

# Supplementary material

## Contents

<b>1 Simulation settings</b>	<b>1</b>
1.1 Base-case scenario . . . . .	2
1.2 Deviations from base-case . . . . .	2
1.3 Risk modeling . . . . .	2
1.4 Approaches to individualize benefit predictions . . . . .	3
<b>2 Adaptive model selection frequencies</b>	<b>19</b>
<b>3 Discrimination and calibration for benefit</b>	<b>22</b>
<b>4 Strong relative treatment effect</b>	<b>26</b>
<b>5 Treatment interactions</b>	<b>29</b>
<b>6 Empirical illustration</b>	<b>34</b>

## 1 Simulation settings

For all patients we observe covariates  $x_1, \dots, x_8$ , of which 4 are continuous and 4 are binary. More specifically,

$$\begin{aligned} x_1, \dots, x_4 &\sim N(0, 1) \\ x_5, \dots, x_8 &\sim B(1, 0.2) \end{aligned}$$

We first, generate the binary outcomes  $y$  for the untreated patients ( $t_x = 0$ ), based on

$$P(y | \boldsymbol{x}, t_x = 0) = g(\beta_0 + \beta_1 x_1 + \dots + \beta_8 x_8) = g(lp_0), \quad (1)$$

where

$$g(x) = \frac{e^x}{1 + e^x}$$

For treated patients, outcomes are generated from:

$$P(y | \boldsymbol{x}, t_x = 1) = g(lp_1) \quad (2)$$

where

$$lp_1 = \gamma_2(lp_0 - c)^2 + \gamma_1(lp_0 - c) + \gamma_0$$

## 1.1 Base-case scenario

The base-case scenario assumes a constant odds ratio of 0.8 in favor of treatment. The simulated datasets are of size  $n = 4250$ , where treatment is allocated at random using a 50/50 split (80% power for the detection of an unadjusted OR of 0.8, assuming an event rate of 20% in the untreated arm). Outcome incidence in the untreated population is set at 20%. For the development of the prediction model we use the model defined in (1) including a constant treatment effect. When doing predictions,  $t_x$  is set to 0. The value of the true  $\beta$  is such that the above prediction model has an AUC of 0.75.

The previously defined targets are achieved when  $\beta = (-2.08, 0.49, \dots, 0.49)^t$ . For the derivations in the treatment arm we use  $\gamma = (\log(0.8), 1, 0)^t$ .

## 1.2 Deviations from base-case

We deviate from the base-case scenario in two ways. First, we alter the overall target settings of sample size, overall treatment effect and prediction model AUC. In a second stage, we consider settings that violate the assumption of a constant relative treatment effect, using a model-based approach.

For the first part, we consider:

- Sample size:
  - $n = 1064$
  - $n = 17000$
- Overall treatment effect:
  - $OR = 0.5$
  - $OR = 1$
- Prediction performance:
  - $AUC = 0.65$
  - $AUC = 0.85$

We set the true risk model coefficients to be  $\beta = (-1.63, 0.26, \dots, 0.26)^t$  for  $AUC = 0.65$  and  $\beta = (-2.7, 0.82, \dots, 0.82)^t$  for  $AUC = 0.85$ . In both cases,  $\beta_0$  is selected so that an event rate of 20% is maintained in the control arm.

For the second part linear, quadratic and non-monotonic deviations from the assumption of constant relative effect are considered. We also consider different intensity levels of these deviations. Finally, constant absolute treatment-related harms are introduced, i.e. positive ( $0.25 \times$  true average benefit), strong positive ( $0.50 \times$  true average benefit) and negative ( $-0.25 \times$  true average benefit; i.e. constant absolute treatment-related benefit). In case of true absent treatment effects, treatment-related harms are set to 1%, 2% and -1% for positive, strong positive and negative setting, respectively. The settings for these deviations are defined in Table S1.

## 1.3 Risk modeling

Merging treatment arms, we develop prediction models including a constant relative treatment effect:

$$E\{y | x, t_x\} = P(y | x, t_x) = g(\beta_0 + \beta_1 x_1 + \dots + \beta_8 x_8 + \gamma t_x) \quad (3)$$

Individualized predictions are derived setting  $t_x = 0$ .

## 1.4 Approaches to individualize benefit predictions

### 1.4.1 Risk stratification

Derive a prediction model using the same approach as above and divide the population in equally sized risk-based subgroups. Estimate subgroup-specific absolute benefit from the observed absolute differences. Subject-specific benefit predictions are made by attributing to individuals their corresponding subgroup-specific estimate.

### 1.4.2 Constant treatment effect

Assuming a constant relative treatment effect, fit the adjusted model in (1.3). Then, an estimate of absolute benefit can be derived from

$$\hat{f}_{\text{benefit}}(lp \mid \mathbf{x}, \hat{\beta}) = g(lp) - g(lp + \hat{\gamma})$$

### 1.4.3 Linear interaction

The assumption of constant relative treatment effect is relaxed modeling a linear interaction of treatment with the risk linear predictor:

$$E\{y \mid \mathbf{x}, t_x, \hat{\beta}\} = g(lp + (\gamma_0 + \gamma_1 lp)t_x)$$

We predict absolute benefit from

$$\hat{f}_{\text{benefit}}(lp \mid \mathbf{x}, \hat{\beta}) = g(lp) - g(\gamma_0 + (1 + \gamma_1)lp)$$

### 1.4.4 Restricted cubic splines

Finally, we drop the linearity assumption and predict absolute benefit using smoothing with restricted cubic splines with 3, 4, and 5 knots. More specifically, we fit the model:

$$P(y = 1 \mid lp, t_x) = g(\beta_0 + \beta_{t_x} t_x + f_{RCS}(lp) + f_{RCS}(lp) \times t_x)$$

where

$$f_{RCS}(x) = \alpha_0 + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots + \alpha_{k-1} h_{k-1}(x)$$

with

$$h_{j+1}(x) = (x - t_j)^3 - (x - t_{k-1})_+ \frac{t_k - t_j}{t_k - t_{k-1}} + (x - t_k)_+^3 \frac{t_{k-1} - t_j}{t_k - t_{k-1}}$$

and  $t_1, \dots, t_k$  are the selected knots. We predict absolute benefit from

$$\hat{f}_{\text{benefit}}(lp \mid \mathbf{x}, \hat{\beta}) = P(y = 1 \mid lp, t_x = 0) - P(y = 1 \mid lp, t_x = 1)$$

Table S1: Scenario settings of the entire simulation study.

Analysis ID					Baseline risk								True treatment effect				Benefit		
Scenario	Base	N	AUC	Treatment-related harm	b0	b1	b2	b3	b4	b5	b6	b7	b8	g0	g1	g2	c	Before harms	After harms
1	absent	4,250	0.75	moderate-positive	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
2	absent	4,250	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.010
3	absent	4,250	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.020
4	absent	4,250	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.010
5	absent	4,250	0.65	moderate-positive	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.000
6	absent	4,250	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.010
7	absent	4,250	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.020
8	absent	4,250	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.010
9	absent	4,250	0.85	moderate-positive	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.000
10	absent	4,250	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.010
11	absent	4,250	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.020
12	absent	4,250	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.010
13	absent	1,063	0.75	moderate-positive	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
14	absent	1,063	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.010
15	absent	1,063	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.020
16	absent	1,063	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.010
17	absent	1,063	0.65	moderate-positive	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.000
18	absent	1,063	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.010
19	absent	1,063	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.020
20	absent	1,063	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.010
21	absent	1,063	0.85	moderate-positive	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.000
22	absent	1,063	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.010
23	absent	1,063	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.020
24	absent	1,063	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.010
25	absent	17,000	0.75	moderate-positive	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
26	absent	17,000	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.010
27	absent	17,000	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	-0.020
28	absent	17,000	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.010
29	absent	17,000	0.65	moderate-positive	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.000
30	absent	17,000	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.010
31	absent	17,000	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	-0.020
32	absent	17,000	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.000	1.000	0.000	0	0.000	0.010
33	absent	17,000	0.85	moderate-positive	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.000
34	absent	17,000	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.010
35	absent	17,000	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	-0.020
36	absent	17,000	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.000	1.000	0.000	0	0.000	0.010
37	absent	4,250	0.75	moderate-positive	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.000
38	absent	4,250	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.010
39	absent	4,250	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.020
40	absent	4,250	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.010
41	absent	4,250	0.65	moderate-positive	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.000
					absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.000

