ADAPTIVE POPULATION ENRICHMENT DESIGNS

Vlad Dragalin, PhD
VP, Scientific Fellow
Quantitative Sciences

vdragali@its.jnj.com

Design and Analysis of Experiments in Healthcare
Isaac Newton Institute | Cambridge | Monday 6th July 2015 to Friday 10th July 2015
OUTLINE

1. Enrichment Strategies
2. Example
3. Methodology
4. Estimation
Motivation

The aspect of "one size fits all" surrounding the conventional design of clinical trials has been challenged, particularly

- when the disease is considered heterogeneous
- or the experimental therapy is tailored to a specific mechanism of action

One size fits all → **Tailoring** → Targeted Therapy

create diagnostic, prognostic and therapeutic strategies tailored for specific groups of patients
A Paradigm Shift

Empirical Medicine
- Blockbuster drugs targeted at broad population segments
- On average, 50% of patients do not have desired therapeutic outcomes
- Significant adverse events

Precision Medicine
- Drugs targeted at subgroups of patient population
- Genomic profiles determine segmentation and therapy
- Best possible therapeutic outcome with minimal adverse events

Personalized Medicine
- Delivering the right medicine,
- to the right patient,
- at the right dose,
- at the right time

http://www.jyi.org/features/ft.php?id=1047
Potential Benefits

• Patients receive more effective drugs with fewer side effects giving better outcomes

• Avoid time and resources wasted trying unsuitable medicines

• Accelerating the development and availability of new diagnostics, medicines and treatment pathways benefit patients, healthcare providers and business.
Key Concepts of Adaptive PE

• Extension from the conventional single population design objective to an objective that encompasses several possible patient sub-populations

• Allow more informative evaluation in the patients having different degrees of responsiveness to the therapy

• At an interim stage, it is decided which subpopulation is selected for further inference (including all subpopulations, i.e., full population)

• Not only selection procedures, but also other adaptive strategies (e.g., sample size reassessment, stopping rule) can be performed
**Phase 3 Study in HER2- Early Stage BC Patients**

- Assume that one of the experimental drugs has been graduated from the I-SPY 2 trial with the biomarker signature of triple negative breast cancer (TNBC) but also with some promising effect in HER2- biomarker signature.

  - **Option 1**: a confirmatory Phase 3 trial in TNBC patients only
    - Prevalence of TNBC is only 34%

  - **Option 2**: a confirmatory Phase 3 trial in HER2- patients
    - Prevalence of HER2- is 63%

  - **Option 3**: Adaptive enrichment design
    - Run a confirmatory trial with a two-stage enrichment design
    - Starting with the full population (HER2- patients),
    - But with the preplanned option of selecting only the TNBC patients after the 1st stage in case the observed effect is not promising in the HER2- patients with positive hormone-receptor status HR+.

**Acknowledgment**: D. Berry. I-SPY-1 Results
Potential Outcomes in Adaptive Enrichment Design

• Outcome 1 – Broad Label:
  – Significant treatment effect in the overall (HER2-) population only

• Outcome 2 – Restricted Label:
  – Significant treatment effect in the target (TNBC) population only

• Outcome 3 – Enhanced Label:
  – Significant treatment effects in both the overall (HER2-) and target (TNBC) populations

A Statistical Framework for Decision Making in Confirmatory Multipopulation Tailoring Clinical Trials
Brian A. Millen, Alex Dmitrienko, Stephen Ruberg and Lei Shen
Drug Information Journal 2012 46: 647 originally published online 6 August 2012
DOI: 10.1177/0092861512454116
Phase 2/3 Study in HER2- Early Stage BC Patients

• Stage 1 objective
  - Stop for futility/efficacy
  - To continue with HER2- (Full) population – Broad Label (F) or Enhanced Label (F+S)
  - To confirm greater benefit in TNBC Subpopulation – Restricted Label (S)
  - To adjust the sample size

