with $B_A = B_M = 0$ as a function of Run of First Reversal (RFR)

Run of First Reversal Breakdown by Method and Run Order



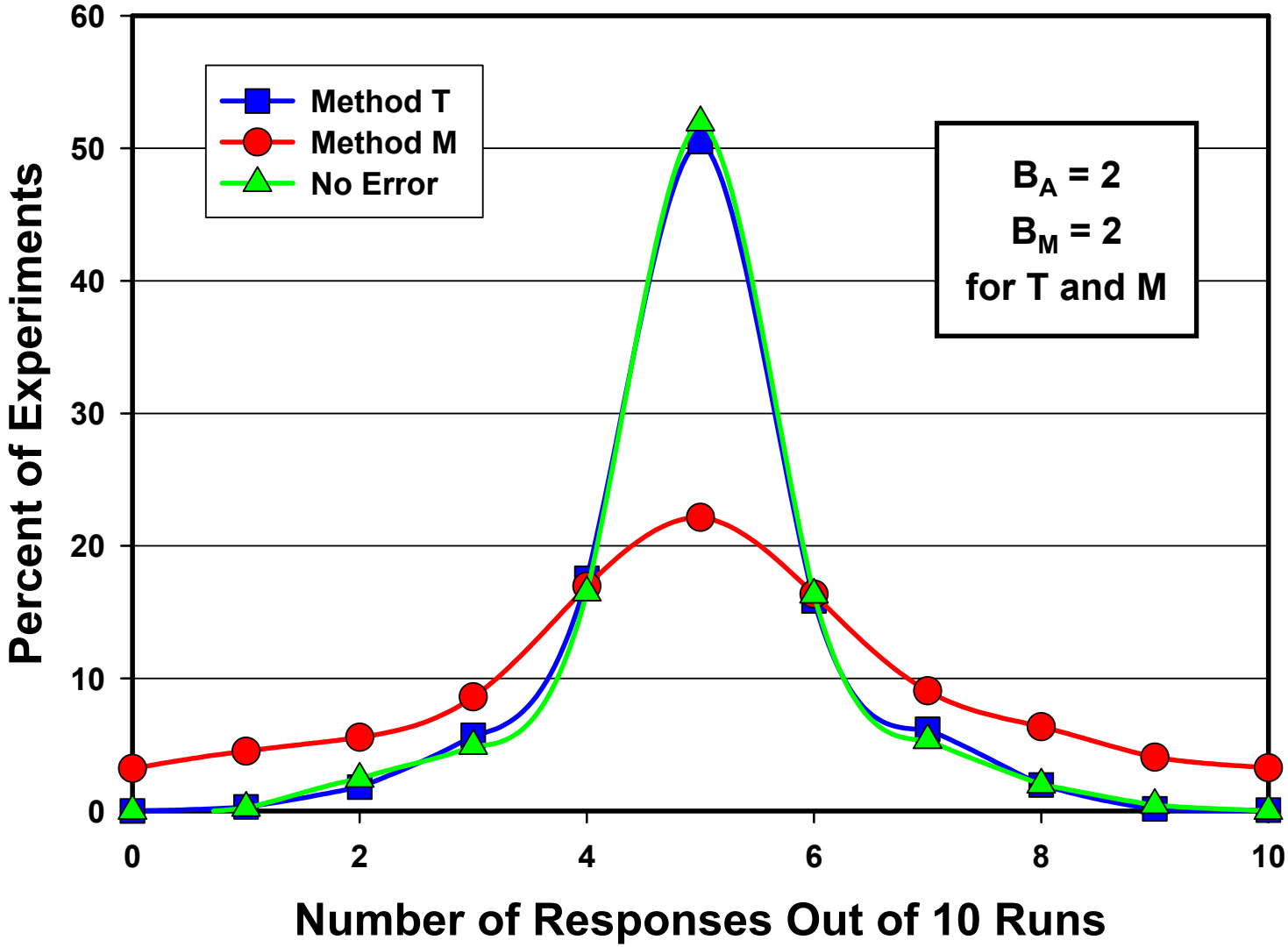
Chance of Non-Reversed (NR) Experiment having a Reversal



