An Adaptive Oncology Dose Finding Trial Using the Time-to-Event Continual Reassessment Method Incorporating Cycle Information

Bo Huang, PhD

Oncology Business Unit
Pfizer Inc.

PhRMA KOL Adaptive Design Seminar

11 November 2011
Acknowledgment

- Michelle Kneissl, PhD (Pfizer)
- Rebecca Rosenstein, PhD (Pfizer)
- Janis Grechko, PhD (Pfizer)
- Christy Chuang-Stein, PhD (Pfizer)
- Camilla Fowst, MD (Pfizer)
- Elena Colajori, MD (Pfizer)
- Ying Kuen Cheung, PhD (Columbia University)
Outline

• Standard 3+3 design versus existing alternative designs
• Time-to-event CRM
• Case study background and motivation
• Adaptive weight function
• Design calibration
• Simulations
• Operational considerations
• Summary and discussion
3+3 Design

- Patients treated in groups of 3
- No statistical estimation involved
- MTD: highest dose with ≤ 1 dose limiting toxicity (DLT) out of 6 evaluable patients
Advantages of 3+3

- Simple, easy to understand
- 3 or 6 patients per dose to account for inter-patient variability
- Extensive clinical experience
Problems of the 3+3 design

- Not flexible (what’s the target rate of toxicity? What’s the definition of MTD?)
- “Memory-less”: using data only from the most recent cohort
- Slow escalation, unable to re-escalate, a large portion of patients potentially treated sub-therapeutically
- The probability of stopping at an incorrect dose level (either sub-therapeutic or toxic) is higher than generally believed (Reiner, Paoletti, O’Quigley 1999)
- Cannot handle complicated situations (e.g. drug combination, multiple endpoints, grade information)

If a dose has true DLT rate of 25%, there is a 60% chance that the algorithm will escalate to a higher dose for the next cohort.
**Alternative designs in the literature**

- **Algorithm-based designs**
  - Up and down designs (Storer, 1989), A+B designs (Lin et al., 2001; Ivanova, 2006)
  - Accelerated titration design (Simon et al., 1997)
  - Biased coin design (Durham, 1997)
  - Isotonic regression (Leung et al., 2001)

- **Model-based designs**
  - Continual reassessment method (CRM) (O’Quigley et al., 1990, 1996) and its modifications and extensions
  - Escalation with Overdose Control (EWOC) design (Babb et al., 1998)
  - Combined outcome (safety+efficacy) designs (e.g. Thall et al., 2004)
  - Other Bayesian model-based designs (multi-population, combination agents, continuous or ordered endpoints etc)

*Most of these designs require a fixed follow-up time (evaluation window) before subsequent patients can be enrolled*
Time-to-event (TITE) CRM Design

• **Reference:** Cheung and Chappell, Biometrics (2000)

• **Motivation:** To account for late-onset toxicities. To shorten the trial duration.

• **Notations**
  - \( T \): fixed DLT evaluation time
  - \( Ti \): time-to-toxicity variable of patient \( i \)
  - \( xi \): dose level of patient \( i \)
  - \( Yi \): \( =I_{\{Ti\leq T&Ti\leq u\}} \) DLT indicator for patient \( i \) at observation time \( u \)

• **Weighted dose response model:** For \( u<T \), note that

\[
\Pr(T_i \leq u) = \Pr(T_i \leq u \mid T_i \leq T) \Pr(T_i \leq T) = w(u, T) F(x_i, \theta)
\]

• **Weighted binomial likelihood**

\[
\prod_{i=1}^{m} \left\{ w_i F(x_i, \theta) \right\}^{Y_i} \left\{ 1 - w_i F(x_i, \theta) \right\}^{1-Y_i}
\]

• **Weight function** \( w_i \): e.g \( w(u, T) = u / T \)

R package: dfCRM
An Oncology Case Study Using the Modified TITE-CRM Method

• **Agent name**: PF-xxxx

• **Modality**: fully human monoclonal antibody

• **FIH Study 1001 Population**:
  - 1001A (single agent): all comers
  - 1001B (in combination with Standard of Care): a subtype of cancer
Background information: FIH study design

1001A: PF- SAMD

1001B: PF- + SOC agent (fixed dose)

Safety – Dose 2 determined to be safe to move to the combination
**Motivation**

- DLT evaluation window: two cycles of PF- (2 x 4 wks)
  - SOC agent: Qwk x 4, cycle 1 only
  - PF- 24 hrs after 2\textsuperscript{nd} weekly dose of SOC
- Minimum 1.5 week window between the first PF dose of patient 1 and the first PF dose of patient 2

```
SOC       SOC+PF       SOC+PF
1wk       3d           1wk  \\
\downarrow   \downarrow    \downarrow
\text{2 cycles of PF-: 8 wks}
```

