A Nonparametric Framework for Treatment Effect Modifier Discovery in High Dimensions

Causal Inference in Randomized Studies - ACIC 2023

Philippe Boileau, UC Berkeley — May 2023

Collaborators

Nima Hejazi
Harvard

Sandrine Dudoit
UC Berkeley

Mark van der Laan
UC Berkeley

Ning Leng
Genentech

Motivation

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

IMmotion 150
Phase 2

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

IMmotion 150
Phase 2

Immune checkpoint inhibitor
vs
Immune checkpoint inhibitor
+ tyrosine kinase inhibitor
vs
Tyrosine kinase inhibitor

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

IMmotion 150
Phase 2

IMmotion 151
Phase 3

Immune checkpoint inhibitor
vs
Immune checkpoint inhibitor
+ tyrosine kinase inhibitor
vs
Tyrosine kinase inhibitor

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

IMmotion 150
Phase 2

IMmotion 151
Phase 3

Immune checkpoint inhibitor
vs
Immune checkpoint inhibitor
+ tyrosine kinase inhibitor
vs
Tyrosine kinase inhibitor

Metastatic Renal Cell Carcinoma

Finding biomarkers predictive of clinical benefit

(Previous) standard of care, tyrosine kinase inhibitors, are ineffective

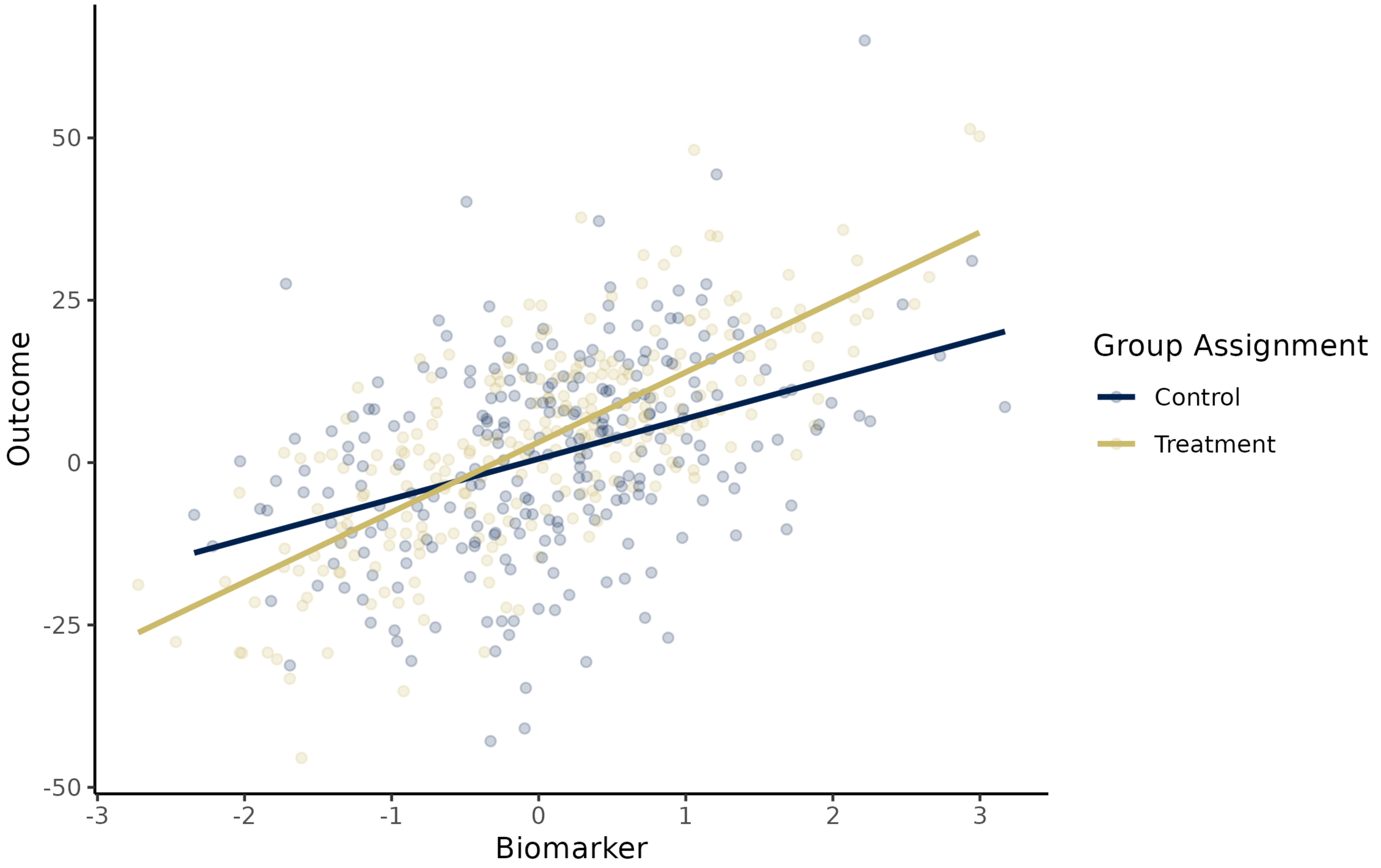
IMmotion 150
Phase 2

Immune checkpoint inhibitor
vs
Immune checkpoint inhibitor
+ tyrosine kinase inhibitor
vs
Tyrosine kinase inhibitor

IMmotion 151
Phase 3

Immune checkpoint inhibitor
+ tyrosine kinase inhibitor
vs
Tyrosine kinase inhibitor

Predictive Biomarkers: Treatment Effect Modifiers



Predictive Biomarker Applications

Predictive biomarkers drive personalized medicine

Predictive Biomarker Applications

Predictive biomarkers drive personalized medicine

- **Diagnostic assay development:** Who benefits most from a therapy?

Predictive Biomarker Applications

Predictive biomarkers drive personalized medicine

- **Diagnostic assay development:** Who benefits most from a therapy?
- **Targeted drug discovery**: What is the biological mechanism of a therapy?

Predictive Biomarker Applications

Predictive biomarkers drive personalized medicine

- **Diagnostic assay development:** Who benefits most from a therapy?
- **Targeted drug discovery**: What is the biological mechanism of a therapy?
- **Refined clinical trials**: Establish a subset of the patient population for which therapy is more efficacious?

Discovering Treatment Effect Modifiers

A variable selection problem

A variable selection problem

- Easy when there are few biomarkers to consider:

A variable selection problem

- Easy when there are few biomarkers to consider:
  - Linear models with treatment-biomarker interaction terms

A variable selection problem

- Easy when there are few biomarkers to consider:
  - Linear models with treatment-biomarker interaction terms
  - Conditional average treatment effect (CATE) estimation

A variable selection problem

- Easy when there are few biomarkers to consider:
  - Linear models with treatment-biomarker interaction terms
  - Conditional average treatment effect (CATE) estimation
- Harder when there are a large number of biomarkers: Penalized versions of the above methods are used.

A variable selection problem

- Easy when there are few biomarkers to consider:
  - Linear models with treatment-biomarker interaction terms
  - Conditional average treatment effect (CATE) estimation
- Harder when there are a large number of biomarkers: Penalized versions of the above methods are used.
- **Bottom line**: Discovery of predictive biomarkers is the byproduct of another inference procedure.

Example: Modified Covariates Approach

Tian et al. (2014): A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates

Example: Modified Covariates Approach

transformed outcome = treatment + treatment x biomarker 1 + … + treatment x biomarker p

Tian et al. (2014): A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates

Example: Modified Covariates Approach

transformed outcome = treatment + treatment x biomarker 1 + … + treatment x biomarker p

Predictive biomarkers have non-zero coefficients

Tian et al. (2014): A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates

Issues with Penalized Regression Methods

Issues with Penalized Regression Methods

Strong assumptions: sparsity and correlation structure.

Issues with Penalized Regression Methods

Strong assumptions: sparsity and correlation structure.

Violations produce to high false positive rates, leading to:

Issues with Penalized Regression Methods

Strong assumptions: sparsity and correlation structure.

Violations produce to high false positive rates, leading to:

- Resources wasted on follow-up experiments and trials

Issues with Penalized Regression Methods

Strong assumptions: sparsity and correlation structure.

