New Biomarker-Adaptive Designs of Clinical Trials for Precision Medicine

Feifang Hu

Department of Statistics

George Washington University

Based on joint research with Wanying Zhao and Fan Wang

Email: feifang@gwu.edu

Sept 14, IMA, Twin Cities, USA

Outline

- Precision Medicine
- Predictive Biomarkers and Prognostic Biomarkers
- A General Framework
- New Covariate-Adaptive Designs
- Some Properties
- Conclusion Remarks

1 Precision Medicine

From Wikipedia:

Precision medicine (PM) (also called personalized medicine) is a medical model that proposes the customization of healthcare, with medical decisions, practices, and/or products being tailored to the individual patient. It usually emphasizes the **systematic use of information about an individual patient** to select or optimize that patient's preventative and therapeutic care.

From NIH:

Precision medicine is an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person.

Precision medicine can broadly be defined as **products and service that leverage the science of genomics and proteomics (directly and indirectly)** and capitalize on the trends toward wellness and consumerism to enable tailored approaches to prevention and care.

Over the past century, medical care has centered on standards of care based on epidemiological studies of large cohorts. However, large cohort studies do not take into account the genetic variability of individuals within a population. Precision medicine (also call **Future medicine**) seeks to provide an objective basis for consideration of such individual differences.

As stated by Margaret Hamburg (2010), commissioner, US Food and Drug administration, MD, USA,
"*However, identifying genes that seem to be linked with a disease is only the beginning of an arduous process.* **New approaches** *to the drug-development paradigm are needed such that new drugs are developed along with the tests that inform their use.* **New designs for clinical trials** *are needed so that genetics or other markers can be used to assist in patient selection.*"

"**New designs for clinical trials** are needed so that genetics or other markers can be used to assist in patient selection."

How to do this?

Our approach:
(i) Different biomarkers could play different roles in clinical studies. Based on roles, we may classify them to *predictive* biomarkers and *prognostic* biomarkers.
(ii) Propose new designs that *predictive* biomarkers can be used to assist in patient selection.
(iii) Study their properties.

Some steps to develop precision medicine:

- From DATA to identify some possible predictive biomarkers and treatments: Bio-informatics, genomics, proteomics, and metabolomics, etc.
- Identify **important predictive biomarkers** from possible predictive biomarkers and treatments with Phase II clinical studies.
- Well designed clinical studies to confirm the significance of biomarkers and treatments, identify suitable (new) treatments (drugs), then approved by FDA.
- Implement to healthcare.

2 Predictive biomarkers and prognostic biomarkers

Nowdays, more covariate information is collected in a clinical trial of precision medicine. The increasing importance of the roles of these covariates can be exemplified by many clinical studies.

In phase II of the "Basket" study of Hyman et al. (2015, NEJM), BRAF V600 Mutant is an important predictive covariate.

In another clinical trial that investigated the efficacy and safety of Nivolumab (Larkin et al., JAMA, 2015):

(i) the covariates (such as age, sex, and metastasis stage) are used for the balancing purpose.

(ii) the covariates (BRAF V600 Mutant and BRAF wild-type) could be considered as predictive covariates.

As indicated in Buyse et al (2011) and Krisam and Kieser (2015), there are two distinct types of biomarkers:

(1) prognostic biomarker: *a biomarker can be used to predict the most likely prognosis of an individual patient*;

(2) predictive biomarker: *a biomarker is likely to predict the response to a specific therapy*.

We accordingly classify covariates into two types:
(i) prognostic covariate Z;
and (ii) predictive covariate X.

Correspondingly, prognostic covariates should be balanced to provide a valid comparison, while predictive covariates may be used for the selection of suitable treatments (precision medicine).

A hypothetical trial: Suppose we have three indicate covariates: Z_1, Z_2 (prognostic) and X (predictive),
where $Z_1 = 1$ (male), or $Z_1 = 2$ (female);
$Z_2 = 1$ (investigation center 1), or $Z_2 = 2$ (investigation center 2);
$X = 0$ (biomarker type 0), or $X = 1$ (biomarker type 1).

Main objective: detect the interaction between treatments and the biomarker X.

