Use of Bayesian Evidence Synthesis Techniques across Phases of Drug Development

Cory R. Heilmann
Eli Lilly and Company
Outline

• Introduction and General Concepts
• Applications to Early Phase Studies
• Applications to Product Decision/Phase III
• Applications to Late Phase
• Conclusions
Introduction
Introduction: Critical Success Factors

• Critical Success Factor (CSF):
  • Attribute(s) that a drug must achieve to be successful from a specific trial or series of trials

• CSF as probability statement requires
  • 1) a measurement e.g. QT interval
  • 2) a statistic e.g. mean QT change from baseline or δ
  • 3) a clinically relevant threshold e.g. δ < 10 ms
  • 4) a probability decision threshold e.g. P(δ < 10 ms) > 90%

• CSF is central to decision making
Introduction: Motivation for Bayesian analysis

- CSF(s) inform future clinical development plan

<table>
<thead>
<tr>
<th>Frequentist</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject or fail to reject hypothesis</td>
<td>Probability that a hypothesis is true</td>
</tr>
<tr>
<td>With enough trials, the chance of false positive &lt; $\alpha$</td>
<td>Probability of success in subsequent trial is $\beta$</td>
</tr>
<tr>
<td>Base decision on single trial or meta-analysis</td>
<td>Flexible in incorporating different sources of data</td>
</tr>
</tbody>
</table>
## Introduction: Probability of Study Success (PrSS)

<table>
<thead>
<tr>
<th>Power</th>
<th>Probability of Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumes a effect size, variability, and dropout</td>
<td>Incorporates uncertainty about effect size, variability, dropout</td>
</tr>
<tr>
<td>Conditional on assumptions, study has x% chance to show superiority/non-inf</td>
<td>Incorporating all data, the probability of the study being successful is x%.</td>
</tr>
</tbody>
</table>
# Introduction: Classical Power vs. PrSS

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Probability of Effect size (Prob)</th>
<th>Conditional Power (Power)</th>
<th>Power* Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>10%</td>
<td>2.5%</td>
<td>0.25%</td>
</tr>
<tr>
<td>0.1</td>
<td>30%</td>
<td>15%</td>
<td>4.5%</td>
</tr>
<tr>
<td>0.2</td>
<td>25%</td>
<td>50%</td>
<td>12.5%</td>
</tr>
<tr>
<td>0.3</td>
<td>20%</td>
<td>80%</td>
<td>16%</td>
</tr>
<tr>
<td>0.4</td>
<td>10%</td>
<td>90%</td>
<td>9%</td>
</tr>
<tr>
<td>0.5</td>
<td>5%</td>
<td>95%</td>
<td>4.75%</td>
</tr>
</tbody>
</table>

- 80% power assuming 0.3 effect size
- 90% power assuming 0.4 effect size
- 48% unconditional probability in this case
“Success” may be measured differently in different phases

In Phase 3/4 a study must hit primary objective to be successful

In early phase, this is not as clear

Consider:

- A study which hits primary objective for a non-commercially viable compound
- A study which fails primary objective and leads to termination of non-commercially viable compound
### Introduction: PrSS

- Can design studies to maximize decision making potential
  - Optimize trade-off between type I and type II error subject to constraints

<table>
<thead>
<tr>
<th></th>
<th>Move to Phase 3</th>
<th>Terminate before Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercially Viable</td>
<td>Pr(True Positive)</td>
<td>Pr(False Negative)</td>
</tr>
<tr>
<td>Not Commercially Viable</td>
<td>Pr(False Positive)</td>
<td>Pr(True Negative)</td>
</tr>
<tr>
<td>Probability of Decision</td>
<td>Pr(Passing CSF)</td>
<td>Pr(Not Passing CSF)</td>
</tr>
</tbody>
</table>
Early Phase - Efficacy

• May have first study in population of interest
• Measure – likely a biomarker for a more robust parameter
• Need large enough magnitude of effect to give confidence of clinically meaningful response in larger study
Early Phase – Efficacy (example)

• Case example: Diabetes drug to lower blood glucose.

