Phase I designs that allow for uncertainty in the attribution of adverse events

Alexia Iasonos and John O'Quigley

Memorial Sloan Kettering Cancer Center, New York,

Universite Pierre et Marie Curie–Paris VI, Paris, France

Define Dose Limiting Toxicities

- Adverse Events
  - Grade (severity); Attribution; Type; Timing

- Define Dose Limiting Toxicities
  - High grade
  - Drug related
  - Timing (cycle 1)

Attribution Description

- **Unrelated**

The AE *is clearly NOT related to the intervention*

- **Unlikely**

doubtfully

- **Possible**

may be related

- **Probable**

likely related

- **Definite**

The AE *is clearly related to the intervention*

Background

Medical Literature

- AE attribution is subjective - err on the side of caution when uncertain (Mukherjee SD, 2011)
- 50% of AE in placebo arm were attributed to trt arm; 36% changed attribution over time (Hillman SL, JCO 2010)
- AE collection (over reporting of AE, noise vs signal)
- FDA. 2010 Investigational new drug safety reporting requirements.

- Zohar S and O'Quigley J (2009)
  - Defined the errors/problem
  - Assessed the impact of errors on 3+3 vs the CRM

Attribution Errors

Observed	True DLT yes	True DLT no
DLT		Type B Flag non DLT
Non DLT	Type A Miss DLTs	

- Type A leads to underestimation of DLTs
- Type B leads to overestimation of DLTs

Zohar, O'Quigley 2009

Concluded Type B has greater influence than Type A error because you can recover from Type A

3+3 is sensitive to Type B errors because of the stopping rule of 2 DLTs

How does CRM recover from errors?

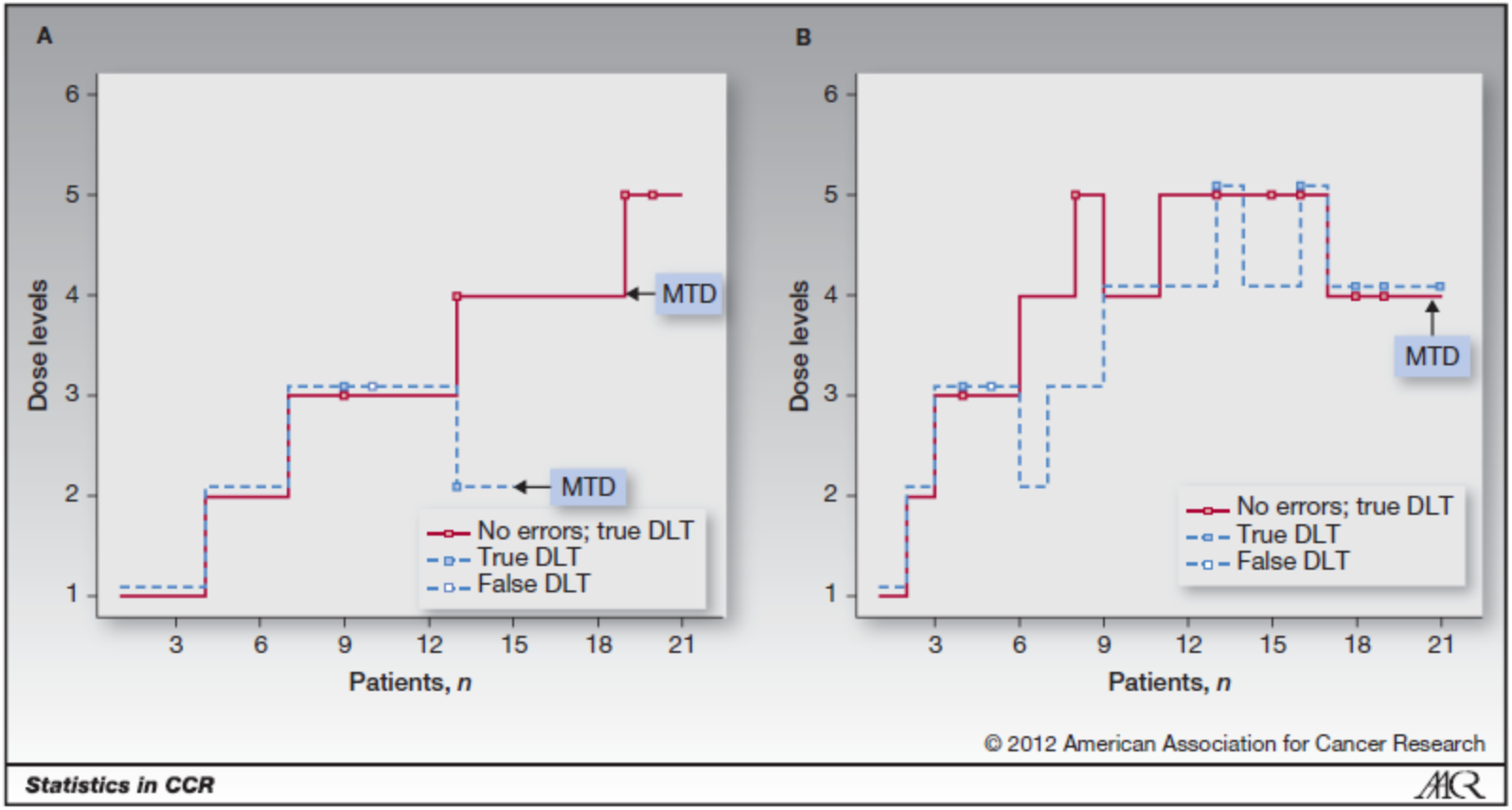


Figure 2. 3+3 (A) and CRM (B) under no error (solid line) and under type B error (dashed line) of incorrectly attributing OCT as DLT.

Research Program (2011-2015)