• Stage 2 data and the relevant groups from Stage 1 data combined
Ballpark Sample Size Calculations

- Primary Endpoint: pathologic complete response (pCR) at surgery
- Power: 90%
- Sign. Level: 0.025
- Control Rate: pCR=0.3
- TRT Effect: 0.2

Possible TRT Effect Range: [0.1 – 0.25]
Population Enrichment Simulation

<table>
<thead>
<tr>
<th>Patient Profile</th>
<th>MP-Her2+ HR+</th>
<th>MP-Her2+ HR-</th>
<th>MP-Her2- HR+</th>
<th>MP-Her2- HR-</th>
<th>MP+ Her2+ HR+</th>
<th>MP+ Her2+ HR-</th>
<th>MP+ Her2- HR+</th>
<th>MP+ Her2- HR-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>16%</td>
<td>7%</td>
<td>23%</td>
<td>6%</td>
<td>4%</td>
<td>10%</td>
<td>6%</td>
<td>28%</td>
</tr>
<tr>
<td>Predicted pCR</td>
<td>47%</td>
<td>67%</td>
<td>25%</td>
<td>43%</td>
<td>35%</td>
<td>55%</td>
<td>17%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Acknowledgment: D. Berry. I-SPY-1 Results

- Prevalence of TNBC in HER2- : 54%
- Control pCR Rate in TNBC: 0.34
- Control pCR Rate in HER2- ∩ HR+: 0.23
- Total of 21 Simulation Scenarios:
  - TRT effect in TNBC: 0 to 0.3 by 0.05
  - TRT effect in HER2- ∩ HR+: 0, 0.1, 0.2

Design
- Total sample size: 300 patients
- Stage 1 sample size: 150 pats
- Testing strategy: inverse normal p-value combination
- Intersection test: Bonferroni
- Selection rule: δ = 0.1 rule
Operating Characteristics:

Legend:
- **Power**
- **P_Reject F**
- **P_Reject S1**

Graph: Power vs. Effect S1

- Power: Blue line
- P_Reject F: Orange line
- P_Reject S1: Red line

Graph parameters:
- pIT Subset2 = 0.230
Operating Characteristics:

![Operating Characteristics Diagram](image_url)
Operating Characteristics:

Legend:
- Blue: Power
- Orange: P_Reject F
- Red: P_Reject S1

Graph: Power | P_Reject F | P_Reject S1 vs. Effect S1

piT Subset 2 = 0.430

Effect S1 vs. Power (0 to 1)
Sample Size Reestimation

• Allow up to a 3-fold sample size increase for Stage 2

• 90% Conditional Power based on observed TRT effect

• Total Sample Size: 300 - 600
Operating Characteristics

Legend
- Power
- P_Reject F
- P_Reject S1

Power | P_Reject F | P_Reject S1 vs. Effect S1

\[ \text{pIT Subset2} = 0.430 \]
Operating Characteristics

![Graph showing Operating Characteristics](image.png)

- **y-Axis:**
  - Blue line: Power
  - Orange line: P_Reject F
  - Red line: P_Reject S1
  - Cyan line: Total ASN

- **Axes:**
  - X-axis: Effect S1
  - Y-axis: Power, P_Reject F, P_Reject S1, Total ASN

- **Legend:**
  - piT Subset2 = 0.430

- **Graph Description:**
  - The graph illustrates the relationship between effect size (Effect S1) and power, as well as the rejection probabilities for different factors (P_Reject F and P_Reject S1), and the total ASN for a specific subset (piT Subset2).
METHODOLOGY
Adaptive Population Enrichment Design

**Stage 1 objective**
- Stop for futility/efficacy
- To continue with Full population or with Sub-population only
- To adjust the sample size

**Stage 2 data and the relevant groups from Stage 1 data combined**
Methodology for Population Enrichment

• Sources for alpha inflation
  – Interim analyses
  – Sample size reassessment
  – Selection from multiple sub-populations

• The procedure is based on the application of the closed test procedure together with combination tests

