Adaptive Bayesian Designs in Dose-finding Studies
VICC Cancer Biostatistics Workshop Series 2008

Leena Choi, Ph.D.
Department of Biostatistics
Vanderbilt University School of Medicine

May 16, 2008
Thomas Bayes – Bayes’ Theorem
Diagnostic Testing

• A woman at age 40 had a positive mammography in a routine screening

• What is the probability that she actually has breast cancer?
Diagnostic Testing

• “The probability that a woman with a positive mammography has breast cancer”

• “The probability that a woman with breast cancer has a positive mammography”
Diagnostic Testing

• “The probability that a woman with a positive mammography has breast cancer”

• “The probability that a woman with breast cancer has a positive mammography”

• 3 pieces of information
  – the prior probability that a woman has breast cancer (prevalence)
  – the probability that a woman with breast cancer gets a positive mammography (sensitivity)
  – the probability that a woman without breast cancer gets a positive mammography (1-specificity)

An Intuitive Explanation of Bayesian Reasoning ©2003 by Eliezer S. Yudkowsky.
Diagnostic Testing

• **Prevalence:** 1% of women at age forty who participate in routine screening have breast cancer
• **Sensitivity:** 80% of women with breast cancer will get positive mammographies
• **1-specificity:** 10% of women without breast cancer will also get positive mammographies
• What is the probability that she actually has breast cancer?

An Intuitive Explanation of Bayesian Reasoning ©2003 by Eliezer S. Yudkowsky.
Diagnostic Testing

• 100 out of 10,000 women at age forty who participate in routine screening have breast cancer
• Before the mammography screening:
  – Breast Cancer
    • 100 women +
    • 9,900 women -
• After the mammography:
  – 100 women +
    • Group A: 80 women with breast cancer, and a positive mammography.
    • Group B: 20 women with breast cancer, and a negative mammography.
  – 9,900 women -
    • Group C: 990 women without breast cancer, and a positive mammography.
    • Group D: 8,910 women without breast cancer, and a negative mammography.
• \( \frac{A}{A + C} = \frac{80}{80 + 990} = \frac{80}{1070} = 7.5\% \)
Bayes’ Rule

\[
P[\text{cancer} | T (+)] = \frac{P[\text{cancer}] \times P[T (+) | \text{cancer}]}{P[\text{cancer}] \times P[T (+) | \text{cancer}] + P[\sim \text{cancer}] \times P[T (+) | \sim \text{cancer}]}
\]

\[
= \frac{0.01 \times 0.8}{0.01 \times 0.8 + (1-0.01) \times (1-0.9)} = 7.5%
\]
Bayes’ Rule

\[ P[\text{cancer}| T (+)] \]

\[
P[\text{cancer}] \times P[T (+)| \text{cancer}] = \frac{P[\text{cancer}] \times P[T (+)| \text{cancer}]}{P[\text{cancer}] \times P[T (+)| \text{cancer}] + P[\sim \text{cancer}] \times P[T (+)| \sim \text{cancer}]}
\]

- **prior probability** - the original proportion of patients with breast cancer.
- **posterior probability** - the estimated probability that a patient has breast cancer, given that we know she has a positive result on her mammography - is known as the **revised probability**
- **Prior** × **Data** → **Posterior**
Bayes’ Rule

\[ P[\text{cancer} \mid T (+)] = \frac{P[\text{cancer}] \times P[T (+) \mid \text{cancer}]}{P[\text{cancer}] \times P[T (+) \mid \text{cancer}] + P[\sim \text{cancer}] \times P[T (+) \mid \sim \text{cancer}]} \]

• Posterior belief in H (hypothesis, cancer) given the data
  \[ \propto \text{Prior belief in H} \times \text{support for H by the data} \]
  – The probability of H before the data (prior)
  – The likelihood of the data given H (likelihood)
  – The probability of H given the data (posterior)

• Posterior is a synthesis of Prior and Data using Bayes’ rule
Prior, Likelihood and Posterior