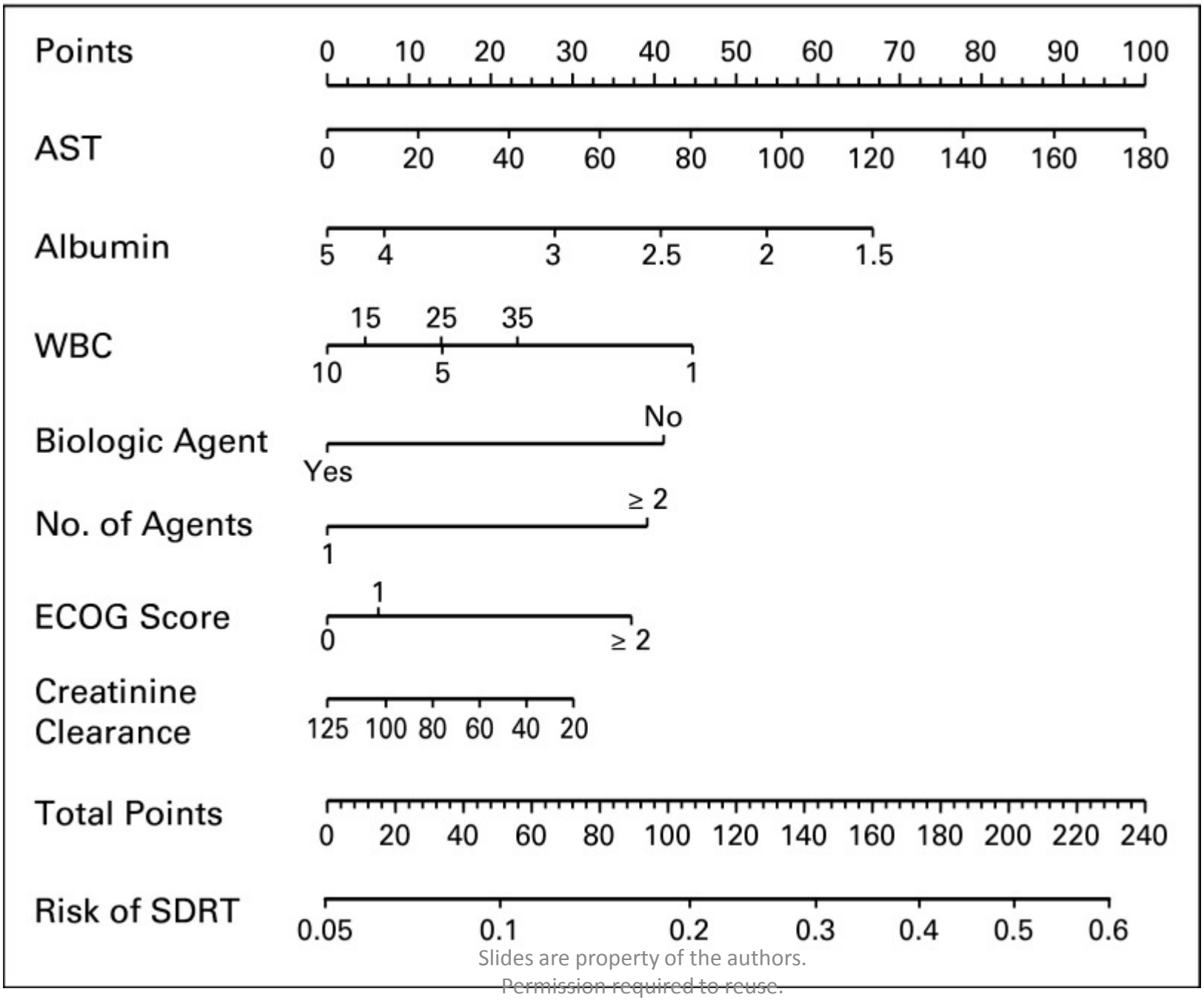
Aim 1: Assess the impact of errors on the estimated MTD

- Implications on drug development (bias in estimated MTD)
- Implications on patients during the trial and future trials
- Iasonos et al CCR 2012

Aim 2: Quantify the impact of errors using NCI Phase I trials

- Hyman DM JCO 2014 (3104 pts, 127 trials); Ballman K, JCO. Editorial.
- Eaton et al CCR 2016; true DLTs show up eventually
  - 30% possibly, 16% probably and definitely
  - Sharma MR, Ratain MJ. Editorial.
- Drilon A et al. Cancer 2016 (dermatologic toxicities)

Aim 3: Propose a solution to this problem



Funding

- The Translational and Integrative Medicine Research Fund at Memorial Sloan Kettering Cancer Center, NY. 2014

Problem

- Errors lead to over or under estimating the rate of DLTs
- Bias in estimation of the MTD
- Inaccurate RP2D

Goal:

- Recover information lost due to errors and Improve accuracy

Notation

- Y = 0, 1
- Z = 0, 1
- True DLT rate: P(Y=1)
- Observed DLT rate P(Z=1)

Types of Errors

A) P(Z=0|Y=1) missclassifying a true DLT

B) P(Z=1|Y=0) erroneously flagging a non DLT

Scores

- If you see an observed DLT, ask the clinician to make a determination whether this is a true DLT (score=1) or what is the likelihood that this is indeed drug related (score 0-1)

What we want to know is P(Y=1| Z=1)

Depends on P(Y=1) and on error rates

Model parameter estimation

$$\mathcal{U}_j(a) = z_j \left\{ s_j \frac{\psi'}{\psi}(x_j, a) + (1 - s_j) \frac{-\psi'}{1-\psi}(x_j, a) \right\} + (1 - z_j) \left\{ \frac{-\psi'}{1-\psi}(x_j, a) \right\}.$$

$$\mathcal{U}_j^*(a) = \sum_{j=1}^{n} \left\{ z_j \int u(s_j, x_j, a) f_j(s) \mathrm{d}s + (1 - z_j) \frac{-\psi'}{1-\psi}(x_j, a) \right\}$$

$$\mathcal{U}_j^*(a) = \sum_{j=1}^{n} \frac{\exp(a) \log(\beta_i)}{1 - \beta_i^{\exp(a)}} \left\{ z_j \int_{s_j^-}^{s_j^+} \frac{s}{c_j} B(g_j, h_j) \mathrm{d}s - \beta_i^{\exp(a)} \right\} \qquad \psi(d_i, a) = \beta_i^{\exp(a)}$$

Quasi Bernoulli Likelihood

Yuan Z, Chappell, Bailey Biometrics 2007

$$\psi\{d(n), s^*(n)\} = \left(R_j^*\right)^{s^*(n)} \left(1 - R_j^*\right)^{1-s^*(n)}.$$

The quasi-Bernoulli likelihood will be updated by

$$L_n = L_{n-1}\psi\{d(n), s^*(n)\},$$

$$U_j(a) = \frac{Z_j\{\lambda_1(x_j) - \lambda_2(x_j)\}\psi'(x_j, a)}{\lambda_1(x_j)\psi(x_j, a) + \lambda_2(x_j)\{1 - \psi(x_j, a)\}} - \frac{(1 - Z_j)\{\lambda_1(x_j) - \lambda_2(x_j)\}\psi'(x_j, a)}{1 - \lambda_1(x_j)\psi(x_j, a) - \lambda_2(x_j)\{1 - \psi(x_j, a)\}}.$$

True vs assigned scores

$P(Y=1|Z=1) = P(Y=1)/(P(Y=1)+P(Z=1|Y=0)P(Y=0))$

If $P(Z=1|Y=1)=1$, $P(Z=1|Y=0)=0.10$

$P(Y=1)=0.23$, $P(Z=1)=0.31$ at MTD

True score= $P(Y=1|Z=1)$ =

0.43, 0.55, **0.75**, 0.88, 0.98, 0.99

If $P(Y=1)$= 0.07, 0.11, 0.23, 0.43, 0.84, 0.98

DEPEND on ERROR RATE and on TRUE DLT RATE

Large sample theory

Well calibrated scores ensure convergence of the dose toxicity parameter to the true value under the conditions in Shen and O'Quigley 1996

Theorem 1. Assume that conditions M_1–M_7 are satisfied. For n sufficiently large, let $\hat{a}_n$ be the maximum likelihood estimate of the parameter a, and $x(n+1)$ the recommended dose level for the next patient. Then, almost surely, $\hat{a}_n \to a_0$ and $x(n+1) \to x_0$.

