Three Methods for Phase I/II Clinical Trials, with Application to Allogeneic Stem Cell Transplantation

Peter F. Thall, PhD
Biostatistics Department
M.D. Anderson Cancer Center

Workshop on Clinical Trial Endpoints for Acute Graft-vs-Host Disease after Allogeneic Hematopoietic Stem Cell Transplantation

May 19, 2009
Financial Conflicts of Interest

• None
Standard “3+3” Phase I Designs for Oncology

- Objective: Identify the maximum tolerated dose (MTD)
- MTD defined by algorithm: *Implicitly* either 17% or 33% grades 3-5 AE
- But the incidence of grade 3 AEs far exceeds 17-33% for BMT patients, so this design is rarely applicable for this population of patients
- Well known to have inferior properties compared to Bayesian adaptive designs
## Three Phase I and Phase I-II Designs

<table>
<thead>
<tr>
<th>Treatment Optimized</th>
<th>Outcome</th>
<th>Decision Criterion</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>Bivariate binary (phase I-II)</td>
<td>Efficacy-toxicity trade-offs</td>
<td>GVHD prophylaxis, anergized cells post allotx, etc.</td>
</tr>
<tr>
<td>Dose and Schedule</td>
<td>Time to toxicity (phase I)</td>
<td>Pr(toxicity by day $t^*$)</td>
<td>Azacitidine post allotx</td>
</tr>
<tr>
<td>Doses of two agents (dose$_1$, dose$_2$)</td>
<td>Bivariate ordinal (phase I-II)</td>
<td>Utility of outcome</td>
<td>Bladder cancer</td>
</tr>
</tbody>
</table>
All methods are

**Bayesian**: Model parameters are considered to be RANDOM quantities

**Sequentially Outcome Adaptive**:  
Choose a treatment  
(dose, dose-schedule, dose pair)  
Treat a cohort of patients  
Observe the patients’ outcomes

Repeat until a stopping rule says “Stop”
### Three Phase I and Phase I-II Designs

<table>
<thead>
<tr>
<th>Treatment Optimized</th>
<th>Outcome</th>
<th>Decision Criterion</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>Bivariate binary (phase I-II)</td>
<td>Efficacy-toxicity trade-offs</td>
<td>GVHD prophylaxis, anergized cells post allotx, etc.</td>
</tr>
<tr>
<td>Dose and Schedule</td>
<td>Time to toxicity (phase I)</td>
<td>Pr(toxicity by day t*)</td>
<td>Azacitidine post allotx</td>
</tr>
<tr>
<td>Doses of two agents</td>
<td>Bivariate ordinal (phase I-II)</td>
<td>Utility of outcome</td>
<td>Bladder cancer</td>
</tr>
</tbody>
</table>
Dose-Finding Based On Efficacy-Toxicity Trade-Offs
(Thall and Cook, 2004; Thall, Cook and Estey, 2006)

Patient Outcome = \{Efficacy, Toxicity\}
- each a binary indicator
\[
\pi_E(x) = \Pr(\text{Efficacy at dose } = x)
\]
\[
\pi_T(x) = \Pr(\text{Toxicity at dose } = x)
\]

MD must specify:
→ A Lower Limit on \( \pi_E(x) \) (minimum response of interest)
→ An Upper Limit on \( \pi_T(x) \) (maximum acceptable toxicity)
→ Three or more equally desirable \( (\pi_E, \pi_T) \) targets…
Two Dose Acceptability Criteria

Target pairs are used to construct an **Efficacy-Toxicity Trade-off Contour**… and a family of Contours each with desirability, $\delta$, for the $(\pi_E, \pi_T)$ pair.
Which of these two \( \pi \) pairs is more desirable?
Trial Conduct

1) The physician chooses the **starting dose**
2) A dose is **Acceptable** if either
   a) it has acceptable $\pi_E$ & $\pi_T$ or
   b) it is the lowest untried dose and has acceptable $\pi_T$
3) **Treat each cohort at the current most desirable dose**
   a) The dose chosen for the next cohort may be *higher than, the same as, or lower than* the current dose
   b) After de-escalation due to *excessive toxicity* or *low efficacy*, if subsequent outcomes at a lower dose are sufficiently safe and efficacious, then *the algorithm may re-escalate*
4) Do not skip untried doses
5) No dose acceptable $\rightarrow$ **Stop the trial**
6) **At the end, select the most desirable dose**
Pentostatin for Graft-Versus-Host Disease

Patients with steroid-refractory GVHD after allotx from an HLA-matched donor
Doses: \( x = 0.25, 0.50, 0.75, \) or \( 1.00 \text{ mg/m}^2 \)
\( N_{\text{max}} = 36, \) cohort size = 3
First cohort treated at \( 0.25 \text{ mg/m}^2 \)

**Toxicity** = \{Infection unresolved by antibiotics, or death, within 2 weeks\}

**Efficacy** = \{ \geq 1 \text{ grade drop in GVHD severity, within 2 weeks}\}

\( .40 = \text{Upper Limit on } \pi_T(x) \)
\( .20 = \text{Lower Limit on } \pi_E(x) \)
Simulation Scenarios for the Pentostatin Trial