• **Prior distribution**: represents the prior information associated with parameter
  – Information available in literature, clinical data base, expert opinion, or any other appropriate source
  – Non-informative representing very little or no relevant information

• **Data distribution (Likelihood)**

• **Posterior distribution**: updating the prior with the likelihood, formally done via Bayes’ theorem
  – Prior is merged with the likelihood to give a final posterior
Example: Placenta Previa

- The proportion of female births in the general population = 0.485 (female births < male births)
- 437 out of total 980 births in placenta previa (437/980=0.446) were females in German

The sex of placenta previa births:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>437</td>
</tr>
<tr>
<td>Male</td>
<td>543</td>
</tr>
<tr>
<td>Total</td>
<td>980</td>
</tr>
</tbody>
</table>

Hypothesis:

the proportion of placenta previa female births is less than 0.485
Prior

Likelihood

Posterior
Posterior distribution

95% credible intervals $[0.425, 0.475]$
Why Bayesian Adaptive Designs?

- Update beliefs in light of new evidence
- Dose finding (dose dropping)
- Stopping early, or late
  - Efficacy
  - Futility
- Seamless phases
- Add arms or drop arms
- Adaptive randomization
- Advantages
  - Smaller trials (usually)
  - More precise conclusions, earlier decision
  - Increase the probabilities of success (faster, better drug development, reduce cost)
  - Better treatment of patients (right drug to right patient at right time)
Challenges in Bayesian Approach

• Need appropriate priors – define what we know based on external evidence (conclusions may depend priors)
• Define complexity of our models – what likelihood to use
• Require good knowledge of probability theory
• Computational challenge
Dose Finding Studies
Garrett-Mayer E. Clin Trials 2006

- Oncology compounds are cytotoxic: the rational in cancer dose-finding trials: to find the highest dose that is also safe for use in a Phase II trial
- Maximum Tolerated Dose (MTD): “optimal” dose, “target”, relatively high dose with manageable side effects
- Dose-limiting toxicities (DLTs): 1 if toxicity occurs, 0 otherwise
- Standard “3+3” dose escalation design starts at dose k with fixed number of ordered dose levels

<table>
<thead>
<tr>
<th>Level</th>
<th>k-1</th>
<th>k</th>
<th>k+1</th>
<th>k+2</th>
<th>k+3</th>
<th>k+4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>100 mg</td>
<td>200 mg</td>
<td>400 mg</td>
<td>600 mg</td>
<td>800 mg</td>
<td>1000 mg</td>
</tr>
</tbody>
</table>

- MTD defined as the highest dose at which 0 or 1 DLTs are observed in 6 patients
- If de-escalation occurs at the first dose level, the study is discontinued
Dose Finding Studies:

Traditional “3 + 3” dose escalation designs

Treat 3 patients at dose k

0 DLTs

Escalate to dose $k+1$: Return to top of flow chart with dose $k+1$

1 DLTs

Treat 3 more Patients at dose $k$

1 DLTs out of 6 patients

Escalate to dose $k+1$: Return to top of flow chart with dose $k+1$

>1 DLTs out of 6 patients

De-escalate to dose $k-1$

If 6 patients already treated at dose $k-1$

If 3 patients already treated at dose $k-1$

Stop Study

Treat 3 more Patients at dose $k-1$

DLT: dose-limiting toxicity

Garrett-Mayer E. Clin Trials 2006
Bayesian Dose-finding Method

Continual Reassessment Method

- Continual Reassessment method (CRM) developed by O’Quigley et al. (1990): a Bayesian Phase I design to estimate MTD
- Assumption: probabilities of both efficacy and toxicity increase with increasing dose
- Use dose-toxicity relationship (dose-efficacy relationship for non-cytotoxic): have rough idea at least
  - Starts with *a priori* dose response curve: parameters chosen based on investigators’ prior belief
  - As data accumulate, the curve completely determined by the data, little like the *a priori* curve
Bayesian Dose-finding Method
Continual Reassessment Method

