Utilities in Clinical Trials

Roger J. Lewis, MD, PhD

Harbor-UCLA Medical Center
Los Angeles Biomedical Research Institute
David Geffen School of Medicine at UCLA
Acknowledgements

• Ari M. Lipsky, MD, MS
• NIH/National Center for Research Resources
  “Multicenter Bayesian Decision-Theoretic Clinical Trials” / 1 F32 RR022167
Clinical Trials: Decisions, Decisions...

- Should we target the disease/population?
- Should we study the drug or device in the selected disease/population?
- What study design should we use?
- Should we start the study we’ve designed?
- When should we stop the clinical trial?
- What should we conclude about the treatments when we stop?
- Should we perform another trial?
Bayesian Decision Theory

- Can coherently and quantitatively incorporate:
  - Goals of estimation
  - Patient outcomes
  - Cost effectiveness
  - Varying patient horizons
- Balancing of priorities is incorporated into the utility, loss, or cost function
Components of a Decision Problem

• Prior beliefs regarding one or more parameters of interest (e.g., treatment effects)
• Data (most of the time)
• A set of possible actions, with goal of choosing the “best” one, e.g., at an interim analysis:
  • Continue the trial
  • Change the allocation proportions or drop arm
  • Stop the trial for futility
  • Stop the trial and draw a conclusion
• A utility function, representing the value of taking a particular action when the parameters of interest have specific values
Broad Components of Clinical Trial Utility

- **Value (positive contributions)**
  - Gain in knowledge (error rates, precision of estimation)
  - Investigator career (publishability, impact)
  - Sponsor (regulatory requirements, marketing)
  - Subjects within trial (outcomes, benefit to class)
  - Future subjects (improved outcomes)

- **Costs (negative contributions)**
  - Financial costs (sponsor, subjects, society)
  - Perception cost (bad press)
  - Subjects: Adverse events, opportunity costs
  - Sponsor: Opportunity costs
Utilities in Confirmatory Clinical Trials

• **Scientific**
  • Accuracy of estimation of treatment effect, $\theta$
  • Accuracy of estimation of side effect rates or differences

• **Subjects** (both within and beyond the trial)
  • Clinical outcomes (may be difficult to quantify)
  • Adverse events or side effects
  • Financial costs (for those beyond the trial)

• **Sponsor**
  • Financial costs
  • Meeting regulatory requirements
  • Clinically compelling $\rightarrow$ Likely to be adopted
  • Financial return (near and long term)
Utility of Clinical Outcomes

• Most outcomes are not binary
• Can get more information from each subject, and increase clinical relevance, by using disease-specific patient utility measures of outcomes, when such measures exist
• Can combine primary clinical outcome measure with measure of side-effect burden, into a single metric, if goal is to balance efficacy and toxicity
Interactions with Clinicians

• “The two main concepts in a decision-theory trial that need to be addressed with a clinical researcher are (i) a criterion such as … maximizing expected successes gained…; and (ii) the number of patients, called the ‘horizon’…

• “In our experience, clinicians do not have trouble with either of these intuitively appealing concepts.”

Interactions with Clinicians

- Clinicians have been conditioned to think in terms of
  - Hypothesis testing/statistical significance
  - Considering treatment effects and side effects separately
    - Exception: Oncology, other fields?
  - Considering patients outside the trial only via publication of results
- *De facto* and unspoken hierarchy of goals: hypothesis testing, estimation, publishability (authors), regulatory requirements (sponsors), patient outcomes in trial, safety, future impact
Some Utilities in Published Work

- Patient outcome over patient horizon (1)
- 0-K loss and N (2)
- Linear, zone of indifference, sampling cost (3)
- Quadratic loss and N (4)
- Patient: Outcome – \( \frac{1}{2} \) Side Effects

Scientific: quadratic loss
Public health: “correct, conclusive decision”

DMC: Weighted sum of the 3 above (5)

- “Adaptive Weights” – “…utility functions which have weights which vary as the trial progresses.” (6)