Scenario 1

Scenario 2

Scenario 3

Scenario 4
Dose Selection Probabilities

Scenario 1

Scenario 2

Scenario 3

Scenario 4
Conclusions

The *Trade-Off-Based Algorithm* reliably

1) Finds Safe Doses having High Efficacy

2) **Stops** if no dose is acceptable

Implementation is Hard Work, but a free computer program is available!
## Three Phase I and Phase I-II Designs

<table>
<thead>
<tr>
<th>Treatment Optimized</th>
<th>Outcome</th>
<th>Decision Criterion</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dose</strong></td>
<td>Bivariate binary (phase I-II)</td>
<td>Efficacy-toxicity trade-offs</td>
<td>GVHD prophylaxis, anergized cells post allotx, etc.</td>
</tr>
<tr>
<td><strong>Dose and Schedule</strong></td>
<td><strong>Time to toxicity (phase I)</strong></td>
<td><strong>Pr</strong>(toxicity by day $t^*$)</td>
<td>Azacitidine post allotx</td>
</tr>
<tr>
<td>Doses of two agents (dose$_1$, dose$_2$)</td>
<td>Bivariate ordinal (phase I-II)</td>
<td>Utility of outcome</td>
<td>Bladder cancer</td>
</tr>
</tbody>
</table>
Goal: Optimize \( (\text{Dose, Schedule}) \) based on \textbf{Time to Toxicity}

\textit{Vidaza\textsuperscript{\textregistered}} (azacitidine) given post allotx in AML pts

- Dose-toxicity profile of \textit{Vidaza\textsuperscript{\textregistered}} unknown
- Cumulative toxicity of repeated administration (multiple 28-day cycles) unknown
Patient Outcome

- \( T = \textit{Time from the start of treatment to toxicity} \)
- Usual “time-to-event” data, as in a survival time analysis. A patient’s outcome consists of
  a) \textit{Time to toxicity} if it occurred, or \textit{Time to last follow up} if toxicity has \textit{not} occurred
  b) An \textit{indicator} of whether toxicity has occurred
- Why is “time-to-event” better than a binary outcome? Using a usual binary (Yes / No) indicator of
  [“Toxicity” within 28 days from the start of therapy]
  - A patient with toxicity at day 27 is scored “Yes”
  - A patient with toxicity at day 29 is scored “No”
  - A patient followed for only 25 days w/o toxicity is inevaluable and \textit{cannot be scored}
Trial Conduct

1) Treat 1st patient at the lowest \((\text{dose, schedule})\)
2) Using current **Time-to-Toxicity data**, treat each patient at the \((\text{dose, schedule})\) pair with \(p_{t\text{o}} = \Pr(\text{Toxicity by day } t^* \mid \text{dose, schedule})\) closest to the target max toxicity rate
3) Do not “skip” untried \((\text{dose, schedule})\) pairs
4) If no \((\text{dose, schedule})\) pair is acceptable \(\rightarrow\) Stop the trial
Hazard of toxicity from 1 cycle

Cumulative hazard of toxicity to day 10

Cumulative hazard of toxicity from multiple cycles

\[ H(10) \]

\[ \text{Prob(Toxicity by day 10)} = 1 - e^{-H(10)} \]
What Actually Happened in the Vidaza® Trial?

- **Treatment parameters**
  - Vidaza doses 8, 16 or 24 mg/m² daily x 5 in each cycle
  - Given for 1, 2, 3 or 4 28-day cycles

- **Definition of toxicity**
  - Severe (NCI grade 3 or 4) kidney, liver, heart, lung or neural toxicity
  - Severe GVHD
  - Systemic infection not resolved by antibiotics within two weeks
  - Severe haematologic toxicity

- **ptoxt = Pr(Toxicity by day 116 | dose, schedule)** closest to the tox target 0.3

- Only 1 toxicity in 27 patients, so 4 more dose levels 32,40,48,56 added

- Optimal dose-schedule identified after 44 patients:
  
  (40 mg/m² x 3 cycles)
Conclusions

The *Dose-Schedule Algorithm* reliably

1) Finds (Dose, Schedule) pairs having specified $\text{Pr}(\text{Toxicity by day } t^*)$

2) *Stops* if no (Dose, Schedule) is acceptable

Implementation is Hard Work, but a free computer program is available!
## Three Phase I and Phase I-II Designs

<table>
<thead>
<tr>
<th>Treatment Optimized</th>
<th>Outcome</th>
<th>Decision Criterion</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>Bivariate binary</td>
<td>Efficacy-toxicity trade-offs</td>
<td>GVHD prophylaxis, anergized cells post allotx, etc.</td>
</tr>
<tr>
<td>Dose and Schedule</td>
<td>Time to toxicity</td>
<td>Pr(toxicity by day $t^*$)</td>
<td>Azacitidine post allotx</td>
</tr>
<tr>
<td>Doses of two agents (dose$_1$, dose$_2$)</td>
<td>Bivariate ordinal</td>
<td>Utility of outcome</td>
<td>Bladder cancer</td>
</tr>
</tbody>
</table>
Optimizing the dose pair of a two-agent combination based on elicited utilities of (Toxicity, Efficacy) outcomes