A good design: (a) balance over covariates (Z_1 and Z_2);
(b) Based on the responses and the covariate X, select the most suitable treatment for patients (precision medicine).

3 A general framework

Consider a clinical trial of n patients, K treatments.
A *randomization sequence* is a random matrix $\boldsymbol{T} = (\boldsymbol{T}_1, ..., \boldsymbol{T}_n)'$ where $\boldsymbol{T}_i = \boldsymbol{e}_j$ for some $j = 1, ...K$, $i = 1, ..., n$ and $\boldsymbol{e}_j$ is a vector of zeros except a 1 in the j-th position, indicating that the patient receives the j-th treatment.

A set of covariate vectors $(\boldsymbol{Z}_1, \boldsymbol{X}_1), ..., (\boldsymbol{Z}_n, \boldsymbol{X}_n)$ are of interest, where $\boldsymbol{Z}_i$ and $\boldsymbol{X}_i$ are the corresponding prognostic covariate and predictive covariate of patient i.

Let $\boldsymbol{Y} = (\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n)'$, where $\boldsymbol{Y}_i = (Y_{i1}, ..., Y_{iK})$, be a random matrix of response variables, where $\boldsymbol{Y}_i$ represents the sequence of responses that would be observed if each treatment were assigned to the i-th patient independently. However, only one element of $\boldsymbol{Y}_i$ is observable.

We will consider probability models for $\boldsymbol{Y}_i$ conditional on $\boldsymbol{T}_i$, $\boldsymbol{Z}_i$ and $\boldsymbol{X}_i$.

Let $\mathcal{T}_n = \sigma\{\boldsymbol{T}_1, ..., \boldsymbol{T}_n\}$; $\mathcal{Y}_n = \sigma\{\boldsymbol{Y}_1, ..., \boldsymbol{Y}_n\}$; $\mathcal{Z}_n = \sigma\{\boldsymbol{Z}_1, ..., \boldsymbol{Z}_n\}$; and $\mathcal{X}_n = \sigma\{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$
Let $\mathcal{F}_n = \mathcal{T}_n \otimes \mathcal{Y}_n \otimes \mathcal{Z}_{n+1} \otimes \mathcal{X}_{n+1}$.

A *randomization procedure* is defined by

$$\boldsymbol{\phi}_n = E(\boldsymbol{T}_n | \mathcal{F}_{n-1}),$$

where $\boldsymbol{\phi}_n$ is $\mathcal{F}_{n-1}$-measurable. One can describe ϕ_n as the conditional probability of assigning treatments $1, ..., K$ to the n-th patient, conditioning on the previous $n-1$ assignments, responses, and covariate vectors, and the current patient's covariate vector.

A superior design of precision medicine relies on whether patients' characteristics (covariates, especially predictive covariates) Therefore, we will focus on
(i)*covariate-adaptive design*

$$\boldsymbol{\phi}_n = E(\boldsymbol{T}_n|\mathcal{F}_{n-1}) = E(\boldsymbol{T}_n|\mathcal{T}_{n-1}, \mathcal{Z}_n, \mathcal{X}_n);$$

and (ii)*covariate-adjusted response-adaptive (CARA) design* if

$$\boldsymbol{\phi}_n = E(\boldsymbol{T}_n|\mathcal{F}_{n-1}) = E(\boldsymbol{T}_n|\mathcal{T}_{n-1}, \mathcal{Y}_{n-1}, \mathcal{Z}_n, \mathcal{X}_n).$$

In the following discussion, for simplicity, we assume that both $\boldsymbol{Z}$ and $\boldsymbol{X}$ are discrete covariates.

4 New CARA procedures

Consider two treatment groups 1 and 2. First, let (Z_j, X_j) be the covariate profile of the j-th patient. Consider that Z has I covariates and m_i levels for the ith covariate, resulting in $m = \prod_{i=1}^{I} m_i$ strata. Let T_j be the assignment of the jth patient, $j = 1, \ldots, n$, i.e., $T_j = 1$ for treatment 1 and $T_j = 0$ for treatment 2. We have $Z_j = (k_1, \ldots, k_I)$ if his or her ith covariate is at level k_i, $1 \leq i \leq I$ and $1 \leq k_i \leq m_i$. For convenience, we use $(k_1, \ldots, k_I)$ to denote the *stratum* formed by patients who possess the same covariate profile $(k_1, \ldots, k_I)$, and use $(i; k_i)$ to denote the *margin* formed by patients whose ith covariate is at level k_i.