Violations produce to high false positive rates, leading to:

- Resources wasted on follow-up experiments and trials
- Decreased signal to noise ratio in diagnostic assays

Issues with Penalized Regression Methods

Strong assumptions: sparsity and correlation structure.

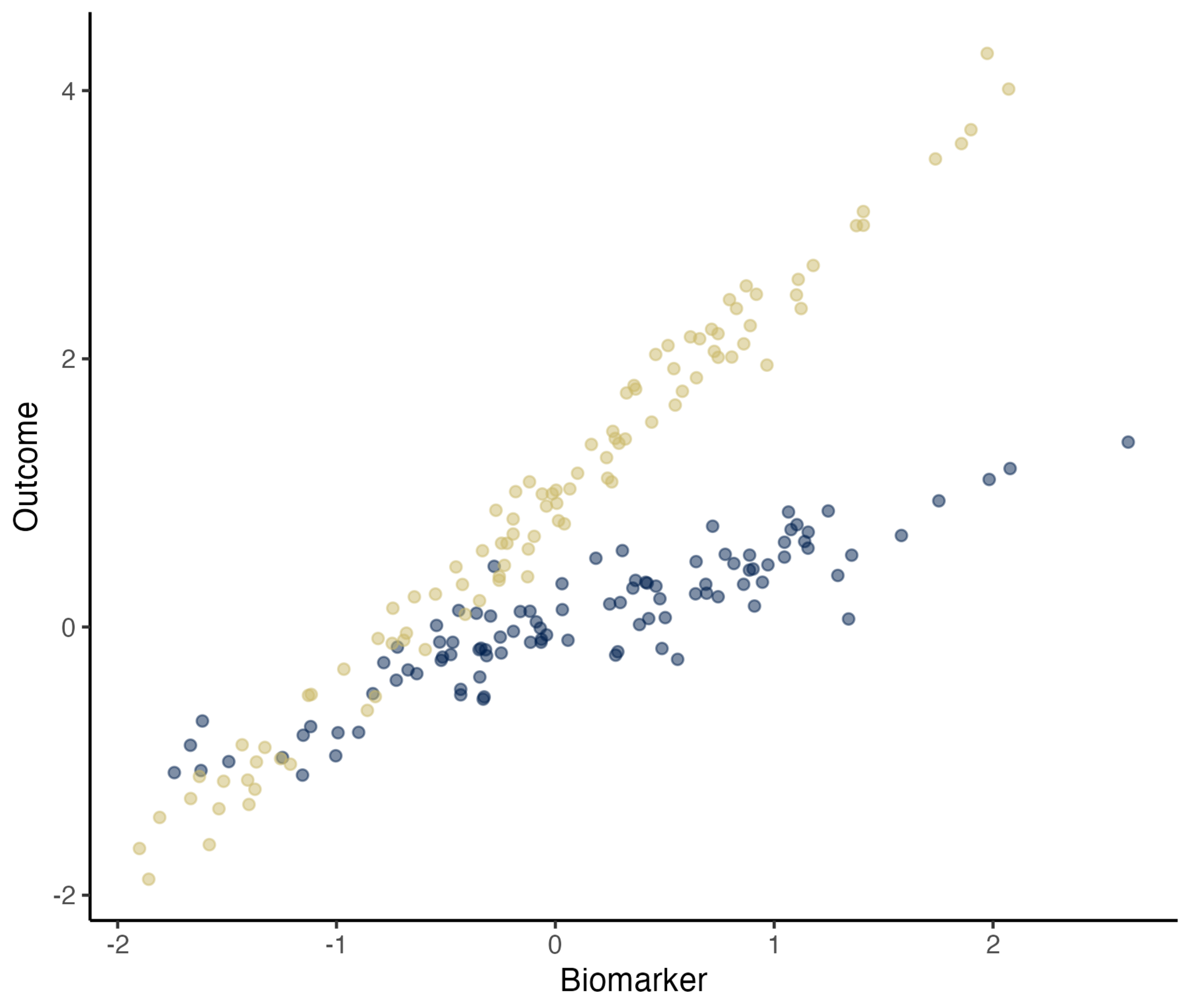
Violations produce to high false positive rates, leading to:

- Resources wasted on follow-up experiments and trials
- Decreased signal to noise ratio in diagnostic assays

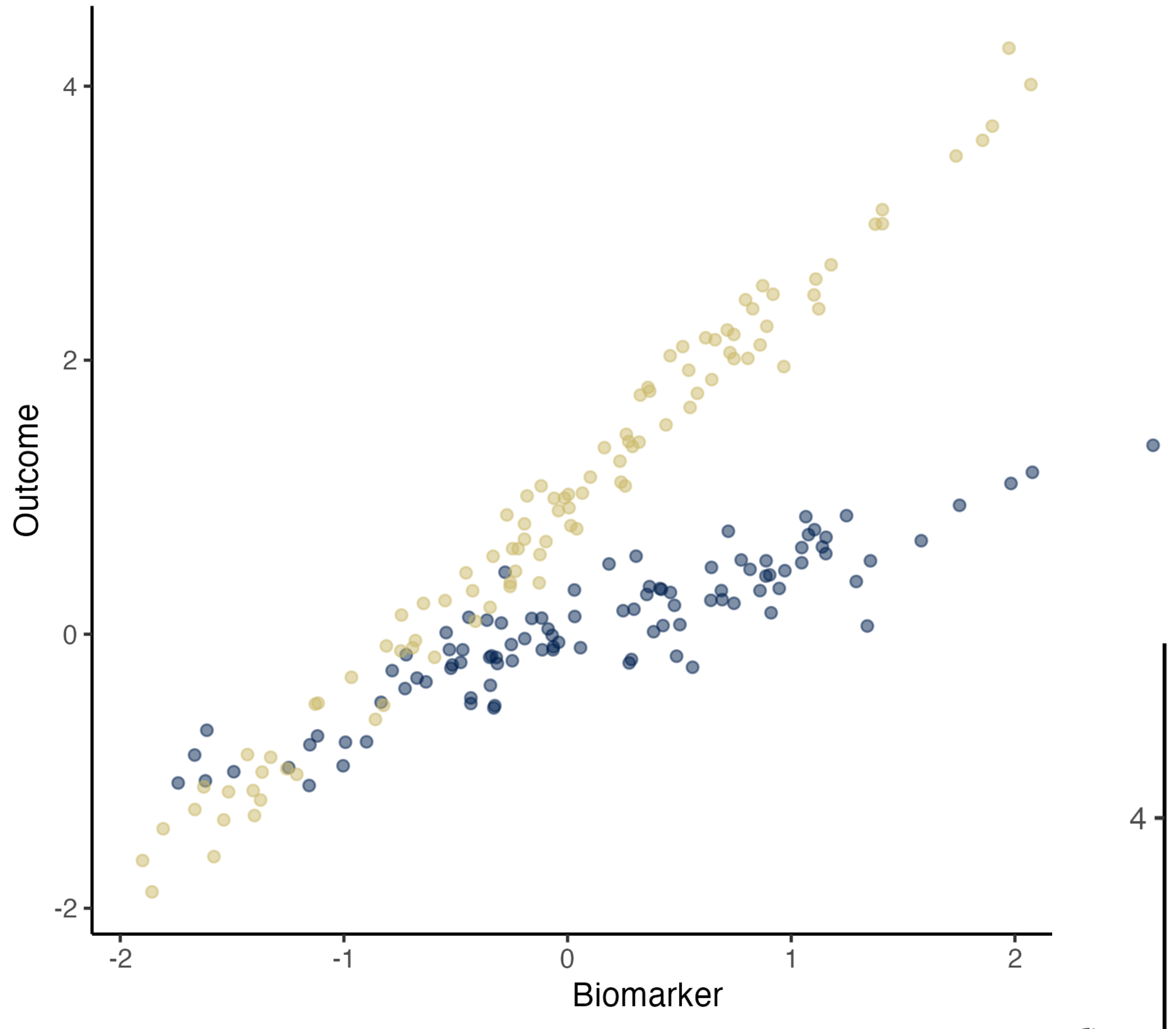
We need to consider alternative problem formulations.