(a) M_1, for each a, the function $\psi(\cdot, a)$ is strictly increasing and it is continuous and strictly monotone in a in the same directions for all x.

(b) M_2, the function

$$u(s,x,a) = s\frac{\psi'}{\psi}(x,a) + (1-s)\frac{-\psi'}{1-\psi}(x,a),$$

for each $0 < s < 1$ and each x, is continuous and strictly monotone in a.

(c) M_3, the parameter a belongs to finite interval $[A, B]$.

(d) M_4, the target dose level is x_0, i.e. $R(x_0) = \theta$.

(e) M_5, the probabilities of toxicity at $x_1, \ldots, x_k$ satisfy $0 < R_1 < \ldots < R_k < 1$.

(f) M_6, for $i = 1, \ldots, k$, $a_i \in S$ where $S = \{a : |\psi(x_0, a) - \theta| < |\psi(x_i, a) - \theta|, \text{ for all } x_i \neq x_0\}$.

(g) M_7, for every dose level i the expected value of assigned scores is $E(s_i) = R(x_i)$.

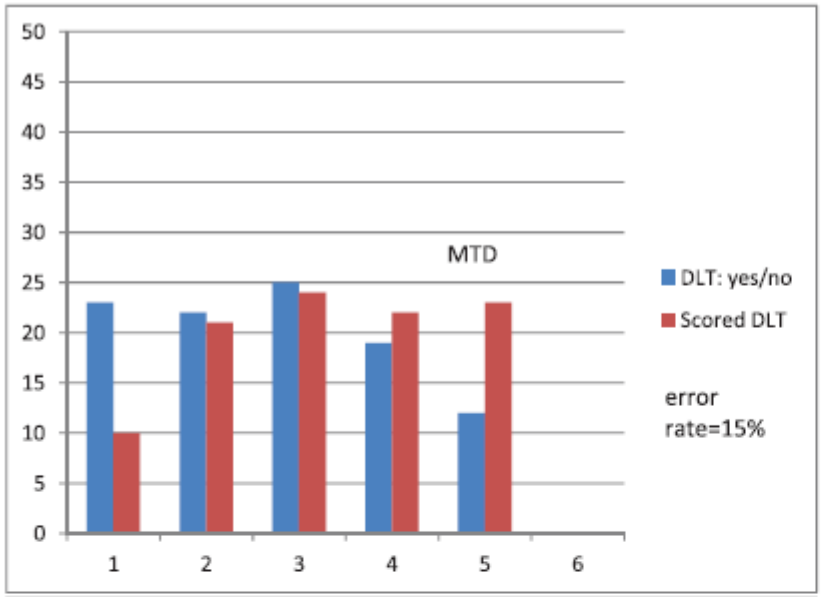
Small sample behavior

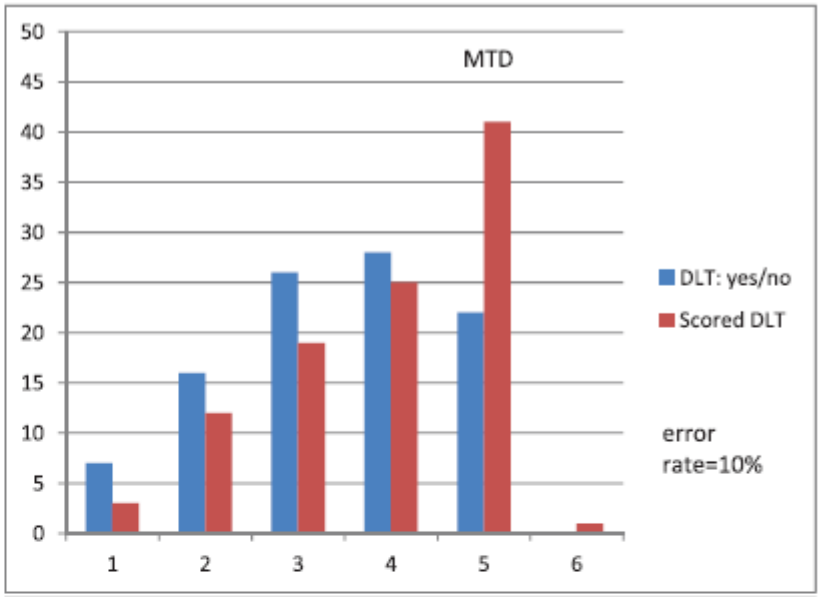
- If scores are right on average we know the method works
- If they are wrong, how wrong can they be while the method works acceptably well

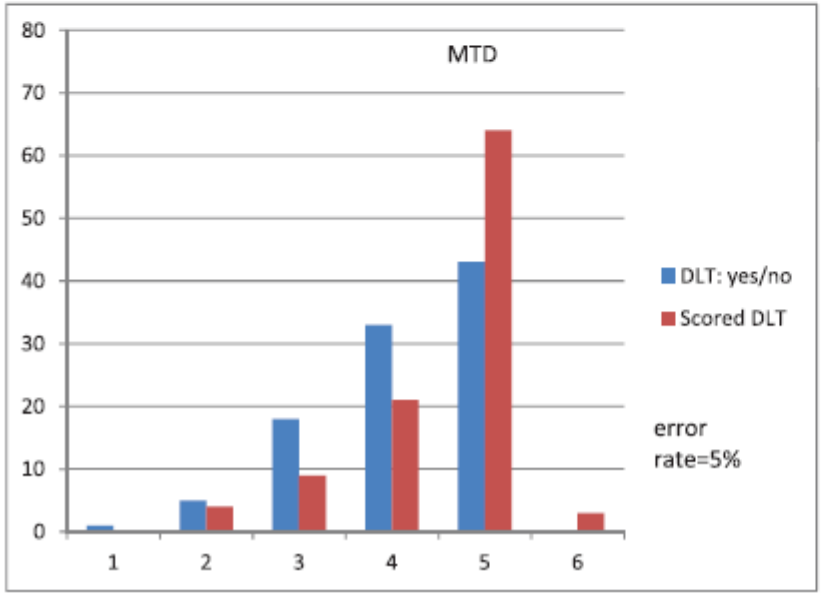
Simulation study

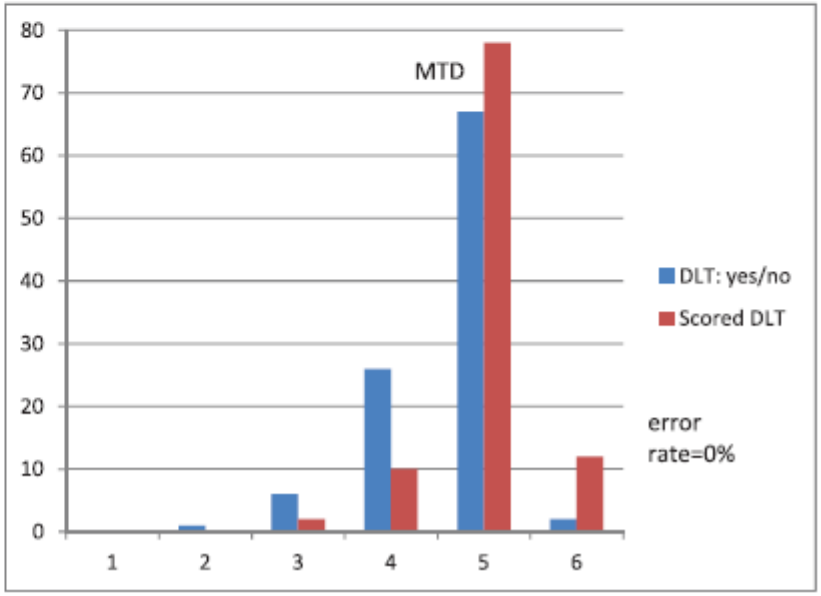
- Scores can be independent of dose
- Scores can go up as dose increases or have a dependency on how far we are from the starting dose
- Interval score (range) as opposed to a point score (DLT with probability 0.5-0.7)

