A Population-finding Design with Non-parametric Bayesian Response Model

Peter Müller, UT Austin

with

Yanxun Xu, UT Austin / JHU; Riten Mitra, U. Louisville

Don Berry and A. Tsimboridou, M.D. Anderson

TRT	TUMOR	PFS	CENS	MUTATIONS								
				m1	m2	m3	m4	m5	m6	m7	m8	...
TT	THYROID	2.6	0	NA	NA	NA	NA	NA	NA	NA	NA	
TT	THYROID	3.6	0	NA	0	0	0	NA	0	NA	NA	
S	OVARIAN	4.2	1	0	NA	0	0	0	0	0	0	
S	MELANOMA	5.8	1	NA	0	0	0	NA	0	0	0	
...	...	...	...									

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)

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)

Secondary objective: find subgroups for

- reporting of winning subgroup for future study

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)

Secondary objective: find subgroups for

- reporting of winning subgroup for future study
- adaptive treatment allocation

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)

Secondary objective: find subgroups for

- reporting of winning subgroup for future study
- adaptive treatment allocation

heterogeneous pat population different mutations; different cancers; basline covs ...

Treatment might be effective in a sub-population

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)

Secondary objective: find subgroups for

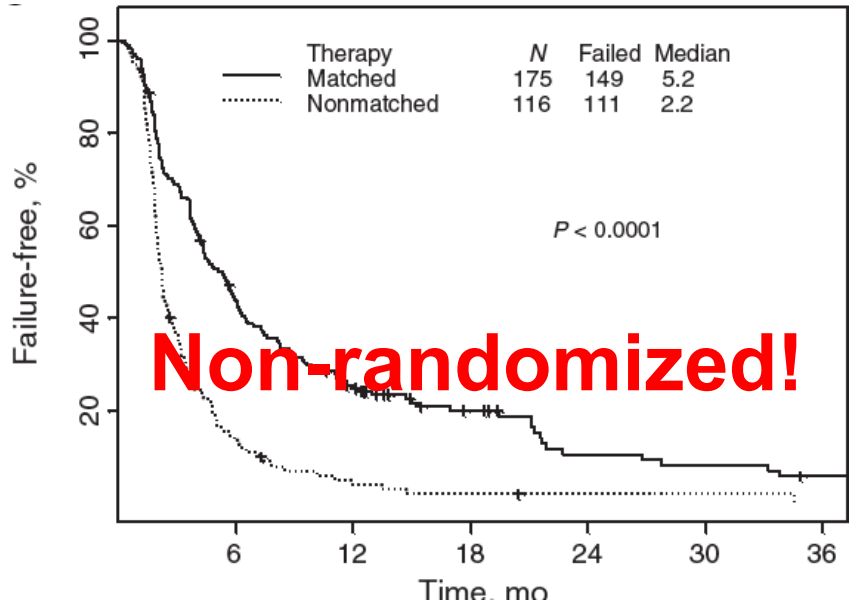
- reporting of winning subgroup for future study
- adaptive treatment allocation

heterogeneous pat population different mutations; different cancers; basline covs ...
Treatment might be effective in a sub-population

