A Likelihood Based Approach to Dichotomizing a Continuous Biomarker in Clinical Trials

Liang Fang & Zheng Su
Genentech Inc
July 2011

Disclaimer: the opinions expressed are our personal opinions
Outline

- Introduction
  - Personalized healthcare (PHC)
  - Dichotomize biomarker

- Method
  - Minimum p-value
  - Maximum Likelihood

- Simulation Study

- Summary
Personalized Healthcare (PHC)

- Personalized medicine is a medical model emphasizing in general the customization of healthcare, with all decisions and practices being tailored to individual patients in whatever ways possible. (www.wikipedia.org)

- The Personalized Health Care Initiative will improve the safety, quality and effectiveness of healthcare for every patient in the US. By using “genomics”, or the identification of genes and how they relate to drug treatment, personalized health care will enable medicine to be tailored to each person’s needs. (www.hhs.gov)

- Core of PHC: find the right population that can benefit from the drug!
Increasing focus on personalized health care (PHC)

- Roche’s vemurafenib has proven one of the stars of the ASCO cancer conference by increasing overall survival in skin cancer patients – www.inpharm.com
  - The drug is being developed alongside a bespoke diagnostic kit - the cobas 4800 BRAF V600 mutation test – to help identity those patients who can benefit from vemurafenib. This is a truly personalised approach that represents the future of oncology treatments, and is being developed in-house by Roche’s diagnostics business.
  - Vemurafenib is the first personalised investigational medicine to have shown a significant overall survival benefit in metastatic melanoma
Another Example of PHC

MetMAb in Combination with Tarceva Doubled the Time People with Lung Cancer Lived without Their Disease Getting Worse
-- Genentech's Investigational Personalized Medicine Showed Promising Final Phase II Results in People With a Form of Lung Cancer --

----- Business Wire

- In the overall population of patients with high and low Met expression, the combination of MetMAb and Tarceva did not show a statistically significant improvement in PFS compared to Tarceva alone (hazard ratio [HR]=1.09, p=0.687, median PFS: 2.2 months for the combination vs. 2.6 months for Tarceva alone).

- In people with high Met tumors, those who received MetMAb plus Tarceva had a statistically significant doubling of PFS compared to those who received Tarceva alone (HR=0.53, p=0.04). The median PFS was improved from 1.5 months to 2.9 months.

- The addition of MetMAb to Tarceva also led to a statistically significant improvement in OS compared to Tarceva alone (HR=0.37, p=0.002) in people with high Met tumors. The improvement in median OS was tripled from 3.8 months to 12.6 months.

"These results support further investigation of MetMAb as a potential personalized medicine for people with lung cancer and we plan to start a Phase III study later this year."
Personalized Healthcare (PHC) is the Future of Oncology Drug Development!
Statistical Innovation is Critical in PHC

- Biomarker development (e.g., prognostic vs predictive, dichotomizing, etc)
- Trial design (e.g., adaptive, enrichment, etc)

Phase I/II Trial in all comers

Statistical Considerations

Phase III Trial in sub-population
Dichotomizing Continuous Biomarker: a Common Problem

Only single arm scenario is considered here!
Dichotomizing Continuous Biomarker

- Why dichotomize continuous biomarker
  - Easy for medical practice
  - Inclusion criteria in protocols
  - Avoid linearity assumption in many statistics

- Challenge: difficult to obtain accurate estimate for small size Phase I/II trials
  - What statistical method to identify the cutoff point?
  - How to validate this cutoff point in clinical trials?
Minimum p-value Approach

- **Binary outcome**
  - The optimal cutoff is the one that maximize the chi-square test statistic
  
  \[ \frac{(\hat{p}_1 - \hat{p}_2)^2}{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2} \]

- Control Type I error rate via correction formulas (Miller and Siegmund, 1982; Altman 1994) or permutation test

- Only search within selection interval (to avoid boundary effect)
Minimum P-value Approach

1. Miller and Siegmund, 1982

\[ p_{ms} = \phi(z) \left( z - \frac{1}{z} \right) \log \left( \frac{\epsilon_{high}(1 - \epsilon_{low})}{\epsilon_{low}(1 - \epsilon_{high})} \right) + 4 \frac{\phi(z)}{z} \]

2. Altman, 1994

\[ p_{alt} = -1.63p_{min}(1 + 2.35 \log(p_{min})) \]

Minimum p-value Approach
pros and cons

- **Pros:** easy to implement

- **Cons:**
  - Choice of selection interval is subjective
  - Assuming all statistics follow the same chi-square distribution
    - Asymptotic distribution in small size trial (such as Phase II oncology trials): hard to justify
  - Loss in power
A Likelihood Approach

- The optimal cutoff point is the one that maximizes the likelihood function
- Control Type I error rate via Permutation test
- No selection interval needs to be specified!
- Can be easily extended to detect multiple cutoff points
Binary Endpoint