C7	42	absent	4,250	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.010
	43	absent	4,250	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.020
C8	44	absent	4,250	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.010
	45	absent	4,250	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.000
	46	absent	4,250	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.010
	47	absent	4,250	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.020
	48	absent	4,250	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.010
	49	absent	1,063	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.000
	50	absent	1,063	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.010
	51	absent	1,063	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.020
	52	absent	1,063	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.010
	53	absent	1,063	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.000
C9	54	absent	1,063	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.010
	55	absent	1,063	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.020
	56	absent	1,063	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.010
	57	absent	1,063	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.000
	58	absent	1,063	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.010
	59	absent	1,063	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.020
	60	absent	1,063	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.010
	61	absent	17,000	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.000
	62	absent	17,000	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.010
	63	absent	17,000	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	-0.020
	64	absent	17,000	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.010
C10	65	absent	17,000	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.000
	66	absent	17,000	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.010
	67	absent	17,000	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	-0.020
	68	absent	17,000	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.080	0.934	0.000	0	0.000	0.010
	69	absent	17,000	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.000
	70	absent	17,000	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.010
	71	absent	17,000	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	-0.020
	72	absent	17,000	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.070	0.930	0.000	0	0.000	0.010
C11	73	absent	4,250	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	0.000
	74	absent	4,250	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	-0.010
	75	absent	4,250	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	-0.020
	76	absent	4,250	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	0.010
	77	absent	4,250	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.290	0.776	0.000	0	0.000	0.000
	78	absent	4,250	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.290	0.776	0.000	0	0.000	-0.010
C12	79	absent	4,250	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.290	0.776	0.000	0	0.000	-0.020
	80	absent	4,250	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.290	0.776	0.000	0	0.000	0.010
	81	absent	4,250	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.220	0.785	0.000	0	0.000	0.000
	82	absent	4,250	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.220	0.785	0.000	0	0.000	-0.010
	83	absent	4,250	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.220	0.785	0.000	0	0.000	-0.020
	84	absent	4,250	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.220	0.785	0.000	0	0.000	0.010
C13	85	absent	1,063	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	0.000
	86	absent	1,063	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.250	0.796	0.000	0	0.000	-0.010















402	moderate	4,250	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.025
403	moderate	4,250	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.016
404	moderate	4,250	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.041
405	moderate	4,250	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.024
406	moderate	4,250	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.018
407	moderate	4,250	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.012
408	moderate	4,250	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.030
409	moderate	1,063	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.029
410	moderate	1,063	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.022
411	moderate	1,063	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.015
412	moderate	1,063	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.036
413	moderate	1,063	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.033
414	moderate	1,063	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.025
415	moderate	1,063	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.016
416	moderate	1,063	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.041
417	moderate	1,063	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.024
418	moderate	1,063	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.018
419	moderate	1,063	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.012
420	moderate	1,063	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.030
421	moderate	17,000	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.029
422	moderate	17,000	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.022
423	moderate	17,000	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.015
424	moderate	17,000	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.173	1.560	0.105	0	0.029	0.036
425	moderate	17,000	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.033
426	moderate	17,000	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.025
427	moderate	17,000	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.016
428	moderate	17,000	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.481	1.783	0.137	0	0.033	0.041
429	moderate	17,000	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.024
430	moderate	17,000	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.018
431	moderate	17,000	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.012
432	moderate	17,000	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.085	1.354	0.074	0	0.024	0.030
433	high	4,250	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.079
434	high	4,250	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.059
435	high	4,250	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.040
436	high	4,250	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.099
437	high	4,250	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.693	1.000	0.000	0	0.089	0.089
438	high	4,250	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.693	1.000	0.000	0	0.089	0.067
439	high	4,250	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.693	1.000	0.000	0	0.089	0.044
440	high	4,250	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.693	1.000	0.000	0	0.089	0.111
441	high	4,250	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.693	1.000	0.000	0	0.069	0.069
442	high	4,250	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.693	1.000	0.000	0	0.069	0.052
443	high	4,250	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.693	1.000	0.000	0	0.069	0.034
444	high	4,250	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.693	1.000	0.000	0	0.069	0.086
445	high	1,063	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.079
446	high	1,063	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.693	1.000	0.000	0	0.079	0.059







582	high	4,250	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.067
583	high	4,250	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.044
584	high	4,250	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.111
585	high	4,250	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.069
586	high	4,250	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.052
587	high	4,250	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.034
588	high	4,250	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.086
589	high	1,063	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.079
590	high	1,063	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.059
591	high	1,063	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.040
592	high	1,063	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.099
593	high	1,063	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.089
594	high	1,063	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.067
595	high	1,063	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.044
596	high	1,063	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.111
597	high	1,063	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.069
598	high	1,063	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.052
599	high	1,063	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.034
600	high	1,063	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.086
601	high	17,000	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.079
602	high	17,000	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.059
603	high	17,000	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.040
604	high	17,000	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-4.860	1.000	-0.052	-5	0.079	0.099
605	high	17,000	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.089
606	high	17,000	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.067
607	high	17,000	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.044
608	high	17,000	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-4.840	1.000	-0.059	-5	0.089	0.111
609	high	17,000	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.069
610	high	17,000	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.052
611	high	17,000	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.034
612	high	17,000	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-4.510	1.000	-0.059	-5	0.069	0.086
613	high	4,250	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.079
614	high	4,250	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.059
615	high	4,250	0.75	strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.040
616	high	4,250	0.75	negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.099
617	high	4,250	0.65	absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.089
618	high	4,250	0.65	moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.067
619	high	4,250	0.65	strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.044
620	high	4,250	0.65	negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.111
621	high	4,250	0.85	absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.069
622	high	4,250	0.85	moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.052
623	high	4,250	0.85	strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.034
624	high	4,250	0.85	negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.086
625	high	1,063	0.75	absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.079
626	high	1,063	0.75	moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.059

627	high	1,063	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.040
628	high	1,063	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.099
629	high	1,063	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.089
630	high	1,063	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.067
631	high	1,063	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.044
632	high	1,063	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.111
633	high	1,063	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.069
634	high	1,063	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.052
635	high	1,063	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.034
636	high	1,063	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.086
637	high	17,000	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.079
638	high	17,000	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.059
639	high	17,000	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.040
640	high	17,000	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.084	2.035	0.210	0	0.079	0.099
641	high	17,000	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.089
642	high	17,000	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.067
643	high	17,000	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.044
644	high	17,000	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.786	2.762	0.321	0	0.089	0.111
645	high	17,000	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.069
646	high	17,000	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.052
647	high	17,000	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.034
648	high	17,000	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.621	1.566	0.138	0	0.069	0.086