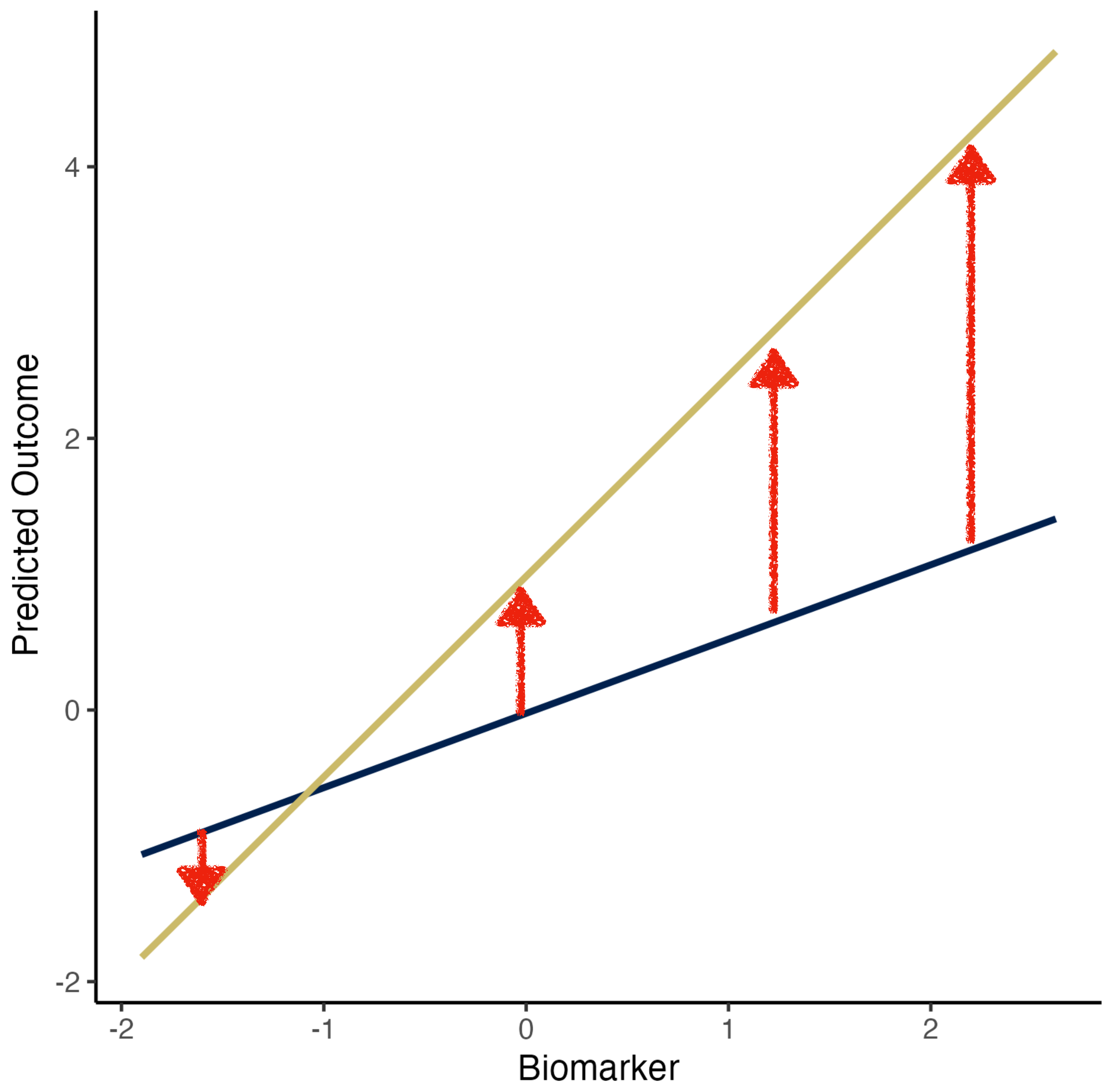
Variable Importance Parameters

Variable Importance Parameters

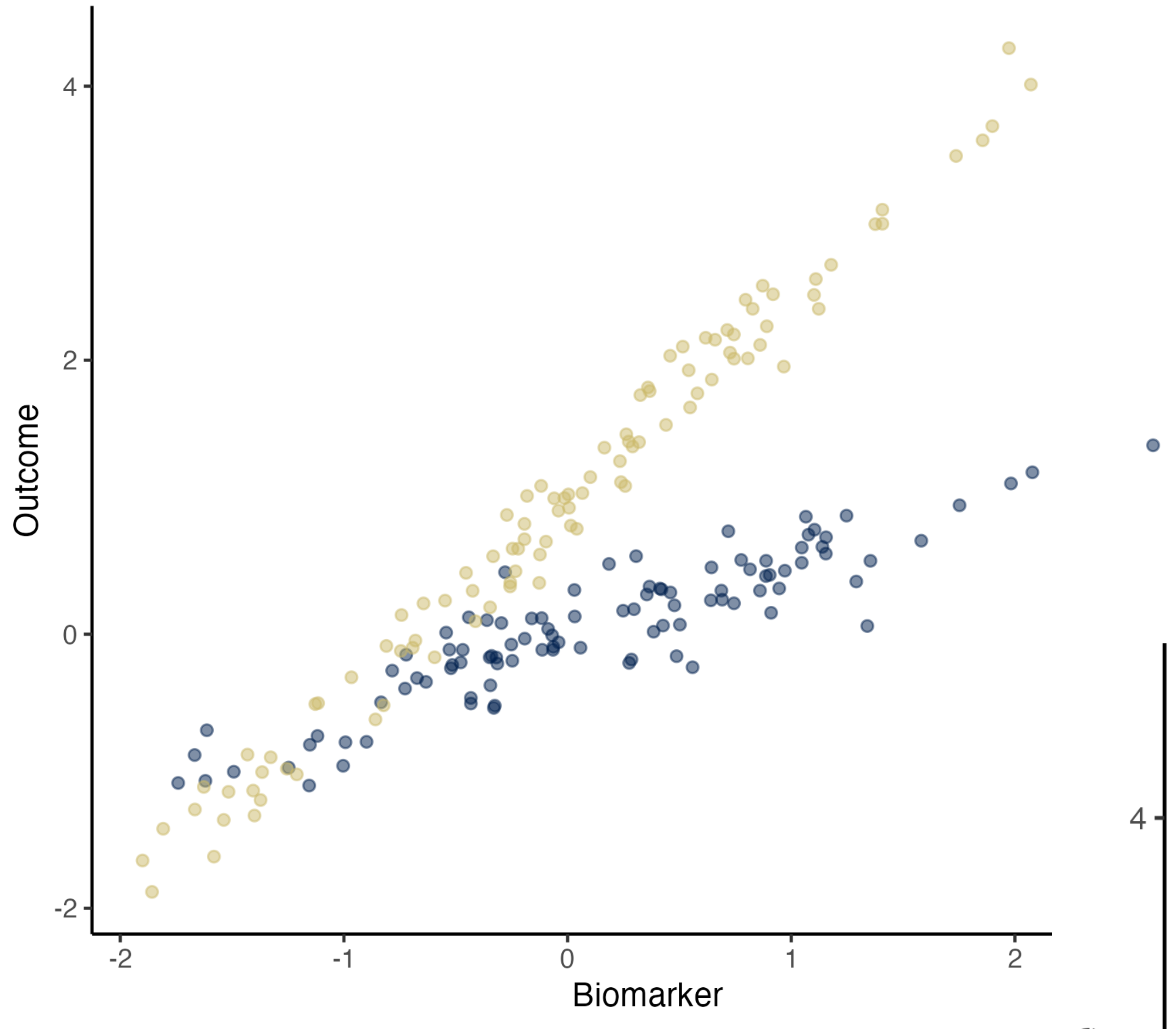


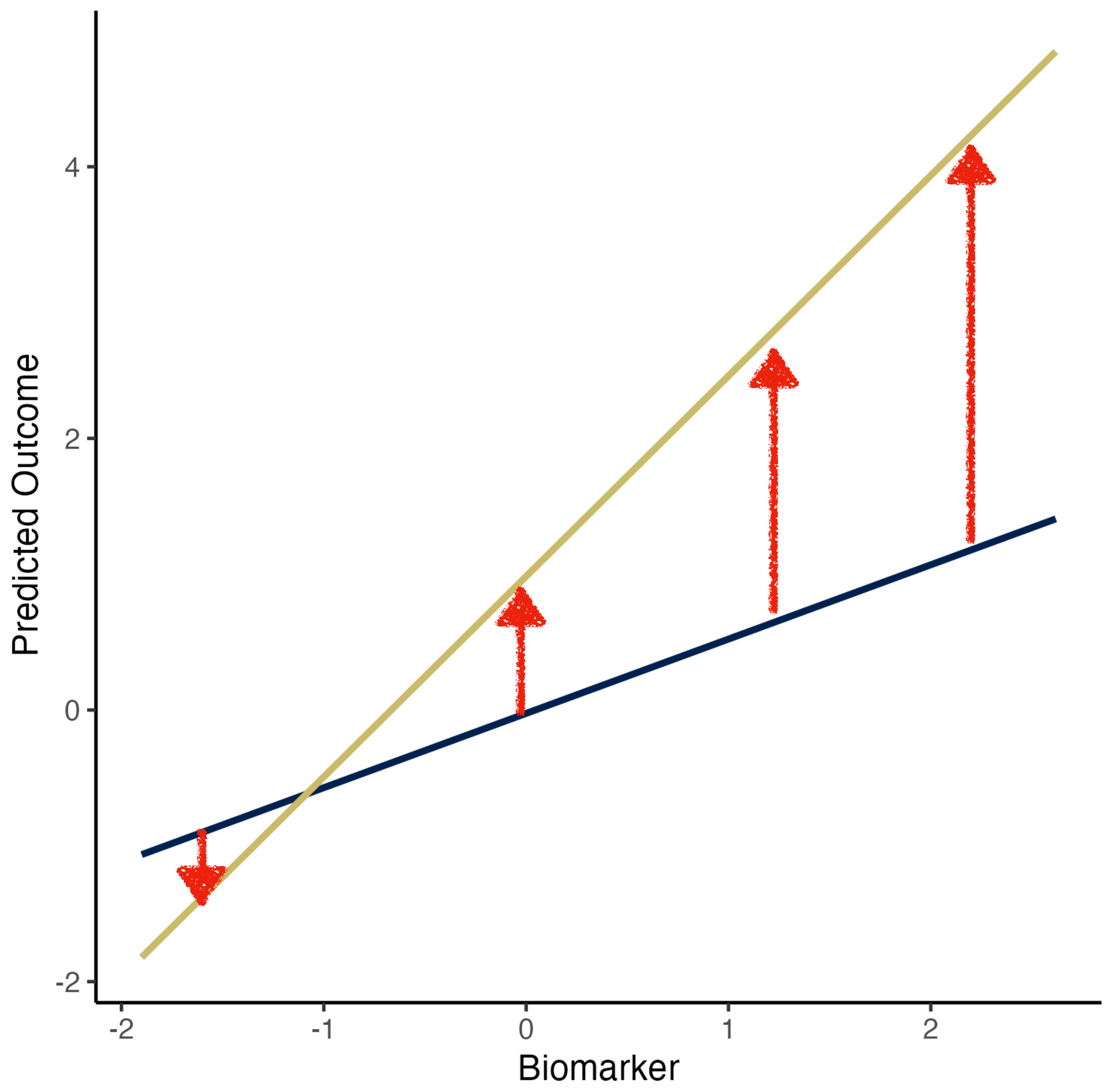
Variable Importance Parameters

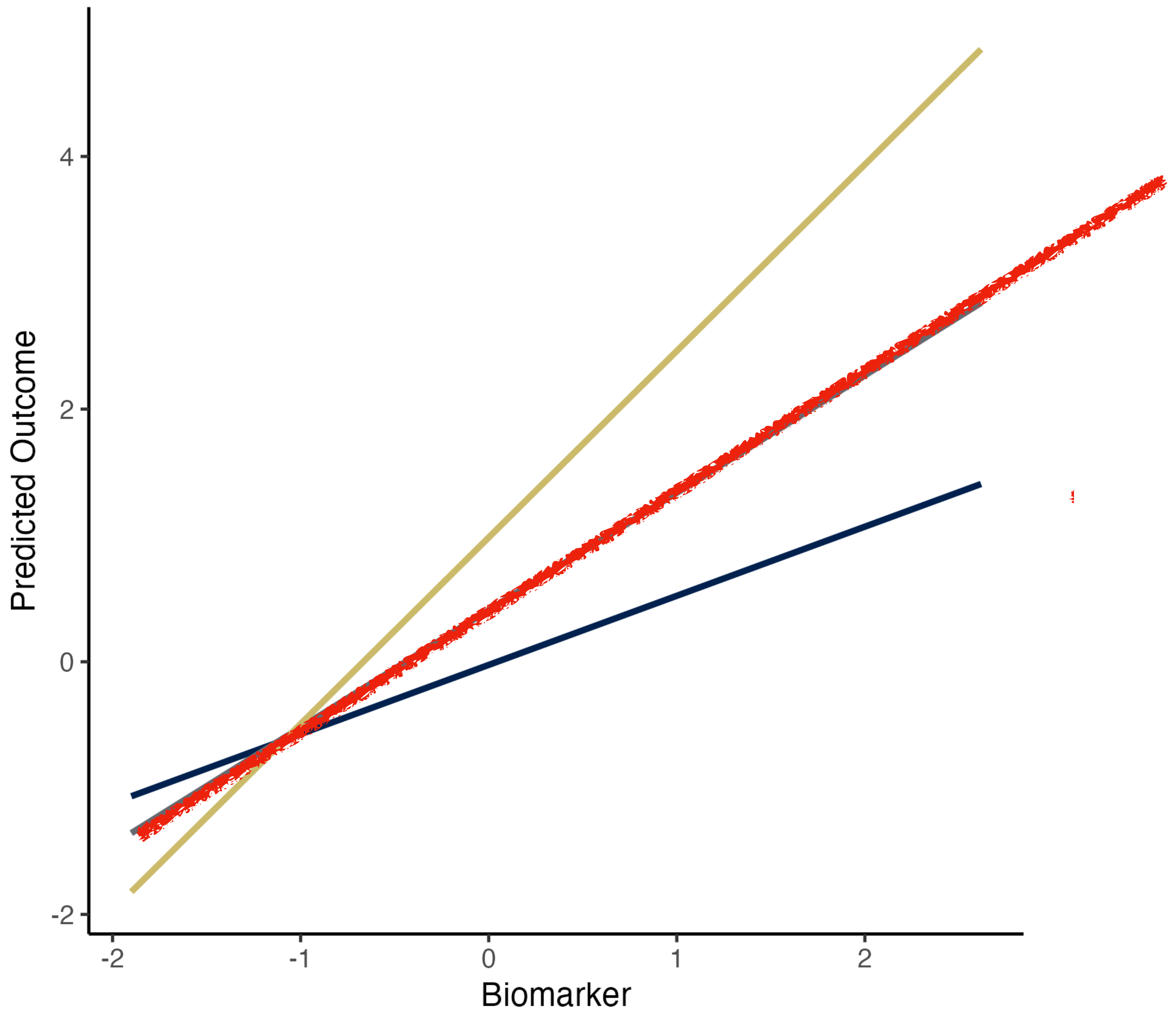




