UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Toxicity-evaluation designs for cancer immunotherapy trials.

Karen Messer
Director of Biostatistics
Moores UCSD Cancer Center
Professor, Division of Biostatistics/Bioinformatics

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Presenter Disclosure Information

Karen Messer

The following relationships exist related to this presentation:

No Relationships to Disclose

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Setting

Phase I/II immunotherapy trials

- Agents with low expected toxicity
  - < 10% DLT rate
  - DLT: Dose Limiting Toxicity
- Expect that

Therapeutic dose < Maximum Tolerated Dose

- Goal is to establish safety of therapeutic dose

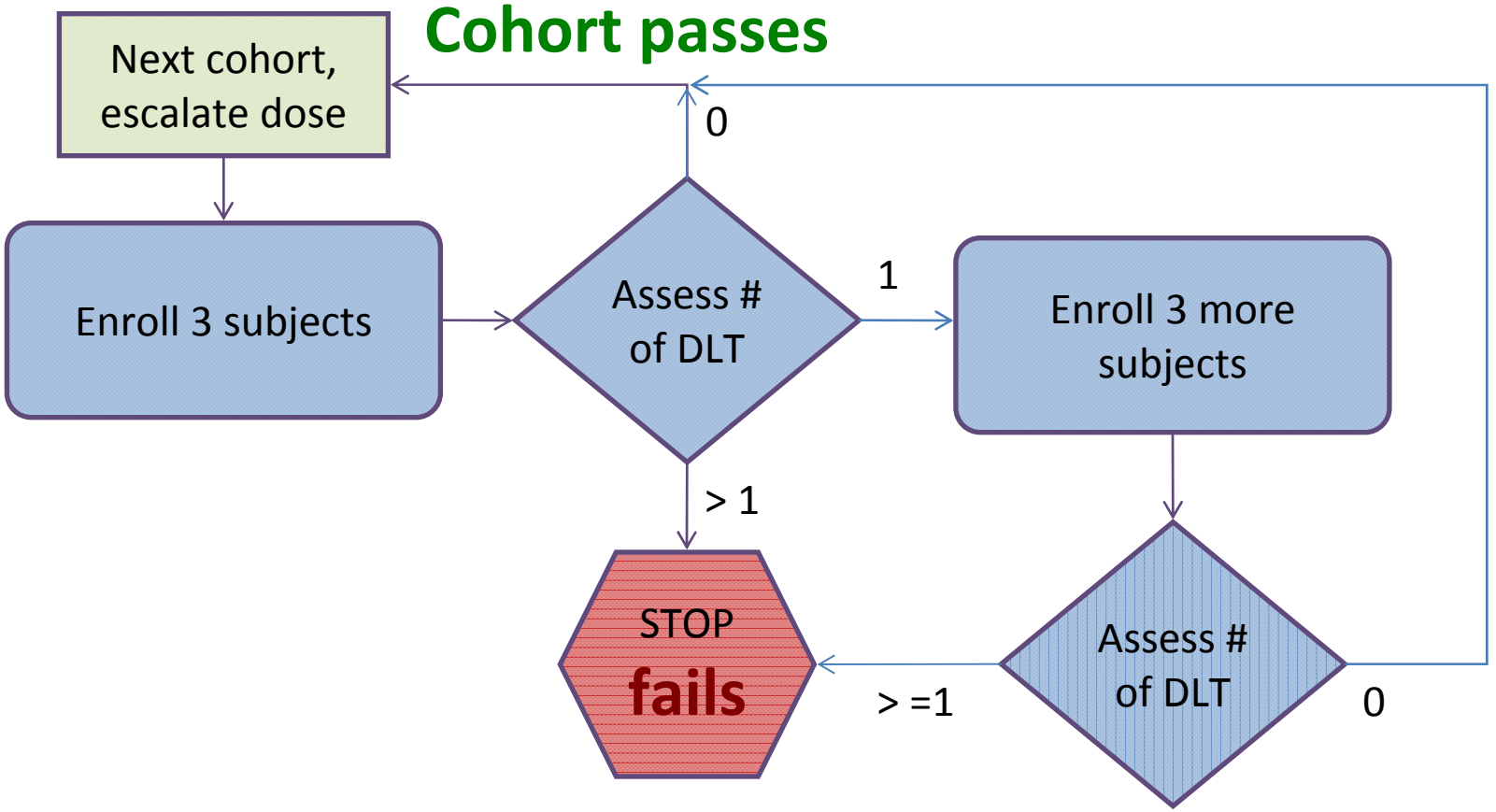
UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Reminder: 3+3 design

- Escalate dose until you see DLT's, then stop
  - Maximum Tolerated Dose (MTD) is one dose below stopping dose
- Commonly used (although inefficient)
- There is a nice theoretical literature
- Rarely compute formal estimates of
  - toxicity rate at MTD
  - Expected sample sizes under high, low toxicity

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

3+3 cohort design



UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Quick lit review

Crowley et al (2006) Handbook of Statistics in Clinical Oncology

Durham, Flournoy, Rosenberger, (1997) *Biometrics*

Gemzu and Flournoy (2006) JSPI review.

