Case Studies in Bayesian Augmented Control Design

Nathan Enas
Ji Lin
Eli Lilly and Company
Outline

• Drivers for innovation in Phase II designs
• Case Study #1 – Pancreatic cancer
  – Study design
  – Analysis
  – Learning
• Case Study #2 – Non-small cell lung cancer
  – Study design
  – Analysis
  – Learning
The Need for Better Phase II Designs

Likelihood of Obtaining FDA Approval by Development Phase

Source: http://www.pharmamedtechbi.com/publications/the-pink-sheet/76/5/rampd-productivity-still-lags-study-shows-success-rates-may-have-been-overestimated
The Need for Better Phase II Designs

Likelihood of Obtaining FDA Approval by Development Phase

Source: http://www.pharmamedtechbi.com/publications/the-pink-sheet/76/5/rampd-productivity-still-lags-study-shows-success-rates-may-have-been-overestimated
The global Phase II challenge

Speed to patients

\[ N \approx 40 \]

Single Arm Trial

Best of both worlds?

\[ \text{VS.} \]

Robust Phase II data package

\[ N \approx 200 \]

Randomized Controlled Trial
Standard Phase II design

Prospective Study

Standard Design

N=160

Pancreatic Cancer Patients

1:1

R

LY arm (LY or LY+SOC)

Control arm (SOC)

N_e = 80

N_c = 80
Standard Phase II design

Prospective Study

Standard Design

N=160

1:1

Pancreatic Cancer Patients

LY arm (LY or LY+SOC)

Control arm (SOC)

N_e = 80

N_c = 80

Historical Studies (SOC patients)
**Novel Phase II design**

- **Nc** = 40
- **Nb** = 40
- **N=120**
- **LY arm** (LY or LY + SOC)
- **Control arm** (SOC)

**Bayesian Augmented Control (BAC) Design**

**Prospective Study**

**Historical Studies (SOC patients)**

“Borrowing Strength”
Novel Phase II design

- Prospective Study
- Bayesian Augmented Control Design
- Historical Studies (SOC patients)

- Reduce patient numbers
  - Without reducing power
- Put more patients on LY arm
- Balance speed and robustness
- Enable better Phase III prediction

Nc = 40
Nv = 40
N = 120
Nv = 80

LY arm (LY or LY+SOC)
Control arm (SOC)

2:1

"Borrowing Strength"
## Using BAC in Lilly Oncology

<table>
<thead>
<tr>
<th>Indication</th>
<th>Design</th>
<th>Endpoint</th>
<th>N</th>
<th>Alloc Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreas</td>
<td>Gem±LY</td>
<td>OS</td>
<td>99</td>
<td>1:2</td>
</tr>
<tr>
<td>NSCLC</td>
<td>Pem-Cis±LY</td>
<td>PFS</td>
<td>100</td>
<td>1:2</td>
</tr>
<tr>
<td>GBM</td>
<td>Temo-RT±LY</td>
<td>PFS</td>
<td>56</td>
<td>1:3</td>
</tr>
<tr>
<td>Pancreas</td>
<td>Gem±LY</td>
<td>OS</td>
<td>150</td>
<td>1:2</td>
</tr>
<tr>
<td>Renal Cell</td>
<td>Sutent±LY</td>
<td>PFS</td>
<td>108</td>
<td>1:2</td>
</tr>
<tr>
<td>Pancreas</td>
<td>Gem±LYlo/hi</td>
<td>OS</td>
<td>120</td>
<td>1:1:1</td>
</tr>
<tr>
<td>GBM</td>
<td>CCNU⊗LY</td>
<td>OS</td>
<td>155</td>
<td>1:1:2</td>
</tr>
<tr>
<td>SCLC</td>
<td>Carbo-Etop±LY</td>
<td>PFS</td>
<td>120</td>
<td>1:2</td>
</tr>
<tr>
<td>Liver</td>
<td>Sorafenib⊗LY</td>
<td>PFS</td>
<td>120</td>
<td>1:1:2</td>
</tr>
</tbody>
</table>
CASE STUDY #1 – PANCREATIC CANCER
Phase 1: LY Dose Escalation in Patients with Advanced or Metastatic Cancer

Phase 2: Randomized Phase in Patients with Metastatic Pancreatic Cancer

Arm A: LY + Gemcitabine
Arm B: Gemcitabine

Overview of Phase 2
- **Population:** Stage II-IV unresectable pancreatic cancer (PS 0-2)
- **Primary Objective:** Overall Survival (OS)
- **Observation time:** 15 months enrollment, 9 months follow-up
- **Success criterion:** \( \text{Prob}(HR < 1 \mid \text{data}) \geq 0.8 \)
Randomized double-blind phase II trial comparing gemcitabine plus LY293111 versus gemcitabine plus placebo in advanced adenocarcinoma of the pancreas.