The target allocation proportion depends on the covariate X. Suppose the target allocation to treatment 1 is $\rho(X)$.

The procedure is defined as follows:

1) The first $2K$ patients are assigned to treatment 1 and 2 with restrict randomization (K patients in each treatment).

2) Suppose $(n-1)$ patients have been assigned to treatments $(n > 2K)$ and we observed their responses. The nth patient has covariate (Z_n, X_n), where Z_n falls within stratum $(k_1^*, \ldots, k_I^*)$.

   Based on $n-1$ responses and X_n, we estimate the target allocation $\hat{\rho}(X_n)$.

3) If the nth patient were assigned to treatment 1,

   - let $D_n^{(1)}(X_n) = N_{n1}(X_n) - \hat{\rho}(X_n)N_n(X_n)$, where $N_n(X_n)$ is the total number of patients with covariate X_n for the first n patients. Further, $N_{n1}(X_n)$ is the number of patients with covariate X_n in treatment group 1.

- similarly, let $D_n^{(1)}(i; k_i^*; X_n) = N_{n1}(i; k_i^*; X_n) - \hat{\rho}(X_n) N_n(i; k_i^*; X_n)$ and $D_n^{(1)}(k_1^*, \ldots, k_I^*; X_n) = N_{n1}(k_1^*, \ldots, k_I^*; X_n) - \hat{\rho}(X_n) N_n(k_1^*, \ldots, k_I^*; X_n)$ be the differences between the target and actual numbers of patients in the treatment 1 on the margin $(i; k_i^*)$, and within the stratum $(k_1^*, \ldots, k_I^*)$, respectively;
- these differences can be positive, negative or zero, and each one is used to measure the *imbalance* at the corresponding level (overall, marginal, or within-stratum).

4) Define an imbalance measure $Imb_n^{(1)}(X_n)$ by

$$Imb_n^{(1)}(X_n) = w_o[D_n^{(1)}(X_n)]^2 + \sum_{i=1}^{I} w_{m,i}[D_n^{(1)}(i; k_i^*; X_n)]^2$$

$$+w_s[D_n^{(1)}(k_1^*, \ldots, k_I^*; X_n)]^2,$$

which is the weighted imbalance that would be caused if the nth patient were assigned to treatment 1. w_o, $w_{m,i}$ $(i = 1, \ldots, I)$ and w_s are nonnegative weights placed on overall, within a covariate margin and within a stratum cell, respectively. Without loss of generality we can assume

$$w_o + w_s + \sum_{i=1}^{I} w_{m,i} = 1.$$

5) If the nth patient were assigned to treatment 2, we can define $Imb_n^{(2)}$ in a similar fashion, that it represents the weighted imbalance that would be caused if the nth patient were assigned to treatment 2.

6) Conditional on the assignments of the first $(n-1)$ patients as well as the covariates' profiles of the first n patients, assign the nth patient to treatment 1 with probability

$$\begin{aligned} &P(T_n = 1 | \mathcal{Z}_{n-1}, Z_n = (k_1^*, \ldots, k_I^*), X_n, \mathcal{T}_{n-1}) \\ &= g\left(Imb_n^{(1)} - Imb_n^{(2)}\right) \end{aligned} \tag{1}$$

where $n > 1$, $g(x)$ is a real function with $0 < g(x) < 1$, $g(-x) = 1 - g(x)$,

$$g(x) \leq 0.5 \text{ when } x \geq 0, \quad \text{and} \quad \limsup_{x \to +\infty} g(x) < 0.5.$$

In practice, one may use the following Efron's biased coin function (2) with $p \in [0.75, 0.95]$ as discussed and suggested in Hu and Hu (2012). That is,

$$g(x) = \begin{cases} q, & \text{if } x > 0, \\ \frac{1}{2}, & \text{if } x = 0, \\ p, & \text{if } x < 0, \end{cases} \tag{2}$$

where $p > 1/2$ and $q + p = 1$. In general, we can define $g(x)$ to be either a continuous function or a discrete function.