- Isotonic regression estimators Leung and Wang (2001) CCT
  Flournoy et al (2003, 2006) Paul and Rosenberger (2004)

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

This paper

Research Article

Statistics in Medicine

Received 29 April 2009, Accepted 20 October 2009 Published online in Wiley Interscience

(www.interscience.wiley.com) DOI: 10.1002/sim.3799

Toxicity-evaluation designs for phase I/II cancer immunotherapy trials

Karen Messer,[a]*† Loki Natarajan,[a] Edward D. Ball[b] and Thomas A. Lane[b]

Adds formal safety estimate to 3+3 design
Can run overlaid on two stage Phase II design

Aims of Tox-Eval design

1. Phase I: formal test of safety hypothesis.
   - The Phase I trial serves as an interim safety analysis.
2. Phase II: confidence interval for DLT rate, at therapeutic dose, combining Phase I/II data
3. Phase I sample size n_1 is the smallest that allows a safety test at 5% significance.
4. Phase II sample size n_2 is the smallest that supports a target margin of error on final conf. interval.
   - Incorporate Phase II efficacy test

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Design characteristics

- Parameters:
  - Expected toxicity rate t_a (≤ 10%)
  - Maximum acceptable safety threshold t_0
- A short run in dose for escalation, then stay at therapeutic dose
- Simple, based on familiar 3+3 design
- Works well at specific toxicity rates t_0 and t_a
  - Somewhat inflexible

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Test of size α

$$H_0 : t \geq t_0 \text{ v.s. } H_a : t < t_0.$$

using *the fewest subjects possible*.

- "Unacceptable toxicity" will stand by default, unless the data compel us to say otherwise.
- Should Phase I succeed, the conclusion will be:
  " Toxicity rate at therapeutic dose $< t_0$, at $(1 - \alpha)100$ confidence."

Test statistic is # of dose cohorts that 'pass' the 3+3 rule

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Group sequential design

- 3+3 dose cohorts
  each at therapeutic dose.
- Assess cohort i prior to enrolling cohort $i + 1$.
- Each cohort *passes* or *fails*

PASS 0 of 3 or 1 of 6 DLT's
FAIL 2 or more DLT's

- Enroll up to $i = m$ cohorts.

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Hypothesis test

$$H_0 : t \geq t_0 \text{ v.s. } H_a : t < t_0.$$

- If m cohorts pass:

  Reject null.
  With 95% confidence, $t < t_0$.

- If fewer than m cohorts pass:

  Fail to reject null.
  There is insufficient evidence to demonstrate safety.

Design properties

- Expected sample size $E[N]$ is determined as the smallest design that will support a test of safety at level (t_0, α)
- Let t_a be actual expected toxicity rate. The design is appropriate only for t_a with adequate power (80%).
- That is, only when the expected toxicity rate t_a is far below the acceptable rate t_0.
- FDA requires ample pre-clinical and clinical evidence that this is the case.

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

"Typical" Phase I test:

For

- $t_0 = 33\%$
- and $\alpha = 0.05, \quad m = 4.$

With $m = 4$ cohorts, in a standard 3+3 design, if all 4 pass then with 95 confidence the rate of DLT is less than 33%

We have then established that $p < 0.33$. Go on to Phase II.

Power ≥ 80% if $t_a \leq 6.5\%$

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Operating characteristics

- If toxicity is low, what is probability that you pass Phase I? (power)
- If toxicity is high, what is expected sample size? (safety)
- How many DLT's do you expect to see?

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

t_0 determines sample size

Table I. Required number of cohorts and alternative toxicity rates to achieve given size and 80 per cent power, 3+3 toxicity-estimation design.

			$\alpha = 0.05$, Power = 80 per cent				
t_0	$p(t_0)$	m	t_a	$E[N\|t_0]$	$E[\text{DLT}\|t_0]$	$E[N\|t_a]$	$E[\text{DLT}\|t_a]$
0.20	0.71	9	0.048	13.6	2.7	27.7	1.3
0.25	0.60	6	0.059	10.2	2.5	19.0	1.1
0.30	0.49	5	0.065	8.3	2.5	16.1	1.1
0.33	0.43	4	0.074	7.4	2.4	13.15	0.97
0.35	0.40	4	0.074	7.0	2.4	13.15	0.97
0.40	0.31	3	0.086	6.0	2.4	10.7	0.87

As safety test becomes more stringent, sample sizes increase

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

t_0 determines feasible $t_a << t_0$

Table I. Required number of cohorts and alternative toxicity rates to achieve given size and 80 per cent power, 3+3 toxicity-estimation design.