Original article

A phase III trial of pemetrexed plus gemcitabine versus gemcitabine in patients with unresectable or metastatic pancreatic cancer.
Published Gemzar studies

Year of publication

Median survival

JEAL

JMES
Piecewise Exponential Model

JMES Survival
(K-M and 10-unequal-piece Piecewise Exponential)

- Kaplan-Meier
- Piecewise Exponential
BAC Model

\[ Y \sim \text{PiecewiseExponential}(\lambda \times m) \]
\[ \log(m) = \beta_{\text{STUDY1}} I(\text{STUDY1}) + \beta_{\text{JEAL}} I(\text{JEAL}) + \beta_{\text{JMES}} I(\text{JMES}) \]
\[ \beta_{\text{STUDY1}} \sim N(\mu, \tau) \{ \text{current study} \} \]
\[ \beta_{\text{JEAL}} \sim N(\mu, \tau) \left\{ \text{historical studies} \right\} \]
\[ \beta_{\text{JMES}} \sim N(\mu, \tau) \]
\[ \mu \sim N(0, \text{precision} = 1.0) \]
\[ \tau = \frac{1}{sd^2} \]
\[ sd \sim \text{Weib}(3.0982, 146.014) \]
\[ \lambda_j \sim \text{Gamma}(0.01, 0.01), j = 1, \ldots, 10 \]

where \( j = j^{\text{th}} \) piece of 10-piece exponential, the vector \( \lambda \) (10 dimensions) is the hazard, and \( m \) is the hazard ratio (constant for each control group).
Choice of prior distributions

• Normal, Gamma priors
  • Chosen to be relatively noninformative, but also allow MCMC to converge

• Weibull prior for sd
  – Weibull dist’n chosen since always > 0 and flexible in shape
  – Need to be informative due to only 3 studies of control info
    • Nc1 = 66, median = 8.3 months (as in JEAL)
    • Nc2 = 278, median = 6.2 months (as in JMES)
    • Nc3 = 33, median = function of Beta random variate
      – Centered at 7 months, min = 5, max = 9 (assumed for STUDY1)
    • Nt = 66, median = function of c3 and another Beta random variate
      – Tends to be 2 months bigger than control

• Simulate 2000 iterations
  – Collect set of 2000 sd’s of random effects from a parametric Weibull regression
  – Best-fit density to these 2000 estimates is Weibull(3,146)
Operating Characteristics

- Posterior probability (HR < 1) ≥ 0.8 at final analysis
- Frequentist OCs obtained by simulation

<table>
<thead>
<tr>
<th>Overall Survival</th>
<th>BAC design</th>
<th>“Power”</th>
<th>“Type I Error”</th>
</tr>
</thead>
</table>
| Gemzar vs. LY+Gemzar, median (mo.) | N = 99  
(N_C = 33, N_T = 66) | 0.660 | 0.052 |
| 6 vs. 8 | | | |
| 7 vs. 9 | | 0.752 | 0.114 |
| 8 vs. 10 | | 0.774 | 0.216 |
Study Results – Baseline

- BAC model takes historic and concurrent control as replicates with only random error, no covariate adjustments

<table>
<thead>
<tr>
<th></th>
<th>Current study</th>
<th>Historical controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STUDY1</td>
<td>JMES</td>
</tr>
<tr>
<td></td>
<td>GEM+LY</td>
<td>GEM</td>
</tr>
<tr>
<td>N</td>
<td>68</td>
<td>39</td>
</tr>
<tr>
<td>ECOG PS, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>42</td>
<td>41</td>
</tr>
<tr>
<td>1</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Disease Stage, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage II</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Stage III</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Stage IV</td>
<td>77</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>278</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>89</td>
</tr>
</tbody>
</table>
KM Curves for Current and Historical

Slide contains fictitious data, for illustration purposes only.
## Frequentist Analyses of OS

<table>
<thead>
<tr>
<th></th>
<th>GEM + LY (N=68)</th>
<th>GEM (N=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kaplan-Meier Estimate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number (%) of Patients</td>
<td>50 (73.5)</td>
<td>23 (59.0)</td>
</tr>
<tr>
<td>with Events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number (%) of Patients</td>
<td>18 (26.5)</td>
<td>16 (41.0)</td>
</tr>
<tr>
<td>Censored</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (95% C.I.)</td>
<td>9.6 (7.3, 12.8)</td>
<td>8.3 (4.8, 14.1)</td>
</tr>
</tbody>
</table>