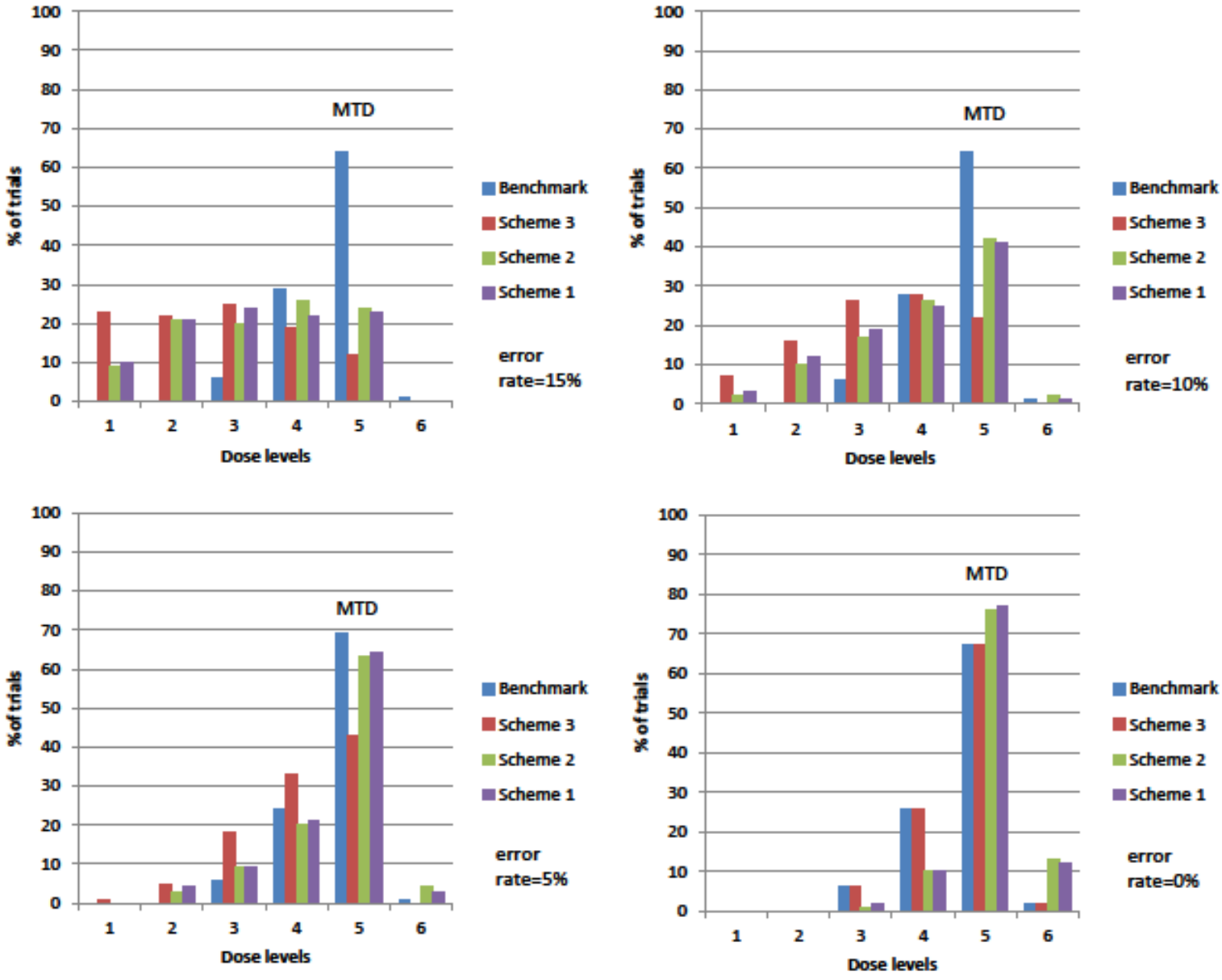
Scores are from Uniform mean=0.75, ± 0.20 (regardless of dose)



