# Generalized Group Lasso for Patient Subgroup Selection

Wenxuan Deng

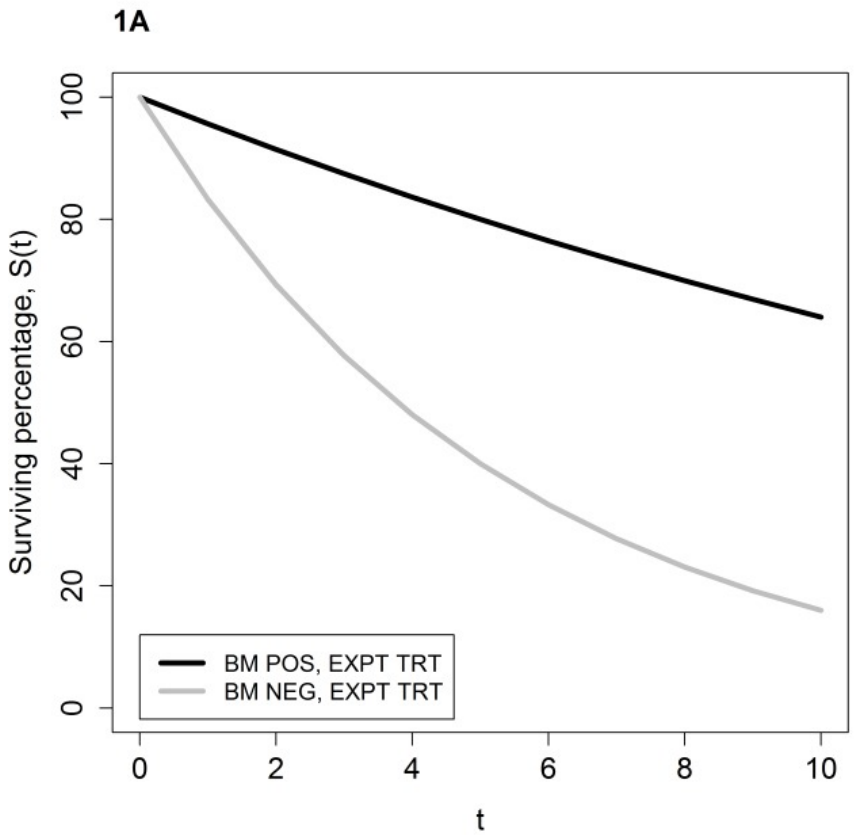
Takeda Pharmaceuticals U.S.A., Inc.