Expected Utilities and Decisions

- Expected utility for an action: the utility function averaged over the uncertainty in the parameters of interest
- Can be continuously reevaluated as information accumulates
- The best decision maximizes the expected utility
- If we are choosing a decision rule, say for all possible interim results, then must use predictive distributions for the possible future data and average over the possible futures
Bayesian Decision-Theoretic Trial Design

- Goal is to find the actions, for all possible data results, which maximize the expectation of the utility function.
- Not all data results are equally likely.
- Expectation is taken with respect to the distribution of unknown parameter(s) of interest and over the distribution of possible data, based on those parameters.
An Example Design

- Two armed trial
- Immediately-available, binary outcome
- Treatment measured by log-odds-ratio

\[ \theta = \log \left[ \frac{p_2(1 - p_1)}{p_1(1 - p_2)} \right] \]
Outcomes and Treatment Effect

• Unknown parameters:
  • $p_1$, the probability of survival in the control group
  • $\theta$, the log odds ratio for survival in test group relative to control group

• Data:
  • $N_1$ and $N_2$ are numbers in control and test groups, respectively
  • $S_1$ and $S_2$ are number of successes in control and test groups, respectively
Adaptive, Group Sequential Design

- Up to $m$ blocks of $n$ subjects enrolled
- Within each block, subjects are randomized between the two arms in a fixed proportion
- After each block, the decision is made:
  - To continue the trial, with a possibly new randomization proportion; or
  - To stop the trial and draw a conclusion
- After the $m^{th}$ block, must stop
Utility Function and Possible Actions

- Quadratic utility for testing of hypotheses regarding treatment effect
- Linear sampling cost: number of subjects
- Negative utility (cost) for each treatment failure
- Possible actions:
  - Non-adaptive: continue 1:1; stop control better; stop equivalence; stop treatment better
  - Adaptive: continue 1:1; continue 1:3; continue 3:1; stop control better; stop equivalence; stop treatment better
Group Sequential Trial
Simplified by assuming $N_1 = N_2$
Picking Optimal Actions

- Assuming one is going to stop enrollment, the optimal action is the one that minimizes the expected loss given the current data and prior.
- To decide whether to stop or continue, must determine the expected cost of continuing.
- This cost of continuing is an average of the cost of the optimal actions for all future possible trial results, weighted by the probability of each possible future result.
- Can be solved by backward induction.
Example Designs

- Up to 4 blocks of 8 patients
- $p_1 \sim \text{beta}(1,1)$, $\theta \sim \text{N}(0,5)$, $K=300$, $R_K=3.5$
- Treatment effect sought is $\theta_0 = 2.2$
  (corresponds to $p_1 = 0.25$ and $p_2 = 0.75$)
- Non-adaptive: continue with 1:1 randomization or stop and draw a conclusion
- Adaptive: continue with 1:3, 1:1, or 3:1 randomization, or stop and draw a conclusion
- Negative utility (cost) for each treatment failure is 0, 10, or 50 (enrollment of one patient is unit of cost)
Optimal Decisions: 
Allowing Adaptive Randomization

<table>
<thead>
<tr>
<th>Control Success</th>
<th>Test Success</th>
<th>Non-Adaptive Randomization</th>
<th>Adaptive Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>6/12</td>
<td>9/12</td>
<td>Continue (1:1)</td>
<td>Continue (1:3)</td>
</tr>
<tr>
<td>6/12</td>
<td>3/12</td>
<td>Continue (1:1)</td>
<td>Continue (1:3)</td>
</tr>
<tr>
<td>7/12</td>
<td>1/12</td>
<td>Stop / Conclude Control Better</td>
<td>Continue (1:3)</td>
</tr>
</tbody>
</table>

- 1:3 Allocation is 25% Control / 75% Test
- Adaptive randomization alone addresses the scientific goal of estimation, but without regard to the number of treatment failures
Optimal Decisions: Negative Utility for Treatment Failure

<table>
<thead>
<tr>
<th>Control Success</th>
<th>Test Success</th>
<th>Non-Adaptive Randomization</th>
<th>Non-Adaptive Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>Success</td>
<td>Ignore Treatment Failures</td>
<td>Include Negative Utility for Failures</td>
</tr>
<tr>
<td>3/12</td>
<td>9/12</td>
<td>Continue (1:1)</td>
<td>Stop / Conclude Test Better</td>
</tr>
</tbody>
</table>