• Advantages:
  – Superior to traditional dose-escalation designs because it "learns" form information gained at early time points
  – Less likely to treat patients at toxic doses, more likely to treat patients at efficacious doses – more ethical
  – Shown that CRM-based designs to be more efficient and safer (MTD is more precise, fewer cases of DLT – will show comparison

• Criticism:
  – *a priori* curve could be dangerous due to large uncertainty
  – Large dose escalation could occur based on little information
  – Practical issue: duration of time for completion of study (original CRM evaluate every patient)

• Modified CRMs
Modified CRM
Fraries, Goodman et al., Möller

- Pre-defined dose levels for escalation as if for a “3+3” design
- Always start at the lowest dose level under consideration
- Any given dose escalation cannot increase by more than one level, although dose de-escalation can be large
- Enrol two or three patients at each prescribed cohort (not one)
- Proceed as a standard dose escalation design in the absence of dose-limiting toxicities

![Graph showing a priori dose-toxicity](image-url)
Bayesian Dose-finding Method
Continual Reassessment Method

- Choice of dose level: practical consideration of preparation and packaging
- Target rate of toxicity (or response)
  - Chemotherapy given short time period: serious side effects, often set 0.2 – 0.3 (efficacy 0.8)
  - Choice not from statistician or single investigator, consultations with colleagues – dramatic difference, beneficial
- Choose a priori dose-toxicity curve
  - Visual display: clear understanding of the implications of various choices
  - One-parameter or two parameter models: hyperbolic tangent, one- or two-parameter logistic models
  - CRM robust in choosing MTD even if the model is incorrect
  - Good choice of model increase efficiency (smaller size of trial)
  - Informative: Pass through doses for high DLT (eg, 90%) and low DLT (eg, 5%)
A *priori* dose-toxicity curve

target DLT rate = 0.3

dose levels:
- DLT=0
- DLT=1
A priori dose-toxicity curve

non-informative

informative

target DLT rate = 0.3

target DLT rate = 0.25
Bayesian Dose-finding Method
Continual Reassessment Method

• Number of patients per dose level (cohort size)
  – 2-3 /cohort
  – Logistical issues: how long, accrual rate, total number of patients available, maximum number of patients, number of dose levels, seriousness of DLT, etc.

• Stopping and sample size: fixed number of patients (continue until total sample size reached for continuous dose levels, infusion), fixed number of patients/dose (6-8 treated at the MTD for discrete dose)
Example: *A priori* curve dose-toxicity with until first DLT observed

Garrett-Mayer E. *Clin Trials* 2006

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Dose (mg)</th>
<th>Outcomes (0 if no DLT, 1 if DLT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0 0</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>0 0</td>
</tr>
<tr>
<td>4</td>
<td>350 mg</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>400 mg</td>
<td>0 1</td>
</tr>
<tr>
<td>4</td>
<td>450 mg</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>500 mg</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>550 mg</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>800</td>
<td></td>
</tr>
</tbody>
</table>

- Cohort size = 2
- Dose levels: 50 – 800 by 50 mg
- DLT rate = 0.25
- *A priori* curve: two parameter logistic model
- Stopping rule: 10 patients treated at the same dose

![Graph showing dose-toxicity relationship](image-url)

Probability of Toxicity

Dose (mg)

target DLT rate = 0.25
Cohort 9

Cohort 10

Cohort 11

Cohort 12

Cohort 13

Cohort 14

MTD

95% CI = [0.1, 0.47]
## Accrual of Data

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Dose</th>
<th>Outcomes (0 if no DLT, 1 if DLT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0 0</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>0 0</td>
</tr>
<tr>
<td>4</td>
<td>400</td>
<td>0 1</td>
</tr>
<tr>
<td>5</td>
<td>350</td>
<td>0 0</td>
</tr>
<tr>
<td>6</td>
<td>400</td>
<td>0 0</td>
</tr>
<tr>
<td>7</td>
<td>550</td>
<td>1 0</td>
</tr>
<tr>
<td>8</td>
<td>450</td>
<td>0 0</td>
</tr>
<tr>
<td>9</td>
<td>500</td>
<td>0 0</td>
</tr>
<tr>
<td>10</td>
<td>550</td>
<td>1 1</td>
</tr>
<tr>
<td>11</td>
<td>450</td>
<td>0 1</td>
</tr>
<tr>
<td>12</td>
<td>450</td>
<td>0 0</td>
</tr>
<tr>
<td>13</td>
<td>450</td>
<td>1 0</td>
</tr>
<tr>
<td>14</td>
<td>450</td>
<td>0 0</td>
</tr>
<tr>
<td>15</td>
<td>450</td>
<td>STOP</td>
</tr>
</tbody>
</table>