## 2 Adaptive model selection frequencies

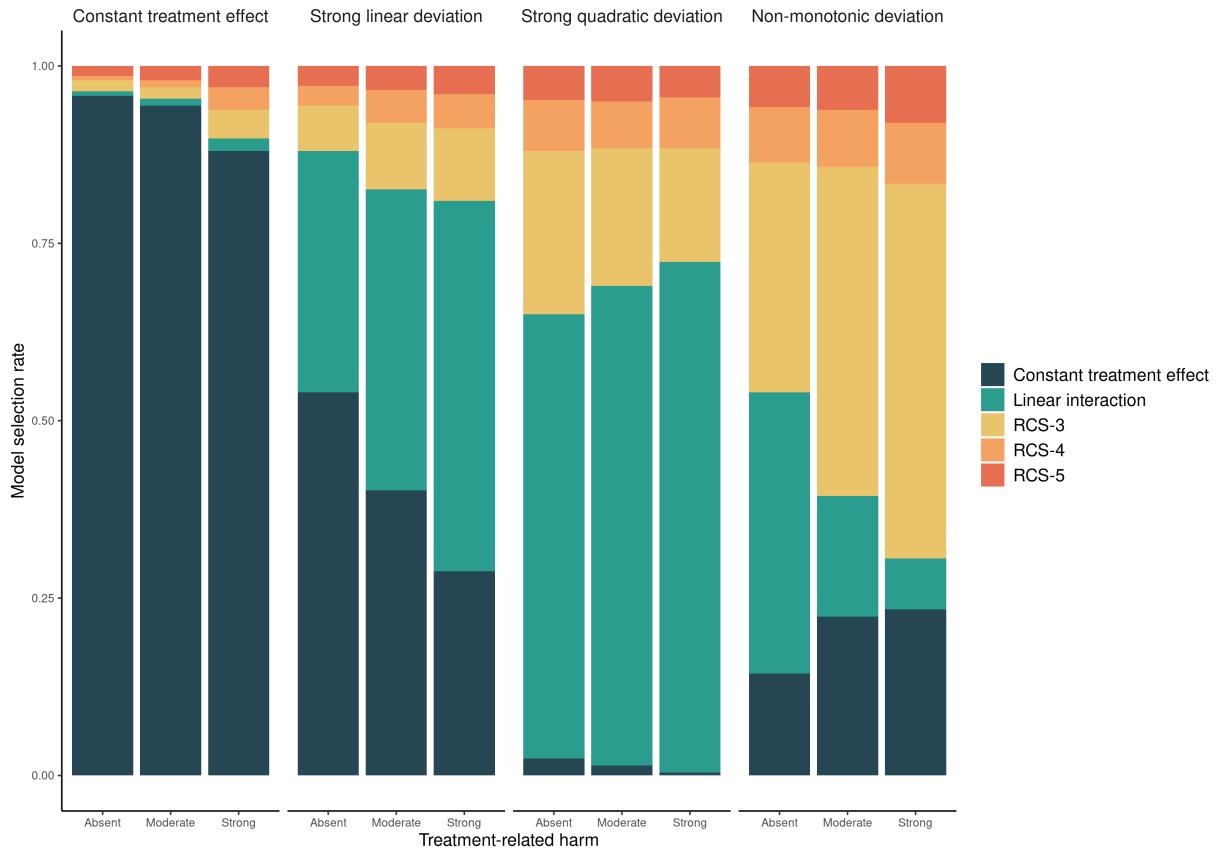


Figure S1: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. The scenario with the true constant relative treatment effect (first panel) had a true prediction AUC of 0.75 and sample size of 4,250.

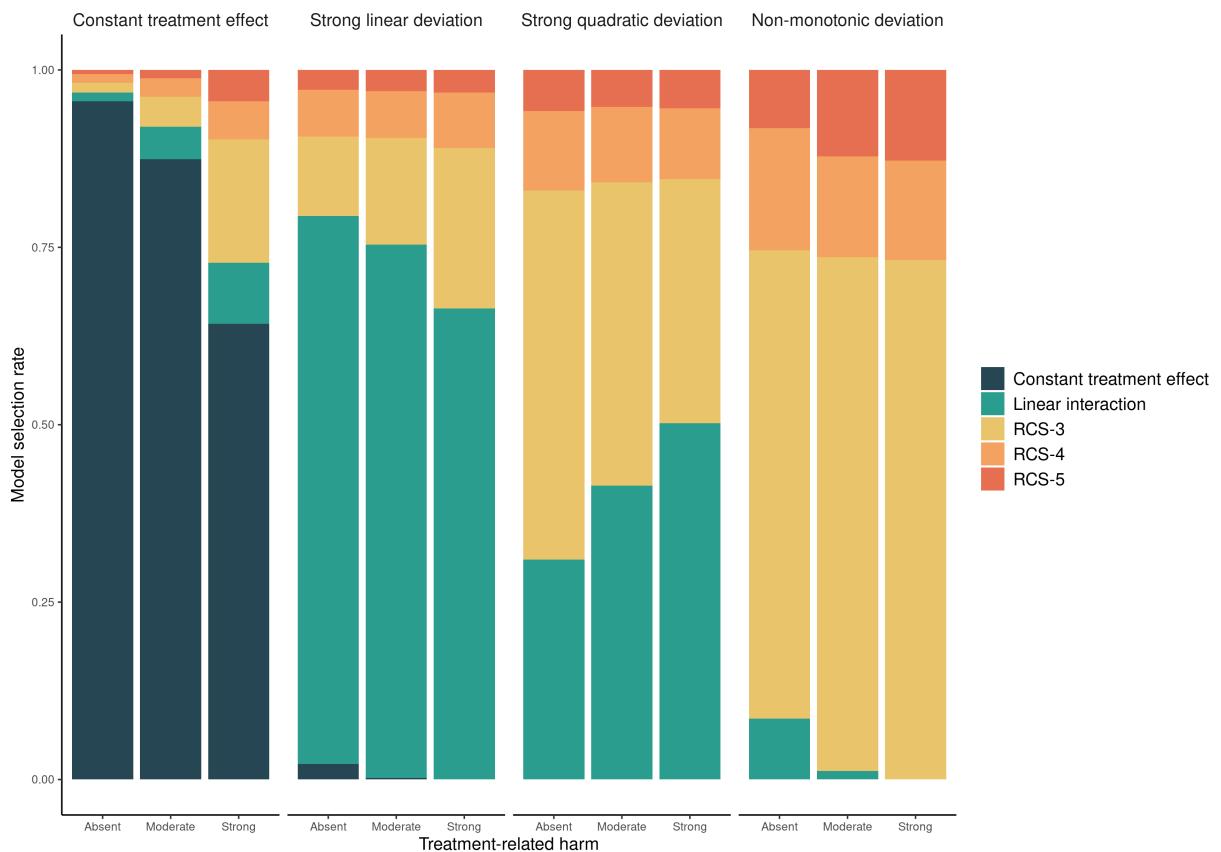


Figure S2: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. Sample size is 17,000 rather than 4,250 in Figure S1

