Two Bayesian designs for first-in-human trials in cancer

Nedjad Losic
March 31, 2017
Agenda

Two Bayesian designs for first-in-human trials in cancer

- Quick intro to first-in-human trials in cancer
- Continual Reassessment Method (CRM)
  - modified CRM (mCRM)
- Bayesian Optimal Interval Design (BOIN)

Credits & Thank You

CRM
Dr Ulf Forssmann, Sr Medical Director, Genmab A/S
- advocate of CRM + significant modifications
Kert Viele & Anna McGlothlin, Berry Consultants Inc.
- Implementation and advice

BOIN
Dr Ulf Forssmann

CRM + BOIN
Henning Friis Andersen, Genmab A/S
Thomas Bayes (1701?-April 7 1761)

• Nonconformist minister
  • Tunbridge wells, 70 km SE of London

• No mathematical/statistical publications
• Unknown/uninfluential on his contemporaries
• Made Fellow of the Royal Society 1741
• Richard Price read his work to the RS on Dec 23 1763
  • “An assay towards solving a Problem in the Doctrine of Chances” (1764)
• One of the most widely known eponyms in Science today

• Laplace, independently, developed same/similar ideas 1774

Dose Limiting Toxicities (DLTs)

Table 6-2: Criteria for defining dose-limiting toxicities

<table>
<thead>
<tr>
<th>Toxicity</th>
<th>Any of the following criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematology</td>
<td>1. CTCAE grade 3 neutropenia (ANC &lt; 1.5 x 10^9/L) 2. CTCAE grade 3 thrombocytopenia (platelets &lt; 10 x 10^9/L) 3. CTCAE grade 3 anemia (Hgb &lt; 9 g/dL)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>1. CTCAE grade 3 fatigue</td>
</tr>
<tr>
<td>Renal</td>
<td>Serum creatinine &gt; 2 x ULN</td>
</tr>
<tr>
<td>Hepatic</td>
<td>3. CTCAE grade 3 total bilirubin (TB) &gt; 1.5 x ULN 2. CTCAE grade 2 total bilirubin and CTCAE grade 2 ALT</td>
</tr>
<tr>
<td>Pancreatic</td>
<td>1. CTCAE grade 2 pancreatitis</td>
</tr>
<tr>
<td>Cardiac</td>
<td>1. CTCAE grade 3</td>
</tr>
<tr>
<td>Dermatologic</td>
<td>0. CTCAE grade 2 photosensitivity</td>
</tr>
<tr>
<td>Other adverse events</td>
<td>Any 2 CTCAE grade 3</td>
</tr>
<tr>
<td>All others</td>
<td>CTCAE grade 3 vomiting or nausea despite optimal anti-emetic therapy</td>
</tr>
<tr>
<td>All others</td>
<td>CTCAE grade 3 diarrhea despite optimal anti-diarrheal treatment</td>
</tr>
<tr>
<td>All others</td>
<td>Any 2 CTCAE grade 3 AE, except for the exclusions noted below</td>
</tr>
</tbody>
</table>

Exceptions to DLT criteria:
CTCAE grade 1 or 2 + eradication of alkaline phosphatase or
CTCAE grade 2, and resolution of all other AEs

Accelerated Titration Design (ATD)

- Just like 3+3, but with one difference
  - Initial cohorts: single-patient cohorts.
  - With or without intra-patient dose-escalation
  - Continue with single-patient cohorts until:
    - DLT (or other relevant toxicity) seen, or
    - reached a "high" dose-level
  - Thereafter, continue as 3+3

![ATD Design Diagram](image)
Trial designs in Phase I Cancer Trials

- Estimate from 2007\(^1\)
  - 1991-2006: 1235 abstracts from **Phase I Cancer Trials**
  - 98.4% step-up-step-down designs
  - 1.6% (n=20) Bayesian adaptive designs

- More recent estimate:
  - 49% 3+3
  - 40% Accelerated Titration Design
  - 10% Bayesian CRM

\(^1\) Rogatko A et al. Translation of Innovative Designs Into Phase I Trials, JCO, 25; 31, pp 4982-4986, 2007

Use of CRM at Novartis

- **Novartis:**
  - Before 2000: the 3+3 design
  - In 2000
    - two trials with CRM
    - both failed (too aggressive dose-recommendations)
  - ~2004: another attempt (2-parameter Bayesian Logistic Regression)
    - Success
  - 2005: CRM is the new Novartis standard
    - Global phase I and Ib: 100%
    - > 60 trials, >30 compounds, >20 FIH