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

Data

Different PFS under TT vs. control,

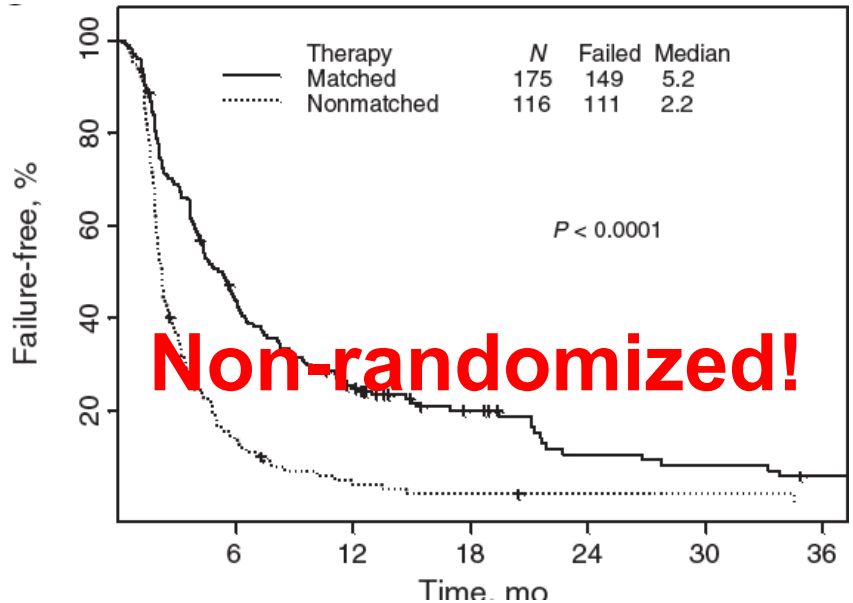


PFS

- mutations are recorded only for small numbers n of patients,

Data

Different PFS under TT vs. control,



PFS

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

Bayesian Subgroup analysis

- Treatment/cov interaction: Dixon and Simon (1991), Simon (2002), Jones et al. (2011)
- Tree based methods: Foster, Taylor & Ruberg (2011)
- Model selection: Berger, Wang and Shen (2014), Sivaganesan et al. (2011)
- Decision problem: next slides...

2. Decision Problem

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

2. Decision Problem

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:

$$\mathbf{a} = (I, \mathbf{x}^\star),$$

Covariates: $I \subset \{1, \ldots, p\}$
Levels: $\mathbf{x}^\star = (x_j^\star,\ j \in I)$.
Population finding: recommend subpop $\{x_j = x_j^\star,\ j \in I\}$

2. Decision Problem

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:

$$\mathbf{a} = (I, \mathbf{x}^\star),$$

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

Levels: $\mathbf{x}^\star = (x_j^\star,\ j \in I)$.

Population finding: recommend subpop $\{x_j = x_j^\star,\ j \in I\}$

Decision problem: separate inference (predicting PFS)

2. Decision Problem

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:

$$\mathbf{a} = (I, \mathbf{x}^\star),$$

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

Levels: $\mathbf{x}^\star = (x_j^\star,\ j \in I)$.

Population finding: recommend subpop $\{x_j = x_j^\star,\ j \in I\}$

Decision problem: separate inference (predicting PFS)

vs.

decision (report subpopulation).

2. Decision Problem

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:

$$\mathbf{a} = (I, \mathbf{x}^\star),$$

Covariates: $I \subset \{1, \ldots, p\}$
Levels: $\mathbf{x}^\star = (x_j^\star,\ j \in I)$.
Population finding: recommend subpop $\{x_j = x_j^\star,\ j \in I\}$

Decision problem: separate inference (predicting PFS)

vs.

decision (report subpopulation).

- no need for multiplicity control
- arbitrary prob model
- disentagle stat significance vs. clinical relevance
- allow for variable # covs.

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

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

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

with $\beta > 0$ a fixed clinically decided threshold and $n(\mathbf{a})$ is the size of the subpopulation.
θ indexes the sampling model

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

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

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

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

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

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

Bayes rule: Report $a^{\star} = \arg\max_a \int u(a, \theta)\, dp(\theta \mid data)$

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

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

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

Bayes rule: Report $a^{\star} = \arg\max_a \int u(a, \theta)\, dp(\theta \mid data)$

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.

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

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

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

Bayes rule: Report $a^{\star} = \arg\max_a \int u(a, \theta)\, dp(\theta \mid data)$

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: Decicsion problem and solution meaningful for any model. For example, we use the following.

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 $m_{is} = 1$ for observed aberration, $m_{is} = -1$ for not observed (and 0 for n/a).
  - other baseline covariates b_i (age, # prior threrapies, etc.)

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 $m_{is} = 1$ for observed aberration, $m_{is} = -1$ for not observed (and 0 for n/a).
  - other baseline covariates b_i (age, # prior threrapies, etc.)

Challenges: prob model needs to allow for

- interacations 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.

Random Partition

$s = (s_1, \ldots, s_n) =$ cluster membership indicators $s_i \in \{1, \ldots, J\}$.
Let $y_j^\star$ and $x_j^\star$ variables by cluster and $S_j = \{i : \ s_i = j\}$.