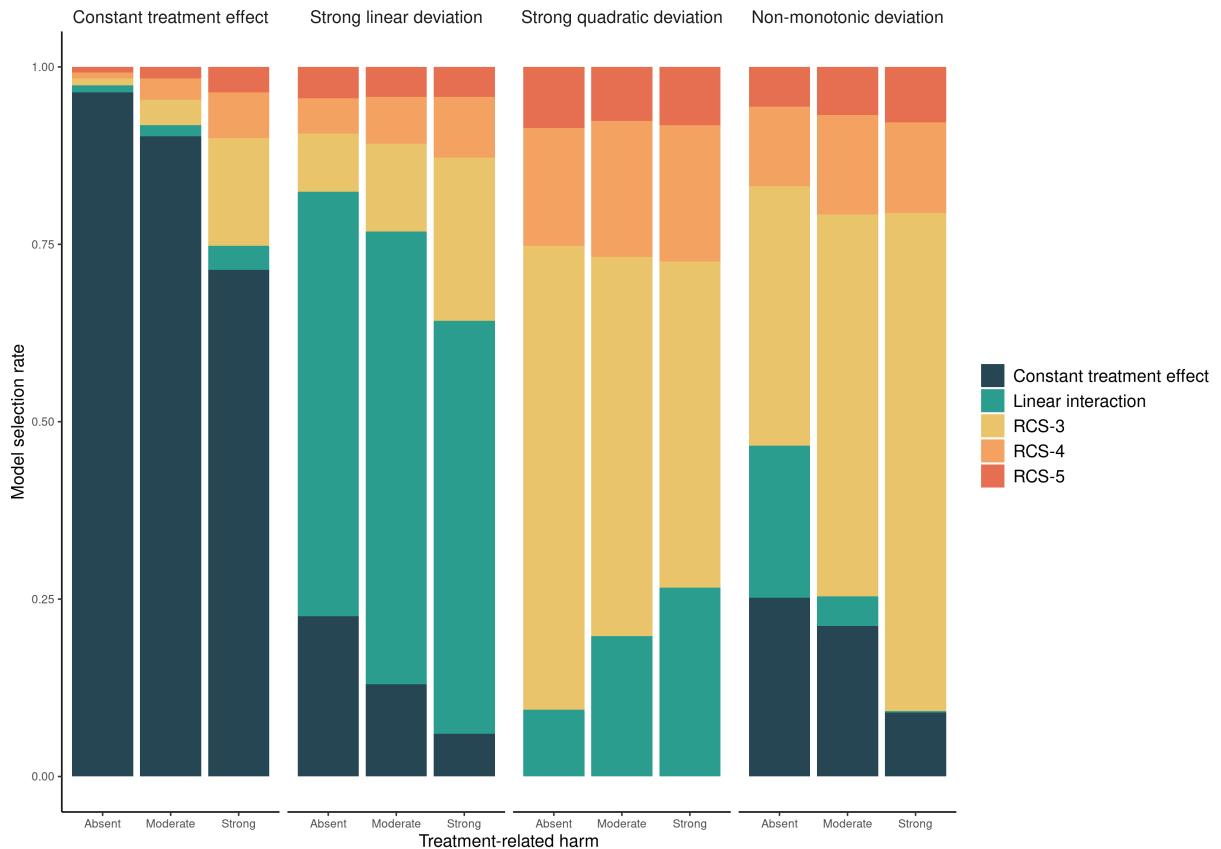


Figure S3: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. AUC is 0.85 rather than 0.75 in Figure S1

### 3 Discrimination and calibration for benefit

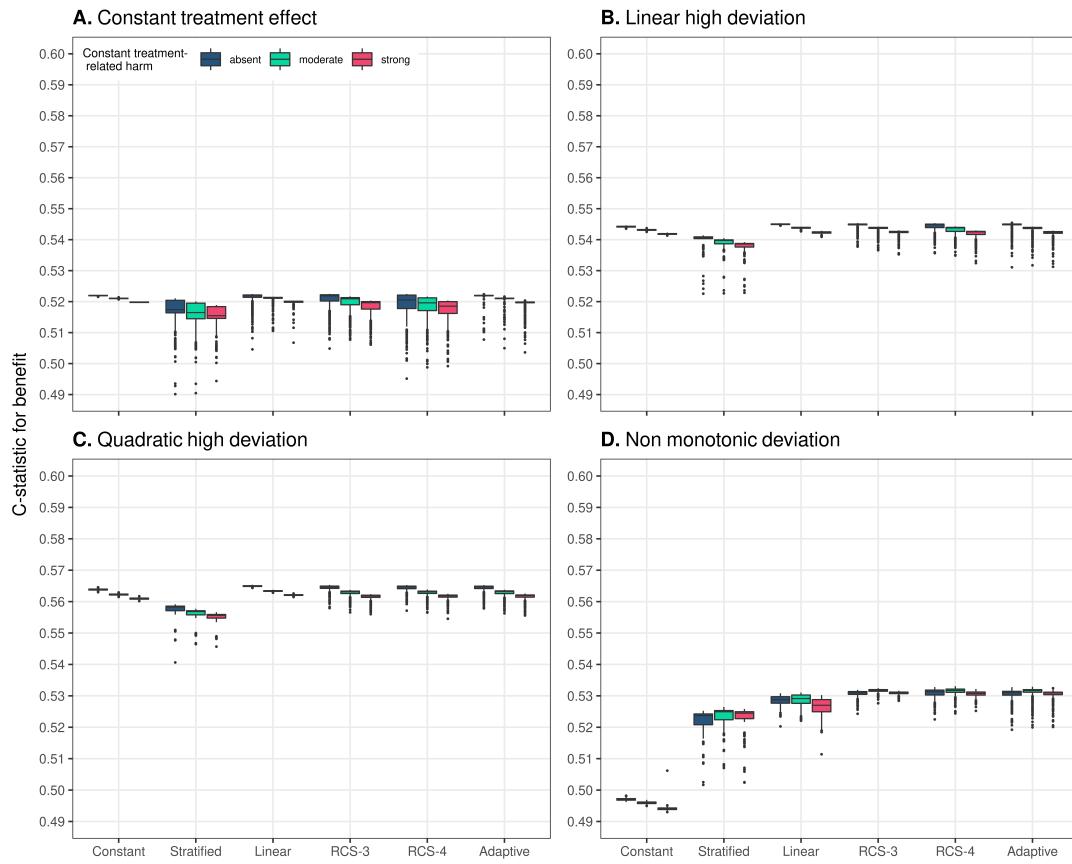


Figure S4: Discrimination for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.75 and sample size of 17,000

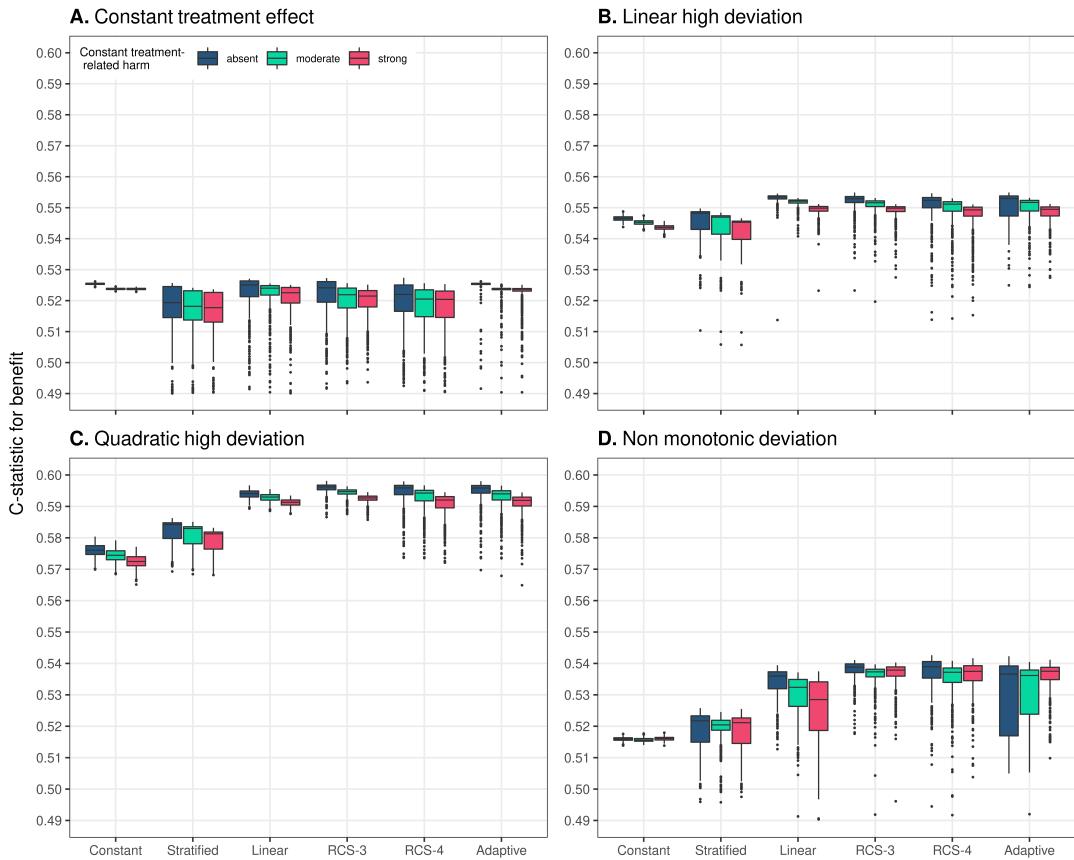


Figure S5: Discrimination for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.85 and sample size of 4,250