• Gold standard for efficacy: HbA1c at 6 months.
  • HbA1c is a measure of average blood glucose over 2-3 months

• Short Phase 2a study
  • 12 week study – not long enough for HbA1c for this drug
  • Measure fasting blood glucose (FBG) at 12 weeks
Early Phase – Efficacy (example)

• Need: Drug to show HbA1c reduction of at least 0.7%
• Equivalent reduction in fasting glucose ~ 21 mg/dL
• Confidence that 21 mg/dL reduction at 12 weeks leads to 0.7% HbA1c reduction at 24 weeks.
Early Phase – Efficacy (example)

• Developing the CSF
  • Statistically significant change in HbA1c may not be useful (may be well below the 21 mg/dL desired).
  • Also need confidence that effect size is clinically relevant
• Naturally leads to Bayesian framework
Early Phase – Efficacy (example)

• Consider:

\[ \Pr(\Delta Y < -21 \text{ mg/dL}) > 70\% \]

• Specifically references the quantity of interest

• Sufficiently high probability
Early Phase – Efficacy (example)

- Model construction:
- \[ \Delta Y = \beta_0 + \beta_1 Y_{BL} + \beta_2 \text{TRT} + \epsilon \]
  - \( \beta_0 \sim N(0, 1000) \)
  - \( \beta_1 \sim N(0, 1000) \)
  - \( \beta_2 \sim N(0, 1000) \)
- If placebo response is consistent, consider informative prior on \( \beta_0 \)
Early Phase – Efficacy (example)

• Considerations
  • Hurdle for efficacy – amount of information versus probability to claim success
Early Phase

- Early phase studies
  - Need to make informed decision
  - Both efficacy and safety are of importance
  - “Failure” can be a success (if failed studies leads to informed decision)
- If compound has clinically relevant effect, want to show that
- If compound does not have clinically relevant effect, want to show that as well
Early Phase

• Optimizing decision rule
• 70% probability of effect size of at least 0.7% sounds good
  • How likely are we to observe this event if the drug works “well”?
  • How likely are we to observe this effect if the drugs does not work “well”?
• Posterior probability does not directly answer this
Early Phase

• If effect size is lower than 15 mg/dL (less than 15 mg/dL decrease relative to placebo), want to terminate
• If effect size is greater than 24 mg/dL (more than 24 mg/dL decrease relative to placebo), want to proceed
• Between 15 mg/dL and 24 mg/dL, the decision is less clear
### Early Phase

<table>
<thead>
<tr>
<th></th>
<th>Move to Phase 3</th>
<th>Terminate Before Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Commercially Viable</strong></td>
<td>True Positive</td>
<td>False Negative</td>
</tr>
<tr>
<td><strong>Unclear</strong></td>
<td>Not certain</td>
<td>Not certain</td>
</tr>
<tr>
<td><strong>Not Commercially Viable</strong></td>
<td>False Positive</td>
<td>True Negative</td>
</tr>
</tbody>
</table>
Early Phase

• Can create decision rules
  • If effect size is 15 mg/dL, what is the probability of terminating? Of advancing?
  • If effect size is 24 mg/dL, what is the probability of terminating? Of advancing?
• This information can complement the probability of clinically relevant effect size – threshold can be optimized based on False Positive/False Negative risk.
Applications to Product Decision/Phase III
## Applications to Product Decision/Phase III

<table>
<thead>
<tr>
<th>Phase I/II</th>
<th>Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need to develop proof of concept (POC)</td>
<td>POC established – need successful study(s)</td>
</tr>
<tr>
<td>Need results for future studies/product decision</td>
<td>Need results to submit for approval</td>
</tr>
<tr>
<td>Interested in probability of effect</td>
<td>Interested in PrSS</td>
</tr>
</tbody>
</table>
Applications to Product Decision/Phase III

- Network meta-analysis – also called mixed treatment comparison
- Typical Phase 2 study may be against placebo
- Phase 2 study may include a competitor
  - Typically, only 1 of many potential comparators
  - Often have low power to separate from (or even show non-inferiority to) competitor
- Need more information to design Phase 3 studies
Applications to Product Decision/Phase III - Example

- Situation: Diabetes compound (Drug X) - have phase 2b study against placebo
- Need to design phase 3 program – consider 4 competitor medications (Drug A, B, C, D)
Applications to Product Decision/Phase III - Example

- First step – literature review
- Crucial step
  - Failure to include studies may lead to bias
  - Including extraneous studies may also bias
  - Often need to consider posters or press releases, not just journal articles
  - May need to involve non-statisticians here
- Recommendation: Include as many studies as possible. Consider sensitivity analyses
Applications to Product Decision/Phase III - Example

• Mixed treatment comparison example
  • Project goal: To assess competitors’ landscape oral diabetic agents and compare with phase 2 compound
  • Endpoint of interest: HbA1c
  • Competitors of interest
Applications to Product Decision/Phase III