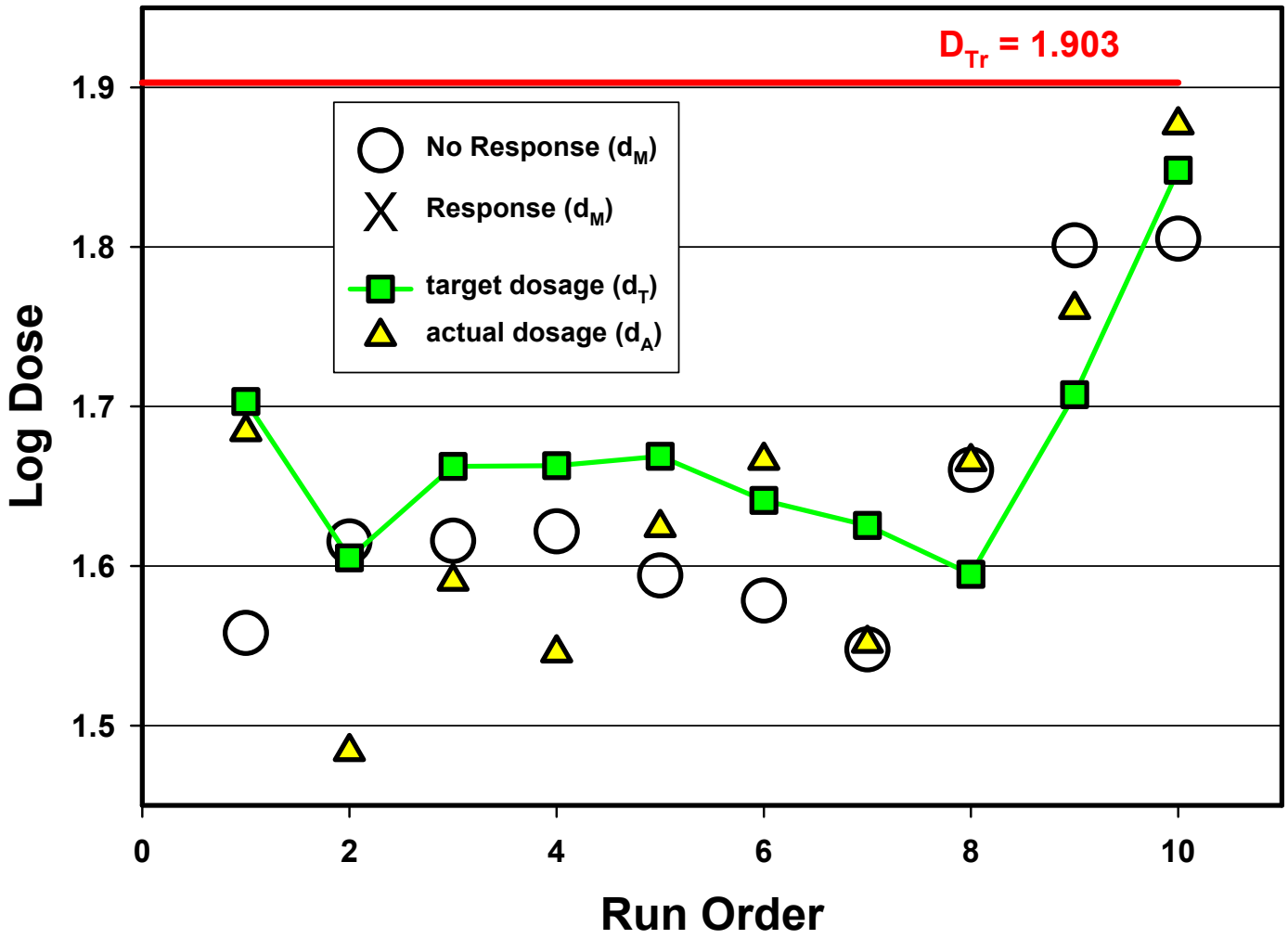
Efficiency Results (Cont.)

- Distribution of Number of Responses
  - Results for Method T (with any dosage errors present) mirrors than of UaD operated with no dosage errors
  - Distribution for Method M becomes shorter and broader as the dosage error increases
- Method T is more robust in recovering from "dead ends" than Method M
  - Target dosages in Method M are more readily influenced by fluctuations resulting from administration (σ_A) and measurement (σ_M) errors
    - Smaller step sizes can be easily overwhelmed
    - Advice on recommended step size ($\Delta d = \sigma_R$ or $A = 1$) may need to be revised