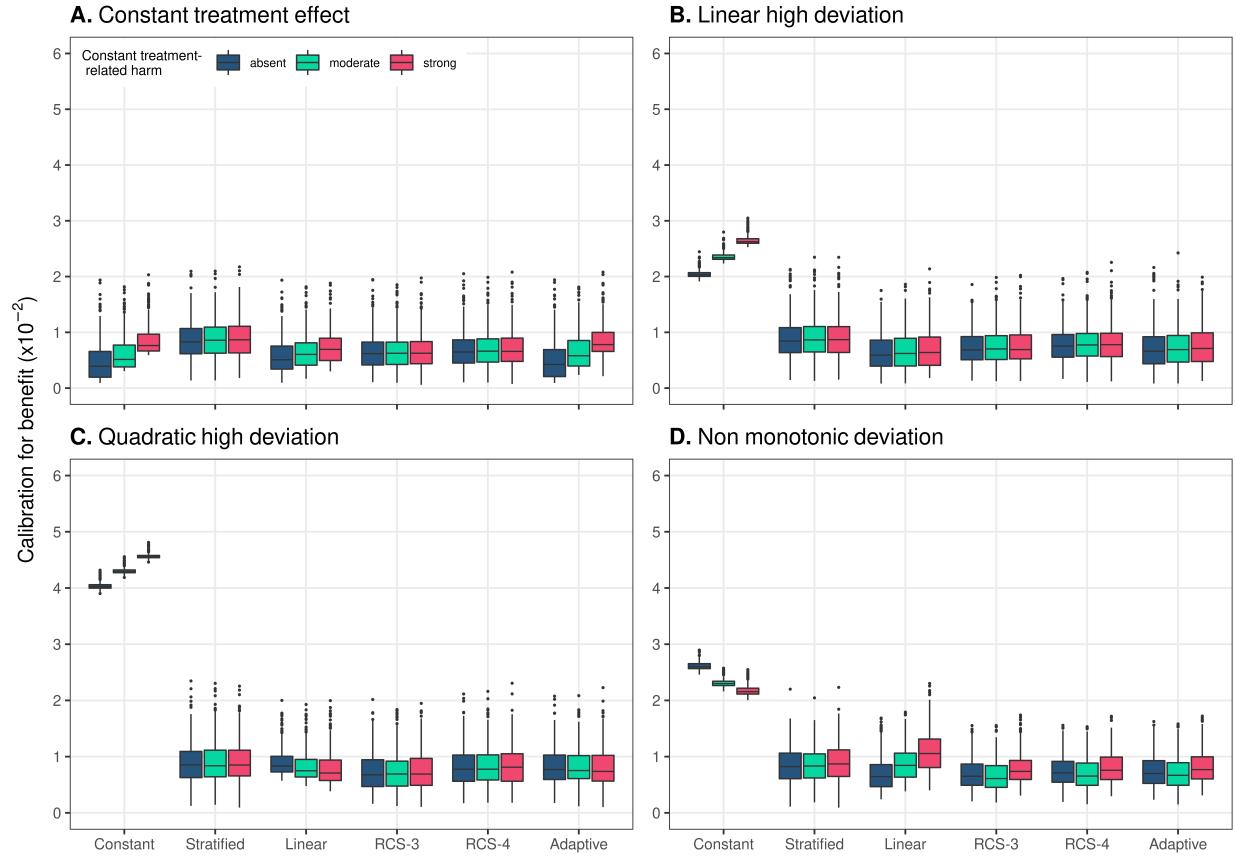


Figure S6: Calibration for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.75 and sample size of 17,000

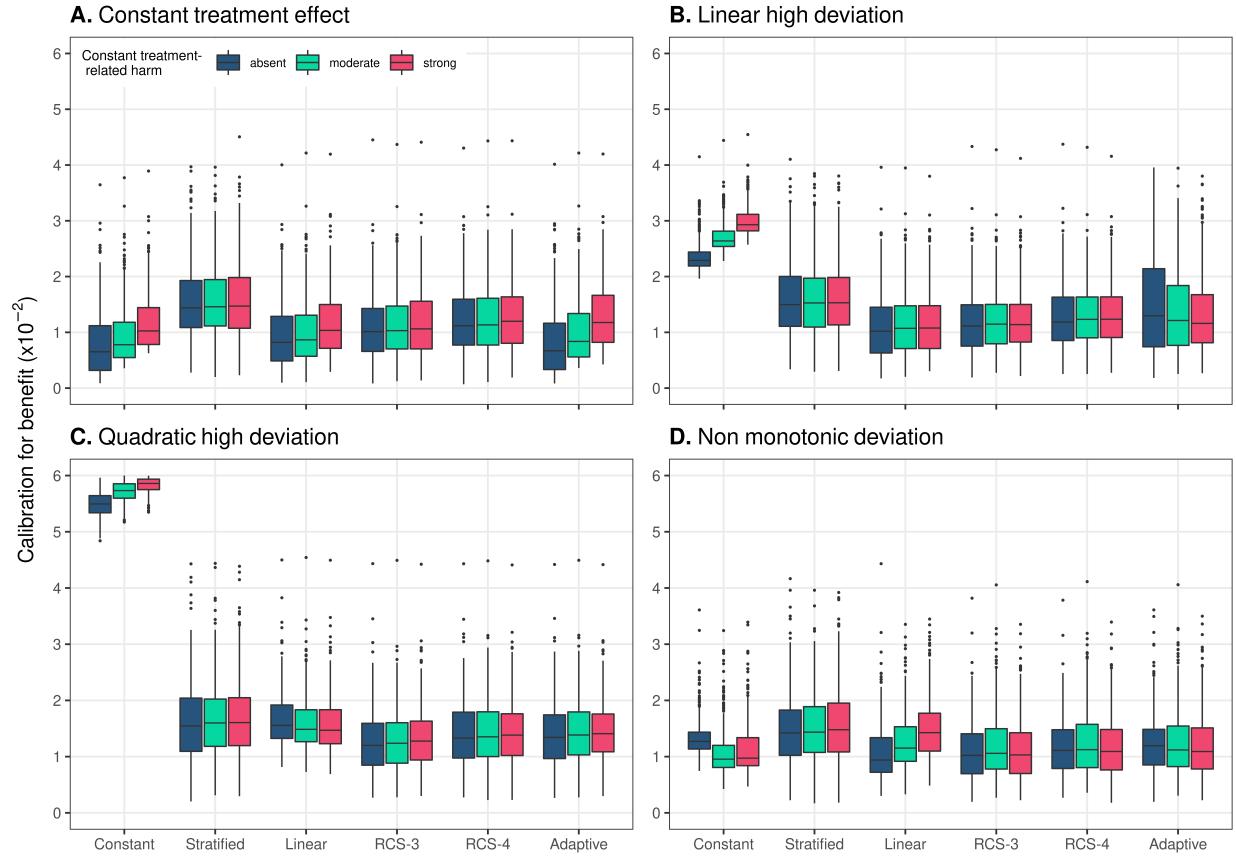


Figure S7: Calibration for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.85 and sample size of 4,250

## 4 Strong relative treatment effect

Here we present the root mean squared error of the considered methods using strong constant relative treatment effect ( $OR = 0.5$ ) as the reference. Again, the same sample size and prediction performance settings were considered along with the same settings for linear, quadratic and non-monotonic deviations from the base case scenario of constant relative treatment effects are considered. All results can be found at [https://arekkas.shinyapps.io/simulation\\_viewer/](https://arekkas.shinyapps.io/simulation_viewer/).