• The adaptive procedure strongly controls the pre-specified family-wise Type I error rate
P-value Combination Test

Stage 1:

\[ 0 \rightarrow \alpha_1 \rightarrow p_1 \rightarrow \alpha_0 \rightarrow 1 \]

- rejection of \( H_0 \)
- acceptance of \( H_0 \)

Stage 2:

\[ 0 \rightarrow c_\alpha \rightarrow p_1 p_2 \rightarrow 1 \]

- rejection of \( H_0 \)
- acceptance of \( H_0 \)
P-value combination method

- Fisher’s combination test combines the separate stage $p$-values $p_1$ and $p_2$, i.e., $C(p_1,p_2) = p_1 p_2$

- Under $H_0$, the $p$-values are stochastically independent, irrespective of the choice of the design for the second stage.

- $H_0$ is rejected after the second stage if $p_1 p_2 \leq c_\alpha = \exp(-1/2 \chi^2_{4,\alpha})$

- Inverse-normal combination function:
  
  $$C(p_1, p_2) = w_1 \Phi^{-1}(1 - p_1) + w_2 \Phi^{-1}(1 - p_2); w_1^2 + w_2^2 = 1$$

  \[
  \Phi^{-1}(1 - p_k) \sim N(0;1) \text{ if } p_k \text{ uniformly distributed on } [0; 1]
  \]
Test strategies

- Combination test:
  - Fisher’s combination test
  - Inverse normal method

- Separate Phase II/III:
  - Phase II only for sub-population selection
  - Phase III is group sequential

- Intersection Tests:
  - Dunnett
  - Bonferroni
  - Sidak
  - Simes
  - Hierarchical
Closed testing procedure

Stage I

\[ H_0^F \cap H_0^{S_1} \cap H_0^{S_2} \]

\[ H_0^F \cap H_0^{S_1} \]

\[ H_0^F \]

\[ H_0^{S_1} \]

\[ H_0^{S_2} \]

Simple “trick”: Test of intersection hypotheses are formally performed as tests for \( H_0^S \).

\( H_0^S \) can be rejected if all combination tests exceed the critical value \( u_2 \).
Closed testing procedure: Stage II

Example $S = S_2$

Stage I

$H_0^F \cap H_0^{S_1} \cap H_0^{S_2}$

Stage II

$H_0^{S_2}$ can be rejected if all combination tests exceed the critical value $u_2$.

The choice of tests for intersection hypotheses is free. One might use Bonferroni, Simes or Sidak tests.

For one subgroup also Dunnett's test can be applied.
Selection Procedure

- Select the (sub)population with the largest effect
- Select $r$ sets with largest effect
- Select sets with effect compared to full population not worse than $\delta$
- Select $i$-th set
- Select a set if effect exceeds a threshold $t$
- Drop a set if $CP < x$
- Effect measured on test statistic or mean effect scale
Different Configurations

- **Configuration 1:**
  - S1: 64%

- **Configuration 2:**
  - S1: 20%
  - S2: 40%

- **Configuration 3:**
  - S1: 80%
  - S2: 25%

- **Configuration 4:**
  - S1: 8%
  - S2: 20%

- **Configuration 5:**
  - S1: 20%
  - S3: 8%
  - S2: 40%

- **Configuration 6:**
  - S1: 20%
  - S3: 20%
  - S2: 20%
ESTIMATION
Estimation in Population Enrichment Design

- The Full Population \( F \) is split: \( F = S_1 + S_2 \)
- \( \theta_1 \) and \( \theta_2 \): the mean of the primary endpoint in \( S_1 \) and \( S_2 \)
- Overall mean:
  \[
  \theta = \lambda \theta_1 + (1 - \lambda) \theta_2
  \]
- Null hypotheses: \( H_0 : \theta \leq 0 \) and \( H_{01} : \theta_1 \leq 0 \)
- \( N \): overall sample size for 90% power at \( \theta = \Delta \)
  with one-sided \( \alpha = 0.025 \)
- \( N = 84 \) for \( \Delta = 0.5 \)
Population Enrichment Design: Interim Decision