Distribution of Number of Response



Example of a Wandering Experiment with Method M



Expt 64040M

Method M

$\Delta d = (0.047)$

$\sigma_R = (1/15)$

$A = 0.707$

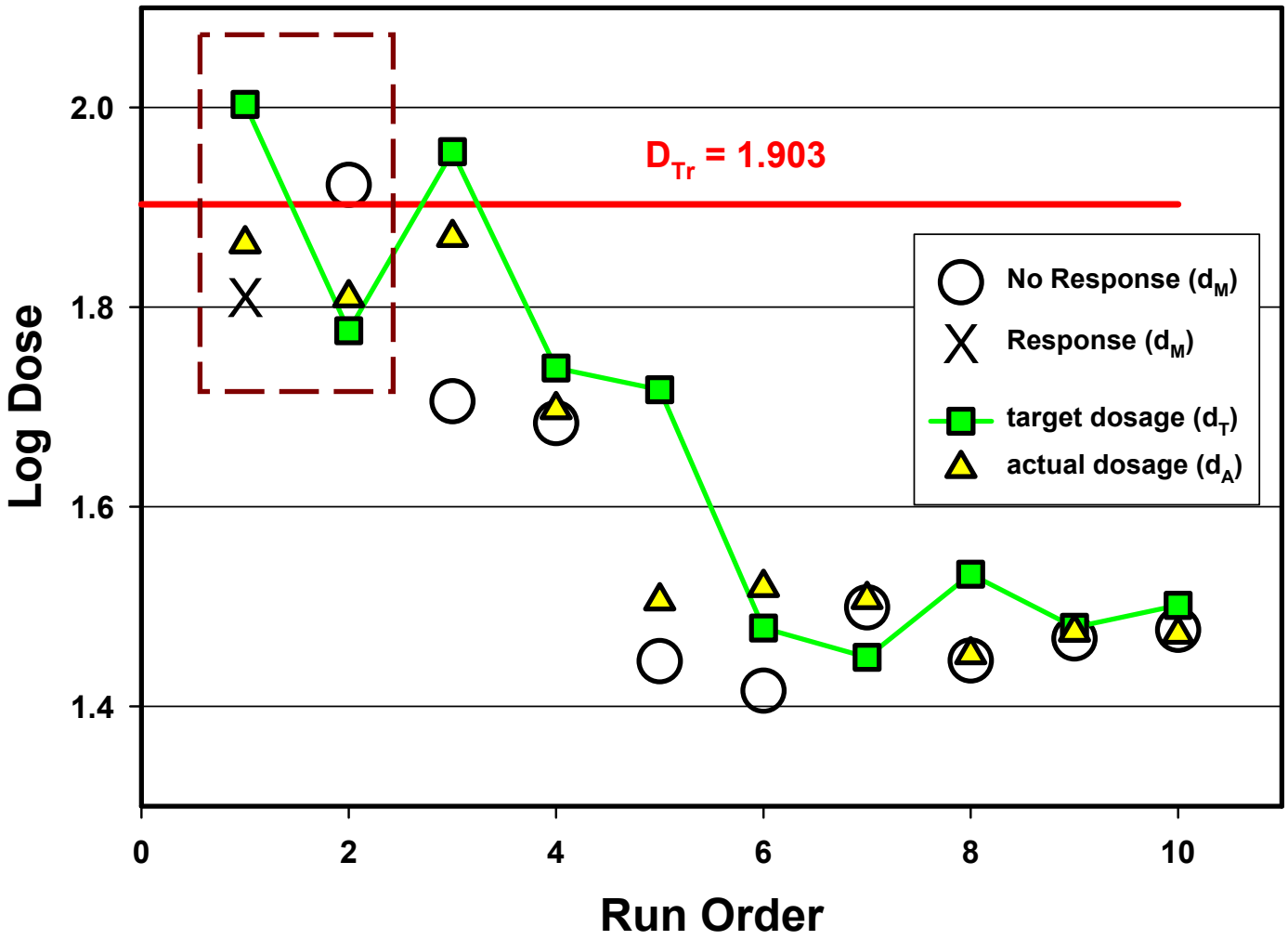
$B_A = 1$

$B_M = 1$

$C = -3$

No Estimate of D_{Tr} Possible

Example of a Negative Step Direction with Method M



Expt 28159M

Method M

$\Delta d = (1/30)$

$\sigma_R = (1/15)$

$A = 0.5$

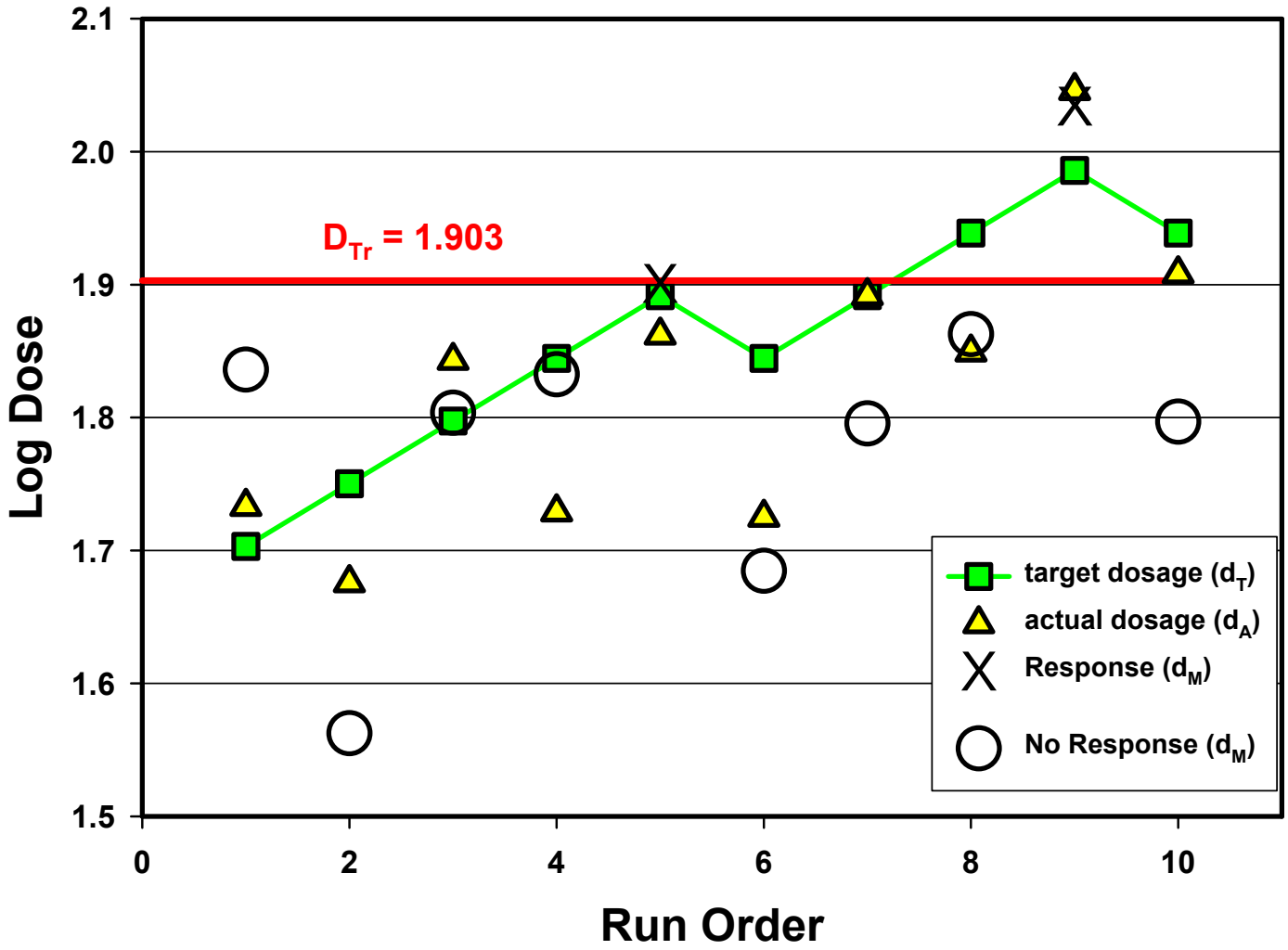
$B_A = 1$

$B_M = 1$

$C = 1$

D_{Est} = 1.866

