Bayesian Optimal Designs for Phase I Clinical Trials in Oncology

Inna Perevozskaya

Merck & Co, Inc.

Collaborators:

William Rosenberger, Linda Haines, Kip Canfield
References:

- Rosenberger WF, Haines LM. Competing designs for phase I clinical trials: a review. *Statistics in Medicine*, 2002; 21; 2757-2770


Outline

- Introduction: phase I clinical trials
- Summary of available methods
- Bayesian optimal design
- Model description and assumptions
- Methodology and theory
- Major considerations
- Building user interface
- Example
- Conclusions
Phase I Clinical Trials

- Typically small, uncontrolled sequential studies
- Designed to determine the maximum tolerated dose (MTD) of the experimental drug
- Design considerations are particularly important in cancer studies (severe side effects of cytotoxic drugs)
- Certain degree of side effects is acceptable
- Accurate determination of the MTD is of grave importance since it is passed for further testing in Phase II clinical trials.
- Balance between individual and collective ethics: maximum information from the minimal number of patients.
Statistical Modeling of Phase I Clinical Trials

- Monotone relationship between the dosage and response

- Two different philosophies in MTD definition:
  1. Risk of toxicity is a sample statistic
  2. Risk of toxicity is a probability (MTD is a quantile of a monotonic dose-response curve).

- Two different approaches in designing phase I clinical trials
Summary of available methods for phase I clinical trials (Rosenberger and Haines, 2002)

1. Conventional (standard) method
2. MTD as a quantile vs. conventional method
   a) Continual reassessment method (CRM)
      O’Quigley, Pepe, Fisher (1990)
   b) Escalation with overdose control (EWOC)
      Babb, Rogatko, Zacks (1998)
   c) Decision-theoretic approaches
      Whitehead and Brunier (1995)
   d) Random walk rules (RWR)
      Durham and Flournoy (1994)
   e) Bayesian sequential optimal design
      Haines, Perevozskaya, Rosenberger (2003)
1. Conventional (standard) method

- Designed under philosophy that MTD is identifiable from the data
- Patients treated in groups of 3
- Designed to screen doses quickly; no estimation involved
- Probability of stopping at incorrect dose level is higher than generally believed (Reiner, Paoletti, O’Quigley; 1999)

First 3 patients treated at initial dose

- If no toxicities, moves to next higher dose
- If $\geq 2$ toxicities, moves to next lower dose
- If 1 toxicity, stays at the current dose

- If 1 toxicity out of 6 treated, moves to next higher dose
- If $\geq 2$ toxicities out of 6 treated, moves to next lower dose
2. The MTD as a quantile

- Dose space: $\Omega_d = \{d_1, \ldots, d_K\}$
- Binary indicators of toxicity: $Y_1, \ldots, Y_n$ for $n$ patients
- $Y=1$ if toxicity, $Y=0$ otherwise

$$P(Y_j = 1 \mid d_i) = p_i = F\left(\frac{d_i - \alpha}{\beta}\right), \quad \beta > 0$$

$$i = 1 \ldots K, \quad j = 1 \ldots n$$

- MTD defined as quantile corresponding to prespecified probability of toxicity $\Gamma \in (0,1)$

$$\mu = \alpha + \beta F^{-1}(\Gamma)$$
2a. The CRM method

- Based on a one-parameter model: \( p_i = \Psi(d_i, a), i = 1, ..., K \)
- Curves cannot cross for different \( a \in (0, \infty) \)
- CRM uses Bayes theorem with accruing data to update a prior distribution of \( a \) based on previous responses
- After each patient’s response, posterior probabilities of a toxic response at each dose \( p_i, i = 1, ..., k \) are updated
- The dose level for next patient is selected as the one with \( p_i \) closest to \( \Gamma \) in some metric
- Procedure stops after \( n \) patients enrolled
- Last patient’s dose selected is taken to be the estimate of MTD
- The method is designed to converge to the MTD
2b. Escalation with overdose control

- Assumes a location-scale family:
  \[ P(Y_{j} = 1|d_i) = p_i = F\left(\frac{d_i - \alpha}{\beta}\right), \quad \beta > 0, \quad i = 1\ldots K, \quad j = 1\ldots n \]