**Dosing schedule with conventional methods**

- To enroll 15 patients, a 3+3 design or a CRM design with cohort size 3 takes more than 1 year
- Early drop-out requires replacement of patients
  - observed at dose level 1 (1/4) of 1001A
PK & PD of Human Efficacious Dose Assumption

Animal Tox study

- Animal tox study with weekly dosing indicates AEs occur after 2\textsuperscript{nd} weekly dose

Simulated PK and PD profiles based on a $X$ mg/kg, 1-hr infusion, Q4W dosing regimen in human. A $X$ mg/kg dose is predicted to maintain the biomarker stimulation needed for efficacy.
Time Savings from the TITE-CRM approach

Assumptions:

- Patients are available when enrollment is open
- Early drop-out rate: 15%
- Time delay due to early drop-out (applied to 3+3 and CRM) ~ N(mean=3wks, sd=1wk)
- TITE-CRM v1-v3: expected delay due to DLT: 1-3 weeks
- Delay due to DLT ~ N(mean=k wks, sd=k/3 wks), k=1,2,3
- CRM (1): cohort size 1
- CRM (3): cohort size 3
Design proposal

• Method: TITE-CRM with adaptive weight function for late-cycle toxicities
  – Bayesian model based dose escalation design
  – Uses all available data on a continuous basis to estimate the dose-toxicity relationship
  – Addresses a number of issues with the 3+3 design
  – Shortens the trial duration by enrolling patients with minimum interruption
  – Capable of incorporating late toxicity information into model
  – An adaptive weight function for the cyclical time-to-toxicity distribution for improved precision and to protect patients from excess toxicity due to late toxicities
Adaptive weight function incorporating cycle information

• Suppose the DLT evaluation window consists of K cycles \((C_1, \ldots, C_K)\), each having duration \(T\)

• For patient \(i\) \((i=1,\ldots,m)\) with DLT, Let \(Z_i=(z_{i1},\ldots,z_{iK})\) indicate the toxicity status at each cycle. \(Z_i \sim \text{Multi}(P, 1)\) where \(P=(p_1,\ldots,p_K)\) is the probability of DLT occurring in each cycle

• Assume \(P \sim \text{Dir}(\alpha_0)\), \(\alpha_0=(\alpha_{01},\ldots,\alpha_{0K})\)
  – When no prior knowledge on time-to-toxicity distribution, \(\alpha_{01}=\ldots=\alpha_{0K}\) (e.g. =1)

\[
P | Z \sim \text{Dir}(\alpha_{01} + \sum_{i}^{m} z_{i1},\ldots, \alpha_{0K} + \sum_{i}^{m} z_{iK})
\]

• \(P\) is estimated as

\[
\hat{P}(m) = \int_{0}^{1} Pf(P | Z)dP
\]
Adaptive weight function incorporating cycle information (cont’d)

• Since $K=2$ here (two-cycle window of PF-), Dirichlet(.) reduces to Beta(.)

• Assume the within-cycle toxicity distribution is fixed: $G(t)$

$$w(t, P(m)) = \begin{cases} p_1(m)G(t) & 0 < t \leq T \\ p_1(m) + p_2(m)G(t-T) & T < t \leq 2T \end{cases}$$

• In the weighted binomial likelihood, derive the posterior estimate of parameter $\theta$

$$\hat{\theta}_m = \int \theta f(\theta | Y, Z, w(P(m)))d\theta$$

• The $(m+1)$th patient receives the model-based MTD estimate for target tox rate $\nu$ (e.g. 0.25)

$$x_{m+1} = \arg\min_{d_i : F(d_i, \hat{\theta}_m) \leq \nu} \left| F(d_i, \hat{\theta}_m) - \nu \right|$$
Some statistical properties

• $\hat{P}(m)$ converges to the underlying true conditional cyclical probability vector $P$ as $m \to \infty$

• $P$ provides useful insights on the safety profiles in association with the drug exposure pattern

• The MTD estimate $x_{m+1}$ converges to the true MTD, under some regularity model conditions

• In small sample dose finding trials, the asymptotic property does not apply. However, failure of consistency is predictive of undesirable performance at moderate samples
TITE-CRM dosing scheme flow chart

1. First cohort of 3 patients treated at the starting dose \(d_1\)

2. DLT data collected in real time on an ongoing basis.

3. Assign weights to patients according to their follow-up time and toxicity status.

4. A statistical model estimates the dose-toxicity curve using all available data. The MTD is estimated accordingly.

5. Dose escalation/de-escalation stops if we have sufficient data to determine the MTD:
   - Have at least 9 subjects been treated at a dose that is predicted to be the MTD?
   - Has the maximum sample size been reached?

