A Bayes Rule for Subgroup Reporting – Adaptive Enrichment Designs  
Peter Müller, UT Austin

## 1 Example: A study for targeted therapy

### Slide 2

1. *A Clinical Trial of Targeted Therapies*  

**Clinical trial**: study of targeted agents in metastatic cancers.

**Patients**: with metastatic cancer (thyroid, ovarian, melano, lung, breast, CRC and other)

**Treatments**: therapy that targets particular molecular aberrations (TT) vs. standard of care (S)

**Objective**: determine whether TT leads to > progression free survival (PFS)

### Slide 3

**Patient population**: patients eligible for non-FDA approved targeted therapy  
record mutations, all cancers

**Population finding**: heterogeneous pat population  
different mutations; different cancers; basline covs ...  
Treatment might be effective in a sub-population  
(subgroup analysis with a purpose)

**Data**: Can use data from similar *observational* study to  
design the trial and evaluate frequentist operating characteristics

### Slide 4

*Data*

Different PFS under TT vs. control,

- mutations are recorded only for small numbers $n$ of patients,
- with varying fraction of observed mutation.

### Slide 5

**Data (ctd.)**

TT effect varies substantially by mutation

TT effect by mutation  
TT effect by tumor  
and by tumor type.

## 2 Decision problem

### Slide 6

**Data**: response $y_i$ (PFS), covariates $x_i = (x_{i1}, \ldots, x_{ip})$.

**Actions**: Report a (i.e., *one*) subgroup of patients who might most benefit from the experimental therapy:  
$$a = (I, x^*)$$

**Covariates**: $I \subset \{1, \ldots, p\}$

**Levels**: $x^* = (x^*_j, j \in I)$.
Population finding: recommend subpop \( \{x_j = x_j^*, j \in I\} \)

use, e.g., in adaptive clinical trial with population finding.

**Decision problem:** separate inference (predicting PFS) vs. decision (report subpopulation).
- no need for multiplicity control
- arbitrary prob model
- disentangle stat significance vs. clinical relevance
- allow for variable # covs.

- other baseline covariates \(b_i\) (age, # prior therapies, etc.)

**Challenges:** prob model needs to allow for
- interactions of \(m_j\)
- many \(m_j\) are not recorded \(\Rightarrow\) var dimension covariate vector \(x_i = (m_{ij}, c_i, b_i)\)
- extrapolation with small # obs.

---

**Utility:** we favor a subpopulation with difference (relative to the overall population) in log hazards ratio (LR), large size and parsimonious description with few covariates

\[
u(a, \theta) = (LR(a, \theta) - \beta) \cdot \frac{n(a)^\alpha}{(|I| + 1)^\gamma}
\]

with \(\beta > 0\) a fixed clinically decided threshold and \(n(a)\) is the size of the subpopulation.
\(\theta\) indexes the sampling model (any model for \(p(y \mid x, \theta)\))

**Alternative utility:** Foster, Taylor & Ruberg (2011, StatMed) use

\[
Q(A) = \text{enhanced treatment effect} - \text{average trt effect}
\]

and sensitivity and specificity to evaluate a reported subpopulation \(A\).

**Model:** Decision problem and solution remain meaningful for any model.

For example, we use the following.

---

### Slide 7

#### Slide 8

**3. Probability Model**

Flexible nonparametric Bayesian model.

**Variables:** for each patient \(i = 1, \ldots, n\)
- Outcome \(y_i\) PFS;
- Covariates \(x_i = (c_i, m_i, b_i)\)
  - tumor type \(c_i \in \{1, \ldots, C\}\) (categorical)
  - molecular aberrations \(m_i = (m_{i1}, \ldots, m_{iM})\)
    with \(m_{ia} = 1\) for observed aberration,
    \(m_{ia} = -1\) for not observed (and 0 for n/a).

**Random Partition**

\(s = (s_1, \ldots, s_n) = \text{cluster membership indicators}\)
\(s_i \in \{1, \ldots, J\}\).

Let \(y_j^*\) and \(x_j^*\) variables by cluster and \(S_j = \{i : s_i = j\}\).

**Random partition:** favor clusters homogeneous in \(x_i\)

\[
p(s \mid x) \propto \prod_{j=1}^{J} c(S_j) \ g(x_j^*)
\]

with \(g(x_j^*)\) scoring “similarity” of \(x_j^* = (x_i; i \in S_j)\).

**Sampling model:** exchangeable within clusters

\[
p(y \mid s, x, \eta) = \prod_{j=1}^{J} \prod_{i \in S_j} p(y_i \mid \eta_j)
\]

**Prediction:** future patient \(i = n + 1\) is
- matched with one of the earlier clusters, on the basis of similar covariates \(x_i = (c_i, m_i, b_i, z_i)\).
- predict similar PFS. That’s all!

---

**Slide 9**

**Random Partition**

\(s = (s_1, \ldots, s_n) = \text{cluster membership indicators}\)
\(s_i \in \{1, \ldots, J\}\).

Let \(y_j^*\) and \(x_j^*\) variables by cluster and \(S_j = \{i : s_i = j\}\).

**Random partition:** favor clusters homogeneous in \(x_i\)

\[
p(s \mid x) \propto \prod_{j=1}^{J} c(S_j) \ g(x_j^*)
\]

with \(g(x_j^*)\) scoring “similarity” of \(x_j^* = (x_i; i \in S_j)\).

**Sampling model:** exchangeable within clusters

\[
p(y \mid s, x, \eta) = \prod_{j=1}^{J} \prod_{i \in S_j} p(y_i \mid \eta_j)
\]

**Prediction:** future patient \(i = n + 1\) is
- matched with one of the earlier clusters, on the basis of similar covariates \(x_i = (c_i, m_i, b_i, z_i)\).
- predict similar PFS. That’s all!

