

Supplementary material

Contents

1	Simulation settings	2
1.1	Base-case scenario	2
1.2	Deviations from base-case	2
2	Plausible scenario settings	19
3	Approaches to individualize benefit predictions	20
3.1	Risk modeling	20
3.2	Risk stratification	20
3.3	Constant treatment effect	20
3.4	Linear interaction	20
3.5	Restricted cubic splines	20
4	Adaptive model selection frequencies	22
5	Discrimination and calibration for benefit	25
6	Strong relative treatment effect	29
7	Treatment interactions	32
8	Empirical illustration	37
9	References	38

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,

$$x_1, \dots, x_4 \sim N(0, 1)$$

$$x_5, \dots, x_8 \sim B(1, 0.2)$$

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

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

where

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

For treated patients, outcomes are generated from:

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

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 β 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, β_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.

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	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
	absent	4,250	0.75		-2.08	0.49	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
	absent	4,250	0.75		-2.08	0.49	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
	absent	4,250	0.75		-2.08	0.49	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	absent moderate-positive strong-positive negative	-1.63	0.26	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
	absent	4,250	0.65		-1.63	0.26	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
	absent	4,250	0.65		-1.63	0.26	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
	absent	4,250	0.65		-1.63	0.26	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	absent moderate-positive strong-positive negative	-2.70	0.82	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
	absent	4,250	0.85		-2.70	0.82	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
	absent	4,250	0.85		-2.70	0.82	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
	absent	4,250	0.85		-2.70	0.82	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	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
	absent	1,063	0.75		-2.08	0.49	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
	absent	1,063	0.75		-2.08	0.49	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
	absent	1,063	0.75		-2.08	0.49	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	absent moderate-positive strong-positive negative	-1.63	0.26	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
	absent	1,063	0.65		-1.63	0.26	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
	absent	1,063	0.65		-1.63	0.26	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
	absent	1,063	0.65		-1.63	0.26	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	absent moderate-positive strong-positive negative	-2.70	0.82	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
	absent	1,063	0.85		-2.70	0.82	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
	absent	1,063	0.85		-2.70	0.82	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
	absent	1,063	0.85		-2.70	0.82	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	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.000	1.000	0.000	0	0.000	0.000
	absent	17,000	0.75		-2.08	0.49	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
	absent	17,000	0.75		-2.08	0.49	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
	absent	17,000	0.75		-2.08	0.49	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	absent moderate-positive strong-positive negative	-1.63	0.26	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
	absent	17,000	0.65		-1.63	0.26	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
	absent	17,000	0.65		-1.63	0.26	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
	absent	17,000	0.65		-1.63	0.26	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	absent moderate-positive strong-positive negative	-2.70	0.82	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
	absent	17,000	0.85		-2.70	0.82	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
	absent	17,000	0.85		-2.70	0.82	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
	absent	17,000	0.85		-2.70	0.82	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	absent moderate-positive strong-positive negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.060	0.947	0.000	0	0.000	0.000
	absent	4,250	0.75		-2.08	0.49	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
	absent	4,250	0.75		-2.08	0.49	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
	absent	4,250	0.75		-2.08	0.49	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	absent	-1.63	0.26	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