Some special cases:

(1) Without the covariate X_n and $\rho = 1/2$, the proposed procedure reduces to the covariate-adaptive designs of balancing covariates: Stratified Permuted Block Randomization; Pocock-Simon procedures (1975); Hu and Hu procedure (2012), etc.

(2) Without the covariate Z_n, the proposed procedure reduces to the CARA designs (Zhang, Hu, Cheung and Chan, 2007; Hu, Zhu and Hu, 2015).

5 Preliminary properties

Numerical studies under a logistic regression.

Assume that the response variable Y_i (ith patient) follows a logistic regression:

$$logit(P(Y_i = 1)) = \mu_1 T_i + \mu_2(1 - T_i) + \beta_1 X_i + \alpha_1 T_i X_i,$$

where $T_i = 1$ (treatment 1) or 0 (treatment 2) as treatment assignment, and X_i takes value 1 (presence of biomarker) and -1 (absence of biomarker). For the two randomization procedures, one will explore whether the imbalances on any of the three levels (within-stratum, marginal and overall) stabilize.

The parameters are specified as follows:

- Z_1, Z_2, X are independently distributed, and take values 1 or -1 with equal probability.
- Biasing probability $p = 0.80$ and $q = 0.2$ for NEW procedures.
- Sample size $n = 200, 400, 1000$; number of simulated trials $N = 5000$.
- NEW: $(w_o, w_{m,1}, w_{m,2}, w_s) = (0.25, 0.25, 0.25, 0.25)$.
- The success probabilities:
  $P(Y = 1|T = 1, X = 1) = 0.36,$
  $P(Y = 1|T = 0, X = 1) = 0.67,$
  $P(Y = 1|T = 1, X = -1) = 0.70,$
  $P(Y = 1|T = 0, X = -1) = 0.33.$

- We use the optimal allocation proposed by Rosenberger *et al* (2001):

$$\rho(X=1) = \frac{\sqrt{P(Y=1|T=1,X=1)}}{\sqrt{P(Y=1|T=0,X=1)}+\sqrt{P(Y=1|T=1,X=1)}}$$

and

$$\rho(X=-1) = \frac{\sqrt{P(Y=1|T=1,X=-1)}}{\sqrt{P(Y=1|T=0,X=-1)}+\sqrt{P(Y=1|T=1,X=-1)}}.$$

In the simulation, we estimate the probabilities sequentially and then estimate $\rho(X = 1)$ and $\rho(X = -1)$.
At the end of the trial, obtain $D_n(X)$ where
$D_n(X) = N_{n1}(X) - \rho(X)N_n(X)$, and $\rho(X)$ is the theoretical value.
Here $\rho(X = 1) = 0.42$ and $\rho(X = -1) = 0.59$ based on the proportion of Rosenberger *et al.* (2001).

Table 1: STD's of $D_n(\cdot)$ of the complete randomization (CR) and the new method (NEW) conditional on $X = 1$ and $T = 1$ under a logistic regression

	SS	$D_n(1;1)$	$D_n(2;2)$	$D_n(1,1)$	$D_n(2,2)$	D_n
CR	200	3.46	3.53	2.53	2.49	4.87
	400	5.11	4.94	3.54	3.45	7.14
	1000	8.07	8.00	5.52	5.61	11.26
NEW	200	1.55	1.51	1.10	1.12	2.51
	400	1.91	2.02	1.22	1.36	3.43
	1000	2.92	2.99	1.65	1.77	5.47

We conduct a preliminary simulation study to investigate the finite sample statistical inference.