---

**Slide 10**

**Probability model with**

**Random partition:** includes a regression on covariates, through \(g(x_j^*)\)

**Variable size regression:** Cluster allocation is possible with available (subset of) covariates -- no problem with variable size cov vector.

**Extrapolation:** restricted to matching with observed patients
4 Simulation

Scenarios: 7 scenarios, \( p = 8 \) covariates (7 mutations, 1 cancer type).
Simulation truth is a log normal regression for \( y_i \in \mathbb{R} \).

True subgroups: Evaluation of (frequentist) error rates requires “true” subgroups.
Defined as a function of the assumed sampling model.

- Evaluate \( u(a, \cdot) \) under the simulation truth using the true log hazards ratios for a subgroup \( a \).
- Repeat for all poss subgroups.
- The top 10 subgroups are labeled as “truth”

Results: next slide.

---

Slide 12

Operating Characteristics: Error Rates

- \( \text{TIE} = p(H_0^c | H_0) \) type-I error
- \( \text{FNR} = p(H_0 | H_1^c) \) false negative rate
- \( \text{TPR} = p(H_1 | H_1) \) true positive r.
- \( \text{FSR} = p(H_a | H_a^c) \) false subgroup r.
- \( \text{TSR} = p(H_a | H_a) \) true subgroup r.
- \( \text{FPR} = p(H_1 | H_a) \) false positive r.

Decision  | \( H_0 \) | \( H_1 \) | \( M_{ih} \) | \( FPR_{ih} \)
---|---|---|---|---
\( H_0 \) | 1-TIE | \( \text{TSR}_{ih} \) | \( \text{FSR} \) | \( \text{TPR} \)
\( H_1 \) | \( \text{FPR}_{ih} \) | \( \text{TPR} \)

- Choose \( c_0 \) to control TIE, and \( c_1 \) to control (average) FSR.
- All but the TIE require additional specification:
  - effect size for FNR, TPR and FSR.
  - TSR and FPR depend on \( i, h \).

---

Slide 13

### Simulation results

<table>
<thead>
<tr>
<th>Scenario</th>
<th>TIE</th>
<th>FNR</th>
<th>TPR</th>
<th>FSR</th>
<th>TSR</th>
<th>FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.00</td>
<td>-</td>
<td>.06</td>
<td>.92</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.00</td>
<td>.04</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3*</td>
<td>.02</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.01</td>
<td>.94</td>
<td>.05</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.03</td>
<td>-</td>
<td>.23</td>
<td>.75</td>
<td>.04</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.01</td>
<td>-</td>
<td>.12</td>
<td>.87</td>
<td>.02</td>
<td></td>
</tr>
</tbody>
</table>

* scenario 3 is true \( H_0 \)
all others are true subgroup and overall effects

---

Slide 14

### Treatment Allocation

| Scenario | \( AP_{trt} \) | \( AP_{ctrl} \) | \( d \)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.92</td>
<td>.85</td>
<td>-.11</td>
</tr>
<tr>
<td>2</td>
<td>.91</td>
<td>.81</td>
<td>-.13</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>.83</td>
<td>.73</td>
<td>-.18</td>
</tr>
</tbody>
</table>

\( AP_t = \text{prob of correct assignment to TT} \)
\( AP_c = \text{prob of correct assignment to C} \)
\( d = \text{bias in estimating succ probs} \)

---

5 Multiple Subgroups

5. Multiple Subgroups


Action set: generalize to a subgroup report with multiple subsets

\( a = \{a_d, d = 1, \ldots, D\} \) with \( a_d = (I_d, x^*_d) \),

- Covariates: \( I_d \subset \{1, \ldots, p\} \)
- Levels: \( x^*_d = (x^*_{dj}, j \in I_d) \).

Utility function: favor subgroups with distinct prediction, large size and parsimonious description:

\[ u(a, \theta) = \prod_d \left\{ \text{LR}(a_d, \theta) - \beta \right\} \cdot \frac{n(a_d)}{(|I_d| + 1)^\gamma} \]

---

Slide 16

### Example: Large ICD study

Study: Large study of implantable cardioverter defibrillators (ICD), to reduce the risk of sudden cardiac death.

Outcome: overall survival

Covariates:

- AGE, coded as age < 65 vs. \( \geq 65 \)
- ISCH, presence of ischemia
• EFCAT, ejection fraction, coded as $< 30$ versus $\geq 30$
• QRS, coded as $\text{QRS} < 120$ vs. $\geq 120$
• MALE
• NYHA, NY heart association class (III vs. IV)

---

**Slide 17**

**Results:** The Bayes rule subgroup report

<table>
<thead>
<tr>
<th>AGE</th>
<th>MALE</th>
<th>NYHA</th>
<th>EFCAT</th>
<th>MI</th>
<th>QRS</th>
<th>ISCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Slide 18**

**Computation**

**Single subset:** with moderate $p$, use full enumeration and posterior MCMC to evaluate expected utilities.

**Multiple covariates:** use inhomogeneous MCMC (variation of simulated annealing).

---

**Slide 19**

**Summary**

• A Bayesian approach to pre-planned subgroup analysis with a sensible strategy to detect subgroup effects.

• Bayes rule (approx)

• Coherent posterior probabilities for subgroup effects.

• Multiplicity control is achieved by
  – choice of priors,
  – by controlling frequentist error rate.

• Report $\geq 1$ subgroup effects (under Bayes rule)

Design: use inference on subgroups for population finding or enrichment design.