Stopping rule: 10 patients treated at the same dose
Comparison of CRM and “3+3” Designs using Simulations

Garrett-Mayer E. Clin Trials 2006

<table>
<thead>
<tr>
<th></th>
<th>CRM I</th>
<th>CRM II</th>
<th>CRM III</th>
<th>“3+3” I</th>
<th>“3+3” II</th>
<th>“3+3” III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Total sample size</td>
<td>30</td>
<td>50</td>
<td>60</td>
<td>27</td>
<td>27</td>
<td>39</td>
</tr>
<tr>
<td>2 Patients per cohort</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3 Number of cohorts</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>9</td>
<td>9</td>
<td>13</td>
</tr>
</tbody>
</table>

Dose (mg) vs. Probability of Toxicity

- Target DLT rate = 0.3
- True MTD = 1656 mg
Comparison of CRM and “3+3” Designs using Simulations Garrett-Mayer E. Clin Trials 2006

<table>
<thead>
<tr>
<th></th>
<th>CRM I</th>
<th>CRM II</th>
<th>CRM III</th>
<th>“3+3” I</th>
<th>“3+3” II</th>
<th>“3+3” III</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>% of trials with recommend. dose within 250 mg of true (1656 mg)</td>
<td>57%</td>
<td>72%</td>
<td>71%</td>
<td>41%</td>
<td>35%</td>
</tr>
<tr>
<td>5</td>
<td>% of trials with recommend. dose within 400 mg of true (1656 mg)</td>
<td>80%</td>
<td>91%</td>
<td>89%</td>
<td>41%</td>
<td>54%</td>
</tr>
<tr>
<td>6</td>
<td>% of trials with recommend. dose DLT rate &gt; 40%</td>
<td>9.5%</td>
<td>5.8%</td>
<td>5.9%</td>
<td>7.1%</td>
<td>21%</td>
</tr>
<tr>
<td>7</td>
<td>% of trials with recommend. dose DLT rate &gt; 50%</td>
<td>0.9%</td>
<td>0.2%</td>
<td>0.6%</td>
<td>7.1%</td>
<td>2.0%</td>
</tr>
<tr>
<td>8</td>
<td>% of trials with recommend. dose DLT rate &lt; 20%</td>
<td>13%</td>
<td>5.7%</td>
<td>6.2%</td>
<td>52%</td>
<td>44%</td>
</tr>
<tr>
<td>9</td>
<td>% of trials with recommend. dose DLT rate &lt; 10%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>11%</td>
<td>16%</td>
</tr>
<tr>
<td>10</td>
<td>Average % of patients treated at doses with 40% or greater DLT rate</td>
<td>7.6%</td>
<td>7.8%</td>
<td>5.7%</td>
<td>17%</td>
<td>23%</td>
</tr>
<tr>
<td>11</td>
<td>Average % of patients treated at doses with 20% or less DLT rate</td>
<td>32%</td>
<td>19%</td>
<td>24%</td>
<td>62%</td>
<td>53%</td>
</tr>
<tr>
<td>12</td>
<td>Average % of patients treated at doses with DLT rate</td>
<td>26%</td>
<td>28%</td>
<td>26%</td>
<td>21%</td>
<td>22%</td>
</tr>
</tbody>
</table>
Discussions

• Real collaboration between clinician and statistician
  – Clinician: dose levels, DLT rate, range of sample sizes, accrual rate
  – Statistician: cohort size, stopping rule
  – Together: dose-response curve