Example of an Experiment with Method T



Expt 64013T

Method T

$\Delta d = (0.047)$

$\sigma_R = (1/15)$

$A = 0.707$

$B_A = 1$

$B_M = 1$

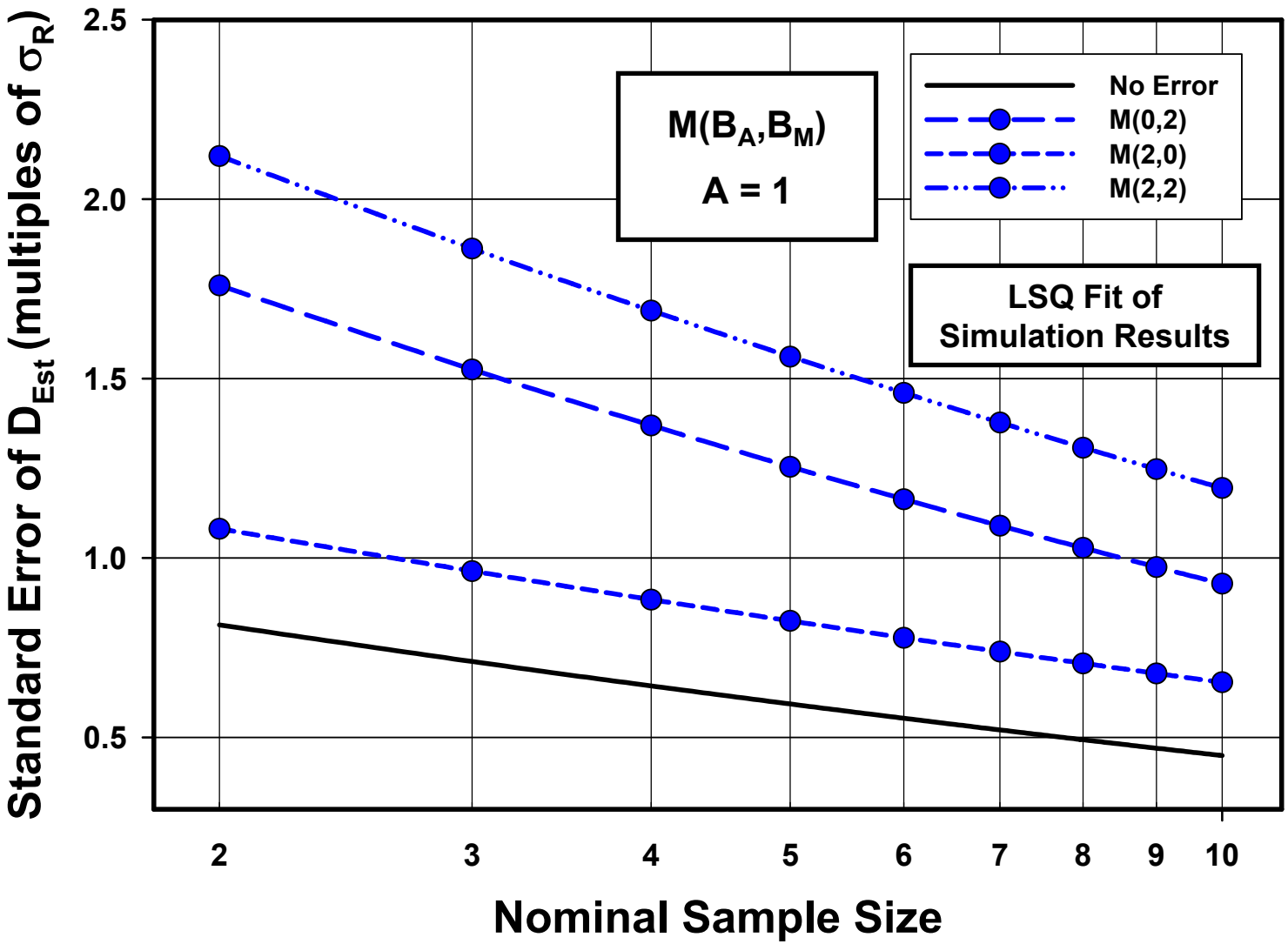
$C = -3$

$D_{Est} = 1.901$

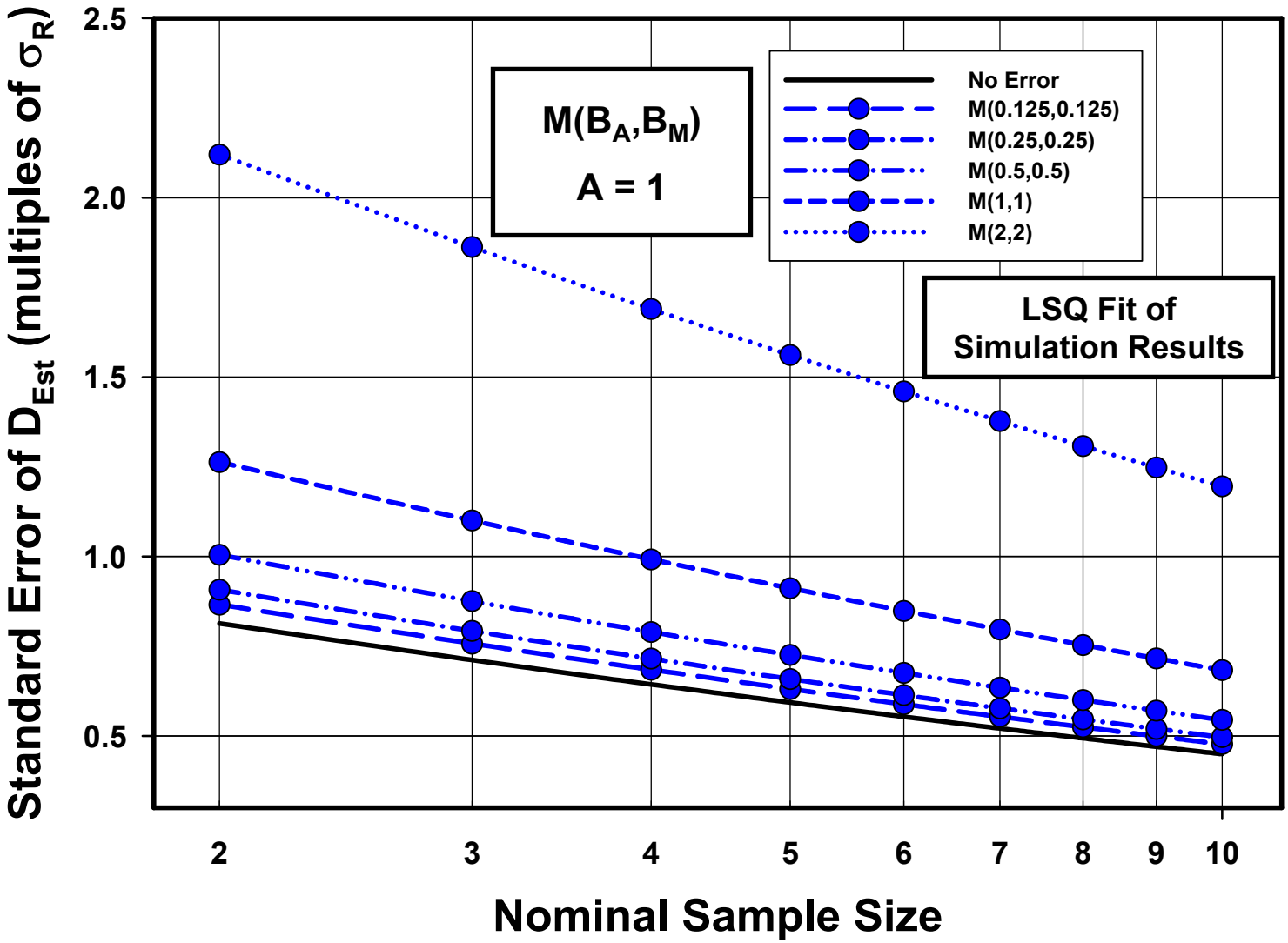
Precision Results

- Additional factor introduced--N (nominal sample size)
  - Accomplished by dividing up the 50 runs per coordinate point into five equal size groups having N_T values of 6, 7, 8, 9 and 10, respectively
    - Knowing RFR for each experiment, N was calculated from N_T and N_{FR}
      - Values of N from 2 to 10 obtained
- Factors in order of decreasing importance: logN, B_M, A, B_A, B_MB_M and B_AB_A
  - Many statistically significant interactions also exist
  - D is only significant when involved in an interaction (ex. AD and B_MD)
  - Dosage errors (B_A and B_M) more important than agent dosage basis (D)
    - This is in contrast to the reverse being true for experiment success rate
- Standard error for D_{Est} is smaller using Method T than with Method M