To compare treatment effects, the working model of inference is

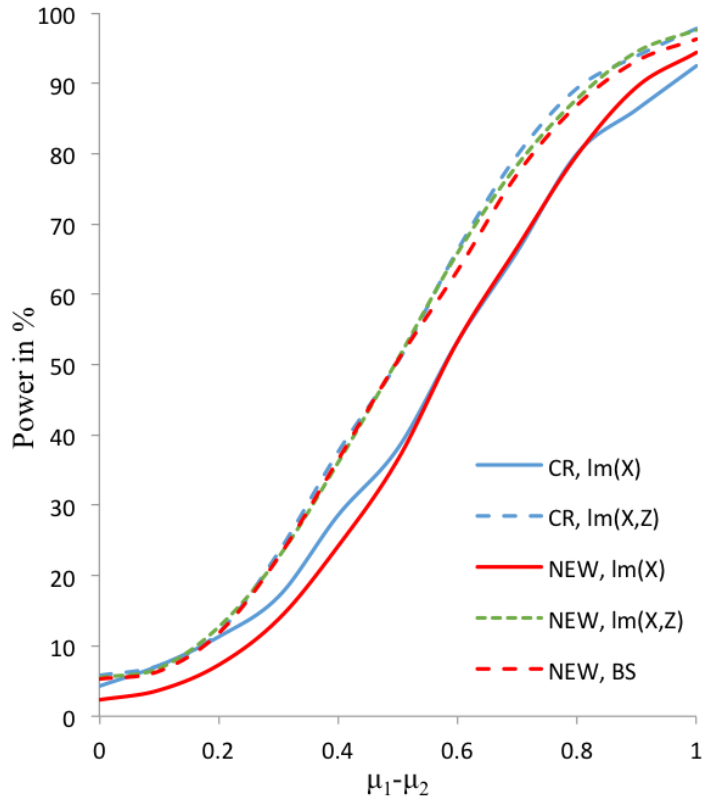
$$logit(P(Y_i = 1)) = \mu_1 T_i + \mu_2(1 - T_i) + \beta_1 X_i + \alpha_1 T_i X_i.$$

To study Type I error under NEW procedure and complete randomization (CR),

- Hypothesis testings: Wald test using working model ($lm(X)$) and corresponding Bootstrap test (BS), Wald test using full model ($lm(X, Z)$).
- Parameters in the logistic model:
  $\mu_1 = \mu_2 = 0.2,\ \beta_1 = 1,\ \alpha_1 = -2,\ \gamma_1 = \gamma_2 = 1.$
- Sample size $n = 200,\ 400,\ 1000$, number of simulation trials $N = 1000$, number of bootstrap $B = 200$.

Table 2: Simulated Type I error for comparing treatment effects under the complete randomization (CR) and the new method (NEW) under a logistic regression

	SS	$lm(X)$	BS	$lm(X, Z)$
CR	200	5.7	-	5.9
	400	5.1	-	5.3
	1000	6.5	-	5.9
NEW	200	2.1	5.1	5.4
	400	2.1	4.9	4.9
	1000	1.8	5.2	5.5



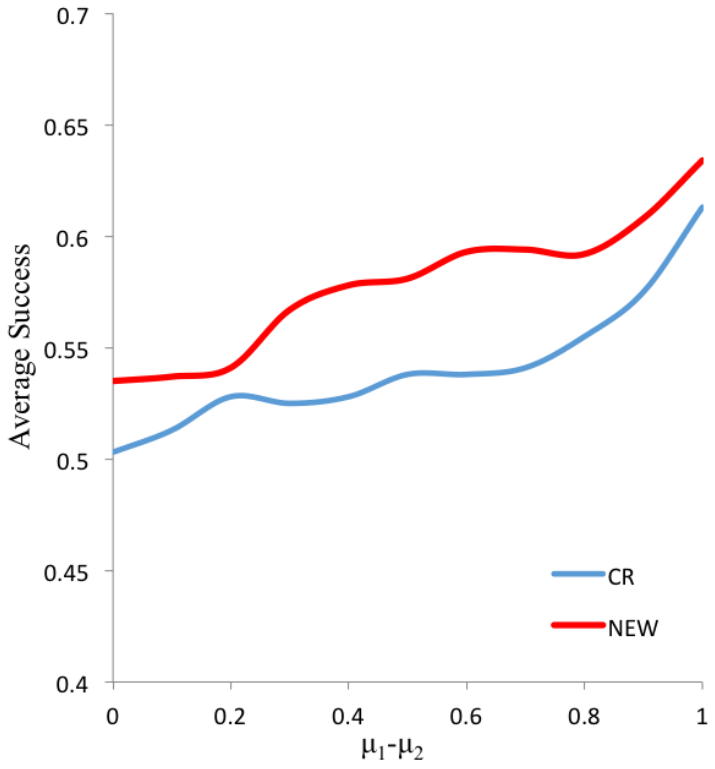


Figure 1: Simulated power (left) and average success (right) for NEW procedure (NEW) and complete randomization (CR) with $\mu_2 = 0$ under a logistic regression.