		$\alpha = 0.05$, Power = 80 per cent					
t_0	$p(t_0)$	m	t_a	$E[N\|t_0]$	$E[\text{DLT}\|t_0]$	$E[N\|t_a]$	$E[\text{DLT}\|t_a]$
0.20	0.71	9	0.048	13.6	2.7	27.7	1.3
0.25	0.60	6	0.059	10.2	2.5	19.0	1.1
0.30	0.49	5	0.065	8.3	2.5	16.1	1.1
0.33	0.43	4	0.074	7.4	2.4	13.15	0.97
0.35	0.40	4	0.074	7.0	2.4	13.15	0.97
0.40	0.31	3	0.086	6.0	2.4	10.7	0.87

As safety test becomes more stringent, expected tox rate must be smaller

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Phase II sample size

Table II. Standard deviation of $\tilde{t}$, the minimum variance estimator of the toxicity rate using the Phase I and Phase II data, for a range of rates t_a and Phase II sample sizes n.

					$n=10$		$n=15$		$n=20$	
t_0	t_a	m	$E[N\|t_a]$	$SD(\hat{t}_1)$	$SD(\hat{t}_2)$	$SD(\hat{t})$	$SD(\hat{t}_2)$	$SD(\hat{t})$	$SD(\hat{t}_2)$	$SD(\hat{t})$
0.20	0.048	9	27.7	0.025	0.068	0.023	0.055	0.022	0.048	0.022
0.25	0.059	6	19.0	0.036	0.075	0.033	0.061	0.031	0.053	0.030
0.30	0.065	5	16.1	0.043	0.078	0.038	0.064	0.036	0.055	0.034
0.33	0.074	4	13.2	0.054	0.083	0.045	0.068	0.042	0.059	0.040
0.35	0.074	4	13.2	0.054	0.083	0.045	0.068	0.042	0.059	0.040
0.40	0.086	3	10.7	0.072	0.089	0.056	0.072	0.051	0.063	0.047

Phase II sample sizes support reasonable confidence limits estimated DLT rate

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Phase II sample size

Table II. Standard deviation of $\tilde{t}$, the minimum variance estimator of the toxicity rate using the Phase I and Phase II data, for a range of rates t_a and Phase II sample sizes n.

					$n=10$		$n=15$		$n=20$	
t_0	t_a	m	$E[N\|t_a]$	$SD(\hat{t}_1)$	$SD(\hat{t}_2)$	$SD(\hat{t})$	$SD(\hat{t}_2)$	$SD(\hat{t})$	$SD(\hat{t}_2)$	$SD(\hat{t})$
0.20	0.048	9	27.7	0.025	0.068	0.023	0.055	0.022	0.048	0.022
0.25	0.059	6	19.0	0.036	0.075	0.033	0.061	0.031	0.053	0.030
0.30	0.065	5	16.1	0.043	0.078	0.038	0.064	0.036	0.055	0.034
0.33	0.074	4	13.2	0.054	0.083	0.045	0.068	0.042	0.059	0.040
0.35	0.074	4	13.2	0.054	0.083	0.045	0.068	0.042	0.059	0.040
0.40	0.086	3	10.7	0.072	0.089	0.056	0.072	0.051	0.063	0.047

Expected combined sample at therapeutic dose is E[N] + n

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Summary, Tox-Eval design

For low toxicity agents:

- Phase I formal test of safety
- Confidence intervals on DLT rate at reasonable Phase II n's
- Implementation is familiar and simple
- Computations are not burdensome
  - Tables of sample sizes available

Works well- we are using this design.

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Statistical refinements

- The Tox-Eval design is not based on a sufficient statistic, hence is necessarily inefficient
  - (But not by much!)
- Covers a restricted set of design paramters
  - i.e. If $t_0 = 30\%$ then $t_a \leq 6.5\%$

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Extension: exact group sequential designs