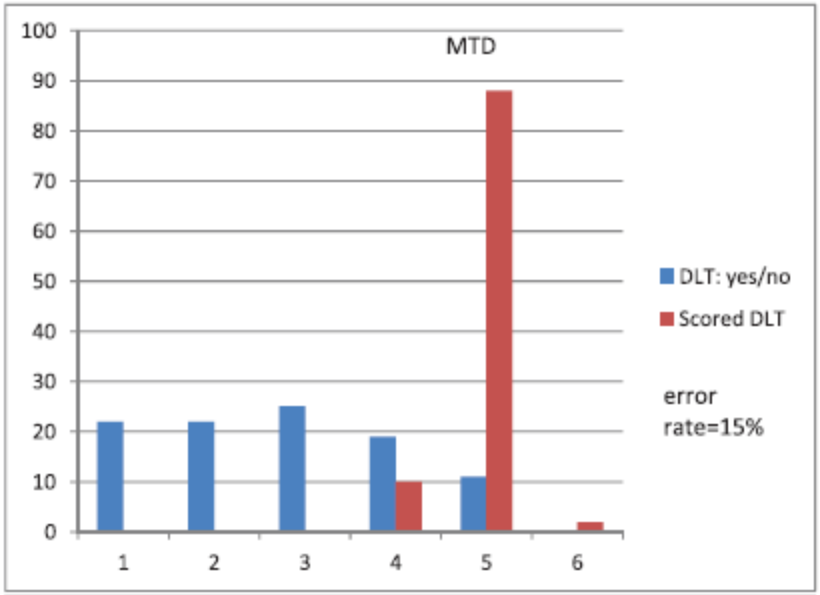


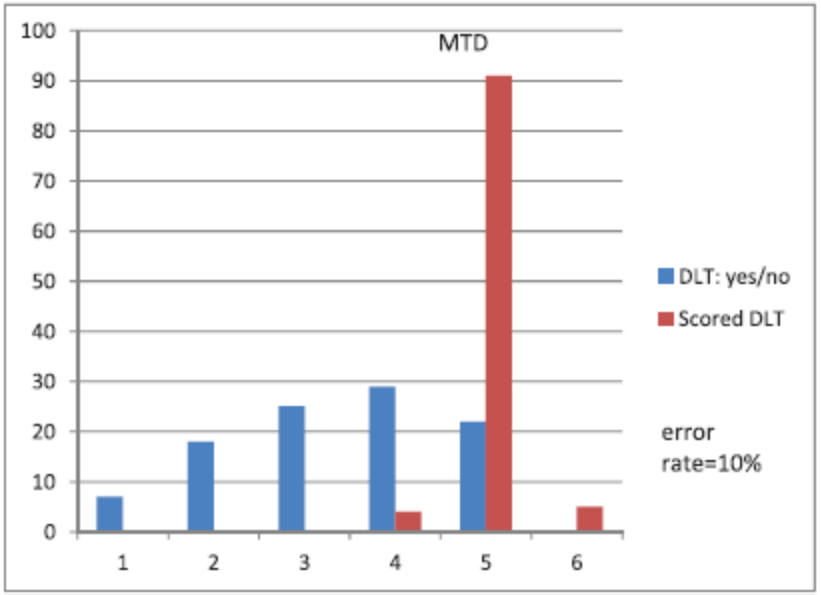


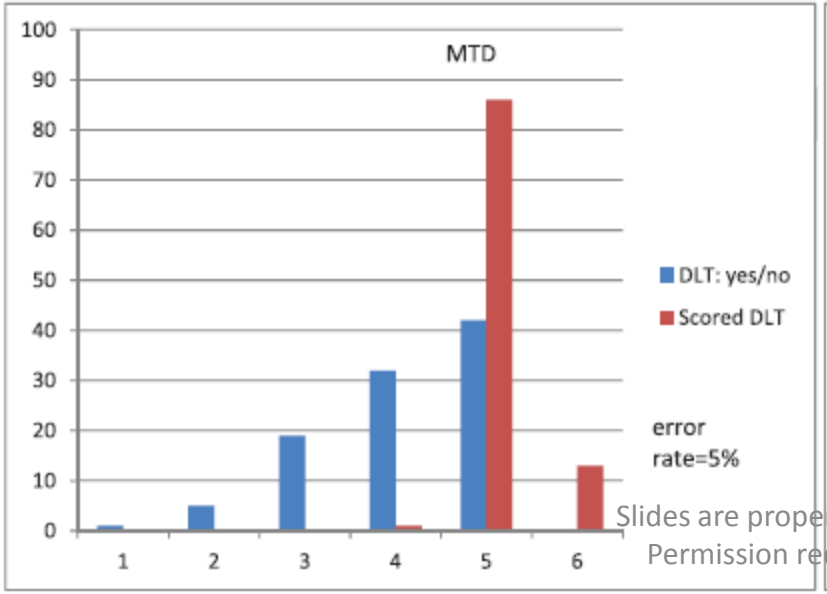


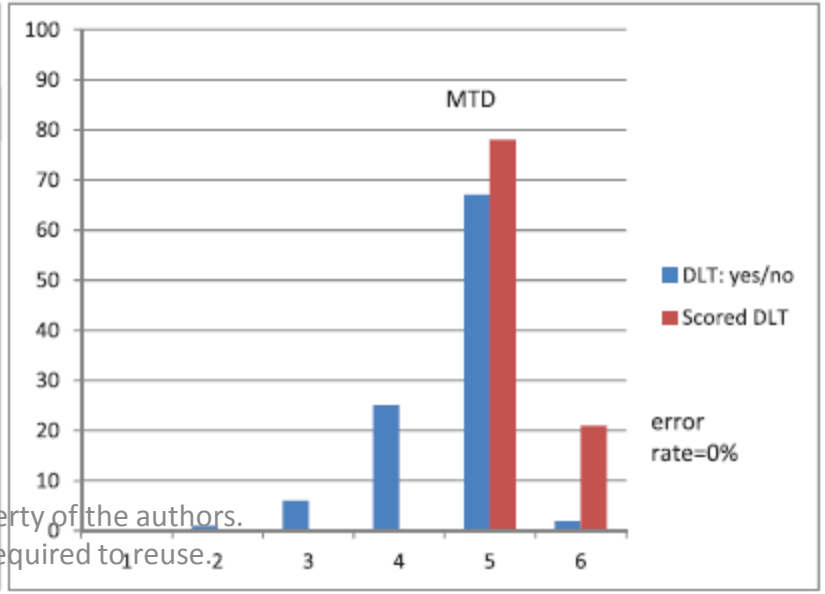


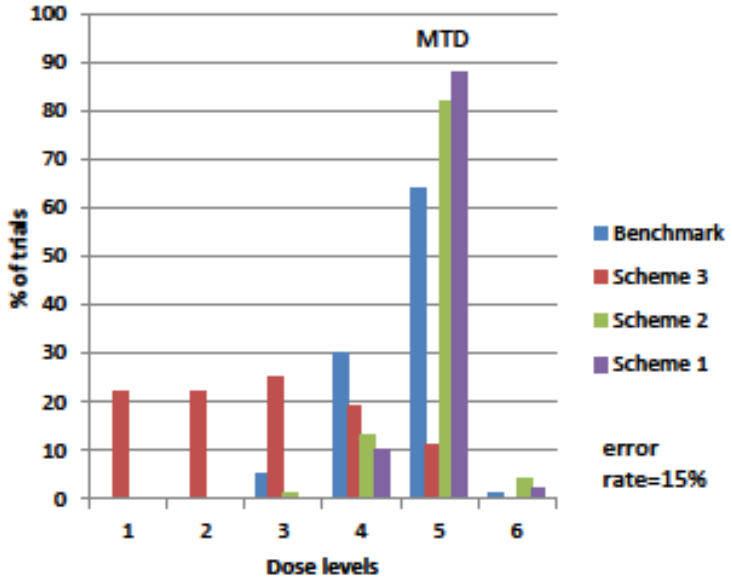
Score=0.1,0.2,0.3,0.4,0.5,0.7
True.rate=0.01,0.05,0.07,0.11,0.2,0.5

