Simulations in Clinical Trial Design

Abhishek Mishra

Overview of Simulation
Operating Characteristics
Decision Criteria
Uncertainty in Simulations
Simulation Plan
Simulation Strategy
FDA’s Opinion
Case Study: Comparing the study design
Simulation

- Imitation of the operation of real-world processes or systems
- CTS is the generation of response for virtual subjects
- Used to compare study designs while coping with notable complexities and uncertainties
  - Uncertainties include minimal prior knowledge of key parameters or study populations etc
  - Complexities range from simple parallel group designs to fully adaptive designs accompanied by multiple decision criteria etc
• Issues that can be addressed through simulations:
  – What sample size?
  – What are the properties of stopping rule / algorithms for dropping doses?
  – If we choose a particular dose from this study, what will be the likely outcome for the following study?
  – How robust is this design to departures from assumptions?
  – Model “validation” / qualification

**Pro’s and Con’s**

• New policies, operating procedures, information flows etc. can be explored
• Simulation study can help in understanding how the drug will work in a given scenario
• “What if” questions can be answered
• Requires special training
• Time consuming and expensive
• Sometimes results can be difficult to interpret
Operating Characteristics

- Operating characteristics describe the sensitivity and specificity of the design, analysis and decision making processes
  - If there is no effect, how often do we find a spurious effect?
  - If there is an effect, how often do we pick it up?
  - Do we make the right choice of dose?
  - If we are estimating something, how close do we get to the “true” value?
  - Type I error & power are examples of operating characteristics for hypothesis testing against a given null hypothesis

Decision Criteria

- Operating characteristics tell us how often do we make the “right decision”
- SO we must construct quantitative metrics for “the right decision”

- Decision criteria:
  - \( P(X > \text{Crit}) > \text{Conf} \)
  - \( X \) = (simulated) “true” effect
  - \( \text{Crit} \) = Critical value (numerical “null” or clinically important difference)
  - \( \text{Conf.} \) = Level of confidence. Can be 80% - 95%, one-sided, two-sided, or just 50% (mean, median).

- E.g., Find the dose such that \( P(\mu_d - \mu_p > 5) > 0.90 \)

- \( \text{Crit.} \) MUST be agreed up-front by the project team. Statistical significance OR clinical significance OR marketable difference?

- There are various other decision criteria’s used in the simulation like stopping boundaries etc
Uncertainty into Simulations

- There are many sources of variability and uncertainty which can lead to differences in simulated trial results

- Prior information (or lack thereof); preclinical information; project team assumptions

- Population level
  - disease characteristics, covariate distributions, special populations

- Model level parameter variability / uncertainty
  - Includes robustness to model choice e.g. Emax versus linear

- Drug level
  - PK, PK/PD effects. Different subjects have different exposure to drug

- Trial level
  - design, dosing regimen, observation schedule, dropouts

Simulation Plan

- A detailed protocol must be produced giving full details of how the study is to be performed, analyzed and reported

- Simulation plan or protocol should consist of:
  - Objective (s), Procedure (s)
  - Methods for generating the data sets
  - Scenarios to be investigated and methods for evaluation
  - Number of simulations required
  - Evaluating the performance of statistical methods for different scenarios
  - Presentation of the simulation results
Simulation Strategy

• We need a clear understanding of the research question(s) which is to be addressed.

• **Technical Success:**
  – Based on the current knowledge of the therapy, is there any dose which will meet the criteria for further development of the therapy?

• **Trial Design**
  – What is the study design which will provide the “best” estimate of the dose which meets the criteria for success given our current estimates of uncertainty and how robust is that design to departures from those assumptions?

• **Trial Success**
  – If we incorporate both uncertainty regarding our current knowledge, and the trial design, how certain can we be that the proposed study will yield a result which meets the criteria for further development?

Simulation Steps – In Practice

• **Generate data**
  – Based on current model.
  – Acknowledge uncertainty in model parameters.

• **Analyse this data.**
  – Using the same / different analysis methods to the data

• **Store the results.**

• **Repeat for each design / scenario to check consistency of results**

• **There are clinical trial simulators available, can also be done using R, SAS, WinBUGS, East etc**
FDA POSITION

• The important attributes for a meaningful simulation practice may include…:
  – Use of a prospective simulation analysis plan
  – A clearly defined scope
  – Clear assumptions
  – A statement about the implementation of the project
  – A team-based approach to model building
  – Evaluation of M&S results as compared to observed data
  – A full team participation in the interpretation of results
  – Specifying predetermined acceptance criteria of the modelling results

Source: FDA Guidance for Industry on Population Pharmacokinetics

---

Case Study: Comparing the Study Designs
Study Details

- **Objective**
  
  - Study: To evaluate the efficacy of a A/B combination compared to A alone in patients with PHN
  
  - Simulation: To compare 2 different study designs – Parallel or Cross-over

- **Primary Endpoint:** Change from baseline to treatment week 2 (for each treatment leg) in weekly average pain score using the 11-point daily pain rating scale.