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

492	high	1,063	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.730	0.930	0.000	0	0.069	0.086
493	high	17,000	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.750	0.947	0.000	0	0.079	0.079
494	high	17,000	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.750	0.947	0.000	0	0.079	0.059
495	high	17,000	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.750	0.947	0.000	0	0.079	0.040
496	high	17,000	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.750	0.947	0.000	0	0.079	0.099
497	high	17,000	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.770	0.934	0.000	0	0.089	0.089
498	high	17,000	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.770	0.934	0.000	0	0.089	0.067
499	high	17,000	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.770	0.934	0.000	0	0.089	0.044
500	high	17,000	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.770	0.934	0.000	0	0.089	0.111
501	high	17,000	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.730	0.930	0.000	0	0.069	0.069
502	high	17,000	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.730	0.930	0.000	0	0.069	0.052
503	high	17,000	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.730	0.930	0.000	0	0.069	0.034
504	high	17,000	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.730	0.930	0.000	0	0.069	0.086
505	high	4,250	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.079
506	high	4,250	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.059
507	high	4,250	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.040
508	high	4,250	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.099
509	high	4,250	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.089
510	high	4,250	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.067
511	high	4,250	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.044
512	high	4,250	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.111
513	high	4,250	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.069
514	high	4,250	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.052
515	high	4,250	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.034
516	high	4,250	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.086
517	high	1,063	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.079
518	high	1,063	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.059
519	high	1,063	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.040
520	high	1,063	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.099
521	high	1,063	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.089
522	high	1,063	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.067
523	high	1,063	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.044
524	high	1,063	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.111
525	high	1,063	0.85		absent	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.069
526	high	1,063	0.85		moderate-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.052
527	high	1,063	0.85		strong-positive	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.034
528	high	1,063	0.85		negative	-2.70	0.82	0.82	0.82	0.82	0.82	0.82	0.82	-0.840	0.785	0.000	0	0.069	0.086
529	high	17,000	0.75		absent	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.079
530	high	17,000	0.75		moderate-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.059
531	high	17,000	0.75		strong-positive	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.040
532	high	17,000	0.75		negative	-2.08	0.49	0.49	0.49	0.49	0.49	0.49	0.49	-0.900	0.796	0.000	0	0.079	0.099
533	high	17,000	0.65		absent	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.089
534	high	17,000	0.65		moderate-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.067
535	high	17,000	0.65		strong-positive	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.044
536	high	17,000	0.65		negative	-1.63	0.26	0.26	0.26	0.26	0.26	0.26	0.26	-0.960	0.776	0.000	0	0.089	0.111

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 Plausible scenario settings

In this section we present specific scenarios from our simulation settings in which evolution of benefit followed similar patterns to [1]. In this case patients were stratified into risk quarters based on their true baseline risk. Within each risk quarter we constructed boxplots of true benefit.

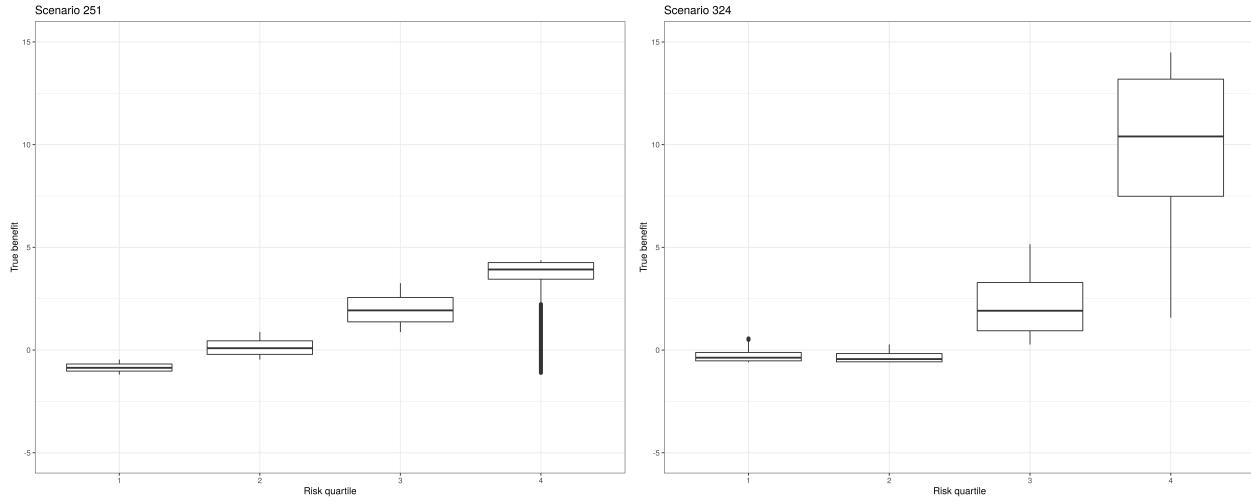


Figure S1: Simulation scenarios that closely follow trials. In this case, we see increasing absolute benefits with increasing baseline risk.