**Goal**: Optimize (Dose of 2 agents) based on **Toxicity** and **Efficacy**

Treatment of bladder cancer with a combination of chemotherapy (c) and a biologic (b) where optimal doses in combination are unknown
### Dose-Combination \((b_x, c_Y)\) Matrix

<table>
<thead>
<tr>
<th></th>
<th>(1,3)</th>
<th>(2,3)</th>
<th>(3,3)</th>
<th>(4,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,2)</td>
<td>(2,2)</td>
<td>(3,2)</td>
<td>(4,2)</td>
<td></td>
</tr>
<tr>
<td>(1,1)</td>
<td>(2,1)</td>
<td>(3,1)</td>
<td>(4,1)</td>
<td></td>
</tr>
</tbody>
</table>

\(b_x = \text{dose of biologic agent}\)

\(c_Y = \text{dose of chemo agent}\)
Patient Outcome is \((\text{Response, Toxicity})\)

<table>
<thead>
<tr>
<th>Response</th>
<th>0 = PD</th>
<th>1 = SD</th>
<th>2 = CR/PR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0, 0)</td>
<td>(0, 1)</td>
<td>(0, 2)</td>
</tr>
<tr>
<td>1</td>
<td>(1, 0)</td>
<td>(1, 1)</td>
<td>(1, 2)</td>
</tr>
<tr>
<td>2</td>
<td>(2, 0)</td>
<td>(2, 1)</td>
<td>(2, 2)</td>
</tr>
</tbody>
</table>

Allows the possibility that Response may be inevaluable
**Elicited Consensus Utilities**

### Response

<table>
<thead>
<tr>
<th></th>
<th>PD</th>
<th>SD</th>
<th>CR/PR</th>
<th>Inevaluable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toxicity</td>
<td>25</td>
<td>76</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>60</td>
<td>82</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>40</td>
<td>52</td>
<td>0</td>
</tr>
</tbody>
</table>
Very Flexible Dose-Outcome Model
Trial Conduct

Choose each cohort’s dose pair to

**Maximize the Posterior Expected Utility**

based on the data observed so far

**Do Not Skip Untried Doses:**

If \((b_1,c_1)\) is the current dose pair, then escalation is allowed to as yet untried pairs \((b_2,c_1), (b_1,c_2),\) or \((b_2,c_2)\)

Stop the trial if all dose pairs are unacceptably toxic
Toxicity =
0 if NO GVHD
1 if grade 1,2 GVHD
2 if grade 3,4 GVHD
or
0 if NO grade 3,4 GVHD
1 if grade 3,4 GVHD but resolved in <2 wks
2 if grade 3,4 GVHD not resolved in < 2 wks
Application to Trials Monitoring GVHD

Efficacy =
0 if dead, or alive but no response at day 100
1 if alive and engrafted with PR at day 100
2 if alive and engrafted with CR at day 100
   (e.g. for CLL transplantation trials)

or

0 if dead, or no plt recovery in 100 days
1 if alive with $20 < \text{plt} < 50$ at day 100
2 if alive with $\text{plt} > 50$ by day 100
   (e.g. for cord blood transplantation trials)
Extensive Computer Simulations Show that the Utility-Based Dose-Finding Method is Very Reliable and Very Safe.

Implementation is Hard Work, but a free computer program is available!
## Phase I and I/II Designs for GVHD Trials

<table>
<thead>
<tr>
<th>Design</th>
<th>Objective</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3+3</td>
<td>MTD</td>
<td>Easy to do, poor properties, rarely applicable to BMT patients</td>
</tr>
<tr>
<td>Accelerated titration</td>
<td>MTD</td>
<td>Acceptable for relatively nontoxic agents, but rarely applicable (like 3+3)</td>
</tr>
<tr>
<td>CRM, mCRM</td>
<td>MTD</td>
<td>Stat-intensive, flexible for toxicity target, find dose based on toxicity</td>
</tr>
<tr>
<td>Time-To-Tox</td>
<td>Max tolerated dose and schedule combo</td>
<td>Stat-intensive, flexible for toxicity target, finds dose and schedule</td>
</tr>
<tr>
<td>Eff-Tox</td>
<td>Best dose based on toxicity and efficacy</td>
<td>Stat-intensive, optimizes efficacy and toxicity jointly</td>
</tr>
<tr>
<td>Doublet Studies</td>
<td>Best combo based on toxicity and efficacy</td>
<td>Stat-intensive, optimizes efficacy and toxicity jointly</td>
</tr>
</tbody>
</table>
References


Software at: http://biostatistics.mdanderson.org/SoftwareDownload