Numerical studies under a linear regression.

Assume that the response variable Y_i (ith patient) follows a linear regression:

$$\mathbb{E}(Y_i) = \mu_1 T_i + \mu_2(1 - T_i) + \beta_1 X_i + \alpha_1 T_i X_i + \gamma_1 Z_{1i} + \gamma_2 Z_{2i},$$

where $T_i = 1$ (treatment 1) or 0 (treatment 2) as treatment assignment, and X_i takes value 1 (presence of biomarker) and -1 (absence of biomarker) as the biomarker indicator. For the two randomization procedures, NEW CARA and complete randomization, one will explore whether the imbalances on any of the three levels (within-stratum, marginal and overall) stabilize.

The parameters are specified as follows:

- Parameters in the linear model:

  $\mu_1 = \mu_2 = 0.5,\ \beta_1 = 1,\ \alpha_1 = -0.5,\ \gamma_1 = \gamma_2 = 0.5.$

- We use the optimal allocation proposed by Rosenberger *et al* (2005):

$$\rho(X=1) = \frac{\sqrt{\mathbb{E}(Y|T=1,X=1)}}{\sqrt{\mathbb{E}(Y|T=1,X=1)}+\sqrt{\mathbb{E}(Y|T=0,X=1)}} = 0.4741$$

  and

$$\rho(X=-1) = \frac{\sqrt{\mathbb{E}(Y|T=1,X=-1)}}{\sqrt{\mathbb{E}(Y|T=1,X=-1)}+\sqrt{\mathbb{E}(Y|T=0,X=-1)}} = 0.6184.$$

Table 3: STD's of $D_n(\cdot)$ of the complete randomization (CR) and the new method (NEW) conditional on $X = 1$ and $T = 1$ under a linear regression

	SS	$D_n(1;1)$	$D_n(2;2)$	$D_n(1,1)$	$D_n(2,2)$	D_n
CR	200	3.56	3.54	2.51	2.50	5.03
	400	5.03	5.01	3.56	3.51	7.18
	1000	7.96	7.95	5.48	5.56	11.32
NEW	200	0.98	0.94	0.82	0.81	1.42
	400	1.16	1.10	0.86	0.85	1.84
	1000	1.52	1.52	0.98	1.00	2.73

We conduct a preliminary simulation study to investigate the finite sample statistical inference.

To compare treatment effects, the working model of inference is

$$\mathbb{E}(Y_i) = \mu_1 T_i + \mu_2(1 - T_i) + \beta_1 X_i + \alpha_1 T_i X_i.$$

To study Type I error under NEW procedure and complete randomization (CR),

- Hypothesis testings: Wald test using working model ($lm(X)$) and corresponding Bootstrap test (BS), Wald test using full model ($lm(X, Z)$).
- Parameters in the linear model:
  $\mu_1 = \mu_2 = 0.5,\ \beta_1 = 1,\ \alpha_1 = -0.5,\ \gamma_1 = \gamma_2 = 0.5.$
- Sample size $n = 200,\ 400,\ 1000$, number of simulation trials $N = 1000$, number of bootstrap $B = 200$.