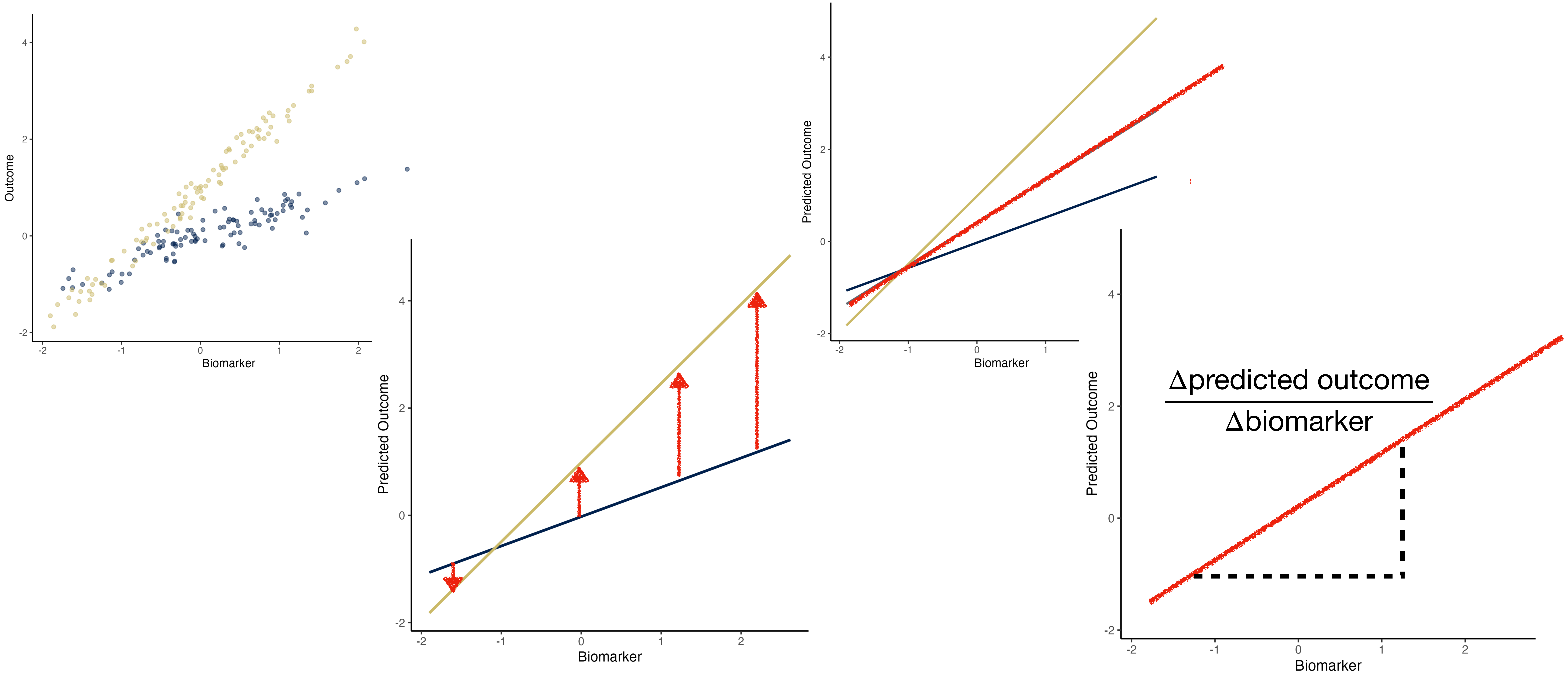
Variable Importance Parameters







Variable Importance Parameters



Treatment Effect Modifier Variable Importance Parameters

Treatment Effect Modifier Variable Importance Parameters

Pathwise differentiable treatment effect parameter