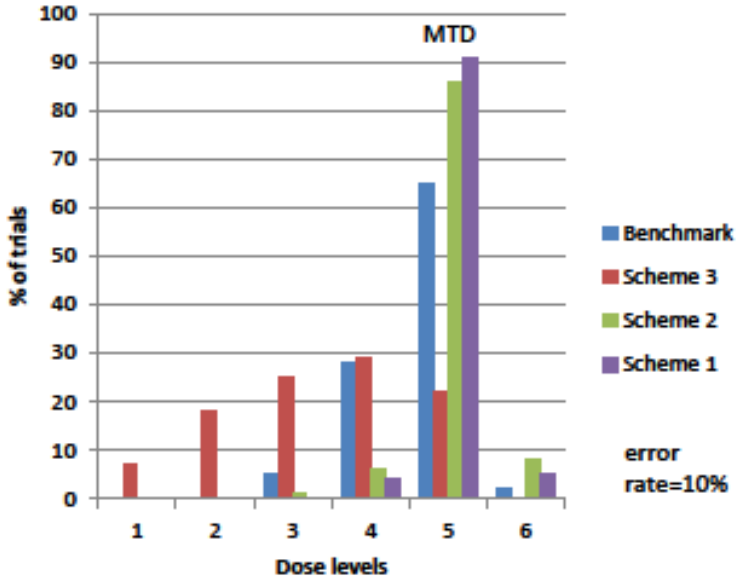


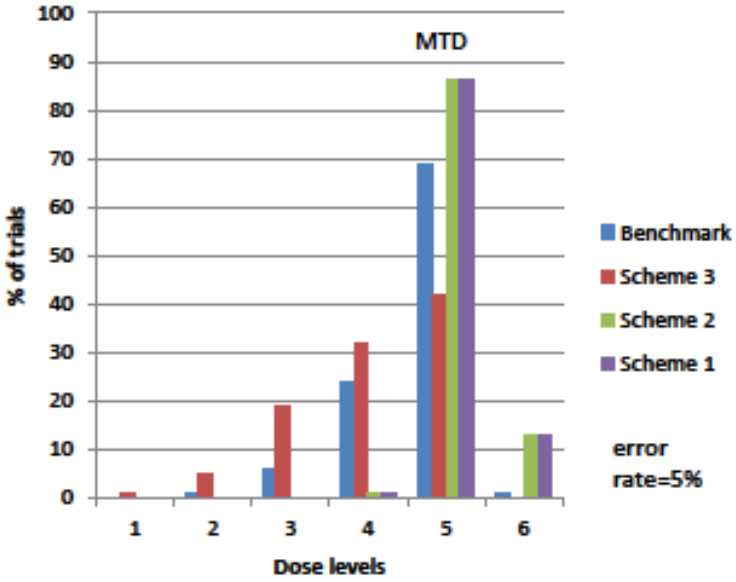


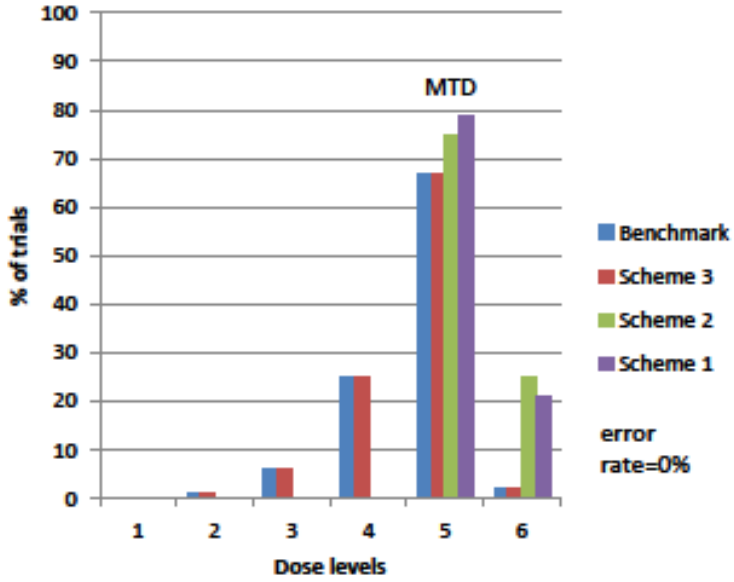




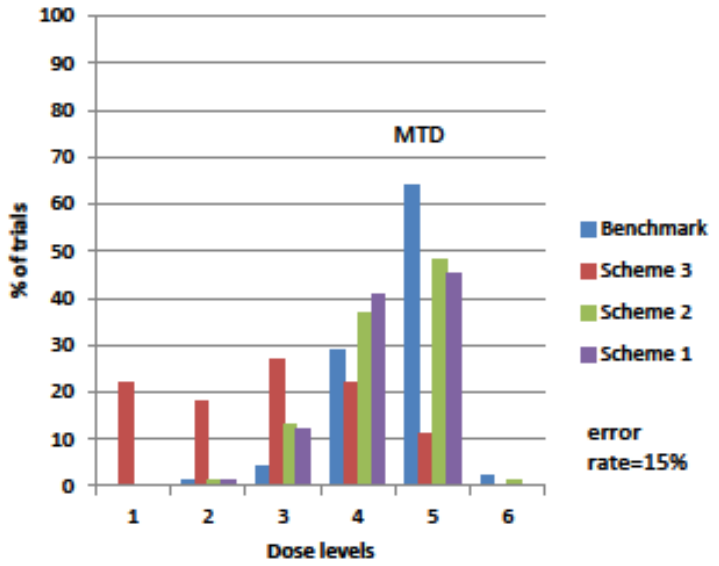


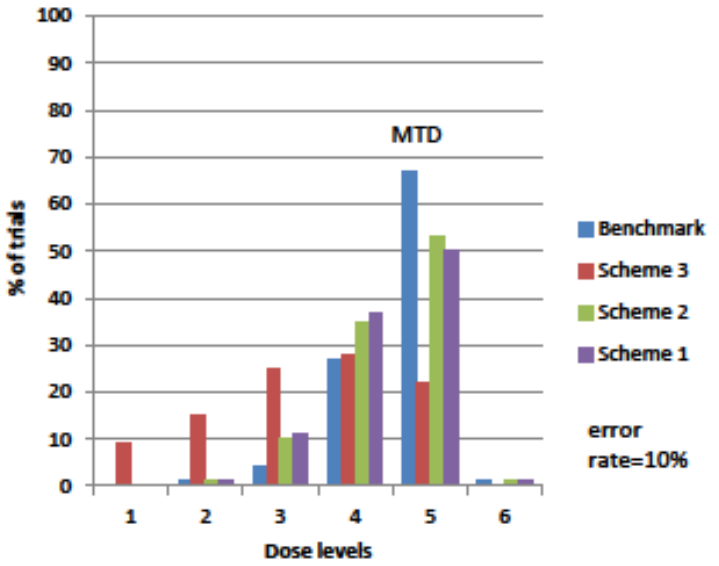


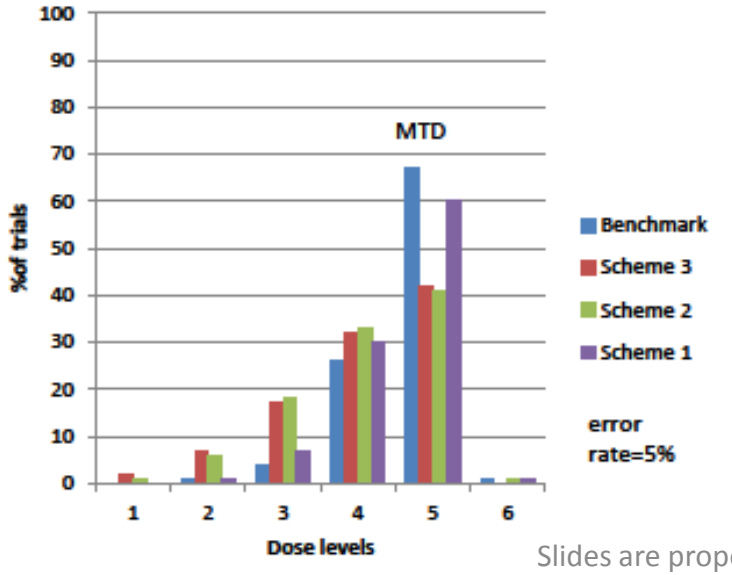


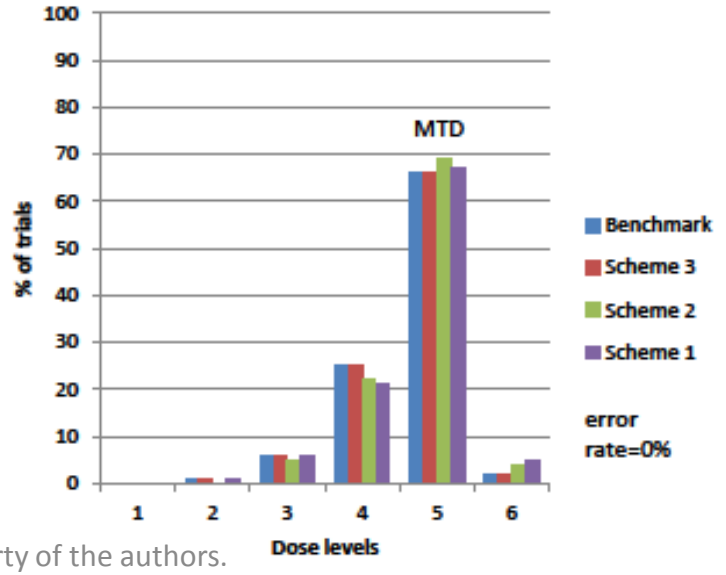


Systematically overestimating DLT risk (high scores)









Sensitivity Analysis

- Compared the interval vs point scores
- Simulated scores from beta with various means and variances
- Overestimating vs underestimating risk systematically (ie for all pts at all levels)
- Different location of the MTD

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Applied Statistics
Series C

Appl. Statist. (2017)

Phase I designs that allow for uncertainty in the attribution of adverse events

Alexia Iasonos

Memorial Sloan Kettering Cancer Center, New York, USA

and John O'Quigley

Université Pierre et Marie Curie–Paris VI, Paris, France

[Received May 2016. Revised September 2016]

Summary. In determining dose limiting toxicities in phase I studies, it is necessary to attribute adverse events to being drug related or not. Such determination is subjective and may introduce bias. We develop methods for removing or at least diminishing the effect of this bias on the estimation of the maximum tolerated dose. The approach