- Two Stages:

- Stage 1 after $m=n_1+n_2$ subjects:
  - If $\frac{X_1(n_1)}{n_1} > \frac{X_2(n_2)}{n_2} + \delta$ continue with $S_1$ only
  - Otherwise, continue with full population

- At Stage 2:

  $T_1 = \left\{ \begin{array}{l}
  n_1 + N - m \\
  N_1
  \end{array} \right. \\
  \hat{\theta}_1 = \left\{ \begin{array}{l}
  \frac{X_1(n_1+N-m)}{n_1+N-m} \\
  \frac{X_1(N_1)}{N_1}
  \end{array} \right.$

  $T_2 = \left\{ \begin{array}{l}
  n_2 \\
  N_2
  \end{array} \right. \\
  \hat{\theta}_2 = \left\{ \begin{array}{l}
  \frac{X_2(n_2)/n_2}{X_2(N_2)/N_2}
  \end{array} \right.$

  $\hat{\theta} = \lambda \hat{\theta}_1 + (1 - \lambda) \hat{\theta}_2$
Fundamental Identity

Let

- \( g(T, X(T)) \) be a statistic with finite mean
- \( T \) a stopping time: \( P_\theta(T < \infty) = 1 \ \forall \ \theta \).

\[
\frac{\partial}{\partial \theta} E_\theta \left[ \frac{g(T, X(T))}{T} \right] = E_0 \left[ \frac{g(T, X(T))}{T} \frac{\partial}{\partial \theta} \exp\{\theta X(T) - \frac{1}{2} \theta^2 T\} \right]
\]

\[
= E_\theta \left[ \frac{g(T, X(T))}{T} (X(T) - \theta T) \right]
\]

\[
= E_\theta \left[ \frac{X(T)}{T} g(T, X(T)) \right] - \theta E_\theta [g(T, X(T))]
\]
Bias of the MLE Estimator

\[
E_\theta \left[ \frac{X(T)}{T} g(T, X(T)) \right] = \theta E_\theta \left[ g(T, X(T)) \right]
+ \frac{\partial}{\partial \theta} E_\theta \left[ \frac{g(T, X(T))}{T} \right]
\]

\[
E_\theta[\hat{\theta}_T] = \theta + \frac{\partial}{\partial \theta} E_\theta \left[ \frac{1}{T} \right]
\]
Bias of the MLE Estimator

\[ E_\theta[\hat{\theta}_T] = \theta + \frac{\partial}{\partial \theta} E_\theta \left[ \frac{1}{T} \right] \]

\[ P_\theta \left( \frac{X_1(n_1)}{n_1} > \frac{X_2(n_2)}{n_2} + \delta \right) = \Phi \left( \frac{\theta_1 - \theta_2 - \delta}{\sqrt{1/n_1 + 1/n_2}} \right) \]

\[ b_1(\theta) = \frac{\phi(u)}{\sqrt{1/n_1 + 1/n_2}} \left[ \frac{1}{n_1 + N - m} - \frac{1}{N_1} \right], \]

\[ b_2(\theta) = \frac{-\phi(u)}{\sqrt{1/n_1 + 1/n_2}} \left[ \frac{1}{n_2} - \frac{1}{N_2} \right], \]

\[ u = \frac{\theta_1 - \theta_2 - \delta}{\sqrt{1/n_1 + 1/n_2}} \]
Contour Plot of Bias of $\hat{\theta}_1$
Bias of $\hat{\theta}_1$
MSE of the MLE Estimator

\[
E_\theta[\hat{\theta}_T^2] = \theta^2 + E_\theta \left[ \frac{1}{T} \right] + 2\theta \frac{\partial}{\partial \theta} E_\theta \left[ \frac{1}{T} \right] + \frac{\partial^2}{\partial \theta^2} E_\theta \left[ \frac{1}{T^2} \right]
\]