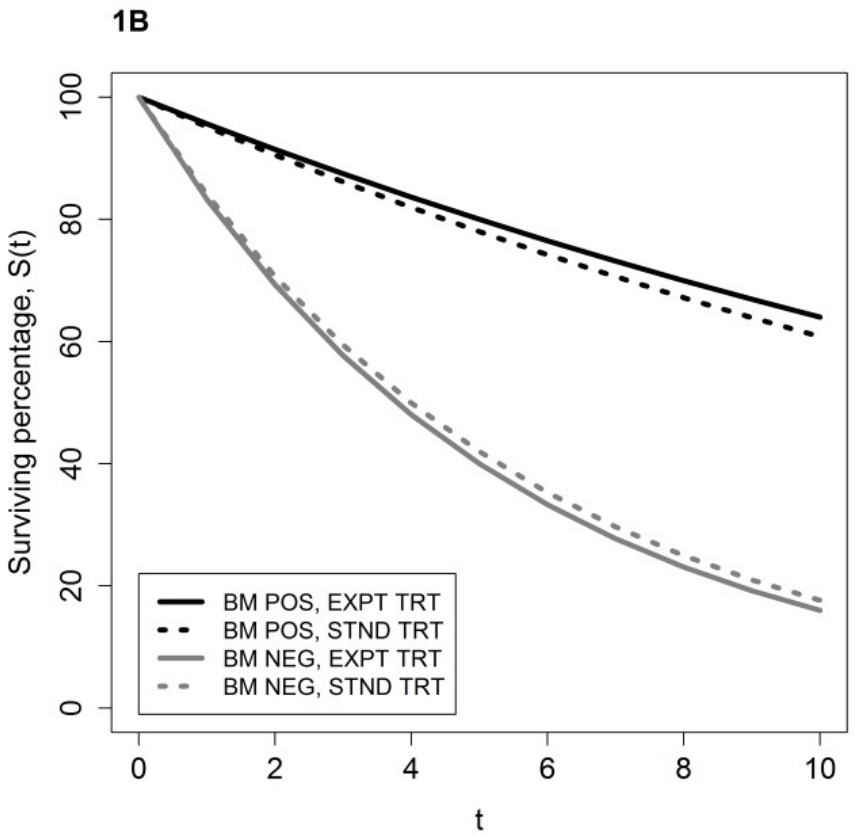
*Wenxuan.Deng@takeda.com*

July 16, 2018

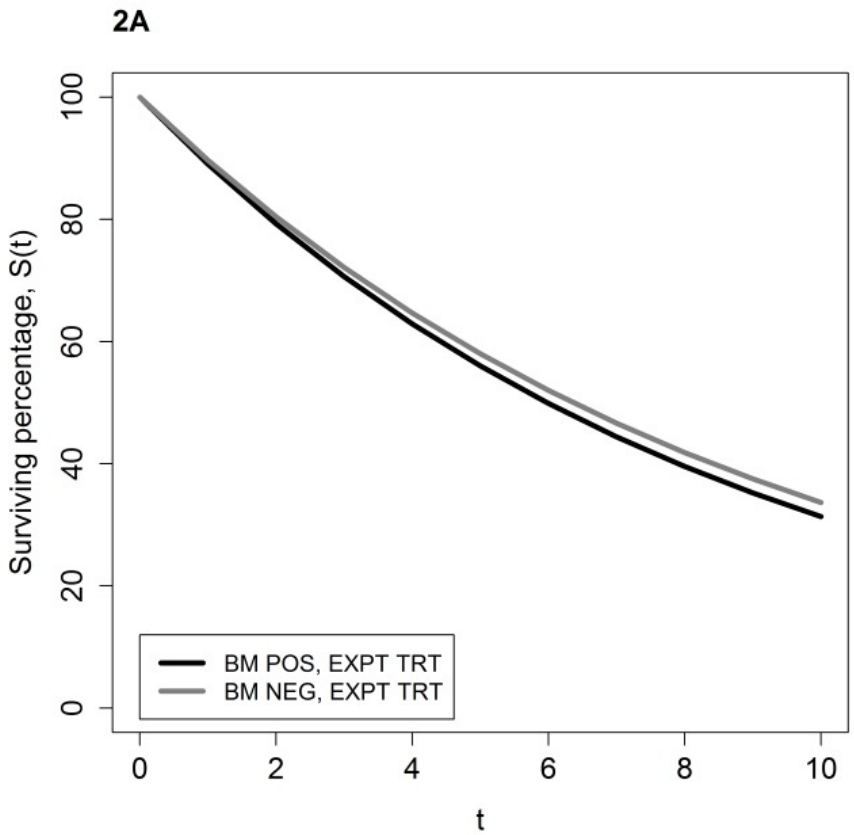
# Overview

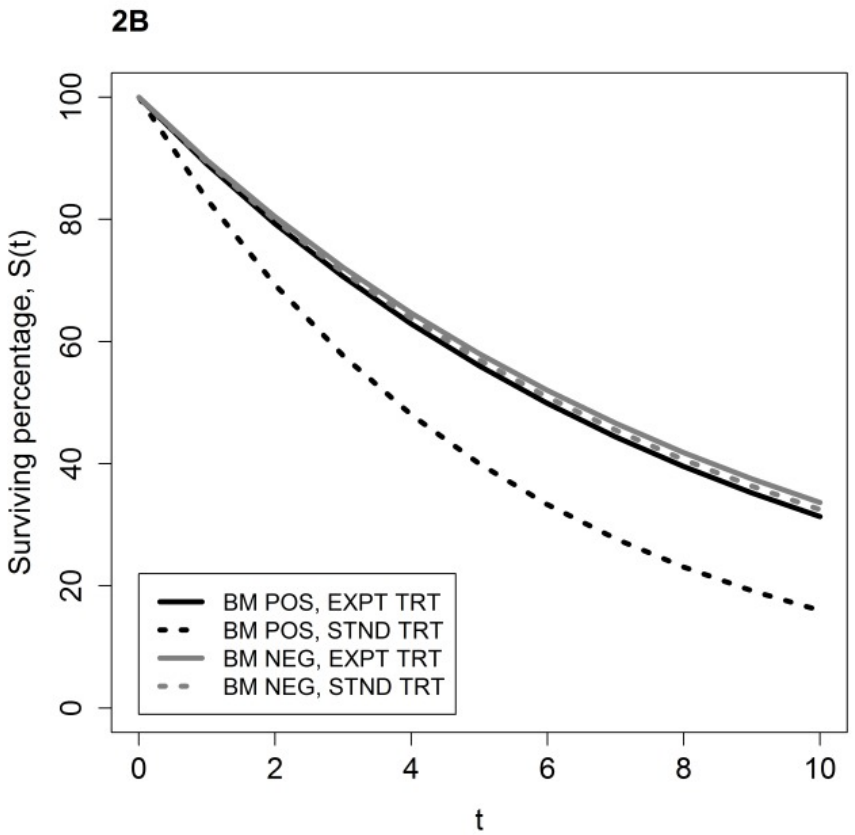
# Prognostic Biomarkers



# Predictive Biomarkers



# Tree-based Methods

Regression trees GUIDE[Loh, 2018]:

- piecewise-linear Model
- examine residual patterns for each treatment level

Cannot repeat even the most naive simulation in GUIDE paper.

Reason: limited sample size. Even two splits will results in small sample size in each branch. The results would be highly unstable. Tree-based method is not appropriate to clinical trial dataset and identify prognostic and predictive biomarkers.

# Ordinary Linear Model

$$Y = X\beta + W\tau + G\alpha + W \otimes G\gamma + \epsilon$$

- $X$: Baseline variables
- $W$: Treatment variables
- $G$: Main effects of genes, i.e. expression levels, SNP or mutation
- $W \otimes G$: Interaction effects of genes and treatment
- $\epsilon$: Random errors

# Group lasso

We choose group lasso for its ability to

- handle high dimensional data
- allow hierarchical structure

However, the current group lasso based methods

- penalize on all parameters
- have no efficient adaptive penalty weights
- do not specifically target on patients treatment subgroup identification

# Loss Function

We assume the hierarchical relationship between prognostic biomarkers and predictive biomarkers, that is a predictive biomarker should be a prognostic biomarker.

The loss function is

$$\min_{\theta} f(\theta|Y, X, W, G) + \lambda \sum_i \eta_i^I |\gamma_i| + \lambda \sum_i \eta_i^M \sqrt{\alpha_i^2 + \gamma_i^2}$$