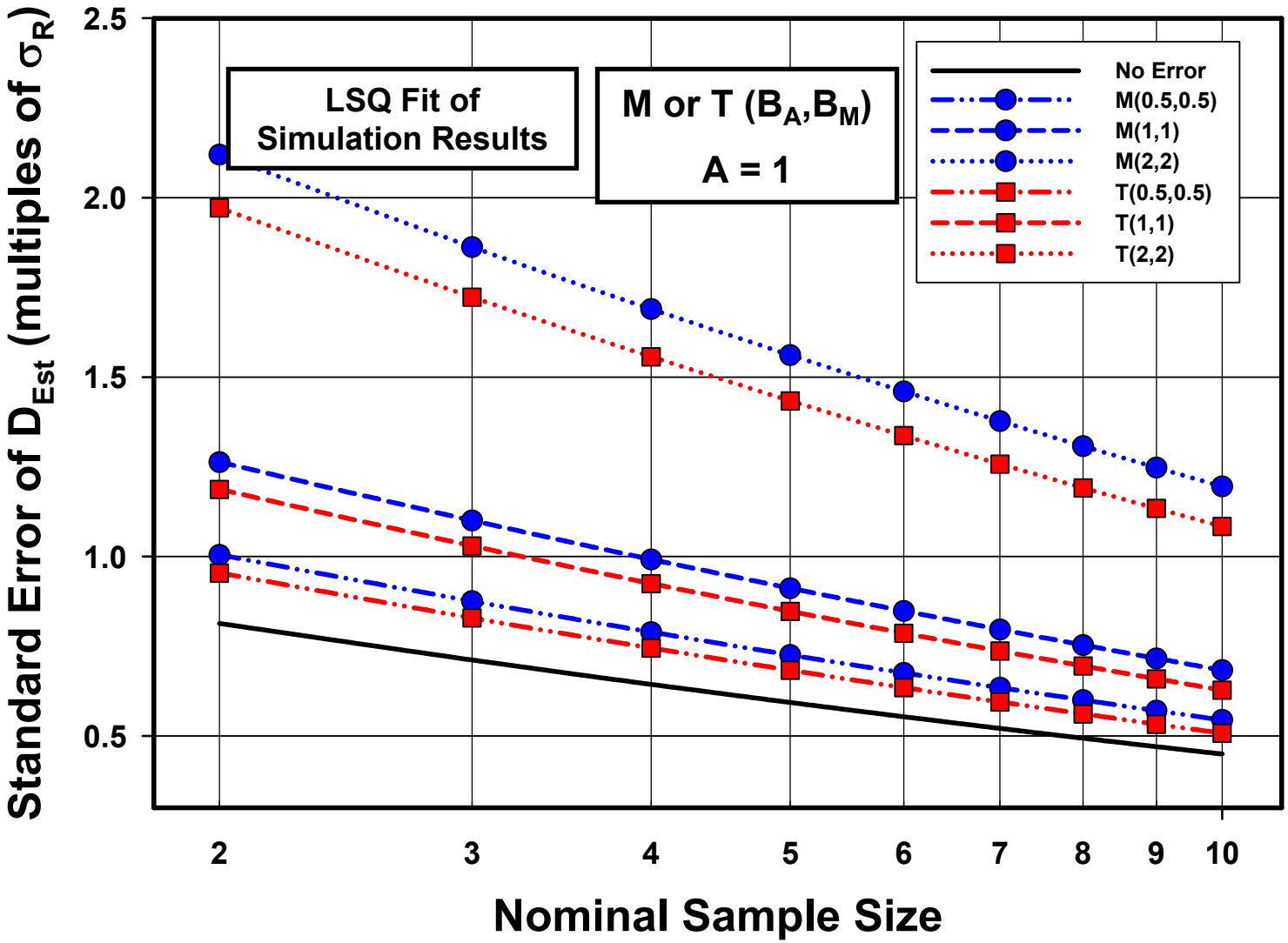
Standard Error of D_{Est} for Method M as a Function of N, B_A and B_M



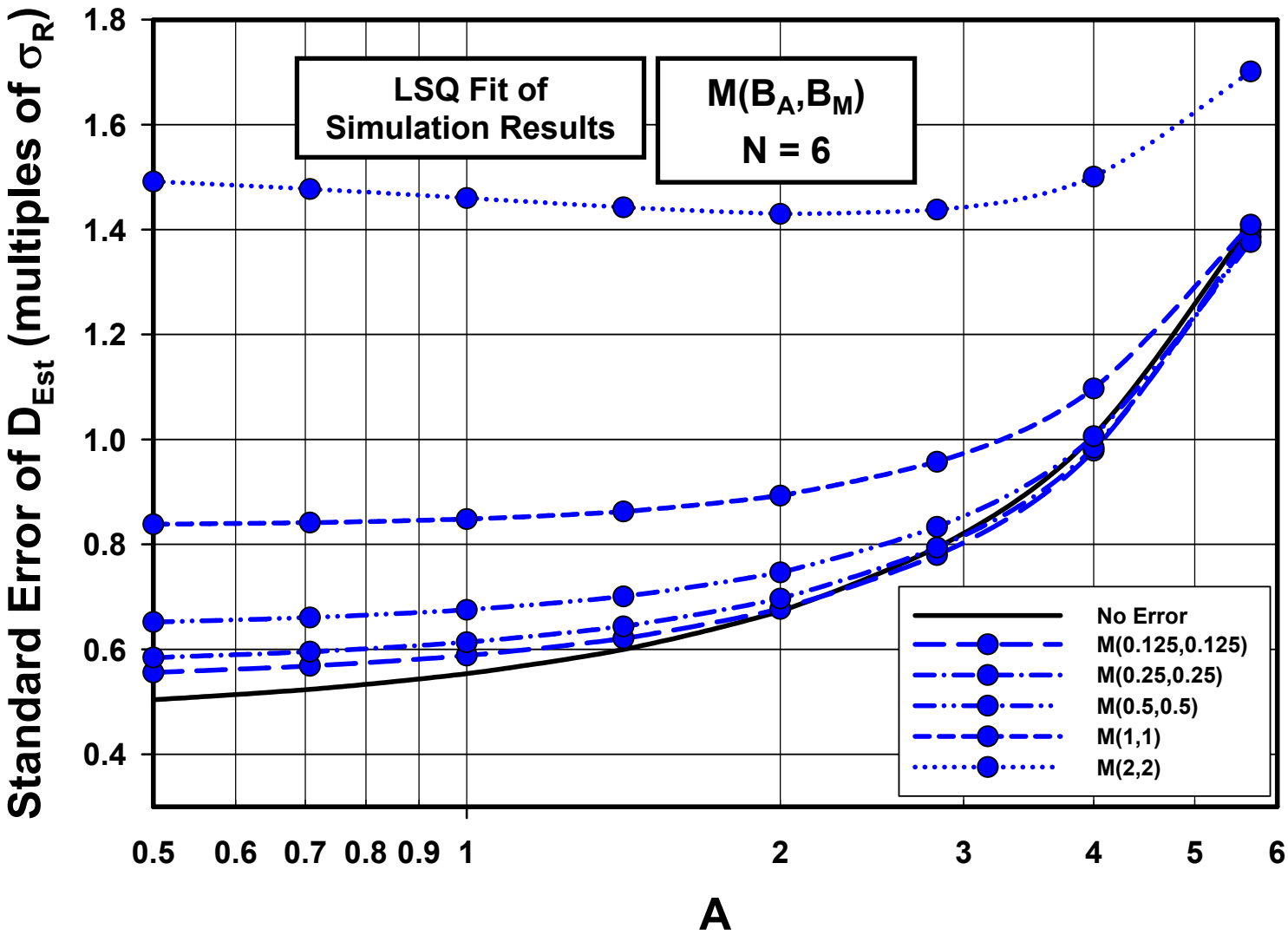
Standard Error of D_{Est} for Method M as a Function of N and (B_A + B_M)