- Non-adaptive trial is limited to 1:1 randomization, so any future enrollment risks randomization to poorly performing control arm
- This risk only matters if one includes the number of treatment failures in the utility function
Optimal Decisions: Adaptive Randomization and Treatment Failures in Utility

<table>
<thead>
<tr>
<th>Control Success</th>
<th>Test Success</th>
<th>Adaptive Randomization</th>
<th>Adaptive Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>Success</td>
<td>Ignore Treatment Failures</td>
<td>Include Negative Utility for Failures</td>
</tr>
<tr>
<td>1/8</td>
<td>4/8</td>
<td>Continue (3:1)</td>
<td>Continue (1:3)</td>
</tr>
<tr>
<td>3/12</td>
<td>9/12</td>
<td>Continue (1:3)</td>
<td>Stop / Conclude Test Better</td>
</tr>
</tbody>
</table>

- 1:3 Allocation is 25% Control / 75% Test
- Including a “penalty” for treatment failures balances the scientific goal of estimation with that of minimizing the number of treatment failures
Optimal Decisions: Adaptive Randomization and Treatment Failures in Utility

<table>
<thead>
<tr>
<th>Control Success</th>
<th>Test Success</th>
<th>Adaptive Randomization Treatment Failure = -10</th>
<th>Adaptive Randomization Treatment Failure = -50</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/12</td>
<td>3/12</td>
<td>Continue (1:3)</td>
<td>Continue (3:1)</td>
</tr>
<tr>
<td>7/12</td>
<td>2/12</td>
<td>Continue (1:3)</td>
<td>Stop / Conclude Control Better</td>
</tr>
</tbody>
</table>

- 1:3 Allocation is 25% Control / 75% Test
- Balance between gaining information and preventing treatment failures is determined by the magnitudes of different utility function components
The Patient Horizon

- N patients total
  - \( n \) within the trial
  - \( m = N - n \); number treated after the trial
- “When a treatment is selected as superior...the selected treatment will be used for the patient horizon of size \( m \).”
- Utility function includes a term
  - \( \sim m \times \Delta \) Outcome with Selected Treatment
  - \( \sim P(Adoption | \Delta, \ldots) \times m \times \Delta \) Outcome with Selected Treatment

Blocks in Translational Research

Conclusions

- One should write down and define what one is trying to accomplish with a clinical trial
- The utility function defines the goals of the trial
- Priors are not sufficient to determine the best action to take for a given set of data, or even the best trial design; one needs a utility function
- Different utility functions lead to different actions
Supplementary Slides
References

• Lewis RJ, Berry DA. Group sequential clinical trial designs: A frequentist evaluation of Bayesian decision-theoretic designs. JASA 1994;89:1528-1534.
• Lewis RJ, Lipsky AM, Berry DA. Bayesian decision-theoretic group sequential clinical trial design based on a quadratic loss function: A frequentist evaluation. Clinical Trials 2007;4:5-14.
Prior

\[ \pi_0(p, \theta | a, b, \theta_0, \sigma_0) = \]

\[
\frac{(a + b - 1)!}{(a - 1)!(b - 1)!} \cdot p^{a-1} (1 - p)^{b-1} \cdot \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[ -\frac{(\theta - \theta_0)^2}{2\sigma_0^2} \right]
\]

\[ p \sim \text{Beta}(a, b) \quad \text{and} \quad \theta \sim \text{N}(\theta_0, \sigma_0) \]
Likelihood Function

\[ L(n_1, s_1, n_2, s_2 \mid p, \theta) = \]

\[
\binom{n_1}{s_1} \binom{n_2}{s_2} p^{s_1 + s_2} (1 - p)^{n_1 + n_2 - s_1 - s_2} e^{s_2 \theta}
\]