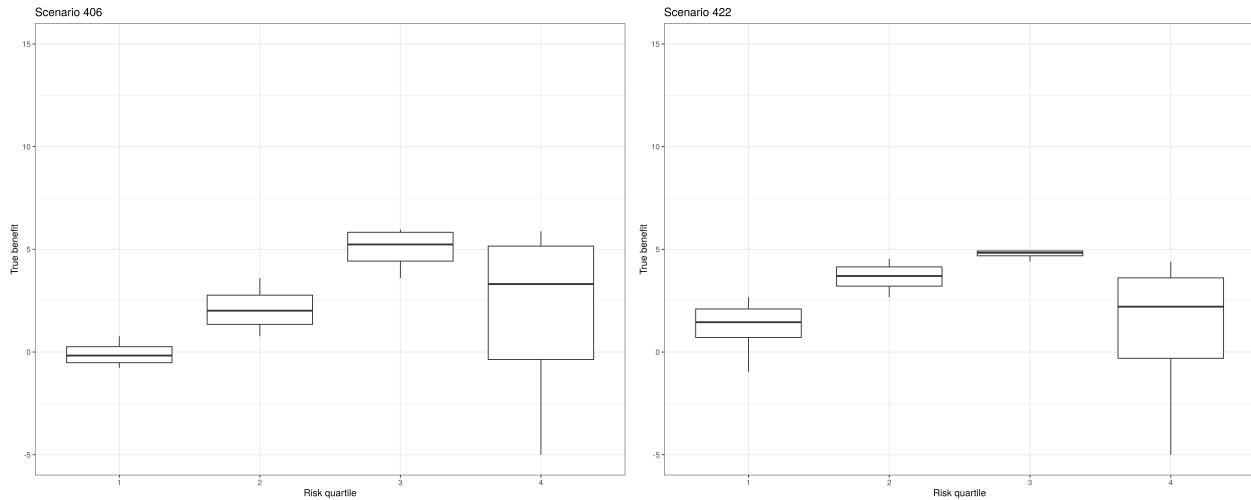


Figure S2: Simulation scenarios that closely follow trials. In this case, we see increasing absolute benefits with increasing baseline risk up to the third risk quarter. In the fourth risk quarter this trend is interrupted and benefits are diminished.

3 Approaches to individualize benefit predictions

3.1 Risk modeling

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

$$P(y = 1 | \mathbf{x}, t_x) = g(\mathbf{x}^t \boldsymbol{\beta} + \gamma t_x) \quad (3)$$

The risk linear predictor is then $lp(\mathbf{x}; \boldsymbol{\beta}) = \mathbf{x}^t \boldsymbol{\beta}$. We derive baseline risk predictions for patients by setting $t_x = 0$ in (3.1). All methods for individualizing benefit predictions are 2-stage methods, that start by fitting a model for predicting baseline risk. The estimated linear predictor of this model is

$$\hat{lp} = lp(\mathbf{x}; \hat{\boldsymbol{\beta}}) = \mathbf{x}^t \hat{\boldsymbol{\beta}}$$

3.2 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.

3.3 Constant treatment effect

Assuming a constant relative treatment effect, fit the adjusted model in (3.1). Then, an estimate of absolute benefit can be derived from

$$\tau(lp; \hat{\boldsymbol{\beta}}, \hat{\gamma}) = g(\hat{lp}) - g(\hat{lp} + \hat{\gamma})$$

3.4 Linear interaction

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

$$P(y = 1 | lp, t_x; \hat{\boldsymbol{\beta}}) = g(\delta_0 + \delta_1 \hat{lp} + \gamma_0 t_x + \gamma_1 \hat{lp} t_x)$$

Then the predicted absolute benefit is

$$\tau(lp; \hat{\boldsymbol{\beta}}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = g(\delta_0 + \delta_1 \hat{lp}) - g((\delta_0 + \gamma_0) + (\delta_1 + \gamma_1) \hat{lp})$$

3.5 Restricted cubic splines

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

$$P(y = 1 | lp, t_x; \hat{\boldsymbol{\beta}}, \boldsymbol{\delta}, f_{RCS}) = g(\delta_0 + \gamma t_x + f_{RCS}(\hat{lp}) \times t_x)$$

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 $h_1(x) = x$ and for $j = 2, \dots, k - 2$

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

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

$$\tau(lp; \hat{\beta}, \delta_0, f_{RCS}) = P(y = 1 | \hat{lp}, t_x = 0) - P(y = 1 | \hat{lp}, t_x = 1)$$