FDA-Industry Workshop 2015, Roychoudhury, Neuenschwander, Wandel, Bayesian Adaptive Phase I Oncology Trials, September 2015
CRM 1/2

- Start by assuming a functional relationship between Dose and DLT:
  \[ \log\left(\frac{p_{DLT}}{1-p_{DLT}}\right) = \alpha + \beta \cdot \text{dose} \]
- Bayesian logistic regression: \[ \alpha, \beta \text{ are not fix parameters but have distributions} \]
- Early: \( \alpha \) fixed (e.g. 3) ⇒ one-parameter logistic regression model
- NB: not actual doses used in model: “dose labels” or “standardized doses” used
- Define Target Toxicity Level (TTL): e.g. 17%-33%
  - The aim is to have TTL DLT-rate on MTD

CRM 2/2

The Continual Re-assessment Method

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assume prior for ( \beta )</td>
</tr>
<tr>
<td>2</td>
<td>Treat 1 patient, at dose level closest to current estimate of the MTD</td>
</tr>
<tr>
<td>3</td>
<td>Observe DLT outcome</td>
</tr>
</tbody>
</table>
| 4    | Compute posterior and update \( \hat{\beta} \): Treat the next patient at the level closest to the updated estimate of the MTD, based on posterior distribution of \( \hat{\beta} \).  
Treat the next patient at model-based MTD estimate:  
\[ d_{i+1} = \text{argmax}_{d} \left\{ p(d, \hat{\beta}) - \text{TTL} \right\} \]  
where \( p(d, \beta) \) is probability of DLT on dose-level \( d \),  
\[ \hat{\beta}_{i} = \frac{\int L_{i}(\beta; d, y) dF(\beta)}{\int L_{i}(\beta; d, y) dF(\beta)} \]  
as well as  
\[ L_{i}(\beta; d, y) = \prod_{j=1}^{i} p(d_{j}, \beta) p(DLT_{j} | 1 - p(d, \beta)) \]  
\( F(\beta) \): a priori distribution for \( \beta \), \( d_{j} \): dose level for patient \( j \), DLT\(_{j} \): DLT outcome (0, 1) for patient \( j \) |
| 5    | Repeat Steps 1-5 until sufficient precision in estimate of \( \hat{\beta} \), or \( N_{\text{max}} \) reached.  
MTD= the dose that would have been given to the \( (N+1) \)st patient.  
MCMC: method for numerical integration, e.g. MCMC |
Safety concerns with the original CRM

Safety concerns:
• Starts at the expected MTD
• Goes straight for the MTD

Modifications proposed: modified CRM

Escalation With Overdose Control (EWOC)

Limit the posterior probability of choosing a dose that exceeds the MTD

Modified CRM in GCT1021-01

- First-in-Human: Do not start at MTD, start low

- 3 patients per cohort
  - However, single-patient cohorts the first 2 dose-levels

- Main dose-levels and intermediate dose-levels
  - Escalate on main dose-levels until DLT observed, then intermediate dose-levels available

- Restricted dose-allocation: escalate one main dose-level at the time

- Escalation With Overdose Control
GCT1021-01 –before we start

• Assumed DLT-rates at dose-levels
  • 8 different scenarios

• Target Toxicity Level (on MTD): 22%

• Escalation with overdose control (EWOC)
  • Escalate to a “safe” dose level; level safe if $P_{DLT}(d) < 22% > 40%$.

• Total $N_{max} = 41$,
  • need 20-30 patients for CRM to work

• Prior distribution....