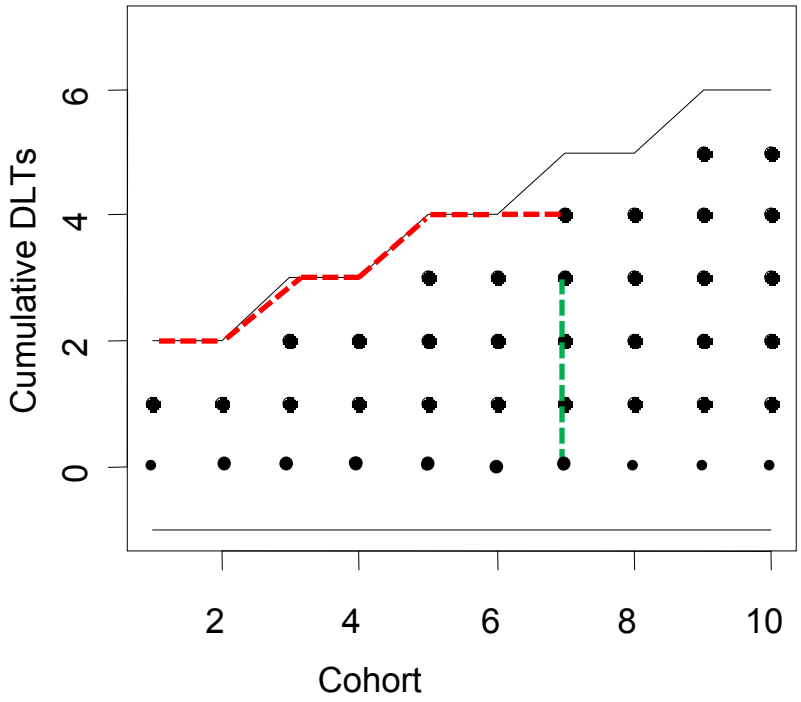
Example: a trial of stem cell therapy in stroke

- t_0 = 30% and t_a = 10%
- test of size α = 10%

The corresponding 3+3 design :

- M=4 cohorts
- Expected sample sizes
  - Null (toxic): 8 subjects
  - Alternative (safe) : 13 subjects
- *Power* 82% at t_a = 7% ; only 67 % at t_a = 10%

An exact group sequential design



One sided:

- Stop early for toxicity
- Declare 'safe" (t < 30%)

Moral: the GS design with comparable size and power is very similar, given feasible t_0 and t_a

M=4 cohorts

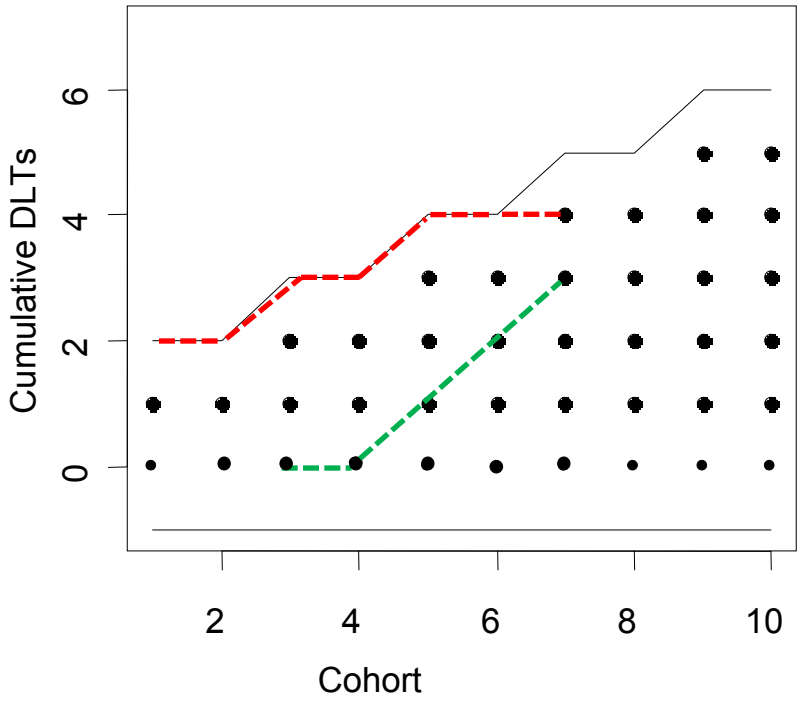
Expected sample sizes

Null (toxic): ~~8~~ 7.2 subjects

Alternative (safe) : ~~13~~ xx subjects

Power ~~82%~~ 80 % at t_a = 7% ; only ~~67 %~~ 66% at t_a = 10%

An exact group sequential design



Two sided:

- Stop early for toxicity
- Declare 'safe" (t < 30%)

Moral: Exact GS designs are more complex,
but also more flexible in terms of feasible t_0 and t_a

M=7 cohorts

Expected sample sizes

Null (toxic): ~~8~~ 8.6 subjects

Alternative (safe) : ~~13~~ xx subjects

Power ~~82%~~ at t_a = 7% ; only ~~67 %~~ 79% at t_a = 10%

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Exact GS designs

- Are more complex
- Do not add much, if there is a Tox-Eval design that fits
- However, cover a wider range of possible t_0, t_a
  - We have code to implement these, but it is not yet published.

UNIVERSITY of CALIFORNIA
SAN DIEGO
MEDICAL CENTER
MOORES CANCER CENTER

Thank you!

Collaborators:

- Loki Natarajan
- Colleen Kelly
- Ted Ball
- Tom Lane