4 Adaptive model selection frequencies

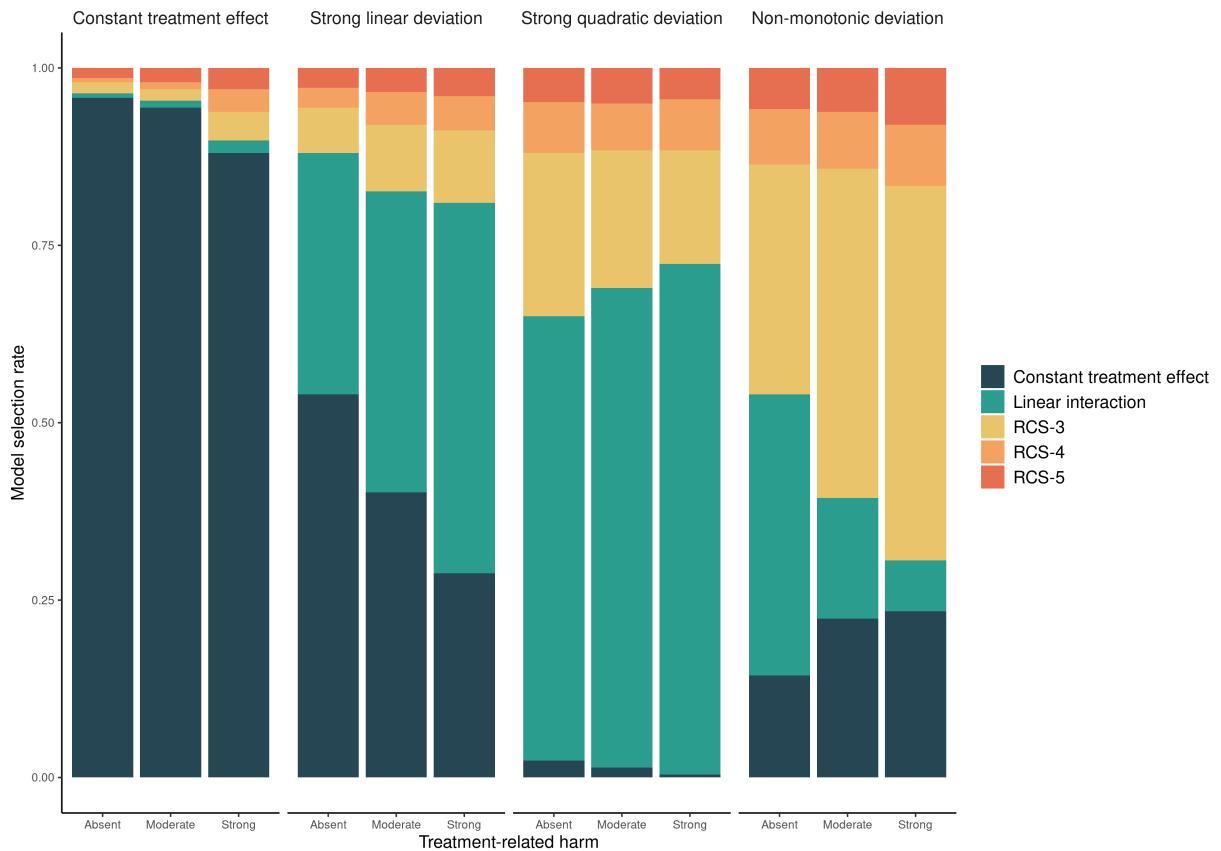


Figure S3: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. The scenario with the true constant relative treatment effect (first panel) had a true prediction AUC of 0.75 and sample size of 4,250.