6. If not: Continue to accrue patients to the trial.

7. If yes: Continue to accrue patients to the trial and assign new patients to the estimated MTD.

MTD: highest dose level with DLT rate \(\leq 0.25\)
Practical considerations

- Starting dose is d1
- For patients to be treated at an untested dose level, at least 3 patients should have been treated at each of the previous lower dose levels
  - Dose skipping to untested level is not allowed
  - To allow for some data collected at each level to account for patient heterogeneity for safety and PK assessment
Stopping rules

a) At least 9 patients have been treated at a dose that is predicted to be the MTD

b) Maximum sample size (36) has been reached

c) All doses appear to be overly toxic and the MTD cannot be determined in the current trial
In the outcome of the pre-IND meeting with FDA, the following FDA comments on the dose finding design were provided:

- Include a rule that will prohibit dose escalation if a particular cohort exhibits ≥ 33% DLT rate at any time.
- Limit dose level k+1 to be the next higher dose above dose level k.
• Include a rule that will prohibit dose escalation if a particular cohort exhibits ≥ 33% DLT rate at any time.
  - Since the target DLT rate is 25%, this restriction is guaranteed by the Coherence Principle (Cheung, Biometrika, 2005; Huang et al., SIM, 2008)

• Limit dose level k+1 to be the next higher dose above dose level k.
  - Restriction applied to the case when dose level k+1 has not been tested (Goodman et al., SIM, 1995; Moller, SIM, 1995)
Design calibration – Clinical parameters

- **Target rate $\nu$: 25%**
  - MTD defined as the highest dose with DLT rate ≤ 25%
  - “More is better” may not be applicable to targeted agents

- **Number of test doses $D$: 6**
  - Denoted as $d_1$, $d_2$, …, $d_6$

- **Sample size $N$: 36 (minimum size ≥ 3 at each dose)**
  - Empirical estimate by simulation performance
  - A sample size of 36 ensures the estimates of any binary variable (e.g. objective response) have a 95% CI of width ≤ 0.33.
  - A sample size of 36 also enables us to detect any unexpected toxicity that occurs at 5% rate with a probability of 0.85, and that occurs at 10% rate with a probability of 0.98.

$$\frac{N - 3(D - 2)}{2} > \max\left(\frac{1}{\nu}, 9\right)$$
$$\Rightarrow N > 30$$
Design calibration – Initial guess of toxicity (skeleton)

• A poorly calibrated CRM can lead to poor operating characteristics (Shen and O’Quigley, Biometrika, 1996)

• Ideally, initial guesses ($\pi_1, \ldots, \pi_6$) of toxicity probability reflects the clinical belief regarding the test doses
  – Not available in practice (FIH trial) or misleading

• Treat $\pi_i$ as a model parameter so that the CRM will yield good operating characteristics based on the model sensitivity

• Model sensitivity in the CRM depends more on the initial skeleton than the functional form $F(., \theta)$ (Lee and Cheung, Clinical Trials, 2009)
Based on the techniques in Lee and Cheung (2009), with a target rate of 0.25, the prior toxicity values at each dose are selected as \((\pi_1, \ldots, \pi_6) = (0.08, 0.15, 0.25, 0.40, 0.55, 0.70)\). The trial will eventually choose a dose with toxicity probability within \((0.16, 0.33)\).
Design calibration – Model and Prior

- Functional form of dose-toxicity curve
  - A power function modeling DLT rate at each dose $d_i (i=1, \ldots, 6)$ expressed as $\Pr(\text{DLT} | d_i) = F_i(\theta)$ is used
    $$F_i(\theta) = \pi_i \exp(\theta)$$
    where $\pi_i$ is the prior estimate of DLT rate at dose level $d_i$, and $\pi_1 \leq \pi_2 \leq \ldots \leq \pi_6$ (on previous slide) and $\theta \sim N(0, \sigma_0^2)$,
    $$\sigma_0 = 0.97$$

- Apply the proposed adaptive weight function with number of cycles $K=2$, and linear weight function within cycle
  $$G(t) = \frac{t}{T}, 0 \leq t \leq T$$
Simulations – Set Up