Random Partition

$s = (s_1, \ldots, s_n) =$ cluster membership indicators $s_i \in \{1, \ldots, J\}$.
Let $y_j^\star$ and $x_j^\star$ variables by cluster and $S_j = \{i : \ s_i = j\}$.

Random partition: favor clusters homogeneous in x_i

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

Random Partition

$s = (s_1, \ldots, s_n) =$ cluster membership indicators $s_i \in \{1, \ldots, J\}$.
Let $y_j^\star$ and $x_j^\star$ variables by cluster and $S_j = \{i : \ s_i = j\}$.

Random partition: favor clusters homogeneous in x_i

$$p(s \mid x) \propto \prod_{j=1}^{J} c(S_j)\, g(x_j^\star)$$

with $g(x_j^\star)$ scoring "similarity" of $x_j^\star = (x_i;\ i \in S_j)$.

Random Partition

$s = (s_1, \ldots, s_n)$ = cluster membership indicators $s_i \in \{1, \ldots, J\}$.
Let $y_j^\star$ and $x_j^\star$ variables by cluster and $S_j = \{i : \ s_i = j\}$.

Random partition: favor clusters homogeneous in x_i

$$p(s \mid x) \propto \prod_{j=1}^{J} c(S_j)\, g(x_j^\star)$$

with $g(x_j^\star)$ scoring "similarity" of $x_j^\star = (x_i;\ i \in S_j)$.

Similarity function: over observed covariates only:

$$g(x_j^\star) = \prod_{\ell=1}^{p} g_\ell(\{x_{i\ell},\ i \in S_j \text{ and } x_{i\ell} \text{ observed } \}$$

Sampling model: exchangeable within clusters

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

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!

4. Simulation

	Overall	Interactions (effect size)	
1	0.1	TP53*z (0.25)	BRCA*z (0.2)
2	0	none	
3	0.1	KRAS*Colon*z (0.4)	
4	0	KRAS*Lung*z (0.2)	KRAS*Colon*z(0.4)
5	0.1	TP53*z (0.2)	KRAS*Lung*z (0.15)
6	0	KRAS*Lung*z (0.2)	KRAS*Colon*z(0.4)
7	0.3	KRAS*Lung*z (0.25)	TP53*Lung*z (0.3)
8	0	BRAF*z (-0.4)	BRCA*z (0.35)
9	0.3	none	
10	0	TP53*KRAS*Lung*z (0.5)	

$p = 7$ covariates (5 mutations, 1 cancer type, trt).
Siumlation truth is a log normal regression for $y_i \in \Re$.

True subgroups: Evaluation of (frequntist) 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"

True subgroups: Evaluation of (frequntist) 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.

Operating Characteristics: Error Rates

TIE $= p(H_0^c \mid H_0)$ type-I error

TPR $= p(H_1 \mid H_1)$ true positive r.

TSR $= p(H_{\mathbf{a}} \mid H_{\mathbf{a}})$ true subgroup r.

FNR $= p(H_0 \mid H_0^c)$ false negative rate

FSR $= p(H_{\mathbf{a}} \mid H_{\mathbf{a}}^c)$ false subgroup r.

FPR $= p(H_1 \mid H_{\mathbf{a}})$ false positive r.

Operating Characteristics: Error Rates

TIE $= p(H_0^c \mid H_0)$ type-I error
TPR $= p(H_1 \mid H_1)$ true positive r.
TSR $= p(H_\mathbf{a} \mid H_\mathbf{a})$ true subgroup r.
FNR $= p(H_0 \mid H_0^c)$ false negative rate
FSR $= p(H_\mathbf{a} \mid H_\mathbf{a}^c)$ false subgroup r.
FPR $= p(H_1 \mid H_\mathbf{a})$ false positive r.

$p(A \mid B)$ = frequentist probability of A over repeat simulation under truth B.

	Truth		
	Subgroup Effect		
Decision	H_0	H_a	H_1
H_0	1-TIE		FNR
H_a		TSR_a	FSR
H_1		FPR_a	TPR

Operating Characteristics: Error Rates

TIE $= p(H_0^c \mid H_0)$ type-I error
TPR $= p(H_1 \mid H_1)$ true positive r.
TSR $= p(H_\mathbf{a} \mid H_\mathbf{a})$ true subgroup r.
FNR $= p(H_0 \mid H_0^c)$ false negative rate
FSR $= p(H_\mathbf{a} \mid H_\mathbf{a}^c)$ false subgroup r.
FPR $= p(H_1 \mid H_\mathbf{a})$ false positive r.