Treatment Effect Modifier Variable Importance Parameters

Pathwise differentiable treatment effect parameter

→

Define univariate conditional parameter

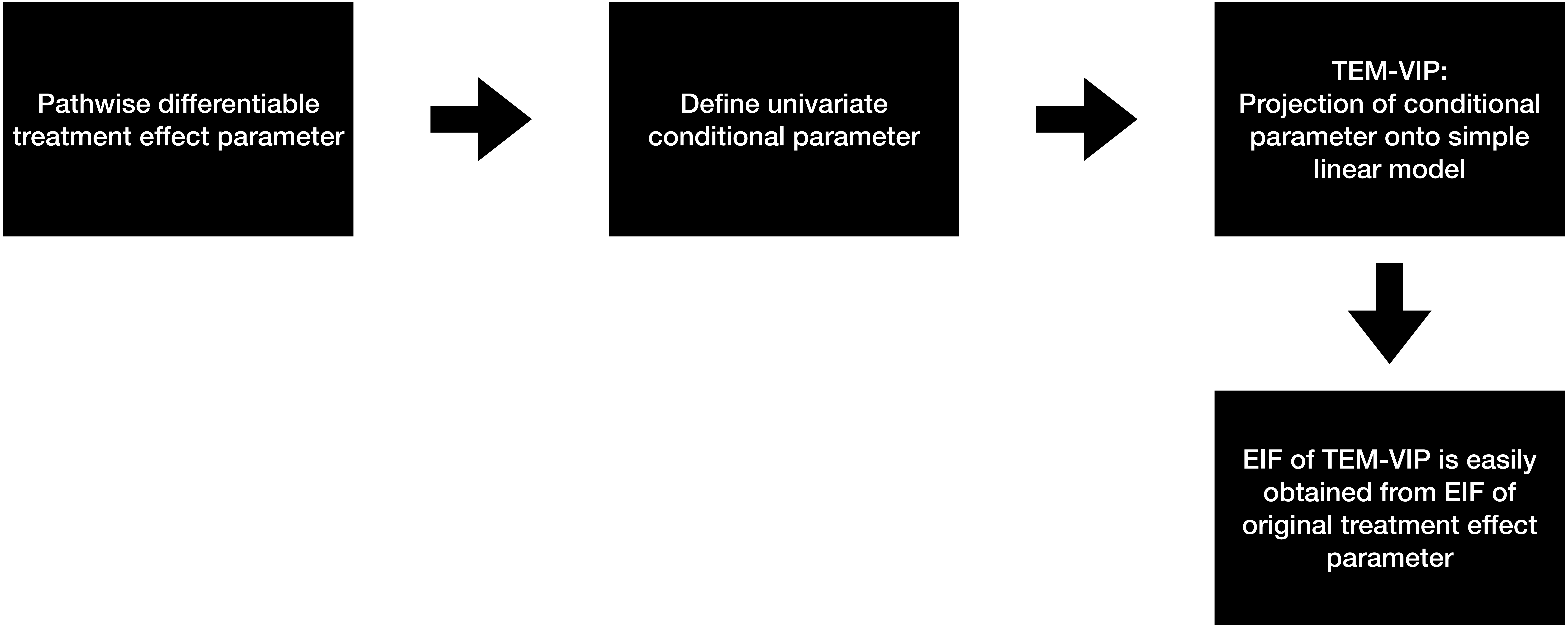
Treatment Effect Modifier Variable Importance Parameters