Log-rank P-value for Comparison of Arms: 0.60

*Slide contains fictitious data, for illustration purposes only.*
Posterior Distribution of Median OS

Effective Sample Size
borrowed = 18
Posterior OS Analyses and Sensitivity

• Sensitivity analyses performed around the SD of the prior distribution of $\beta$

$$sd \sim \text{Weib}(3,146)$$

- Set $sd = 100$ for little borrowing
- Set $sd = 0.1$ for heavy borrowing

<table>
<thead>
<tr>
<th>Historical Data</th>
<th>Prob of HR &lt; 1</th>
<th>Prob of HR &lt; 0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Little borrowing</td>
<td>0.578</td>
<td>0.106</td>
</tr>
<tr>
<td>Borrowing per</td>
<td>0.775</td>
<td>0.171</td>
</tr>
<tr>
<td>Protocol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy borrowing</td>
<td>0.992</td>
<td>0.461</td>
</tr>
</tbody>
</table>

Slide contains fictitious data, for illustration purposes only.
Learning Points

• BAC design enables randomized controlled trials with smaller sample sizes, yet can maintain power
• BAC justifies increased enrollment to the experimental arm
• Posterior plots and probabilities can help understand data and borrowing
• Similarity and sample size of historical controls both play a significant role in borrowing
• Sensitivity analysis, either Bayesian or frequentist, can help interpret study results
CASE STUDY #2 – NON-SMALL CELL LUNG CANCER
Original Study Concept

Phase 1:
LY Dose Escalation in Patients with Advanced or Metastatic Cancer

Phase 2:
Randomized Phase in Patients with Nonsquamous Non-small Cell Lung Cancer (NSCLC)

Overview of Phase 2
- **Population**: Stage IV nonsquamous non-small cell lung cancer
- **Primary Endpoint**: Objective Response Rate (ORR)

MTD

$3 + 3$ Escalation

Pem-Cis (std. dosing) + LY

N=50

LY + Pemetrexed - Cisplatin
Overview of Phase 2

- **Population**: Stage IV nonsquamous non-small cell lung cancer
- **Primary Endpoint**: Progression-free Survival (PFS)
- **Observation time**: 12 months enrollment, 6 months follow-up
- **Success criterion**: \( \text{Prob}(HR < 1 \mid \text{data}) \geq 0.85 \)
**Historical data**

**Original study**

**Phase III Study Comparing Cisplatin Plus Gemcitabine With Cisplatin Plus Pemetrexed in Chemotherapy-Naive Patients With Advanced-Stage Non–Small-Cell Lung Cancer**


---

**Table 1. Baseline Patient and Disease Characteristics for Randomly Assigned Patients**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cisplatin/ Pemetrexed (n = 862)</th>
<th>Cisplatin/ Gemcitabine (n = 863)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Patients</td>
<td>%</td>
</tr>
<tr>
<td>Histologic type*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>436</td>
<td>50.6</td>
</tr>
<tr>
<td>Large-cell carcinoma</td>
<td>76</td>
<td>8.8</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>244</td>
<td>28.3</td>
</tr>
<tr>
<td>Other: NSCLC, NOS</td>
<td>106</td>
<td>12.3</td>
</tr>
</tbody>
</table>

---

**Graph**

- Median; 95% CI
- CP: 5.3; 4.8, 5.7
- CG: 4.7; 4.4, 5.4
- CP v CG: Adjusted HR; 95% CI 0.90; 0.79, 1.02

**PFS Probability vs. PFS (months) in Patients With Nonsquamous Histology**
Using FACTS for simulations

- The hazard rate for the treatment arm is modeled as piecewise exponential.
- Survival time is divided into $S$ segments (maximum allowable $S$ is 10). The hazard rate is assumed to be constant in each segment.
- Proportional hazards model is used to model the hazard rate in segment $s$: 

\[ \lambda_{st} = \lambda_s \exp(\gamma_t) \]

\[ \gamma_t \sim N(\mu_\gamma, \tau_\gamma^2) \]

\[ \tau_\gamma^2 \sim IG(a_\gamma, b_\gamma) \]

\[ \mu_\gamma \sim N(m_\gamma, t_\gamma^2) \]

- User must specify:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_s$</td>
<td>hazard rate in segment $s$ for control arm in current trial</td>
</tr>
<tr>
<td>$\lambda_{st}$</td>
<td>hazard rate in segment $s$ for previous trial $t$</td>
</tr>
<tr>
<td>$\gamma_t$</td>
<td>log hazard ratio</td>
</tr>
</tbody>
</table>