- The dose-response curve can be uniquely defined by two quantiles: \( \mu = \text{MTD} \) and \( \rho = \Pr \text{(toxicity at dose } d_1) \)

- one-to-one transformation \( (\mu, \rho) \leftrightarrow (\alpha, \beta) \)

- EWOC updates posterior distribution of \( \mu \) based on two-parameter model

- Introduces overdose control: predicted probability of next assignment exceeding \( \mu \) is equal to \( \varepsilon \) (Bayesian feasible design)

- EWOC is optimal in the class of the feasible designs
2c. Decision theoretic approaches

- First introduced by Whitehead and Brunier (1995)
- Incorporates elements of Bayesian Decision Theory
- Actions: assigned dose levels \( d_1, \ldots, d_K \)
- Two-parameter model with priors on \((\alpha, \beta)\)
- Loss function: minimizing asymptotic variance of
  \[
  \mu = \alpha + \beta \, F^{-1}(\Gamma)
  \]
  \[
  \text{var}\left(\hat{\mu}|x_1,\ldots,x_{j-1},d\right) = \text{var}(\hat{\alpha}) + \text{var}(\hat{\beta}) \left\{ F^{-1}(\Gamma) \right\}^2 + 2 \, \text{cov}(\hat{\alpha}, \hat{\beta}) \, F^{-1}(\Gamma)
  \]
- Posterior means of \(\alpha\) and \(\beta\) are substituted into the above equation
2d. Random Walk Rules

- Nonparametric approach;
- Generalizes the up-and-down approach of the conventional method;
- Creates a unimodal distribution around the target quantile;
- Consequently, some patients will be assigned above the MTD.
Bayesian Optimal Sequential Design

- The methodology is similar to decision-theoretic approach, i.e. principally concerned with efficiency of estimation
- Based on formal theory of optimal design (Atkinson and Donev, 1992)
- Similar to EWOC, a constraint is added to address the ethical dilemma of avoiding extremely toxic doses
- General methodology developed for the case when the dose space is unknown (continuous dose space)
- Case when doses are fixed in advance is particularly important in practice (discrete dose space)
- Sequential procedure is developed based on discrete designs
Model and Assumptions

- $Y_1, \ldots, Y_n$ binary indicators of toxicity for $n$ patients
- $d_1, \ldots, d_K$ distinct doses of administered drug
- A two parameter model is used with logistic link function defining the dose response curve:

$$p_i = \Pr\{Y_j = 1 \mid d_i\} = F \left( \frac{d_i - \alpha}{\beta} \right) = \frac{1}{1 + e^{-\frac{d_i - \alpha}{\beta}}}, \beta > 0,$$

$$i = 1, \ldots, K \quad j = 1, \ldots, n$$
Model and Assumptions (cont.)

- Quantile of interest:

\[
\text{MTD} : Pr\{Y = 1|\mu\} = \Gamma, \quad \Gamma \in (0,1), \quad \mu = \alpha + \beta \logit(\Gamma)
\]

- Design: \(n\) patients assigned to \(K\) distinct doses \(d_1,\ldots, d_k \in \Omega_d\)

\[
\xi = \{d_1,\ldots, d_k, w_1,\ldots, w_K\}
\]

\(N_i\) – number assigned to each dose, \(i = 1\ldots K\)

\(w_i = N_i / N\) – design weights
Information Matrix and Optimality Criterion

- Fisher’s information matrix for \( \theta = (\alpha, \beta) \) at a design point \( d \):
  \[
  I(d, \theta) = \frac{e^z}{\beta^2(1+e^z)^2} \begin{bmatrix} 1 & z \\ z & z^2 \end{bmatrix}, \quad z = \frac{d - \alpha}{\beta}
  \]

- Full design information matrix:
  \[
  M(\xi, \theta) = \sum_{i=1}^{K} w_i I(d_i, \theta)
  \]

- Optimization criterion:
  \[
  \phi_D(\xi) = E[\log \det M(\xi, \theta)] = \int_{\Theta} \log \det M(\xi, \theta) g(\theta) \, d\theta
  \]
Constraints