that we suggest takes into account the subjectivity in the attribution of adverse events by using model-based dose escalation designs. The results show that gains can be achieved in terms of accuracy by recovering information lost to biases. These biases are a result of ignoring the errors in toxicity attribution.

Keywords: Clinical trials; Continual reassessment method; Dose finding algorithms; Dose limiting toxicity; Phase I trials; Sequential monitoring

- Supplemental Material includes extensive simulation studies based on error rates, true dose toxicity curve and how far the scores are from the true well calibrated scores
- Proofs
- Code

http://rss.onlinelibrary.wiley.com/hub/journal/10.1111/(ISSN)1467-9876/

Results - summary

- Calibrated scores recover significant bias in the estimated MTD as a result of attribution errors
- Scores need to be correct on average, not on any individual patient; the scores can be inexact for every patient.
- Systematic biased scores, i.e. underestimating DLT risk will lead to a higher dose and overestimating risk systematically will lead to a lower dose

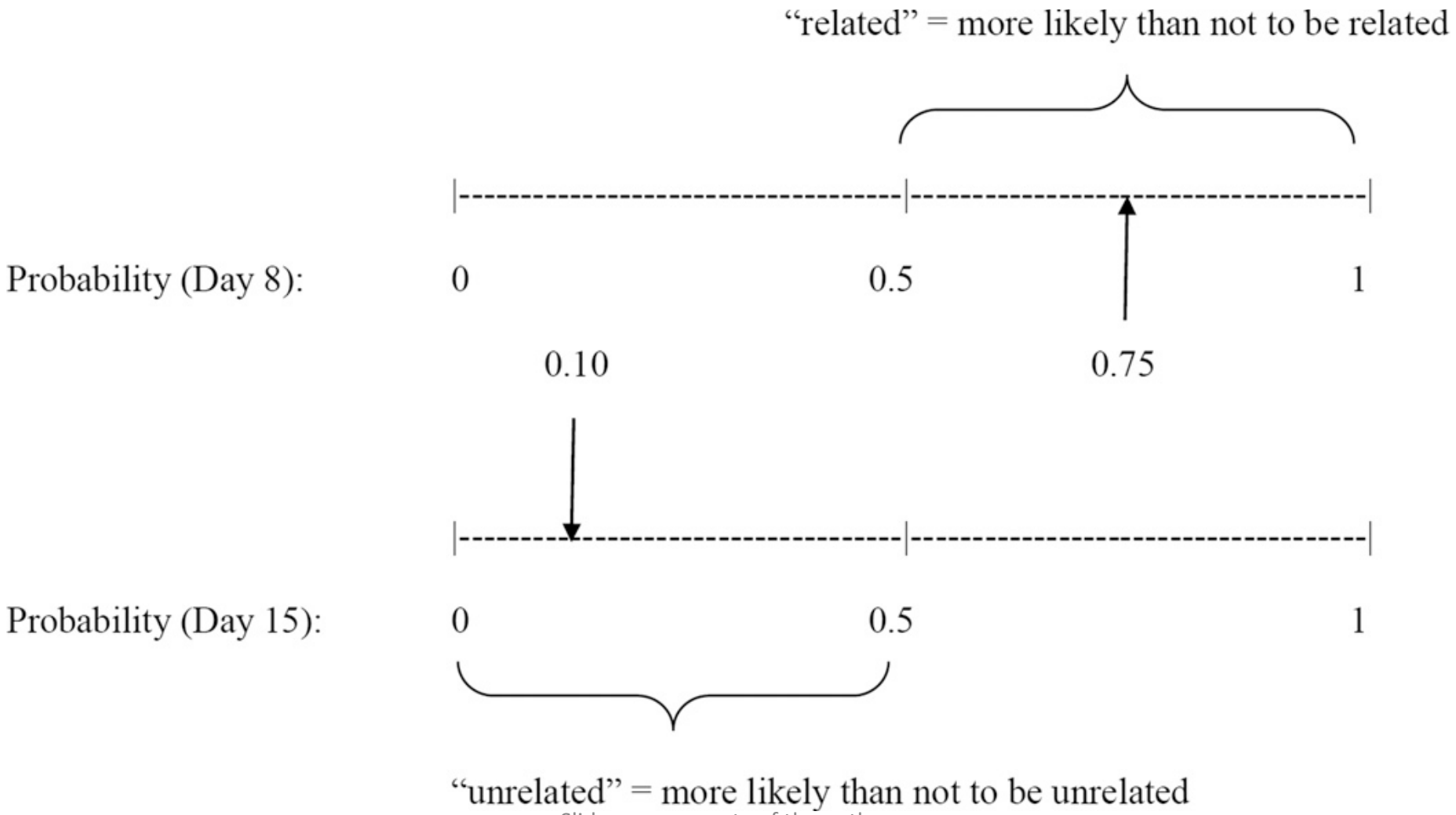
Conclusions

- We can incorporate the subjectivity involved in toxicity attribution into Phase I designs, such that the final recommended dose reflects this uncertainty.