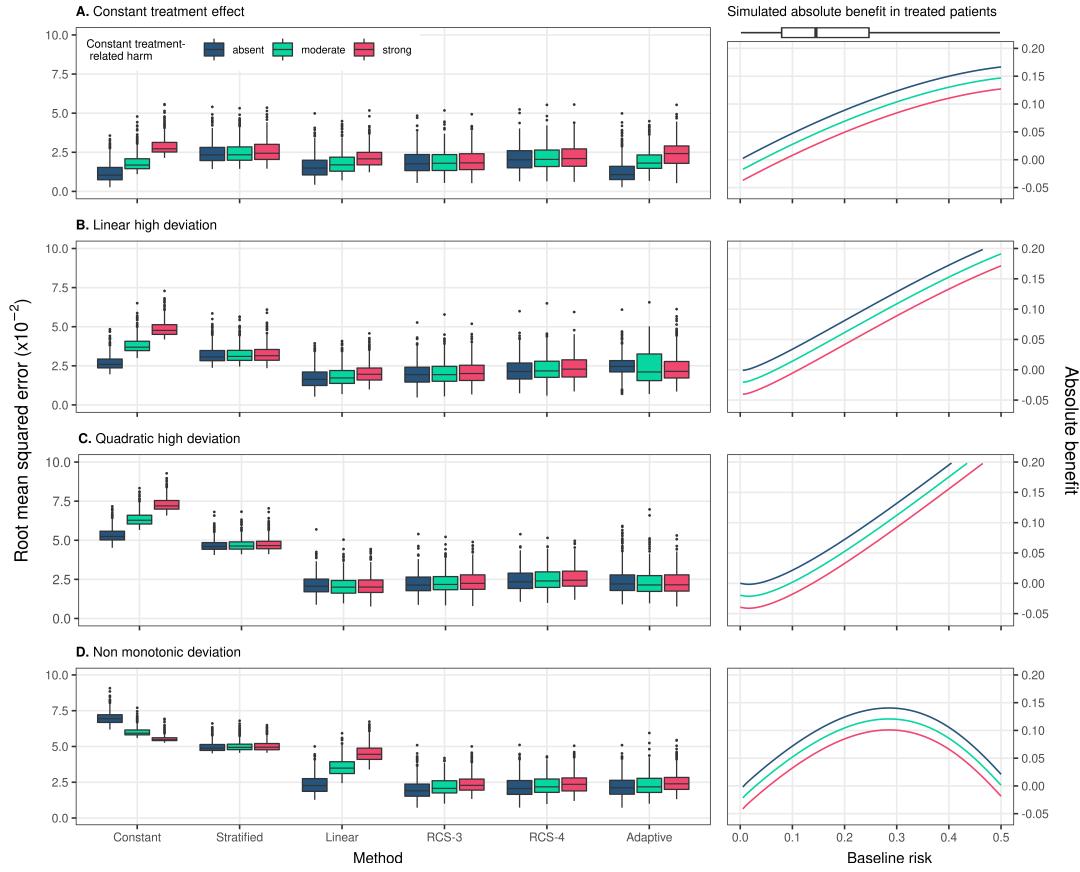


Figure S8: RMSE of the considered methods across 500 replications calculated in a simulated super-population of size 500,000. The scenario with true constant relative treatment effect (panel A) had a true prediction AUC of 0.75 and sample size of 4,250. The RMSE is also presented for strong linear (panel B), strong quadratic (panel C), and non-monotonic (panel D) deviations from constant relative treatment effects. Panels on the right side present the true relationship between baseline risk (x-axis) and absolute treatment benefit (y-axis). The 2.5, 25, 75 and 97.5 percentiles of the risk distribution are expressed in the boxplot.

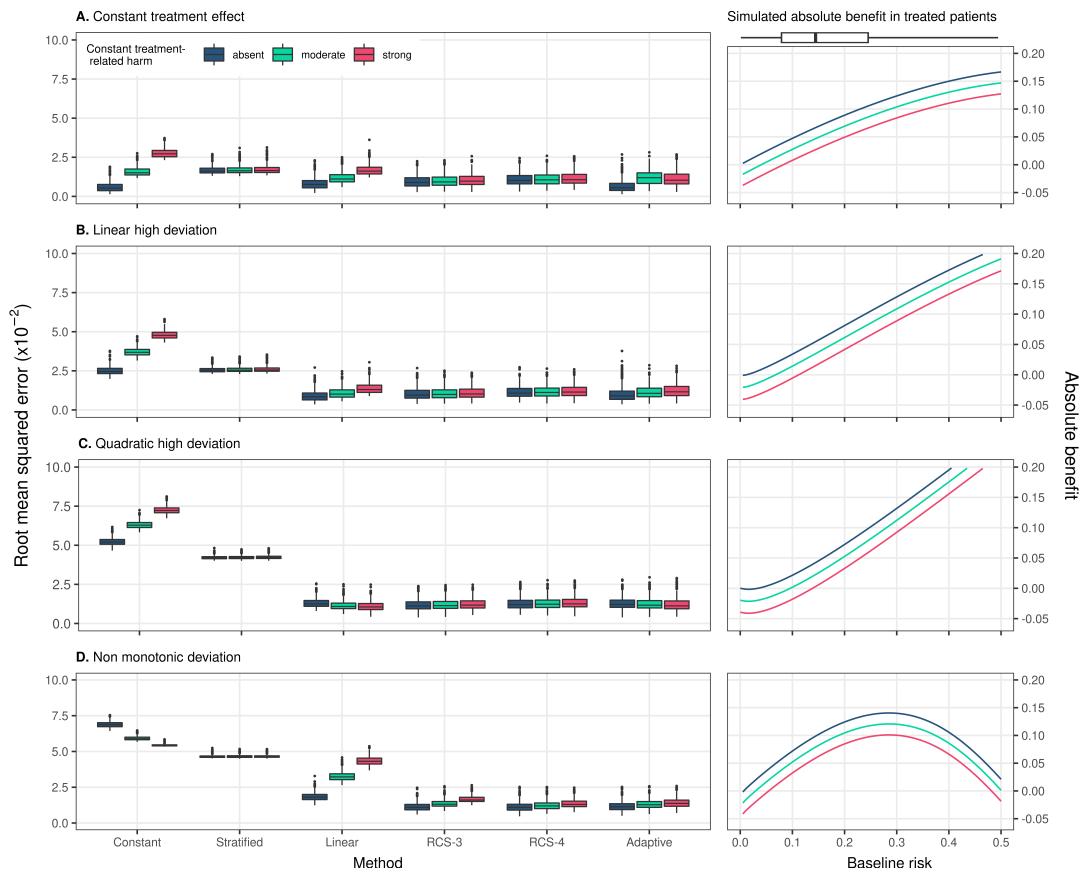


Figure S9: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000. Sample size is 17,000 rather than 4,250 in Figure S8.