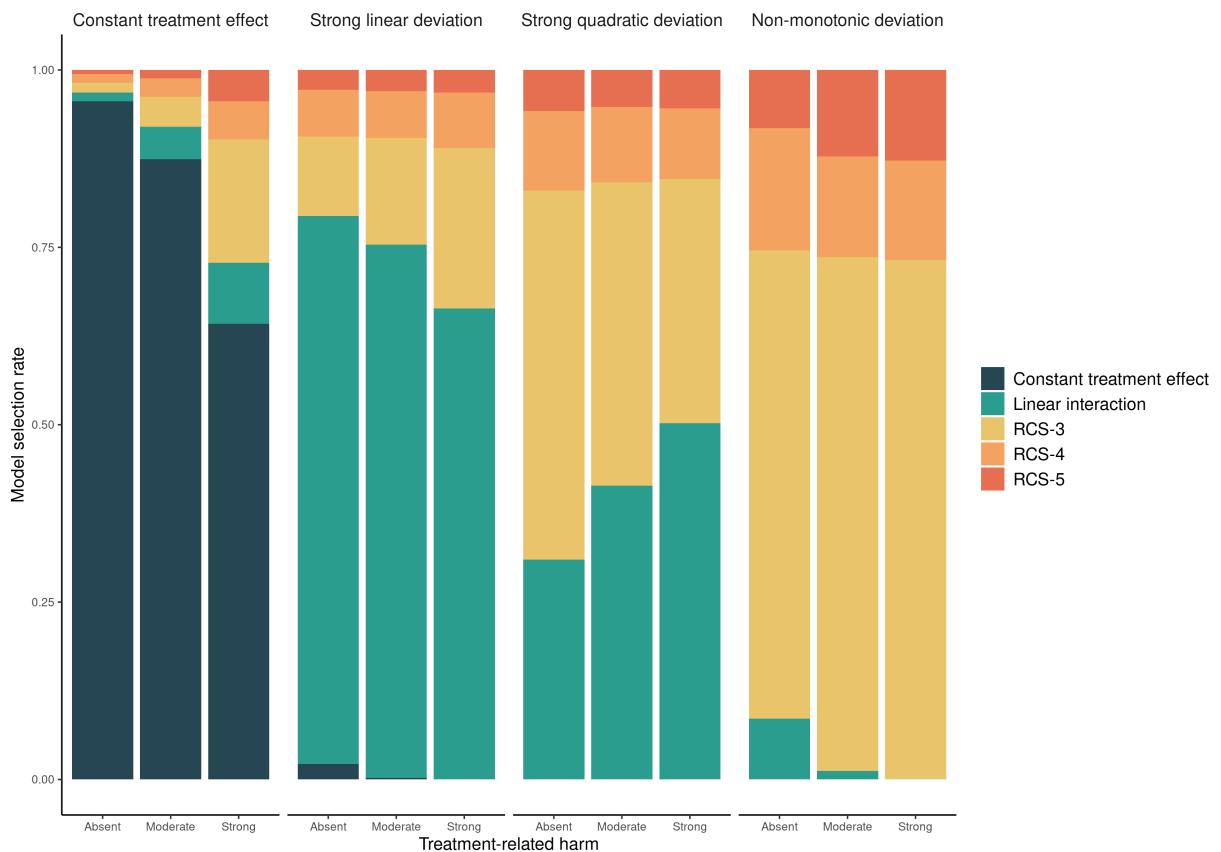


Figure S4: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. Sample size is 17,000 rather than 4,250 in Figure S3