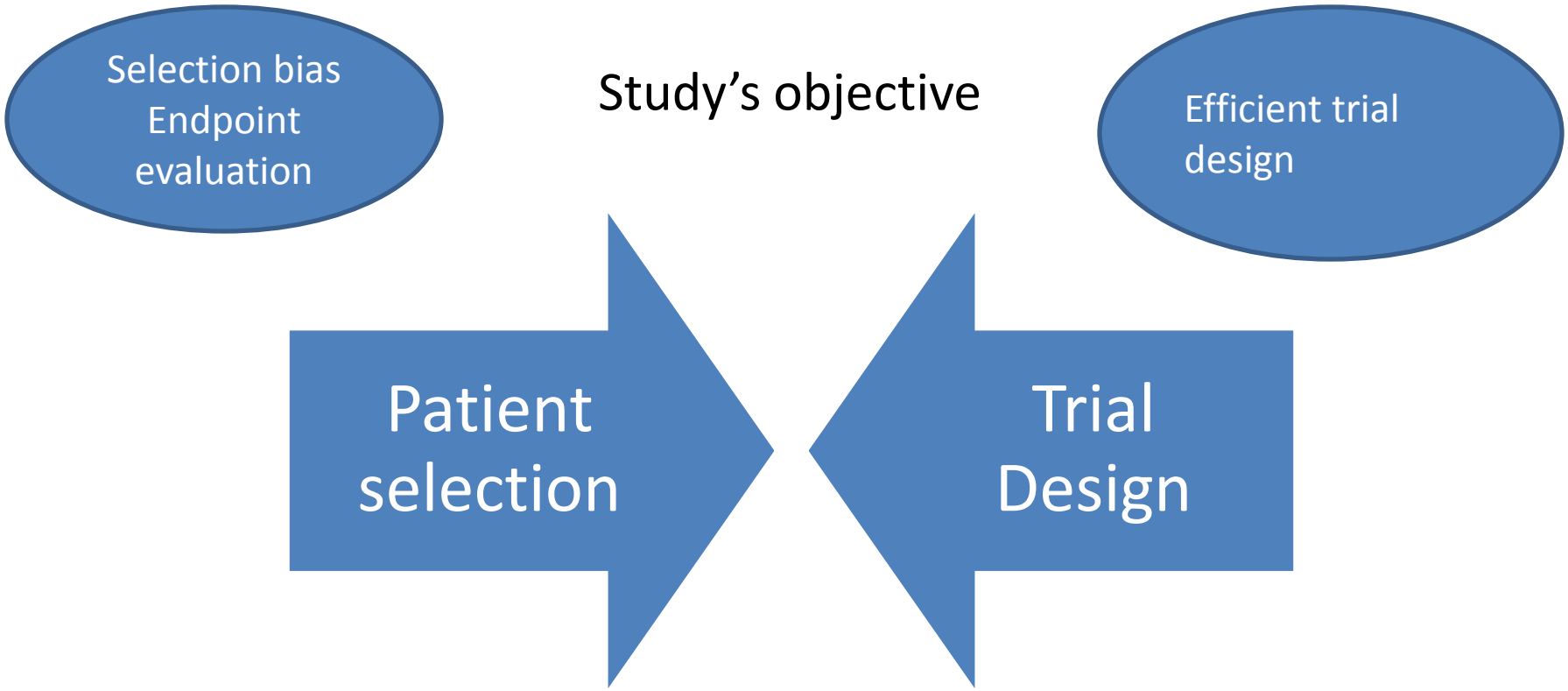
  Operationally:

  strictly adhere to the protocol by recording all DLTs and, yet, still allow expert opinion to prevent AE that are most likely not drug related from seriously compromising the identification of the correct MTD.

Taking a Measured Approach to Toxicity Data in Phase I Oncology Clinical Trials, Sharma and Ratain 2016



Patient selection and Trial Design are complementary and need to be optimized



Phase I trial improvement: a question of patient selection, trial design, or both?
Ballman KV. J Clin Oncol. 2014. Hyman et al. JCO 2014. Patient selection

References

8. Crowe, B.J., et al., *Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team.* Clin Trials, 2009. **6**(5): p. 430-40.
9. Ellenberg, S.S., T.R. Fleming, and D. DL., *Data monitoring committees in clinical trials: A practical perspective*. 2003, West Sussex, England: John Wiley & Sons Ltd.
10. **Hillman, S.L., et al., *Evaluation of the value of attribution in the interpretation of adverse event data: a North Central Cancer Treatment Group and American College of Surgeons Oncology Group investigation.* J Clin Oncol, 2010. 28(18): p. 3002-7.**
11. Arimone, Y., et al., *Agreement of expert judgment in causality assessment of adverse drug reactions.* Eur J Clin Pharmacol, 2005. **61**(3): p. 169-73.
12. Rothenberg, M.L., et al., *Mortality associated with irinotecan plus bolus fluorouracil/leucovorin: summary findings of an independent panel.* J Clin Oncol, 2001. **19**(18): p. 3801-7.
13. Thomas, E.J., et al., *The reliability of medical record review for estimating adverse event rates.* Ann Intern Med, 2002. **136**(11): p. 812-6.
14. Kaiser, L.D., et al., *Optimizing collection of adverse event data in cancer clinical trials supporting supplemental indications.* J Clin Oncol, 2010. **28**(34): p. 5046-53.
15. Mahoney, M.R., et al., *Dealing with a deluge of data: an assessment of adverse event data on North Central Cancer Treatment Group trials.* J Clin Oncol, 2005. **23**(36): p. 9275-81.
16. Sargent, D.J. and S.L. George, *Clinical trials data collection: when less is more.* J Clin Oncol, 2010. **28**(34): p. 5019-21.
17. **Mukherjee, S.D., et al., *A qualitative study evaluating causality attribution for serious adverse events during early phase oncology clinical trials.* Invest new drugs, 2011. 29(5): p. 1013-20.**
18. Sherman, R.B., et al., *New FDA regulation to improve safety reporting in clinical trials.* N Engl J Med, 2011. **365**(1): p. 3-5.
19. FDA *Guideline for industry: clinical safety data management: definitions and standards for expedited reporting.* 1995.
20. FDA, *Investigational new drug safety reporting requirements for human drug and biological products and safety reporting requirements for bioavailability and bioequivalence studies in humans. Final rule.* Fed Regist, 2010. **75**(188): p. 59935-63.
21. Iasonos A, et al., *The impact of non-drug-related toxicities on the estimation of the maximum tolerated dose in phase I trials.* Clin Cancer Res, 2012. **18**(19): p. 5179-87.