Table 4: Simulated Type I error for comparing treatment effects under the complete randomization (CR) and the new method (NEW) under a linear regression

	SS	$lm(X)$	BS	$lm(X, Z)$
CR	200	5.0	-	5.4
	400	5.3	-	5.0
	1000	5.0	-	4.8
NEW	200	2.2	5.4	5.3
	400	1.9	5.4	5.1
	1000	1.5	4.8	4.7

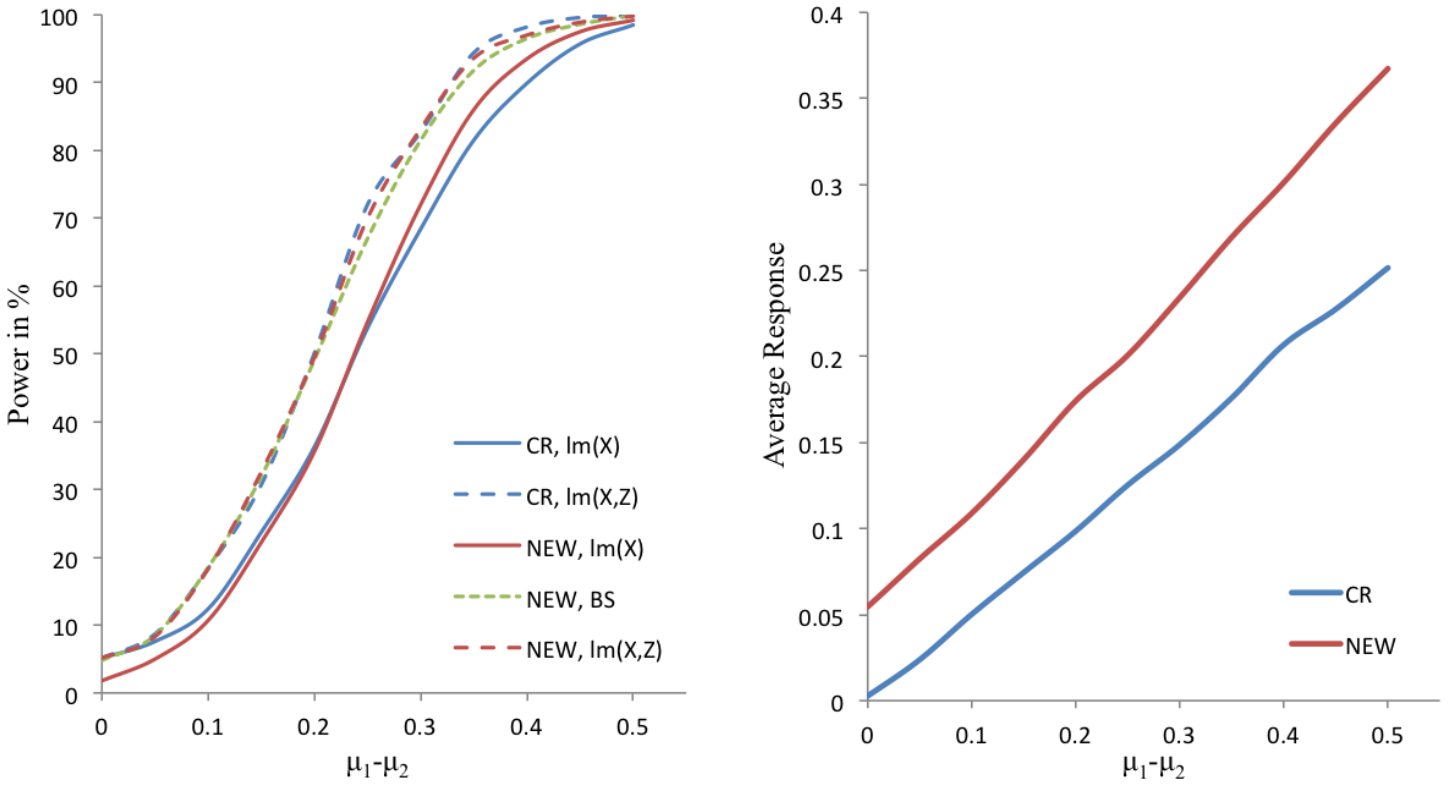


Figure 2: Simulated power (left) and average response (right) for NEW procedure (NEW) and complete randomization (CR) with $\mu_2 = 0$ under a linear regression.

6 Conclusions Remarks

- Prognostic covariate Z and predictive covariate X.
- New and suitable CARA designs for precision medicine.
- Precision Medicine, many new problems about designs and inference.

Thank you!