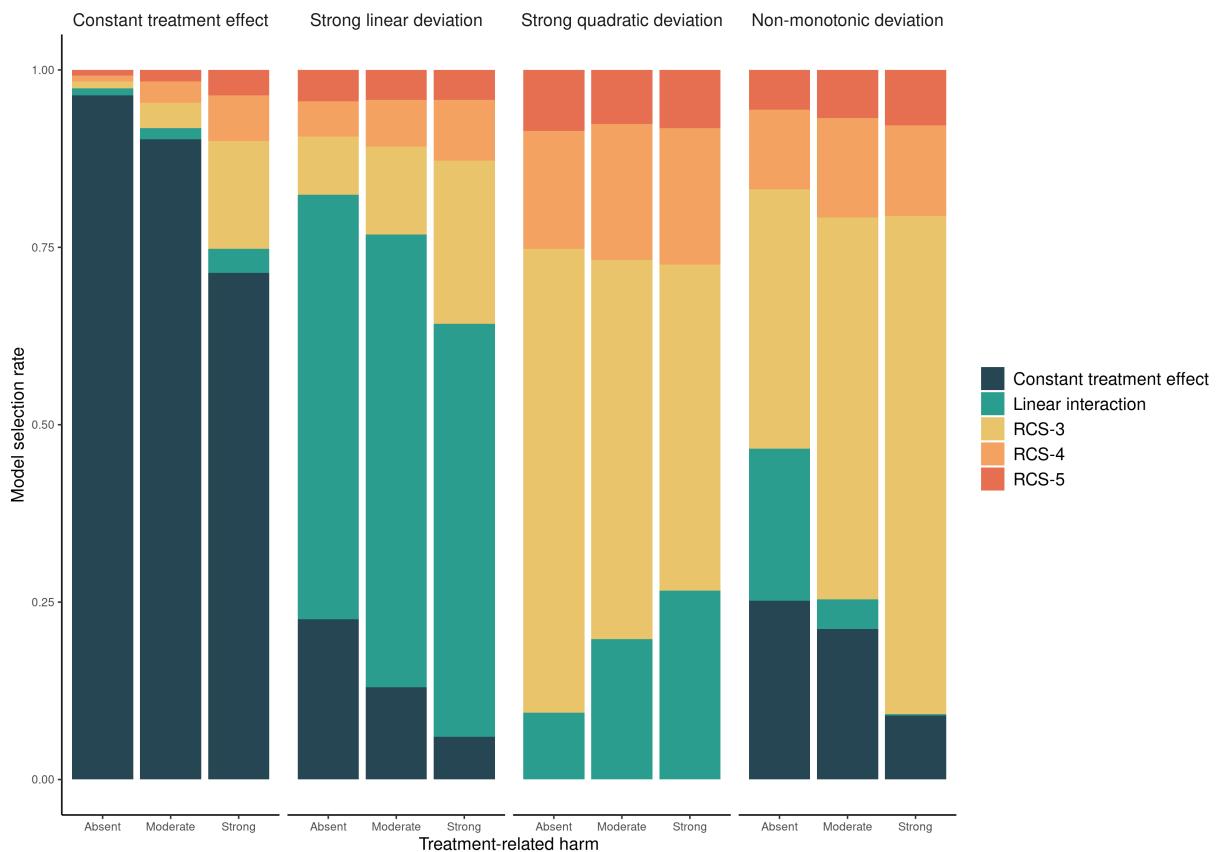


Figure S5: Model selection frequencies of the adaptive approach based on Akaike's Information Criterion across 500 replications. AUC is 0.85 rather than 0.75 in Figure S3