\[
Var_\theta[\hat{\theta}_T] = E_\theta \left[ \frac{1}{T} \right] - \left\{ \frac{\partial}{\partial \theta} E_\theta \left[ \frac{1}{T} \right] \right\}^2 + \frac{\partial^2}{\partial \theta^2} E_\theta \left[ \frac{1}{T^2} \right]
\]

\[
MSE_{E_\theta}[\hat{\theta}_T] = E_\theta \left[ \frac{1}{T} \right] + \frac{\partial^2}{\partial \theta^2} E_\theta \left[ \frac{1}{T^2} \right]
\]
MSE of $\hat{\theta}_1$
MSE of $\hat{\theta}_2$
MSE of $\hat{\theta}$
Bias Adjusted MLE

\[ E_\theta[\tilde{\theta}_T] = \theta + b(\theta) \]

\[ \tilde{\theta}_T = \hat{\theta}_T - b(\hat{\theta}_T) \]

- Newton-Raphson iterative procedure

\[ \tilde{\theta}_{T,i+1} = \tilde{\theta}_{T,i} + \frac{(\hat{\theta}_T - \tilde{\theta}_{T,i}) - b(\tilde{\theta}_{T,i})}{1 + b'_\theta(\tilde{\theta}_{T,i})} \]

\[ \tilde{\theta}_{T,1} = \hat{\theta}_T - \frac{b(\hat{\theta}_T)}{1 + b'_\theta(\hat{\theta}_T)} \]
## Bias and MSE of BAMLE

<table>
<thead>
<tr>
<th>$(\theta_1, \theta_2)$</th>
<th>BIAS</th>
<th>MSE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.5, 0.5)</td>
<td>-0.0078</td>
<td>-0.0213</td>
<td>-0.0145</td>
</tr>
<tr>
<td></td>
<td>-0.0010</td>
<td>-0.0005</td>
<td>-0.0008</td>
</tr>
<tr>
<td>(0.5, 0.25)</td>
<td>-0.0100</td>
<td>-0.0303</td>
<td>-0.0202</td>
</tr>
<tr>
<td></td>
<td>-0.0019</td>
<td>-0.0058</td>
<td>-0.0039</td>
</tr>
<tr>
<td>(0.5, 0.0)</td>
<td>-0.0079</td>
<td>-0.0208</td>
<td>-0.0144</td>
</tr>
<tr>
<td></td>
<td>-0.0014</td>
<td>-0.0017</td>
<td>-0.0016</td>
</tr>
<tr>
<td>(0.25, 0.0)</td>
<td>-0.0103</td>
<td>-0.0298</td>
<td>-0.0200</td>
</tr>
<tr>
<td></td>
<td>-0.0021</td>
<td>-0.0053</td>
<td>-0.0037</td>
</tr>
<tr>
<td>(0.0, 0.0)</td>
<td>-0.0070</td>
<td>-0.0218</td>
<td>-0.0144</td>
</tr>
<tr>
<td></td>
<td>-0.0002</td>
<td>-0.0010</td>
<td>-0.0006</td>
</tr>
</tbody>
</table>
Conclusion

- Naïve MLE in Population Enrichment Design has a negative bias
- Bias Adjusted MLE considerably reduces the bias
- And has comparable MSE
Summary

• Attractive and general procedure for adaptive confirmatory design that controls Type I error rate

• The “rules” for adaptation and stopping for futility
  ▪ Do not need to be pre-specified
  ▪ Adaptations may depend on all interim data including secondary and safety endpoints.
  ▪ Can make use of Bayesian principles integrating all information available, also external to the study.
  ▪ Should be evaluated (e.g. via simulations) and preferred version recommended, e.g., in the Simulation Report or DMC Charter.

• Comparison of different strategies and options for analyses is mandatory. The role of simulation becomes increasingly important.
References