where $f(\theta|Y, X, W, G)$ is L-2 loss function, i.e. sum of squared errors for ordinary linear model.

$\theta = (\beta, \tau, \alpha, \gamma)$ is parameter vector.

# Loss function for ordinary linear model

$$\min_{\theta} \parallel Y-(X\beta+W\tau+G\alpha+W\otimes G\gamma) \parallel^2 +\lambda\sum_i \eta_i^I|\gamma_i|+\lambda\sum_i \eta_i^M\sqrt{\alpha_i^2+\gamma_i^2}$$

Denote $X^{(I)} = [G_I, W \otimes G_I]$ is the $I$th group of the main and interaction effects of gene $I$. Then we let

$$\eta_i^I = \parallel X^{(i)} \parallel_2$$

# Blocks of Highlighted Text

## Block 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

## Block 2

Pellentesque sed tellus purus. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vestibulum quis magna at risus dictum tempor eu vitae velit.

## Block 3

Suspendisse tincidunt sagittis gravida. Curabitur condimentum, enim sed venenatis rutrum, ipsum neque consectetur orci, sed blandit justo nisi ac lacus.

# Multiple Columns

**Heading**

1. Statement
2. Explanation
3. Example

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

# Table

| Treatments | Response 1 | Response 2 |
|---|---|---|
| Treatment 1 | 0.0003262 | 0.562 |
| Treatment 2 | 0.0015681 | 0.910 |
| Treatment 3 | 0.0009271 | 0.296 |

Table: Table caption

# Theorem

## Theorem (Mass–energy equivalence)

$E = mc^2$

# Verbatim

## Example (Theorem Slide Code)

```
\begin{frame}
\frametitle{Theorem}
\begin{theorem}[Mass--energy equivalence]
$E = mc^2$
\end{theorem}
\end{frame}
```

# References

Loh, WeiYin, Michael Man, and Shuaicheng Wang.

"Subgroups from regression trees with adjustment for prognostic effects and postselection inference."

*Statistics in medicine* (2018).

# The End