$p(A \mid B)$ = frequentist probability of A over repeat simulation under truth B.

	Truth		
	Subgroup Effect		
Decision	H_0	H_a	H_1
H_0	1-TIE		FNR
H_a		TSR_a	FSR
H_1		FPR_a	TPR

- Choose c_0 to control TIE, and c_1 to control (average) FSR.

Operating Characteristics: Error Rates

TIE $= p(H_0^c \mid H_0)$ type-I error
TPR $= p(H_1 \mid H_1)$ true positive r.
TSR $= p(H_\mathbf{a} \mid H_\mathbf{a})$ true subgroup r.
FNR $= p(H_0 \mid H_0^c)$ false negative rate
FSR $= p(H_\mathbf{a} \mid H_\mathbf{a}^c)$ false subgroup r.
FPR $= p(H_1 \mid H_\mathbf{a})$ false positive r.

$p(A \mid B)$ = frequentist probability of A over repeat simulation under truth B.

	Truth		
	Subgroup Effect		
Decision	H_0	H_a	H_1
H_0	1-TIE		FNR
H_a		TSR_a	FSR
H_1		FPR_a	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 specific subgroup a.

Simulation results

Scenario	TIE	TSR	TPR	FSR	FNR	FPR
1	-	.74	-	.02	.00	.00
2	0.05	-	-	-	-	-
3	-	.98	-	.00	.01	.00
4	-	.93	-	.00	.00	.00
5	-	.81	-	.02	.01	.00
6	-	.91	-	.00	.01	.00
7	-	.77	-	.00	.00	.03
8	-	.62	-	.00	.01	.00
9	-	-	.89	-	-	-
10	-	.66	-	.00	.00	.00

TIE $= \Pr(H_0^c \mid H_0)$; TSR $= \Pr(a \mid a)$; TPR $= \Pr(H_1 \mid H_1)$; FSR $= \Pr(a \mid a^c)$; FNR $= \Pr(H_0 \mid H_0^c)$; and FPR $= \Pr(H_1 \mid H_1^c)$.

Simulation results

Scenario	TIE	TSR	TPR	FSR	FNR	FPR
1	-	.74	-	.02	.00	.00
2	0.05	-	-	-	-	-
3	-	.98	-	.00	.01	.00
4	-	.93	-	.00	.00	.00
5	-	.81	-	.02	.01	.00
6	-	.91	-	.00	.01	.00
7	-	.77	-	.00	.00	.03
8	-	.62	-	.00	.01	.00
9	-	-	.89	-	-	-
10	-	.66	-	.00	.00	.00

TIE $= \Pr(H_0^c \mid H_0)$; TSR $= \Pr(a \mid a)$; TPR $= \Pr(H_1 \mid H_1)$; FSR $= \Pr(a \mid a^c)$; FNR $= \Pr(H_0 \mid H_0^c)$; and FPR $= \Pr(H_1 \mid H_1^c)$.

* scenario 2 is true H_0 – others are true subgroup & overall effects

Treatment Allocation

Scenario	AP_{TT}	AP_C	$\bar{d}$
1	.78	-	.12
2	-	-	-
3	.62	-	.11
4	.68	-	.08
5	.67	-	.10
6	.68	-	.08
7	.85	-	.07
8	.79	.63	.13
9	.81	-	.08
10	.76	-	.08

AP_t = prob of correct assignment to TT.
AP_c = prob of correct assignment to C.
$\bar{d}$ = bias in estimating succ probs.

Summary

- A Bayesian approach to pre-planned subgroup analysis with a sensible strategy to detect subgroup effects.
- Bayes rule (approx)
- Coherent posterior probabilites for subgroup effects.
- Multiplicity control is achieved by
  - choice of priors,
  - by controlling frequentist error rate.
- Report ≥ 1 subgroup effects (under Bayes rule)

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