- **Likelihood function**

\[
L(p_1, p_2, \tau) = \prod_{t=1}^{n} (p_1 I\{b_t \leq \tau\} + p_2 I\{b_t > \tau\})^{x_t} (1 - (p_1 I\{b_t \leq \tau\} + p_2 I\{b_t > \tau\}))^{(1-x_t)}
\]
Simulation Study 1: Type I Error Rate

Table 1: Type I error rates for the proposed approach, the minimum p-value approach with a permutation test, and the two asymptotic approaches with corrected p-values

<table>
<thead>
<tr>
<th></th>
<th>$n = 100$</th>
<th>$n = 200$</th>
<th>$n = 300$</th>
<th>$n = 400$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>10.1%</td>
<td>7.8%</td>
<td>6.3%</td>
<td>6.2%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Altman</td>
<td>10.8%</td>
<td>7.6%</td>
<td>6.4%</td>
<td>6.3%</td>
<td>5.4%</td>
</tr>
<tr>
<td>min p-value</td>
<td>5.1%</td>
<td>4.8%</td>
<td>4.9%</td>
<td>5.2%</td>
<td>4.9%</td>
</tr>
<tr>
<td>max likelihood</td>
<td>4.9%</td>
<td>5.1%</td>
<td>5.0%</td>
<td>4.8%</td>
<td>5.1%</td>
</tr>
</tbody>
</table>
Simulation Study 2

- 100 patients with biomarker values, $X$, randomly generated from $U(0, 1)$
- If $X \leq 0.5$, then response is randomly generated from Bernoulli ($p_1$); else response is randomly generated from Bernoulli ($p_2$)
- Both methods (min p-value vs max likelihood) were used to search the cutoff point
- 2000 simulations were run
Power Study

\[ \Delta p = p_1 - p_2 \]

![Graph showing power study results](image-url)
### Median(Q1,Q3) of Detected Cutoff Points (true cutoff = 0.5)

<table>
<thead>
<tr>
<th>(\Delta p) (p1, p2)</th>
<th>Min p-value</th>
<th>Max Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 (0.45, 0.55)</td>
<td>0.48 (0.20, 0.82)</td>
<td>0.50 (0.35, 0.63)</td>
</tr>
<tr>
<td>0.2 (0.4, 0.6)</td>
<td>0.50 (0.35, 0.67)</td>
<td>0.50 (0.42, 0.58)</td>
</tr>
<tr>
<td>0.3 (0.35, 0.65)</td>
<td>0.50 (0.44, 0.57)</td>
<td>0.50 (0.46, 0.54)</td>
</tr>
<tr>
<td>0.4 (0.3, 0.7)</td>
<td>0.50 (0.46, 0.54)</td>
<td>0.50 (0.48, 0.52)</td>
</tr>
</tbody>
</table>
Time to Event Endpoint

- Log Partial Likelihood function

\[
l(\beta, \tau) = \sum_{t=1}^{n} \delta_t \{\beta I\{b_t > \tau\} - \log(\sum_{j: T_j \geq T_t} e^{\beta I\{b_t > \tau\}})\}
\]

Cutoff Point, \(\tau\) \quad \beta = \frac{\lambda_1}{\lambda_2}

\(\lambda_1\) \quad \downarrow \quad \lambda_2

Biomarker Value, \(b\)
Simulation Study 3

- 100 patients with biomarker values, X, randomly generated from $U(0, 1)$
- If $X \leq 0.4$, then time to event is randomly generated from exponential ($\lambda_1$); else time to event is randomly generated from exponential ($\lambda_2$)
- Non-informative censoring of 10%
- Both methods (min p-value vs max likelihood) were used to search the cutoff point
- 2000 simulations were run
Power Study

The graph shows the relationship between power and the hazard ratio. The graph includes two lines:
- Blue line: min p-value
- Pink line: max likelihood

The x-axis represents the hazard ratio, ranging from 1 to 0.4, while the y-axis represents power, ranging from 0 to 100.
# Median(Q1,Q3) of Detected Cutoff Points (true cutoff = 0.4)

<table>
<thead>
<tr>
<th>HR ($\lambda_1$, $\lambda_2$)</th>
<th>Min p-value</th>
<th>Max Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7 (0.1, 0.07)</td>
<td>0.33 (0.20, 0.40)</td>
<td>0.41 (0.35, 0.49)</td>
</tr>
<tr>
<td>0.6 (0.1, 0.06)</td>
<td>0.36 (0.30, 0.41)</td>
<td>0.40 (0.36, 0.47)</td>
</tr>
<tr>
<td>0.5 (0.1, 0.05)</td>
<td>0.37 (0.30, 0.40)</td>
<td>0.40 (0.37, 0.44)</td>
</tr>
<tr>
<td>0.4 (0.1, 0.04)</td>
<td>0.38 (0.34, 0.40)</td>
<td>0.40 (0.38, 0.43)</td>
</tr>
</tbody>
</table>
Summary

- Personalized healthcare (PHC) is the future of oncology drug development
- Dichotomizing a continuous biomarker is a common practice for research in PHC
- Maximum likelihood approach has statistical advantages and more flexibility, compared to min p-value method, especially for trials with small sample size
- Future research needed for trials with multiple treatment arms