- Six dose-toxicity scenarios
- Underlying dose-toxicity model as calibrated (Scenario 5)
- Three simulated cycle-toxicity patterns
  - $p_1 = p_2 = \frac{1}{2}$
  - $p_1 = \frac{1}{3}, p_2 = \frac{2}{3}$
  - $p_1 = \frac{1}{5}, p_2 = \frac{4}{5}$
- Prior cycle pattern: $p_1 = p_2 = 0.5$ ($\alpha_{01} = \alpha_{02}$); Uniform time-to-event distribution within cycle
- Patients enrolled continuously with a 3-week window between cohorts ($n=3$)
- Operating characteristics
  - Accuracy (MTD and MTD-1)
  - Dose allocation (ethical)
  - Over-dose control and safety
Simulations – Results (accuracy)

- Prob. of choosing the correct MTD
- Prob. of choosing the MTD and (MTD−1)

Scenarios 1–6

- 3+3
- TITE-CRM (p1=p2=1/2)
- TITE-CRM (p1=1/3, p2=2/3)
- TITE-CRM (p1=1/5, p2=4/5)

p1=Pr(DLT in C1|DLT occurs); p2=Pr(DLT in C2|DLT occurs)
Simulations – Results (dose allocation)

Prob. of treating patients at the MTD

- 3+3
- TITE-CRM (p1=p2=1/2)
- TITE-CRM (p1=1/3, p2=2/3)
- TITE-CRM (p1=1/5, p2=4/5)

p1=Pr(DLT in C1|DLT occurs); p2=Pr(DLT in C2|DLT occurs)
Simulations – Results (over-dose control and safety)

- **Graph 1**: Prob. of choosing toxic doses (>33%) across scenarios 1–6.
  - **3+3**
  - TITE-CRM (p1=p2=1/2)
  - TITE-CRM (p1=1/3, p2=2/3)
  - TITE-CRM (p1=1/5, p2=4/5)

- **Graph 2**: Mean prob. of toxicity across doses.

p1 = Pr(DLT in C1|DLT occurs); p2 = Pr(DLT in C2|DLT occurs)
Safety data collected and collated on a weekly basis and reported to the clinician and statistician in real time.

Study statistician updates the dose-toxicity model and estimates the dose-tox curve and the MTD.

When TITE-CRM makes a dose escalation decision, the statistician meets with clinician and the team (clinPharm, translational oncologist) to discuss dose recommendation.

Clinician and SM communicate with the participating Site and recommend the dose to be administered to the next enrolled patient.
Key trial conduct considerations

• **Real-time data** (may not be QCed) to make timely and most informed dose administration

• TITE-CRM dose recommendation is the primary decision driver
  – However, AE type, AE grade, PK and PD all may play a role
  – E.g. dose escalation may be overruled or enrollment halted in case of multiple grade 2 AEs and erratic PK data in order to collect more data at the current dose level

• Create log files of statistical analyses and meeting minutes to document each dose escalation decision, and store at a central location that all key participants can access
Key trial conduct considerations (cont’d)

• Timely communication between the site and clinical team becomes more important
  – Part of Good Clinical Practice (GCP)
  – Weekly communication on patients’ status and time on treatment

• Very slow enrollment will largely reduce the time saving from the TITE-CRM. However, other benefits (such as improved accuracy and efficiency) are retained
Summary and discussion

- Primary goals of oncology phase I trials: patient safety, accuracy of dose finding, treating more patients at potentially effective doses

- TITE-CRM with adaptive weight:
  - Shortens the trial dramatically;
  - Allows a flexible definition of the MTD per the biologic mechanism;
  - Improves accuracy in MTD estimate;
  - Assigns more patients at the therapeutic zone to better inform a RP2D selection;
  - Handles patient drop-out
  - Avoids over-dosing through adaptive weight function and design calibrations

- The proposed approach can be applied to other biologic agents with cyclical injection

- Bayesian model-based phase I designs establish a framework for adaptive seamless phase I/II trials
Final Notes

• Statisticians can and should play a more important role in the design and conduct of Phase I clinical trials

• Like in most phase II and phase III studies, the study team should explore thoroughly different design options and select the optimal dose-finding approach for their study

• The time invested will be paid back with enhanced clinical trial design and improved probability of technical and regulatory success in the long run

The worst thing you can do is a weak (biased) early study, because it sets you off in wrong direction, wastes time and money, and takes so long to figure out what was wrong
Reference


• Huang B and Cheung YK . Incorporating cycle and grade information using the time-to-event continual reassessment method (in preparation)


Reference