Operating Characteristics

- The percentage of simulated trials which demonstrate statistical significant difference.

- The percentage of simulated trials which stop at the interim.

- The mean/median sample size used.

- The mean/median treatment effect size for the studies which stop at the interim.

- The percentage of simulated trials which have a treatment effect less than -0.5 and -0.7.

- The percentage of trials where the \[P(\text{treatment diff}<-0.5) > 0.6, 0.7 \text{ and } 0.8\].
Option 1: Parallel Group

Where, A is approved worldwide for several indications including central and peripheral neuropathic pain, fibromyalgia, as an adjunctive therapy in the treatment of partial seizures, and for generalized anxiety disorder.

Interim for futility

- Assumptions
  - 80 patients per group based on 1-sided 10% significance level and 80% power (not including dropouts or placebo group)
  - Between standard deviation of 2
  - 1:1 randomisation ratio
  - No sample size re-estimation

- 1000 simulations for each scenario, can be increased

- Treatment differences were to be 0 to -1.0.
Stopping Rule

- Scenarios
  - Stopping Rule
    - Stop if the conditional power < 0.2, 0.4, 0.6, 0.7 and 0.8
  - Timing of the interim (33% vs. 50%)

- Simulations were done in SAS

Summary of Simulation Results

<table>
<thead>
<tr>
<th>Treatment difference</th>
<th>Power %</th>
<th>Mean Total Number of Patients*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>108</td>
</tr>
<tr>
<td>-0.5</td>
<td>57</td>
<td>19 (140)</td>
</tr>
<tr>
<td>-0.7</td>
<td>76</td>
<td>11 (148)</td>
</tr>
<tr>
<td>-1.0</td>
<td>93</td>
<td>5 (177)</td>
</tr>
</tbody>
</table>

*not including placebo arm or dropouts
Option 2: Cross-over Design

Cross-over Design

- Assumptions
  - Testing will be carried out at the 1-sided 10% significance level (critical value of 1.28155).
  - The initial sample size will be that based on 80% power calculations and simulations were run for 10-30 patients in total for part (1).
  - The initial sample size will be that based on 80% power calculations and simulations which gave ~10 patients in total for part (2).
  - There will be no interim analyses.
  - However there will be an sample size re-estimation.
Cross-over Design

- **Scenarios**

  - A within patient standard deviation of 1.9 and between patient standard deviation of 2.0 will be assumed.
  
  - A correlation of 0.78 (optimistic estimate) and 0.55 (realistic estimate) respectively was considered.
  
  - Treatment differences were to be 0 to -1.0 for part (1).
  
  - Treatment differences were to be -1.0 or greater for part (2).

### Summary of Simulation Results

<table>
<thead>
<tr>
<th>N total</th>
<th>N (1)</th>
<th>N (2)</th>
<th>Treatment difference</th>
<th>Power % (1)</th>
<th>Power % (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>25</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>-0.5</td>
<td></td>
<td>51</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.7</td>
<td></td>
<td>69</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.0</td>
<td></td>
<td>92</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>-0.5</td>
<td></td>
<td>57</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.7</td>
<td></td>
<td>77</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.0</td>
<td></td>
<td>93</td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>

*not including dropouts
Overall Summary

<table>
<thead>
<tr>
<th>Design</th>
<th>Treatment difference</th>
<th>Power</th>
<th>Average total N (including dropouts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel group with interim (CP &lt; 0.6)</td>
<td>0</td>
<td>10</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>-0.7</td>
<td>76</td>
<td>165</td>
</tr>
<tr>
<td>Cross over study (4 sequences)</td>
<td>0</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>-0.7</td>
<td>75</td>
<td>50</td>
</tr>
</tbody>
</table>

Results

• Cross-over design was selected after simulations

• In order to assess the validity of the estimate of within patient standard deviation, a blinded sample size re-estimation was to be carried out after 33% of patients have completed both treatment periods.

• Depending upon the observed value of the within patient standard deviation, the final study size may be decreased, or increased to a maximum number decided by the team
Thank You