Hyperparameters $a_\gamma$, $b_\gamma$, $m_\gamma$, $t_\gamma$

For each segment in each historical control arm,
- $X_{st}$ = exposure time
- $E_{st}$ = number of events
Operating Characteristics

Heavy Borrowing

Little Borrowing

Power (%) vs True Hazard Ratio for different Hazard values.
Change in standard of care

New historical study

Maintenance therapy with pemetrexed plus best supportive care versus placebo plus best supportive care after induction therapy with pemetrexed plus cisplatin for advanced non-squamous non-small-cell lung cancer (PARAMOUNT): a double-blind, phase 3, randomised controlled trial

Luis Paz-Ares, Filippo de Marinis, Mircea Dediu, Michael Thomas, Jean-Louis Pujol, Paolo Bidoli, Olivier Molinier, Tarini Prasad Sahoo, Eckart Lasock, Martin Reck, Jesús Corral, Symantha Melemed, William John, Nadia Chouaki, Annemaria H Zimmermann, Carla Vissereen-Grut, Cesare Gridelli
Updated BAC Study Design

**Phase 1:**
LY Dose Escalation in Patients with Advanced or Metastatic Cancer

- **MTD**
- 3 + 3 Escalation
- Pem-Cis (std. dosing) + LY

**Phase 2:**
Randomized Phase in Patients with Nonsquamous Non-small Cell Lung Cancer (NSCLC)

- **Arm A:** LY + Pemetrexed - Cisplatin
- **Arm B:** Pemetrexed - Cisplatin

Overview of Phase 2

- **Population:** Stage IV nonsquamous non-small cell lung cancer
- **Primary Endpoint:** Progression-free Survival (PFS)
- **Observation time:** 12 months enrollment, 6 months follow-up
- **Success criterion:** Prob(HR < 1 | data) > 0.85
New Operating Characteristics

Under the assumption that the current control arm has the same PFS hazard rate as the historical controls (0.031), this study will have about 73% power to detect HR = 0.70 with a Type I error rate of about 16%.
Effective sample size borrowed

Note how borrowing is modulated by the hierarchical model when the current control hazard differs from the historical ("true") control hazard.
KM Curves for Current and Historical

Slide contains fictitious data, for illustration purposes only.
Frequentist Analyses of OS

<table>
<thead>
<tr>
<th></th>
<th>GEM + LY (N=39)</th>
<th>GEM (N=23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaplan-Meier Estimate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number (%) of Patients with Events</td>
<td>30 (76.9)</td>
<td>19 (82.6)</td>
</tr>
<tr>
<td>Number (%) of Patients Censored</td>
<td>9 (23.1)</td>
<td>4 (17.4)</td>
</tr>
<tr>
<td>Median (95% C.I.)</td>
<td>32.3 (22.1, 42.3)</td>
<td>14.4 (9.9, 21.4)</td>
</tr>
</tbody>
</table>

Log-rank P-value for Comparison of Arms: 0.001
Posterior Distribution of Median PFS

Effective Sample Size borrowed = 17

Slide contains fictitious data, for illustration purposes only.
Primary and Sensitivity Analysis

- The primary endpoint is defined as the Posterior Probability of Superiority (Hazard Ratio: LY+Pem+Cis / Pem+Cis < 1.0)
- A secondary endpoint could be considered as the Posterior Probability of Clinically Meaningful Difference (Hazard Ratio: LY+Pem+Cis / Pem+Cis < 0.7)
- The threshold for the primary analysis is 0.85 at the end of the trial.

<table>
<thead>
<tr>
<th>Historical Data</th>
<th>Prob of HR &lt; 1</th>
<th>Prob of HR &lt; 0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Little borrowing</td>
<td>0.994</td>
<td>0.915</td>
</tr>
<tr>
<td>Borrowing per protocol</td>
<td>0.987</td>
<td>0.765</td>
</tr>
<tr>
<td>Full borrowing</td>
<td>0.982</td>
<td>0.481</td>
</tr>
</tbody>
</table>

Slide contains fictitious data, for illustration purposes only.
Learning Points

• Since this BAC method doesn’t control for covariates, a close look at baseline disease characteristics is necessary
• Patient level historic data is helpful for matching covariates between historical and current controls
• BAC requires careful use of historical data, and monitoring of shifts in “standard of care”
• Sensitivity analyses are useful for interpreting study results, including frequentist analyses
Acknowledgements

• Andy Mugglin
• Baoguang Han
• Berry Consultants
THANKS!