\[
\frac{1}{(1 - p + pe^\theta)^{n_2}}
\]
Posterior PDF

\[
\pi(p, \theta | n_1, s_1, n_2, s_2, a, b, \theta_0, \sigma_0) \propto p^{a+s_1+s_2-1} (1 - p)^{b+n_1+n_2-s_1-s_2-1} \times \exp \left[ -\frac{(\theta - \theta_0)^2}{2\sigma_0^2} \right] \frac{e^{s_2\theta}}{(1 - p + pe^{\theta})^{n_2}}
\]
Decision Part of Loss Function

\[ L(\text{stop}) = \min\{E_\pi(L(\theta, H_-)), E_\pi(L(\theta, H_0)), E_\pi(L(\theta, H_+))\} \]

\[ L(\theta, A) = \begin{cases} 
K_0\theta^2 & A = H_0 \\
K_1(\theta - \Delta)^2 & \theta < \Delta, A = H_+ \\
0 & \theta > \Delta, A = H_+ \\
K_1(\theta + \Delta)^2 & \theta > -\Delta, A = H_- \\
0 & \theta < -\Delta, A = H_- 
\end{cases} \]

\[ R_K \equiv \frac{K_1}{K_0} \]
Loss Function

- Plus total number of subjects
- Plus constant times number of treatment failures
Picking the Optimal Local Action

• Assuming one is going to stop enrollment, the optimal action is the one that minimizes the expected loss.

\[
L(\text{stop}) = \min \left\{ E_{\pi} (L(\theta, H_-)), E_{\pi} (L(\theta, H_0)), E_{\pi} (L(\theta, H_+)) \right\}
\]
Deciding Whether to Continue

- To decide whether to stop or continue, must determine the expected cost of continuing.
- This cost is an average of the cost of the optimal actions for all future possible trial results, weighted by the probability of each possible future result.
- Can be solved by backward induction.
Looking Forward
Looking Forward

Block 0  Block 1  Block 2  Block 3
Back Induction
Simplified by assuming $N_1 = N_2$
Back Induction
Simplified by assuming $N_1 = N_2$
Back Induction
Simplified by assuming $N_1 = N_2$
Back Induction
Simplified by assuming $N_1 = N_2$
Example Designs

• Up to 4 blocks of 8 patients
• \( p_1 \sim \text{beta}(1,1), \theta \sim \text{N}(0,5), K=300, R_K=3.5 \)
• Treatment effect sought is \( \theta_0 = 2.2 \)
• Type I error at \( p = 0.5 \) and \( \theta = 0 \)
• Type II error evaluated at \( p_1 = 0.25, \theta = 2.2 \) (implies \( p_2 = 0.75 \))

<table>
<thead>
<tr>
<th>Adaptive</th>
<th>Randomization</th>
<th>Add’l Cost of Failure</th>
<th>Exp N</th>
<th>Exp Cost</th>
<th>Type I Error Rate</th>
<th>Type II Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1:1</td>
<td>0</td>
<td>26.2</td>
<td>400</td>
<td>0.040</td>
<td>0.215</td>
</tr>
<tr>
<td>Yes</td>
<td>1:3 1:1 3:1</td>
<td>0</td>
<td>25.8</td>
<td>354</td>
<td>0.028</td>
<td>0.239</td>
</tr>
<tr>
<td>Yes</td>
<td>1:3 1:1 3:1</td>
<td>50</td>
<td>15.6</td>
<td>764</td>
<td>0.080</td>
<td>0.301</td>
</tr>
</tbody>
</table>
## Bayesian and Frequentist Properties

<table>
<thead>
<tr>
<th></th>
<th>Random Simulation</th>
<th>Fixed Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp Cost</td>
<td>Exp N</td>
</tr>
<tr>
<td>Base</td>
<td>400</td>
<td>26.2</td>
</tr>
<tr>
<td>Adaptive</td>
<td>354</td>
<td>25.8</td>
</tr>
<tr>
<td>Fail10</td>
<td>514</td>
<td>22.1</td>
</tr>
<tr>
<td>Fail50</td>
<td>845</td>
<td>16.0</td>
</tr>
<tr>
<td>Adaptive/Fail10</td>
<td>463</td>
<td>21.5</td>
</tr>
<tr>
<td>Adaptive/Fail50</td>
<td>764</td>
<td>15.6</td>
</tr>
</tbody>
</table>