- Let $\mu_R$ be quantile corresponding to undesirable probability of toxicity $\Gamma_R$
- Dose space restriction: $\Omega_R = \{d : d \leq \mu_R\}$
- Constraint function:

$$\phi_R(\xi) = \sum_{i=1}^{K} w_i \Pr(\mu_R \leq d_i) \leq \varepsilon$$
Constrained Bayesian Optimal Design

- Maximize
  \[
  \phi_D(\xi) = E[\log \det M(\xi, \theta)] = \int_\Theta \log \det M(\xi, \theta) g(\theta) d\theta
  \]

- Subject to the constraint
  \[
  \phi_R(\xi) = \sum_{i=1}^K w_i \Pr(\mu_R \leq d_i) \leq \varepsilon
  \]

- Pilot phase: allocate first \(n_0\) patients according to the optimal design given prior information

- Rounding algorithm by Pukelsheim (1993) used to achieve integer allocation of \(n_0\) patients to \(K\) doses
Sequential Stage of the Optimal Design

- Pilot phase data $D_0$ (responses and dose assignments) of first $n_0$ patients obtained
- Prior density $g(\theta)$ is updated with posterior density $g(\theta|D_0)$
- Second stage: stepwise allocation of patients to the doses that maximize
  $$\int_{\Theta} \log \det \left( n_0 M(\xi^*_D, \theta) + I(d, \theta) \right) g(\theta|D_0) d\theta$$
- Subject to the constraint $\Pr(\mu_R \leq d) \leq \varepsilon$ evaluated over the posterior density $g(\theta|D_0)$
Major Considerations

- Prior elicitation
- Constraints
- Numerical integration
- Pilot phase of the sequential design
- Creating the user interface
Prior Elicitation and Constraints

- Physician’s pharmacologic and toxicologic knowledge is critical in determining prior distributions
- Our procedure: uniform prior placed either directly on \((\alpha, \beta)\) or on two nominated quantiles \((\mu_1, \mu_2)\) based on physician’s range guess
- Constraints: Bayesian version of \(d \leq \mu_R\): \(\sum_{i=1}^{K} w_i \Pr(\mu_R \leq d_i) \leq \varepsilon\)
- Involves choice of \(\Gamma_R \leftrightarrow \mu_R\) (undesirable toxicity) and \(\varepsilon\) (tolerance level)
- If \(\Gamma_R = \Gamma\), no patients are assigned above MTD \(\Rightarrow\) loss of efficiency
- If \(\Gamma_R > \Gamma\), allows more flexibility:
  - Example: \(\Gamma = 0.25, \Gamma_R = 0.5, \varepsilon = 0.01\)
  - If \(\varepsilon = 0\), all dose assignments are below min of the range of \(\mu_R\)
  - If \(\varepsilon = 0.5\), all dose assignments are below median of \(\mu_R\)
Numerical Integration

- Numerical integration is crucial for implementation of any Bayesian procedure.

- Gaussian quadrature and MCMC are standard methods.
  - Quadrature can be problematic when the denominator integral value is small.
  - MCMC is more accurate and efficient but less straightforward to implement.

- Caveat: In the sequential optimal design procedure both were inappropriate.
  - Multiple integral evaluations fed into optimization routine.
  - Adaptive nature of quadrature and MCMC makes the results highly variable.
  - This variability causes lack of convergence of the optimization routine.

- Easy work-around: discrete prior on a uniformly spaced grid.
Pilot Phase

- The goal of the pilot phase is to provide a starting point for sequential procedure
- Size of the pilot phase $N_0$ is user specified
- Should be as small as possible
- Avoid assigning many patients based on prior which could be wrong
- More patients should be assigned based on updated posterior reflecting accruing information
- Simulation studies: $N_0=5$, $N_0=10$, and $N_0=15$ provided comparable results
Creating User-Friendly Interface: iDose (Interactive Doser) Software

//http://haggis.umbc.edu/cgi-bin/dinteractive/inna1.html

- Web-based application is available to any workstation equipped with a web browser
  - High availability and simple deployment
  - Ease of integration with patient record systems
  - The ability to support long transactions
  - Ease of use for clinicians