Pathwise differentiable treatment effect parameter

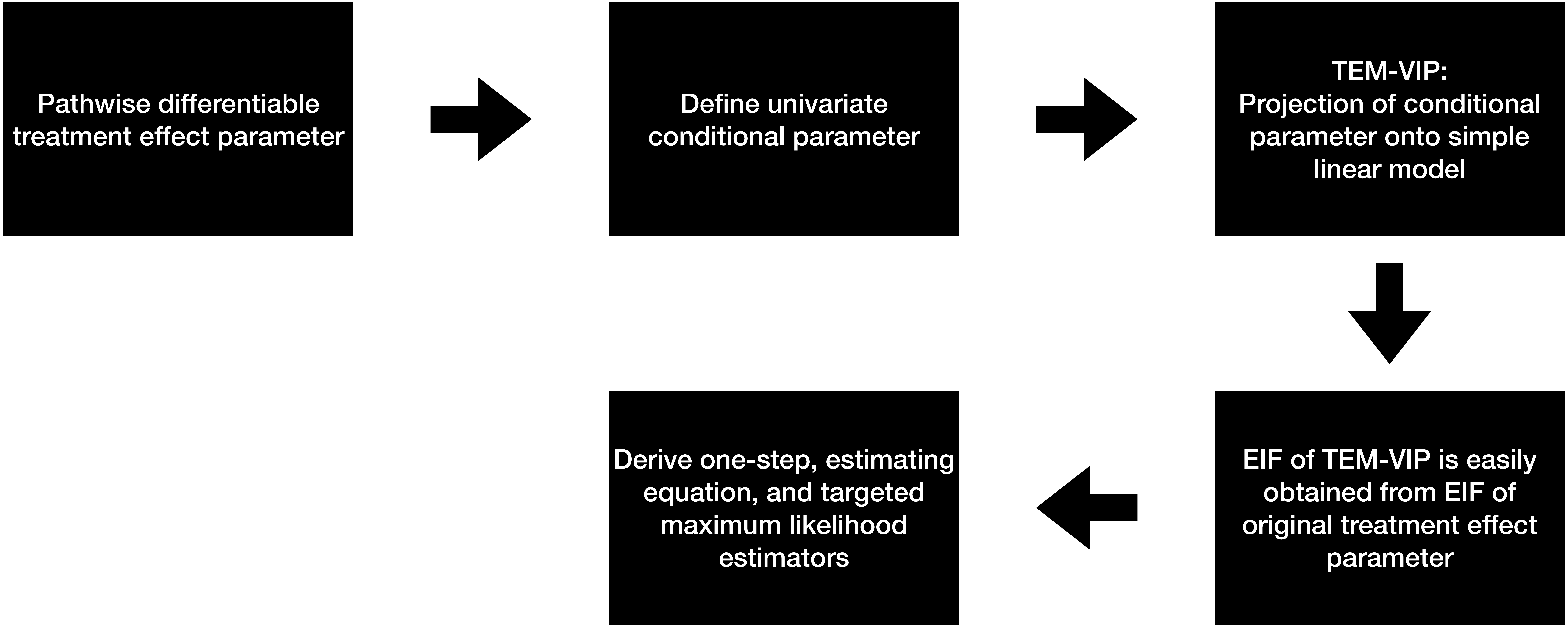
Define univariate conditional parameter

TEM-VIP:
Projection of conditional parameter onto simple linear model

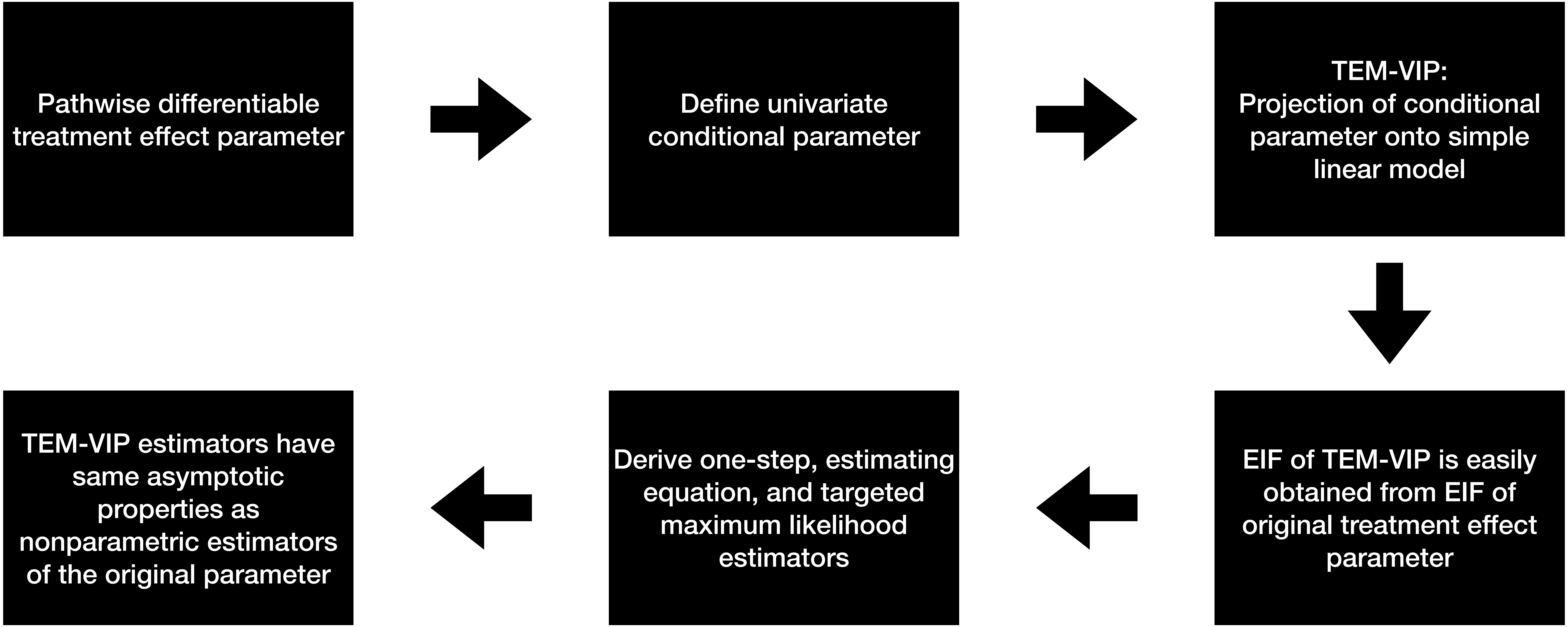
Treatment Effect Modifier Variable Importance Parameters



Treatment Effect Modifier Variable Importance Parameters



Treatment Effect Modifier Variable Importance Parameters



TEM-VIP Examples

Treatment Effect	Effect Type	Outcome Type
Average treatment effect	Absolute	Continuous
Average treatment effect	Absolute	Binary
Average treatment effect	Absolute	Count
Relative risk	Relative	Positive continuous
Relative risk	Relative	Count
Relative risk	Relative	Binary
Restricted mean survival time	Absolute	Right-censored time-to-event
Survival time ratio	Relative	Right-censored time-to-event

ATE-Based Absolute TEM-VIP in Action

ATE-Based Absolute TEM-VIP in Action

```
n ← 200
biomarker_1 ← rnorm(n, mean = 0, sd = 1)
biomarker_2 ← rnorm(n, mean = 0, sd = 1)
biomarker_3 ← rnorm(n, mean = 0, sd = 1)
biomarker_4 ← rnorm(n, mean = 0, sd = 1)
covariate ← rbinom(n, 1, 0.4)
treatment ← rbinom(n, 1, 0.5)
response ← covar + 1 * biomarker_1 * treatment
              + 2 * biomarker_2 * treatment
```

ATE-Based Absolute TEM-VIP in Action

```
n ← 200
biomarker_1 ← rnorm(n, mean = 0, sd = 1)
biomarker_2 ← rnorm(n, mean = 0, sd = 1)
biomarker_3 ← rnorm(n, mean = 0, sd = 1)
biomarker_4 ← rnorm(n, mean = 0, sd = 1)
covariate ← rbinom(n, 1, 0.4)
treatment ← rbinom(n, 1, 0.5)
response ← covar + 1 * biomarker_1 * treatment
              + 2 * biomarker_2 * treatment
```

	Est.	**SE**	**Z-score**	**P**	**P (BH)**
biomarker_2	1.90	0.0768	24.7	3.58E-13	1.43E-13
biomarker_1	0.820	0.129	6.38	1.75E-10	3.51E-10
biomarker_3	0.0482	0.145	0.333	7.39E-01	9.03E-01
biomarker_4	-0.0202	0.166	-0.122	9.03E-01	9.03E-01

ATE-Based Absolute TEM-VIP in Action

```
n ← 200
biomarker_1 ← rnorm(n, mean = 0, sd = 1)
biomarker_2 ← rnorm(n, mean = 0, sd = 1)
biomarker_3 ← rnorm(n, mean = 0, sd = 1)
biomarker_4 ← rnorm(n, mean = 0, sd = 1)
covariate ← rbinom(n, 1, 0.4)
treatment ← rbinom(n, 1, 0.5)
response ← covar + 1 * biomarker_1 * treatment
              + 2 * biomarker_2 * treatment
```

	Est.	SE	Z-score	P	P (BH)
biomarker_2	1.90	0.0768	24.7	3.58E-13	1.43E-13
biomarker_1	0.820	0.129	6.38	1.75E-10	3.51E-10
biomarker_3	0.0482	0.145	0.333	7.39E-01	9.03E-01
biomarker_4	-0.0202	0.166	-0.122	9.03E-01	9.03E-01