mCRM in GCT1021-01 – historical data for the prior

Overview of start dose and MTD in some (MMAE-) ADC Phase 1 trials

<table>
<thead>
<tr>
<th>Drug</th>
<th>Name</th>
<th>Company</th>
<th>Target</th>
<th>Linker</th>
<th>Indication</th>
<th>Ph1 Doses (mg/kg)</th>
<th>Ph1 Regimen</th>
<th>MTD (mg/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adcetris</td>
<td>Brentuximab vedotin</td>
<td>Seattle Genetics</td>
<td>CD30</td>
<td>vc</td>
<td>SK &amp; ALCL</td>
<td>0.1-1.8</td>
<td>Q1W</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.2-2.7*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUK-011</td>
<td>Debatuximab vedotin</td>
<td>Celldex</td>
<td>GFRAMB</td>
<td>vc</td>
<td>Breast</td>
<td>Q1W</td>
<td>Q2/3W</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03-2.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDT12805</td>
<td>Trinituximab vedotin</td>
<td>Genentech/Roche</td>
<td>CE22</td>
<td>PABC</td>
<td>9% &amp; DBLCL</td>
<td>Q3W</td>
<td></td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.3-3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSMA-ADC</td>
<td>Prograna</td>
<td>Prograna</td>
<td>PSMA</td>
<td>vü</td>
<td>mast/ERCC</td>
<td>Q3W</td>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4-2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DC015401A</td>
<td>Trinituximab vedotin</td>
<td>Genentech/Roche</td>
<td>CC79H</td>
<td>PABC</td>
<td>9% &amp; DBLCL</td>
<td>Q3W</td>
<td></td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.3-1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS-1</td>
<td>Agenys</td>
<td>Agenys</td>
<td>SLCA44A</td>
<td>vü</td>
<td>Prostate, gastric &amp; pancreas</td>
<td>Q1W</td>
<td>Q3W</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Astellas)</td>
<td></td>
<td></td>
<td></td>
<td>0.3-1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABN0064</td>
<td>Millennium</td>
<td>Millennium</td>
<td>Guanylyl</td>
<td>NO</td>
<td>GI</td>
<td>Q3W</td>
<td></td>
<td>ND but &gt;1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cyclosine C</td>
<td></td>
<td></td>
<td>0.3-1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HuMax-1F-</td>
<td>Tisotuximab vedotin</td>
<td>Genmab</td>
<td>TF</td>
<td>vc</td>
<td>Solid tumors</td>
<td>Q3W</td>
<td></td>
<td>2.0</td>
</tr>
<tr>
<td>ADC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.3-2.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Refer to:
2. No DLTs in doses up to 1.8 mg/kg in first 10 patients
3. [Comparative clinical pharmacokinetics of antibody-drug conjugates in first-in-human Phase 1 studies, mAbs 6,4:12; July/August 2014](http://ir.progenics.com/irdata.cfm?ItemNumber=irData416585)
4. Decrux, A., Comparative clinical pharmacokinetics of antibody-drug conjugates in first-in-human Phase 1 studies, mAbs 6,4:12; July/August 2014
5. No DLTs in doses up to 1.8 mg/kg in first 10 patients
6. Generics

Mean MTD = 2.1 mg/kg
mCRM– Design Calibration

mCRM in GCT1021-01

- Allows for flexibility in cohort sizes
  - In case of a drop-out: 2 patients, or 5 patients
  - In case of over-recruitment: 4 patients, 7 patients ...

- Better estimate of MTD

- More patients exposed to efficacious dose-levels
  - Efficacy information available earlier, before cohort-expansion
Example from GEN501

Probability of at least one DLT in the next 3 patients, based on 5% DLT-rate: $1-(1-0.05)^3 = 14.3\%$.

Other version of the CRM

- Many modifications of CRM
- TITE-CRM (time to event CRM)
- Pharmacologically guided CRM
- Maximum Likelihood-versions of CRM
- ...
Regulatory guidelines

- **Bayesian statistics**
  - ICH E9
    - Just mentions that it exists
  - FDA
    - "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials".
    - 'Non-medical-device'-divisions (CDER/CBER) refer to it.
  - EMA
    - No specific guidance
    - Mentioned in other guidances, e.g. 'Guideline on clinical trials in small populations': "Such [Bayesian] methods may be advantageous when faced with small datasets, although introducing prior beliefs is often a concern in drug regulation."

Bayesian methods

**Well accepted**
- Exploratory/descriptive: Dose ranging, Dose response relationship
- Confirmatory: Formally compare New vs. Control
- WIT: Sensitivity analyses on the a priori distribution

**Phase**
- Phase I: Descriptive Statistics
- Phase II: Hypothesis test + DS + modelling
- Phase III: DS + modelling + hypothesis tests
- Phase IV: Hypothesis test + DS
Stakeholder interactions