• General mixed treatment comparison model: continuous variable.
  • Let $Y_{jk}$ be outcome of interest for $j^{th}$ treatment and $k^{th}$ study
  • $Y_{jk} \sim N(\varphi_{jk}, \sigma/n_{jk})$ where $\varphi_{jk}=\varphi_k + \theta_j$
  • $\varphi_k$ is a study effect and $\theta_j$ is a treatment effect
  • The unknown parameters are $\varphi$, $\theta$ and $\sigma$
  • We assume exchangeable (random) study effect $\varphi_k \sim N(\mu_\varphi, \tau_\varphi)$
  • We assume independent (fixed) treatment effect $\theta_j \sim N(\mu_j, \sigma_j)$ where $\mu_j$ and $\sigma_j$ are constants

Spiegelhalter et al., Bayesian Approaches to Clinical Trials and Health Care Evaluation
Applications to Product Decision/Phase III

• Model the standard deviation (SD) for $j^{th}$ treatment and $k^{th}$ study, $s_{jk}$, as

$$\frac{(n_{jk} - 1)s_{jk}^2}{\sigma^2} \sim \chi^2_{n_{jk} - 1}$$

• Note that literature data will give summary level information so make sure to convert the standard error (SE) to SD i.e $s_{jk} = \sqrt{N/2} \times SE$

• It follows that

$$(n_{jk} - 1)s_{jk}^2 \sim \Gamma((n_{jk} - 1)/2, 1/(2\sigma^2))$$

Spiegelhalter et al., *Bayesian Approaches to Clinical Trials and Health Care Evaluation*
Applications to Product Decision/Phase III - Example

• Situation: Diabetes compound (Drug X) - have phase 2b study against placebo
• Need to design phase 3 program – consider 4 competitor medications (Drug A, B, C, D)
• Mixed treatment comparison is used, incorporating all Drug A, B, C, D, X studies, and also Drug E, F studies (which appear as comparator compounds)
Applications to Product Decision/Phase III - Example

Drug A
Drug B low
Drug B high
Drug C low
Drug C high
Drug D
Drug E low
Drug E high
Drug F low
Drug F high
Drug X low
Drug X high
Placebo
Applications to Product Decision/Phase III
Applications to Product Decision/Phase III

• PrSS – Probability that study will be successful
• Interested in variable of interest (e.g. effect size) and how likely the phase 3 study will be to show this
• If effect size is favorable, but low chance to show this in phase 3
  • Terminate
  • Consider alternative design
Applications to Product Decision/Phase III

Prior information and phase 2 data

Prior information from compounds in the same class. One Phase 2 Study

Apply Bayes model using prior information and phase 2 data in Bayesian hierarchical model

This estimates the distribution of the parameters of interest.

Simulate future observations in ph 3 study, calculate Pr(CSFs)

This estimates the power of the Ph 3 studies, accounting for uncertainty in the hypotheses. Posterior prediction to inform the gatekeeping strategy.
Applications to Product Decision/Phase III - Example

- Information from mixed treatment comparison is useful, but not sufficient.
- Need to inform future studies
  - Important consideration is probability of study success in phase III
• Simulate phase 3 study from posterior samples given by the MTC
  • MTC shows uncertainty of the true effect size
  • Need to simulate individual studies to incorporate sampling variability
  • May need to incorporate inconsistency of MTC matrix – not just describe it
Applications to Product Decision/Phase III

• Other issues to Phase III programs
  • Phase III studies are pivotal for approval (and costly)
  • Secondary variables may be key to differentiation
  • Benefit/risk may be important
  • Need to consider probability of study success
    - Tailor CSFs accordingly
Applications to Product Decision/Phase III - Example

• Phase 3 studies have primary objective – “gold standard efficacy”
  • Often superiority/non-inferiority with standard margin or effect size
• Need to control Type 1 error rate for label claims
  • Gatekeeping and possible combination with Hochberg (or other) multiplicity adjustments
• Combination of importance of claim and likelihood to observe claim are relevant
Applications to Product Decision/Phase III - Example

• Diabetes example – primary is non-inferiority with margin of 0.3% (HbA1c)
• Consider weight reduction, hypoglycemia, or superiority for secondary objectives
  • Variables may be related
Applications to Phase IV

• Network meta-analyses are key for reimbursement OUS
• NICE has extensive technical documents regarding these
• http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series%282391675%29.htm
Applications to Phase IV

- Safety detection/evaluation (e.g., AERS data).
- Use in observational studies – particularly in unmeasured confounding
Conclusions

- Bayesian techniques are applicable across phases of drug development
- Early phases – internal decision making
- Post approval – greater ability to influence externally
- Crucial to understand assumptions
  - Requires statistical, therapeutic area, and business knowledge