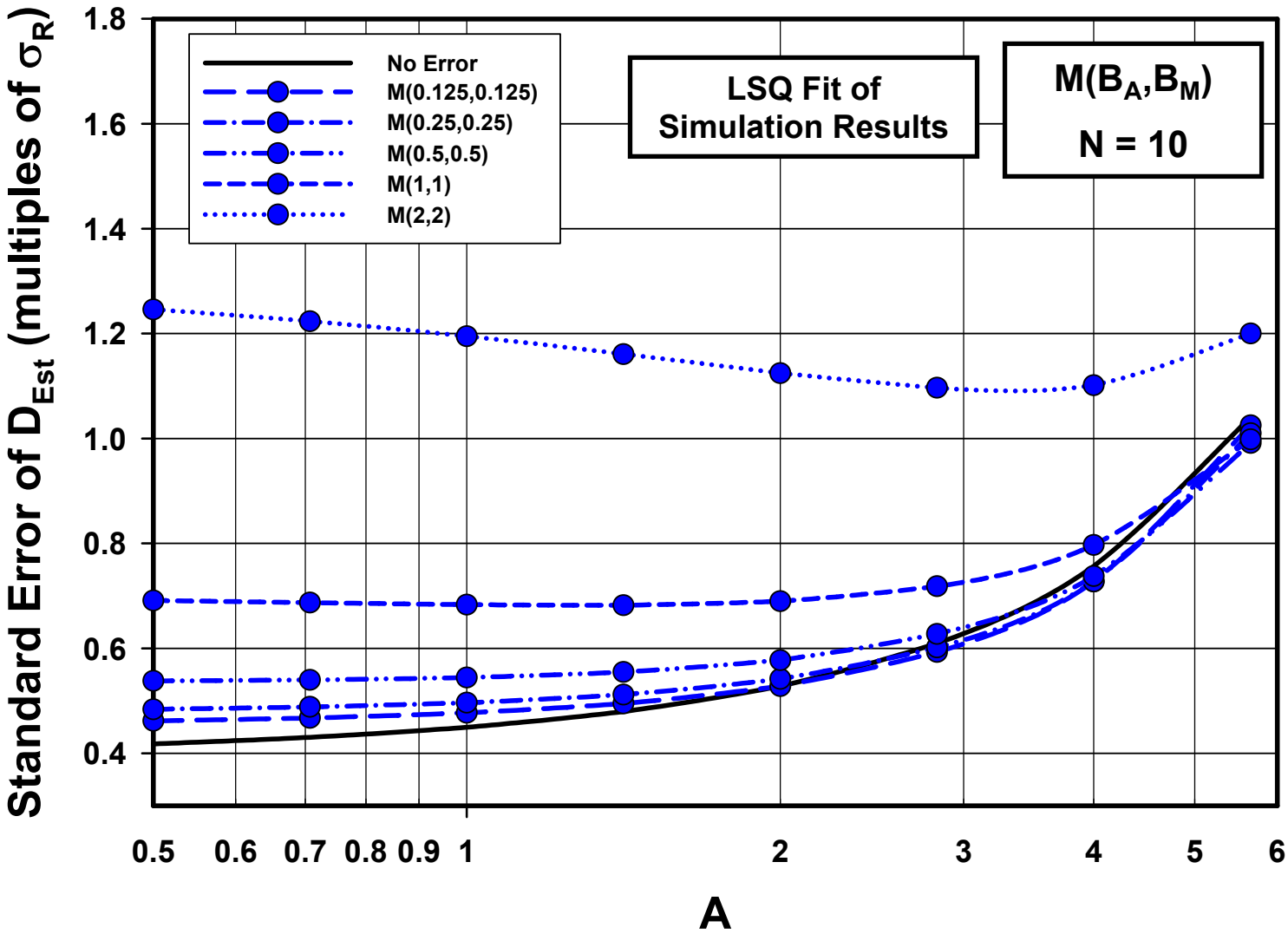
Standard Error of D_{Est} for Methods M and T as a Function of N and (B_A + B_M)