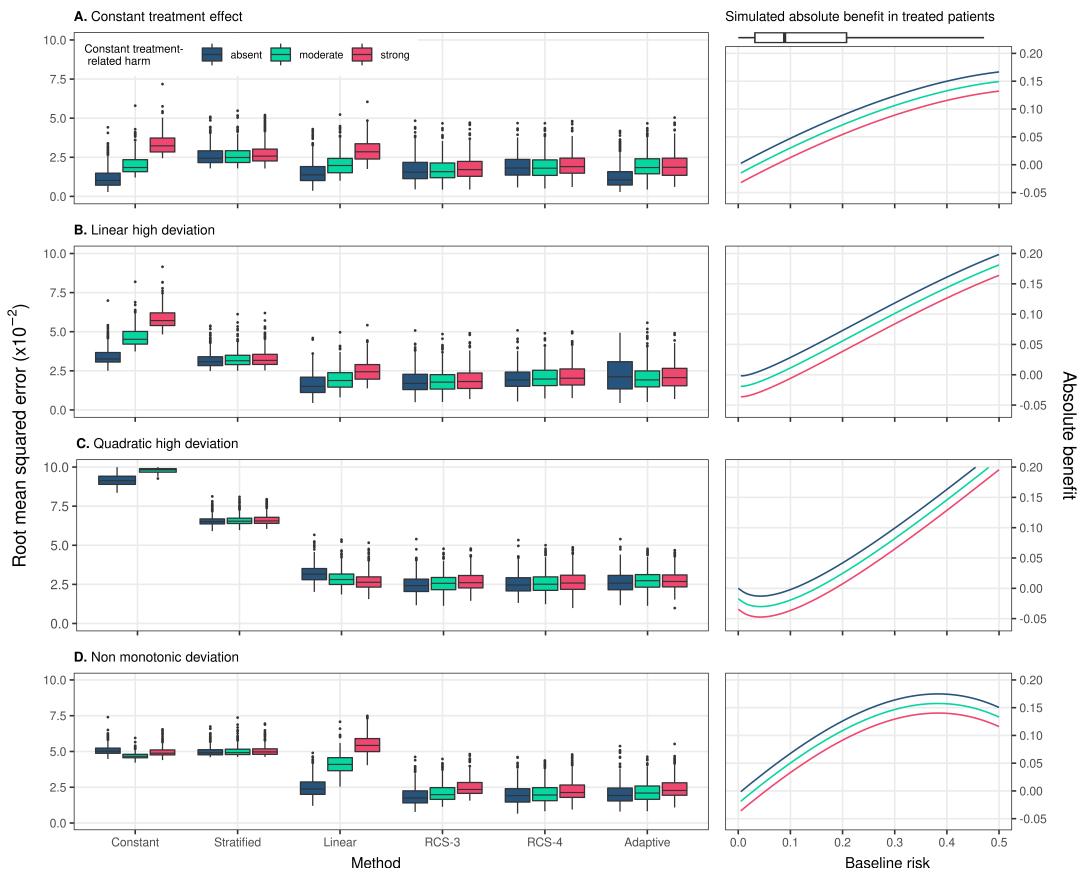


Figure S10: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000. AUC is 0.85 rather than in Figure S8.

## 5 Treatment interactions

We carried out a smaller set of simulations, in which we assumed true treatment-covariate interactions. Sample size was set to 4,250 and the AUC of the true prediction model was set to 0.75. The following scenarios were considered: 1) 4 true weak positive interactions ( $OR_{t_x=1}/OR_{t_x=0} = 0.83$ ); 2) 4 strong positive interactions ( $OR_{t_x=1}/OR_{t_x=0} = 0.61$ ); 3) 2 weak and 2 strong positive interactions; 4) 4 weak negative interactions ( $OR_{t_x=1}/OR_{t_x=0} = 1.17$ ); 5) 4 strong negative interactions ( $OR_{t_x=1}/OR_{t_x=0} = 1.39$ ); 6) 2 weak and 2 strong negative interactions; 7) combined positive and negative strong interactions. We also considered constant treatment-related harms applied on the absolute scale to all treated patients. The exact settings were: 1) absent treatment-related harms; 2) moderate treatment-related harms, defined as 25% of the average true benefit of the scenario without treatment-related harms; 3) strong treatment-related harms defined as 50% of the true average benefit of the scenario without treatment-related harms; 4) negative treatment-related harms (benefit), defined as an absolute risk reduction for treated patients of 50% of the true average benefit of the scenario without treatment-related harms. The exact settings can be found in Table S2.

Table S2: Scenario settings of the entire simulation study.

Analysis ID					Baseline risk								True treatment effect					Benefit		
Scenario	Base	N	AUC	Treatment-related harm	b0	b1	b2	b3	b4	b5	b6	b7	b8	g0	g1	g2	g5	g6	Before harms	After harms
649	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.19	-0.19	-0.19	0.10	0.10
650	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.19	-0.19	-0.19	0.10	0.07
651	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.19	-0.19	-0.19	0.10	0.05
652	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.19	-0.19	-0.19	0.10	0.12
653	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.49	-0.19	-0.49	0.10	0.10
654	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.49	-0.19	-0.49	0.10	0.08
655	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.49	-0.19	-0.49	0.10	0.05
656	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.19	-0.49	-0.19	-0.49	0.10	0.13
657	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	-0.49	-0.49	-0.49	0.11	0.11
658	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	-0.49	-0.49	-0.49	0.11	0.08
659	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	-0.49	-0.49	-0.49	0.11	0.06
660	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	-0.49	-0.49	-0.49	0.11	0.14
661	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.16	0.16	0.16	0.06	0.06
662	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.16	0.16	0.16	0.06	0.05
663	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.16	0.16	0.16	0.06	0.03
664	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.16	0.16	0.16	0.06	0.08
665	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.33	0.16	0.33	0.05	0.05
666	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.33	0.16	0.33	0.05	0.04
667	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.33	0.16	0.33	0.05	0.03
668	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.16	0.33	0.16	0.33	0.05	0.06
669	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.33	0.33	0.33	0.33	0.04	0.04
670	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.33	0.33	0.33	0.33	0.04	0.03
671	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.33	0.33	0.33	0.33	0.04	0.02
672	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	0.33	0.33	0.33	0.33	0.04	0.05
673	interaction	4,250	0.75	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	0.33	-0.49	0.33	0.08	0.08
674	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	0.33	-0.49	0.33	0.08	0.06
675	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	0.33	-0.49	0.33	0.08	0.04
676	interaction	4,250	0.75		-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.69	-0.49	0.33	-0.49	0.33	0.08	0.10

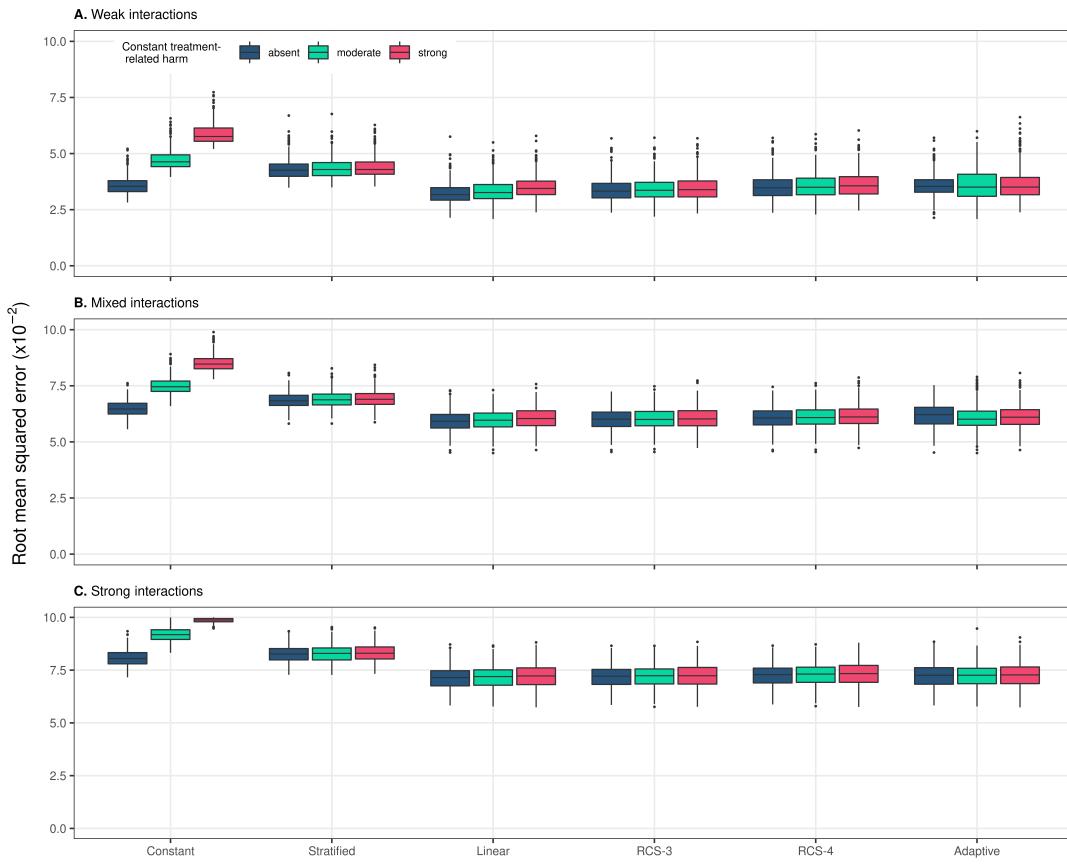


Figure S11: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions all favoring treatment were considered.