- An example: screen shots from actual web application are shown
Figure 1: Initial data entry screen for iDose

Interactive Doser (Demo version)

Bayesian Optimal Dosing for Phase I Clinical Trials

Enter the following values for the design (Mouse over item for help):

- Number of patients for initial phase: 5
- Total Number of Patients: 8
- Dose levels to be used (Enter a space delimited list of doses): 1000 400 600 900 1200
- MTD: Percentile to be estimated: .25
- Range on LD25 d1: 500 d2: 700
- Range on LD50 d3: 800 d4: 886
- Restriction percentile: .5
- Probability of assigning above the restriction percentile: .1

Enter a range for your best guess of LD50, but the range cannot overlap with LD25.
Figure 2: Data entry for the initial phase

Interactive Doser (Demo Version)

- Number of patients for initial phase: 5
- Total Number of Patients: 8
- Dose levels to be used: 100 300 600 900 1200
- Problem Type: D_optimality
- constraintID: restricted
- Gamma Restriction: 0.5
- Tolerance: 0.1
- Range on LD25: [500, 700]
- Range on LD50: [800, 886]
- Prior Distribution: unif
- Delta step for integration: 20

- Dose Levels:
  100 | 300 | 600 | 900 | 1200

- Number of Patients at each Dose:
  0   2   2   0   1

Enter a space delimited list of the toxicities at each of the 5 doses (Mouse over the input for help):
- [0 0 1 0 1]

Enter a space delimited list of the number of toxicities at each dose (like [0 0 1]) inside the brackets.
**Figure 3: Entering the dose for a patient**

Interactive Doser (Demo Version)

**Program Values**

- Number of toxicities at each dose: 0 0 1 0 1
- Dose levels to be used: 100 300 600 900 1200

Continue on to submit the dose for next patient *(Mouse over input for help)*:

- **Next Dose**: 600

You can accept the programmatically suggested dose or choose another from: 100 300 600 900 1200. If you do change the value, the procedure will NOT give an optimal solution.
Figure 4: Final Summary

Interactive Doser (Demo Version) Final Summary

Program Values

- Number of patients for initial phase: 5
- Total Number of Patients: 8
- Dose levels to be used: 100 300 600 900 1200
- Problem Type: D_optimality
- constraintID: restricted
- Gamma Restriction: 0.5
- Tolerance: 0.1
- Range on LD25: [500, 700]
- Range on LD50: [800, 886]
- Prior Distribution: unif
- Delta step for integration: 20
- Number of toxicities at each dose: 0 0 3 0 1
- Number of patients at each dose: 0 2 5 0 1
- New Patient Responses: 0 1 1
- New Patient Assignments: 600 600 600
- New point: 600 with toxicity: 1 at iteration: 3

Summary Statistics

- Posterior mean of the target quantile: 573.749
- Standard deviation of the target quantile: 50.5355
- Summary Graphs (opens in a new window).

- Done with optimal solution for 3 added Patients.
- Start all over again
Figure 5: Final Summary graphic
Conclusions

- A web-based interface implemented
  - Uses powerful Bayesian optimal design theory
  - Patients assigned sequentially according to the procedure allowing efficient estimation of MTD with as few patients as possible.
  - Only guessed ranges of LD25 and LD50 are required to start Bayesian updating
  - Width of ranges should incorporate clinician’s degree of uncertainty as well as best judgment on the dose-response relationship
  - User can override the suggested dose assignment, the procedure will no longer be optimal

- Flexibility of the software:
  - Progress of the trial can be monitored (parameter estimates, prior updates)
  - \((\alpha, \beta)\) estimates provide complete information about the dose-response curve \(\Rightarrow\) Any quantile of interest can be estimated
Conclusions (cont.)

- Patient consent: formal constraint on dose space should be attractive to patients concerned about toxicity

- Patients in the initial phase will be allocated strictly on the basis of the clinicians prior knowledge about dose-response curve

- This allocation should be conservative if there is little prior knowledge about the dose-response relationship