• Internal

• External
  • KOLs
  • Regulatory authorities

Software - CRM

<table>
<thead>
<tr>
<th>SAS</th>
<th>Software - CRM</th>
<th>PROC IML</th>
<th>PROC MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ed. Menon, SM &amp; Zink RC, Modern Approaches to Clinical Trials Using SAS, Classical Adaptive and Bayesian Methods, SAS Institute, 2015</td>
<td>PROC IML</td>
<td>PROC MCMC Example 54.3 for inspiration</td>
<td></td>
</tr>
<tr>
<td>DIY</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R</th>
<th>CRM</th>
<th>1-parameter hyperbolic or 1-parameter logistic CRM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OFCRM</td>
<td>1-parameter logistic CRM</td>
</tr>
<tr>
<td></td>
<td>BCRM</td>
<td>1-parameter hyperbolic or 1-parameter power or 1-parameter logistic or 2-parameter logistic CRM</td>
</tr>
<tr>
<td></td>
<td>POCRM</td>
<td>Partial order CRM – for drug combination trials</td>
</tr>
<tr>
<td></td>
<td>etc.</td>
<td></td>
</tr>
</tbody>
</table>

+ several implementations found online
BOIN flow chart

Start at the lowest dose

Treat a new cohort of patients

Reached max sample size?

Yes

Stop trial and select the MTD

No

\( \hat{\beta} \geq \lambda_c \)

\( \hat{\beta} \)-DLT-rate at the current dose-level

\( \lambda_c < \hat{\beta} < \lambda_h \)

De-escalate the dose

Remain on the same dose-level

Escalate the dose

\( \hat{\beta} \leq \lambda_c \)

Boundaries

Table 1. Dose escalation and de-escalation boundaries

<table>
<thead>
<tr>
<th>Boundary</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
<th>0.25</th>
<th>0.3</th>
<th>0.35</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )-concentration</td>
<td>0.278</td>
<td>0.192</td>
<td>0.197</td>
<td>0.197</td>
<td>0.226</td>
<td>0.278</td>
<td>0.374</td>
</tr>
<tr>
<td>( \beta )-concentration</td>
<td>0.709</td>
<td>0.797</td>
<td>0.289</td>
<td>0.298</td>
<td>0.358</td>
<td>0.419</td>
<td>0.470</td>
</tr>
</tbody>
</table>

Table 2. Dose escalation and de-escalation boundaries for target toxicity rates ≥ 30%

| Action | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|
| Escalate if no. of DLTs ≤ 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 4 | 4 |
| De-escalate if no. of DLTs ≥ 2 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 5 | 5 | 6 | 6 | 6 | 6 | 7 | 7 |

Bayesian Optimal Interval Design - BOIN

- Similar to 3+3, but
- Allows flexible cohort sizes
- May allow re-escalation

BOIN(9,48) boundaries - example

<table>
<thead>
<tr>
<th>Decision, based on the number of patients with DLTs ($N_{DLT}$)</th>
<th>Number of patients evaluable for DLT at the current dose-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Escalate if $N_{DLT} \leq$</td>
<td>0 0 0 0 0 1 1 1 1</td>
</tr>
<tr>
<td>Remain on dose-level if $N_{DLT} =$</td>
<td>- - 1 1 1 2 2 2 2</td>
</tr>
<tr>
<td>De-escalate if $N_{DLT} \geq$</td>
<td>1 1 2 2 2 3 3 3 4</td>
</tr>
<tr>
<td>Disallow dose-level if $N_{DLT} \geq$</td>
<td>NA NA $\geq 3$ $\geq 3$ $\geq 4$ $\geq 5$</td>
</tr>
</tbody>
</table>

- Trial stops when:
  - the maximum sample size has been reached (e.g. $N_{max}=$48), or
  - there are $n$ (e.g. $n=9$) patients evaluable for DLTs on a dose-level, or
  - the lowest dose has been disallowed

- Allows for flexibility in cohort sizes
  - In case of a drop-out: 2 patients, or 5 patients
  - In case of over-recruitment: 4 patients, 7 patients ...
BOIN – Where’s the “Bayes”?