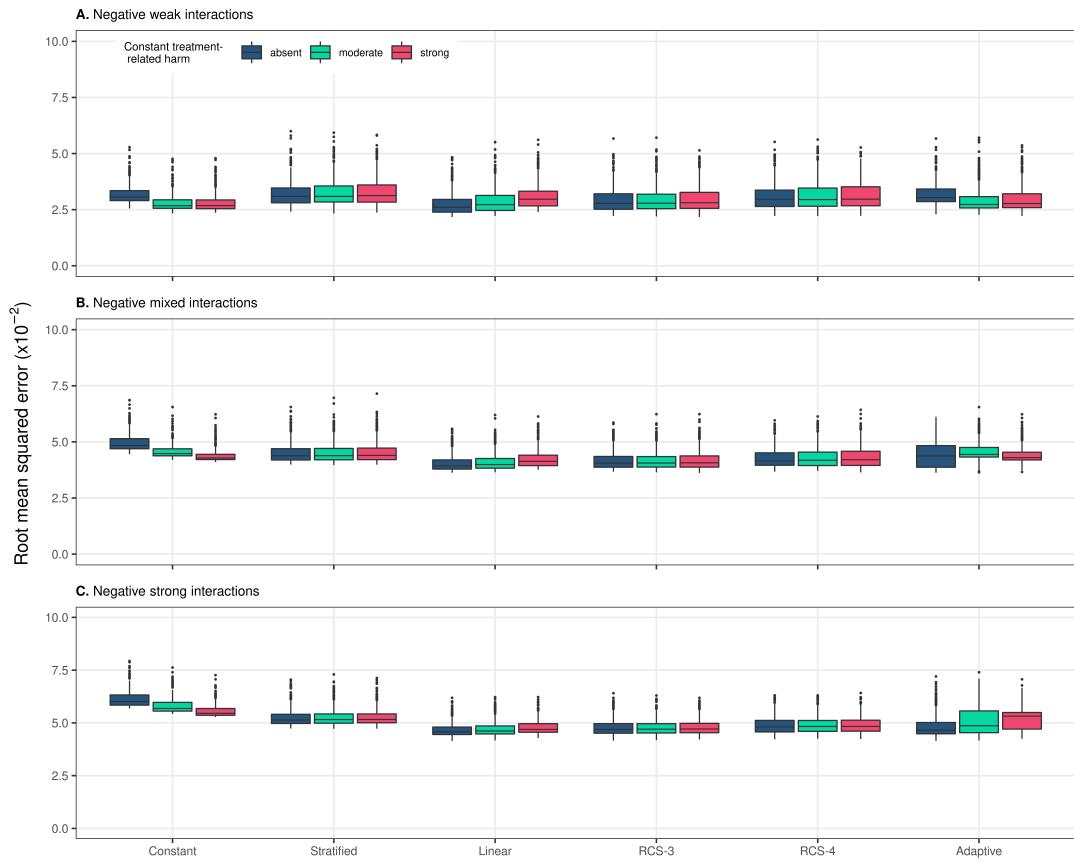


Figure S12: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions all favoring the control were considered.

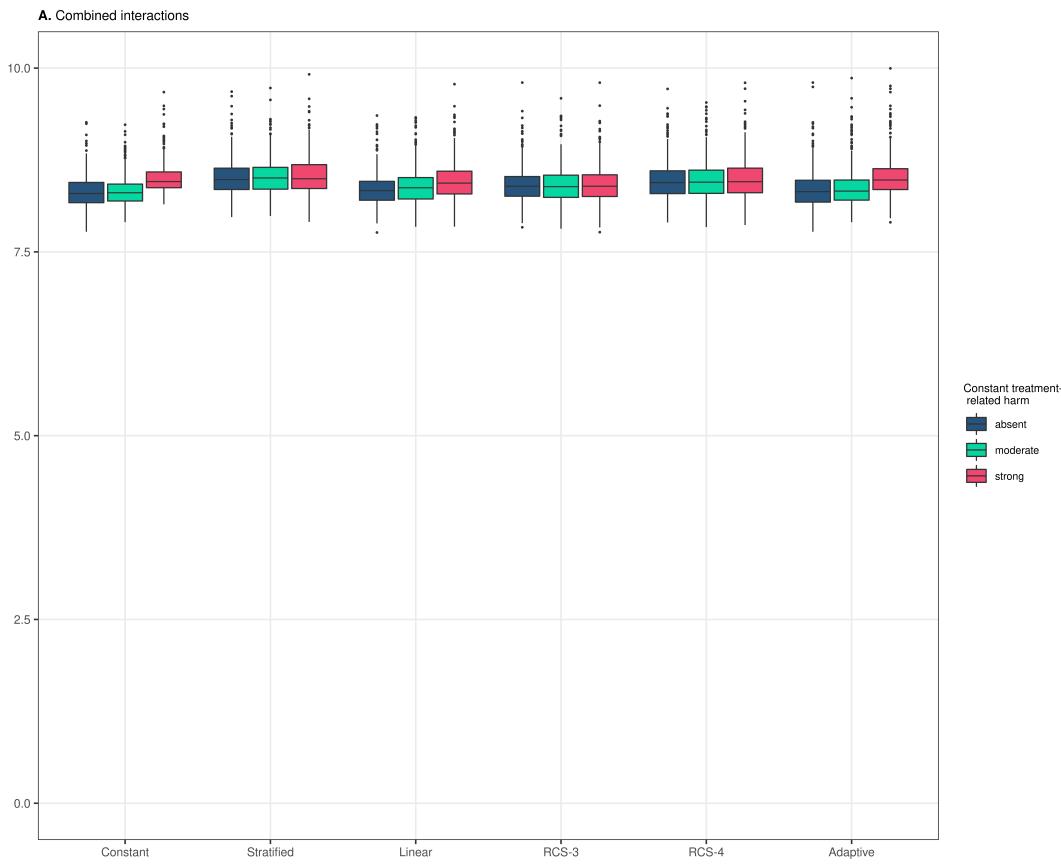


Figure S13: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions 2 favoring treatment and 2 favoring the control were considered.

## 6 Empirical illustration

$$P(\text{outcome} = 1 | X = x) = \text{expit}(lp(x)), \quad (4)$$

where

$$\begin{aligned} lp(x) = & \beta_0 + \beta_1 \text{age} + \beta_2 I(\text{Killip} = II) + \beta_3 I(\text{Killip} = III) + \\ & \beta_4 I(\text{Killip} = IV) + \beta_5 \text{ceil}(\text{sysbp}, 120) + \beta_6 \text{pulse} + \\ & \beta_7 \max(\text{pulse} - 50, 0) + \beta_8 I(\text{pmi} = yes) + \\ & \beta_9 I(\text{miloc} = \text{Anterior}) + \beta_{10} I(\text{miloc} = \text{Other}) + \\ & \gamma \times \text{treatment} \end{aligned} \quad (5)$$

and  $\text{expit}(x) = \frac{e^x}{1+e^x}$

Table S3: Coefficients of the prediction model for 30-day mortality, based on the data from GUSTO-I trial.

Variable	Estimate	stderror	zvalue	pvalue
Intercept	-3.020	0.797	-3.788	0.000
Age	-0.208	0.053	-3.935	0.000
Killip class = II	0.077	0.002	31.280	0.000
Killip class = III	0.614	0.059	10.423	0.000
Killip class = IV	1.161	0.121	9.566	0.000
Systolic blood pressure	1.921	0.162	11.872	0.000
Pulse rate (1)	-0.039	0.002	-20.332	0.000
Pulse rate (2)	-0.024	0.016	-1.521	0.128
Previous MI (yes)	0.043	0.016	2.675	0.007
MI location (Other)	0.447	0.056	7.964	0.000
MI location (Anterior)	0.286	0.135	2.126	0.033
Treatment	0.543	0.051	10.625	0.000