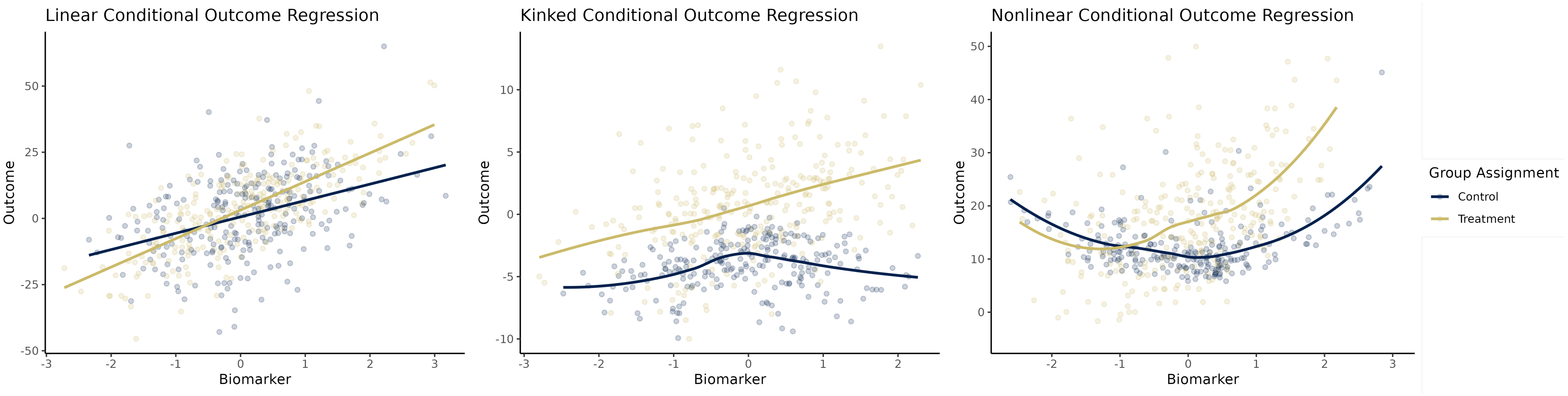
ATE-Based Absolute TEM-VIP in Action

```
n ← 200
biomarker_1 ← rnorm(n, mean = 0, sd = 1)
biomarker_2 ← rnorm(n, mean = 0, sd = 1)
biomarker_3 ← rnorm(n, mean = 0, sd = 1)
biomarker_4 ← rnorm(n, mean = 0, sd = 1)
covariate ← rbinom(n, 1, 0.4)
treatment ← rbinom(n, 1, 0.5)
response ← covar + 1 * biomarker_1 * treatment
            + 2 * biomarker_2 * treatment
```

	Est.	SE	Z-score	P	P (BH)
biomarker_2	1.90	0.0768	24.7	3.58E-13	1.43E-13
biomarker_1	0.820	0.129	6.38	1.75E-10	3.51E-10
biomarker_3	0.0482	0.145	0.333	7.39E-01	9.03E-01
biomarker_4	-0.0202	0.166	-0.122	9.03E-01	9.03E-01

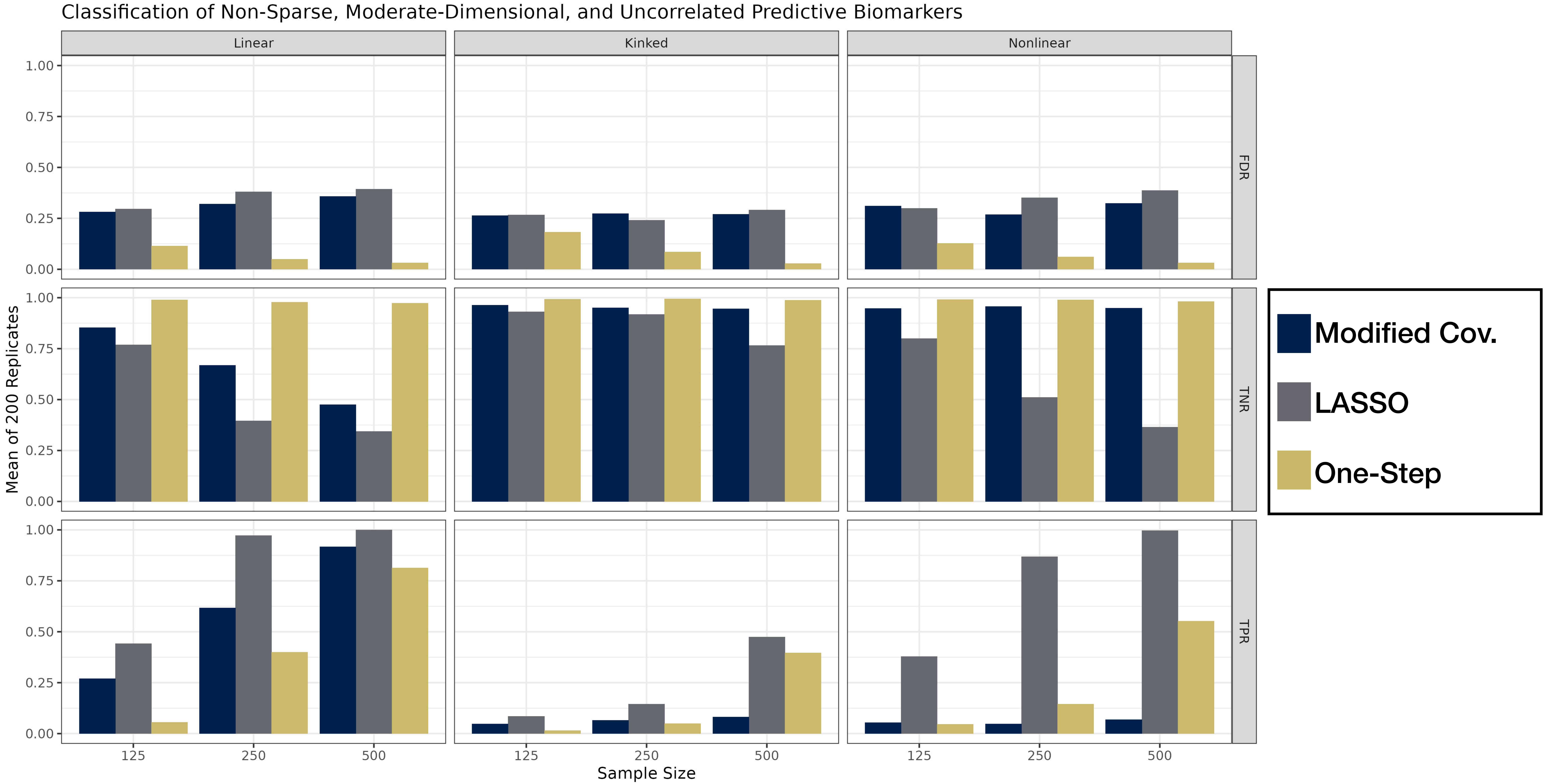
Simulation Study: ATE-Based Absolute TEM-VIP