Closer look at the BOIN lambdas

- $\theta =$ target toxicity level, $\theta_1 =$ lower boundary, $\theta_2 =$ upper boundary
- Authors propose, as default, $\theta_1 = 0.6 \cdot \theta$ and $\theta_2 = 1.4 \cdot \theta$ (e.g. $\theta = 0.3$, $\theta_1 = 0.18$, $\theta_2 = 0.42$)

$$
\lambda_{v,j} = \frac{\log(\frac{1-\theta_1}{1-\theta})}{\log(\frac{\theta_1}{\theta_1(1-\theta)})}, \lambda_{d,j} = \frac{\log(\frac{1-\theta_2}{1-\theta})}{\log(\frac{\theta_2}{\theta_2(1-\theta)})}
$$

- Let $p_j =$ true toxicity probability for dose-level $j$.
- Formulate 3 hypotheses: $H_{0j}: p_j = \theta$, $H_{1j}: p_j = \theta_1$, $H_{2j}: p_j = \theta_2$.

$$
\pi_{k,j} = Pr(H_{k,j}) \text{ i.e. the a priori probability of hypothesis } k \text{ being true}
$$

- Assign equal a priori probabilities: $\pi_{0j} = \pi_{1j} = \pi_{2j} = 1/3$
  - Renders $\lambda_{v,j}$ and $\lambda_{d,j}$ invariant to $j$ (the dose level)
  - Renders $\lambda_{v,j}$ and $\lambda_{d,j}$ invariant to $n_j$ (sample size on dose level $j$)

31-Mar-2017
BOIN something about Nmax

- Stopping criterion for the trial
- Allocate enough patients to allow the trial to find MTD

Software - BOIN

|  | BOIN
---|---|---
| R | Developed by authors et al. Implementation includes features not included in the paper, and vice versa. |
|   | + can handle combination trials. |
Operational characteristics

- Expected total number of patients
- Expected number of cohorts
- Expected Number of DLTs per dose-level
- Estimated MTD

Some Trial Designs for Next Trial

- 3+3: the standard traditional 3+3 design.

- ATD (accelerated titration design): a version of 3+3
  - Stage 1: Single patient cohorts in first 2 dose levels or until relevant toxicity observed, thereafter Stage 2
  - Stage 2: Standard traditional 3+3

- aBOIN(9,48): Same as BOIN(9,48) except for single patient cohorts in first 2 dose levels

- BOIN(9,48): BOIN that stops after 9 patients on dose-level or 48 patients in Total. Patients allocated in cohorts of 3 patients
Number of Patients

- "Standard" vs. "accelerated": standard designs (3+3 and BOIN) require 3-4 patients more than accelerated designs (ATD & aBOIN)
- "BOIN" vs. "3+3": BOIN / aBOIN designs require 2-4 patients more than 3+3 / ATD

Number of Cohorts

BOIN / aBOIN designs require 0.5-1 cohort more than 3+3 / ATD
Number of DLTs

Expected MTD

The MTD on the 3+3s are ~1/2 dose-level lower than the BOINS
Expected MTD

The 3+3s find MTD earlier than the BOINs

Comparison of a mCRM, aBOIN(9,48) and ATD: mean number of patients
In summary

<table>
<thead>
<tr>
<th></th>
<th>mCRM</th>
<th>ATD</th>
<th>aBOIN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimate of MTD</strong></td>
<td>Estimates of or near actual MTD.</td>
<td>Under-estimation of MTD by design, not as bad as 3+3.</td>
<td>Estimates at or near actual MTD.</td>
</tr>
<tr>
<td><strong>Number of patients</strong></td>
<td>In line with aBOIN</td>
<td>Smallest sample size, 1-5 patients less than the others</td>
<td>In line with mCRM</td>
</tr>
<tr>
<td><strong>Number of patients on different dose levels</strong></td>
<td>More patients on higher (near MTD) dose levels</td>
<td>Stop earlier: more patients on lower dose levels</td>
<td>More patients on higher dose levels</td>
</tr>
<tr>
<td><strong>Number of patients with DLT</strong></td>
<td>More patients with DLTs (~1)</td>
<td>Less patients with DLTs</td>
<td>More patients with DLTs (~1)</td>
</tr>
<tr>
<td><strong>Pros</strong></td>
<td>Better estimate of MTD (accuracy &amp; precision)</td>
<td>Straightforward</td>
<td>Better estimate of MTD (accuracy &amp; precision)</td>
</tr>
<tr>
<td></td>
<td>Flexibility (cohort sizes may vary)</td>
<td>Nearly memoryless</td>
<td>Flexibility (cohort sizes may vary)</td>
</tr>
<tr>
<td></td>
<td>Uses information available before and during trial</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>Can &quot;go wrong&quot;* if not set-up correctly</td>
<td>Rigid &quot;3+3&quot; &amp; more biased and uncertain</td>
<td></td>
</tr>
</tbody>
</table>