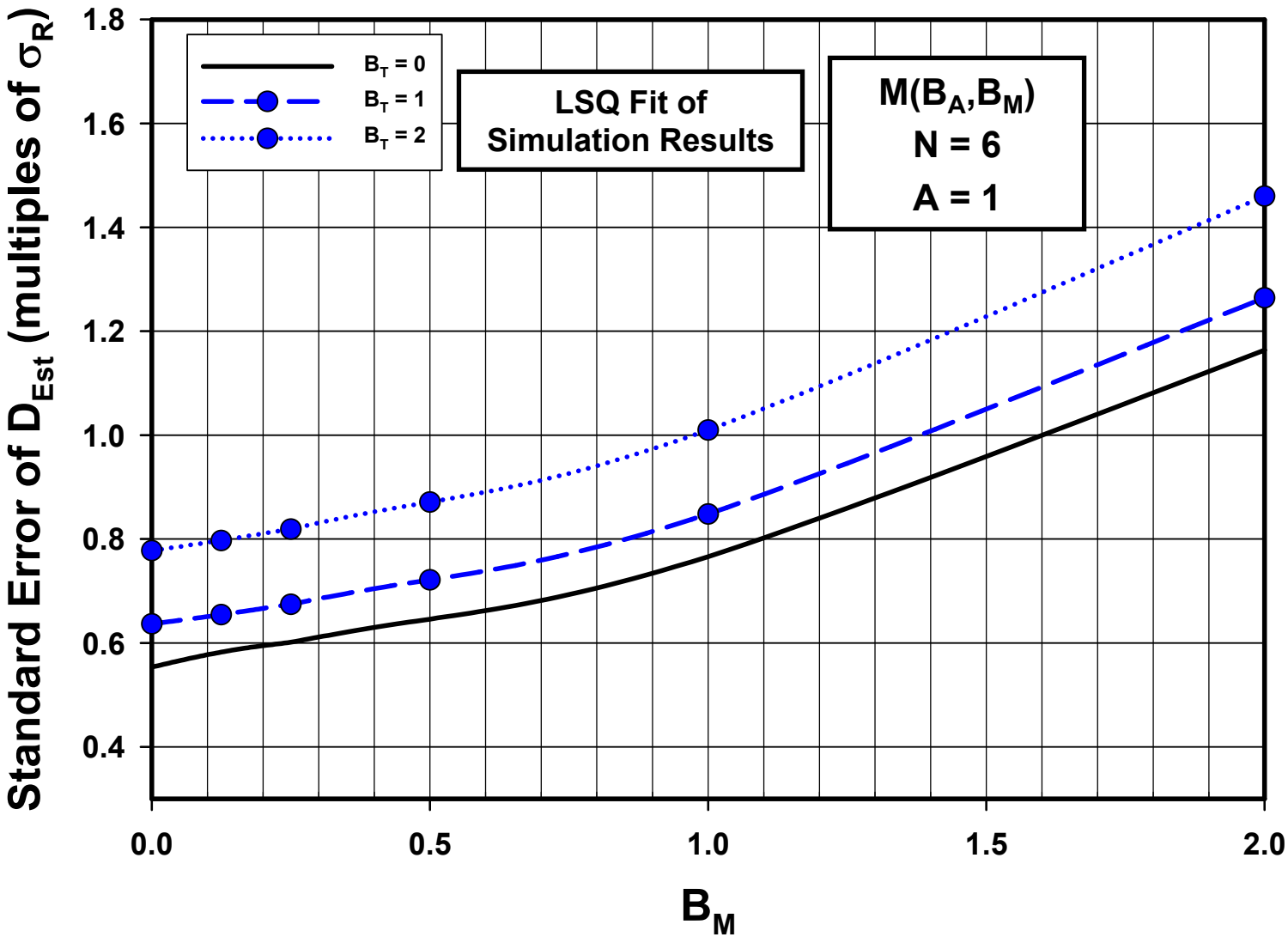
Standard Error of D_{Est} for Method M as a Function of A and (B_A+ B_M)



Standard Error of D_{Est} for Method M as a Function of A and (B_A+ B_M)



Standard Error of D_{Est} for Method M as a Function of B_A and B_M



Summary

- Witn non-zero dosage errors (B_A & B_M), UaD Method is robust with respect to efficiency & precision in the following parameter space
  - $-3 \leq C$ (initial dosage) ≤ 3
  - Neither B_A and/or B_M exceeds roughly 0.25 to 0.5
- When selecting next target dosage (Factor D), Method T is better than Method M overall
  - Difference between methods becomes more pronounced as:
    - A decreases
    - B_A and/or B_M increases
    - |C| increases
  - Method T is more able to resist effects of large values of B_A & B_M

Summary (Cont.)

- Method T produces a distribution of number of responses and Run of First Reversal that is essentially identical to what is produced by traditional UaD (no dosage errors)
- Method T produces more precise D_{Est} values than Method M
  - However, dosage errors (B_A & B_M) have larger impact on precision than Factor D (Methods T and M) (the reverse is true for efficiency)

Conclusions

- Dosage administration and measurements errors do have an effect on the efficiency and precision of the Up-and-Down Method
- The effect of dosage errors on efficiency is significant overall
  - Can be practically eliminated by using Method T (basing next step on the previous target dosage) instead of Method M (basing next step on the previous measured dosage)
- The greater precision afforded by smaller step sizes is increasingly eroded by increasing dosage errors
  - Dosage measurement error has greater impact than both step size and dosage administration error on precision