Simulations: Assess Finite Sample Behaviour

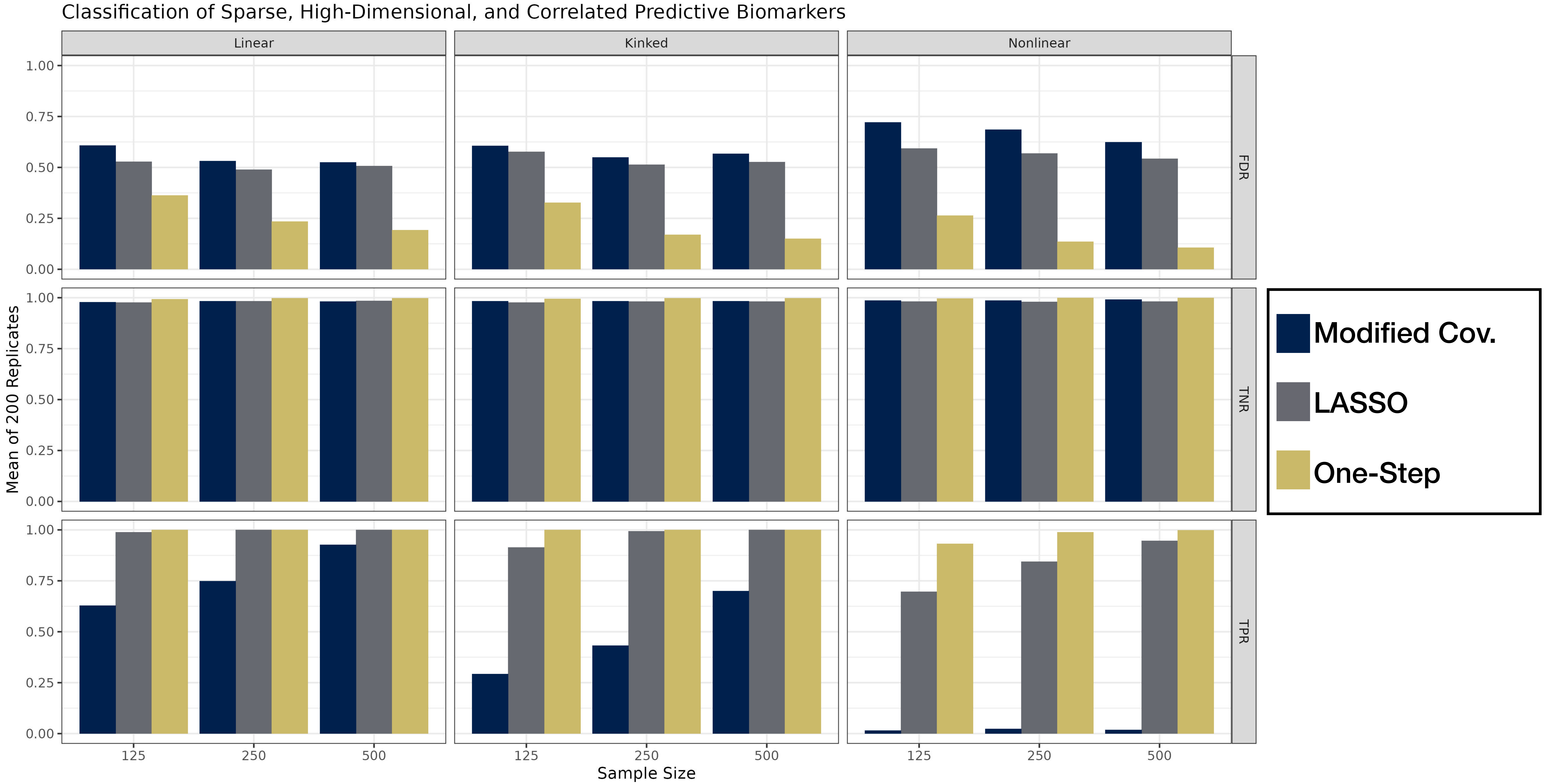


Difficulty

One-Step Controls False Positive Rates



One-Step *Still* Controls False Positive Rates



Application to IMmotion 150/151

Application to IMmotion 150

Results align with recent findings in similar medication

Application to IMmotion 150

Results align with recent findings in similar medication

1. Only patients with tumor RNA-seq data in the sunitinib (n=71) and atezolizumab + bevacizumab (n=77) arms were considered.

Application to IMmotion 150

Results align with recent findings in similar medication

1. Only patients with tumor RNA-seq data in the sunitinib (n=71) and atezolizumab + bevacizumab (n=77) arms were considered.

2. Selected the 500 most variable, log-transformed genes as biomarkers.

Application to IMmotion 150

Results align with recent findings in similar medication

1. Only patients with tumor RNA-seq data in the sunitinib (n=71) and atezolizumab + bevacizumab (n=77) arms were considered.

2. Selected the 500 most variable, log-transformed genes as biomarkers.

3. Objective response was used as the response variable.

Application to IMmotion 150

Results align with recent findings in similar medication

1. Only patients with tumor RNA-seq data in the sunitinib (n=71) and atezolizumab + bevacizumab (n=77) arms were considered.

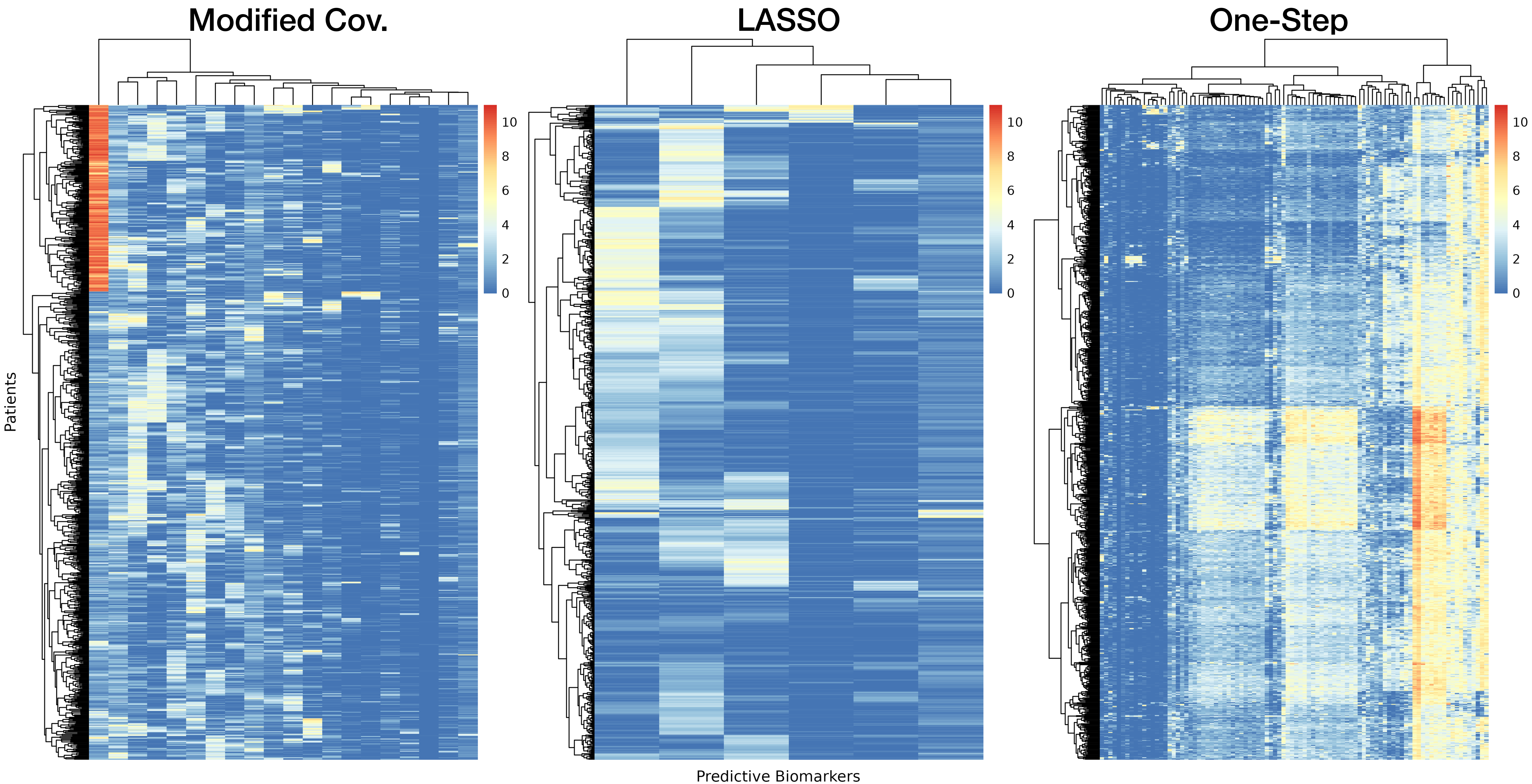
2. Selected the 500 most variable, log-transformed genes as biomarkers.

3. Objective response was used as the response variable.

92 genes were identified as predictive using a 5% FDR cutoff. They are associated with immune responses, including those mediated by B cells and lymphocytes.

Validation on IMmotion 151

One-step identifies meaningful predictive biomarkers



Papers, software and contact info at **pboileau.ca**