5 Discrimination and calibration for benefit

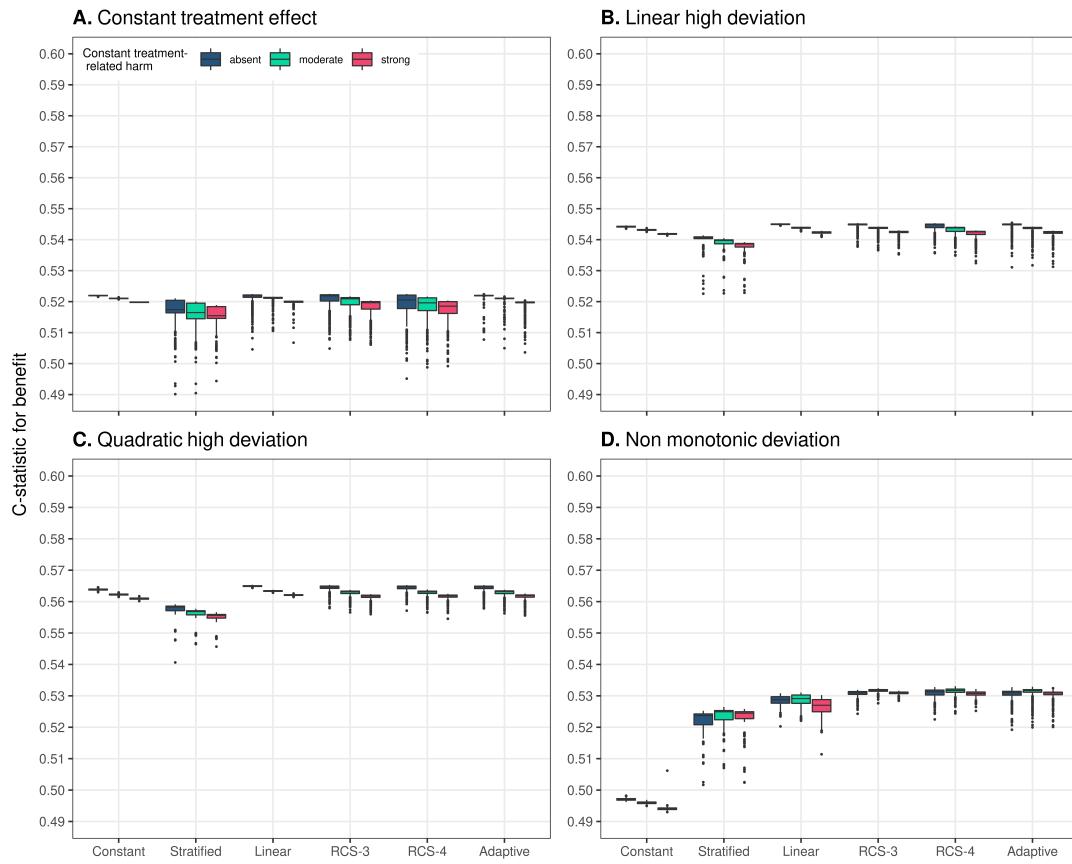


Figure S6: Discrimination for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.75 and sample size of 17,000

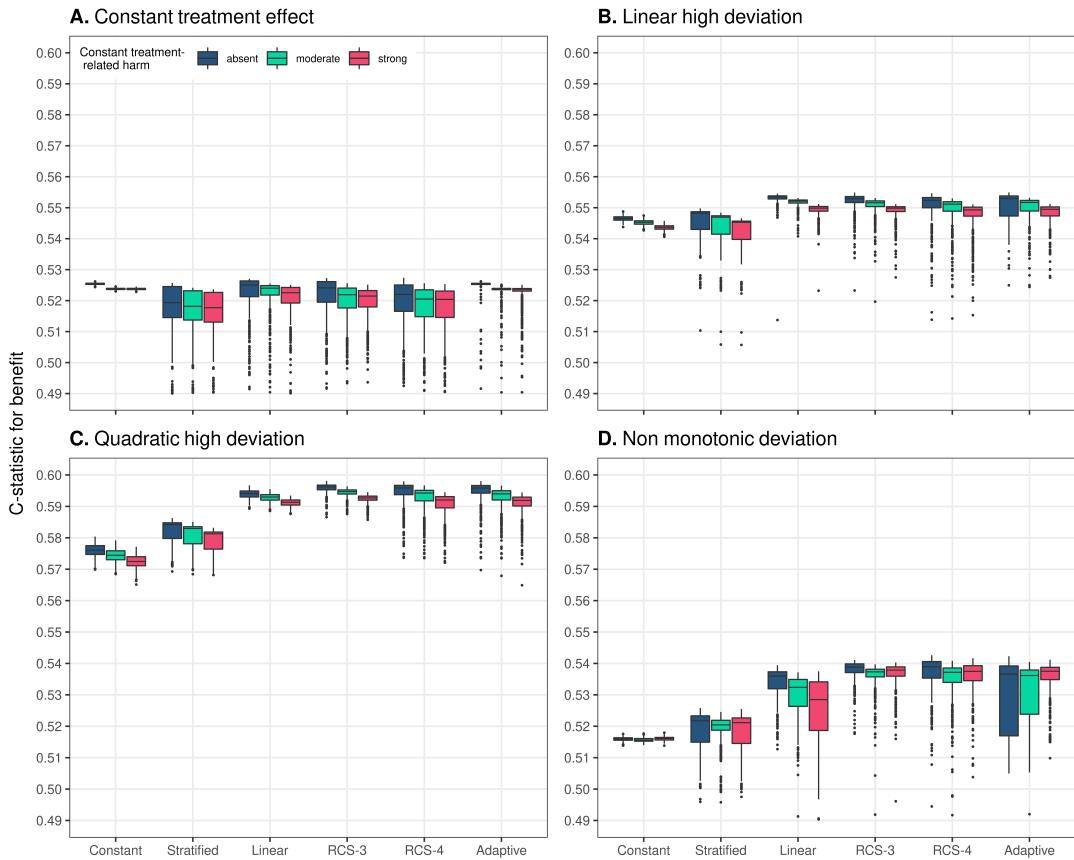


Figure S7: Discrimination for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.85 and sample size of 4,250

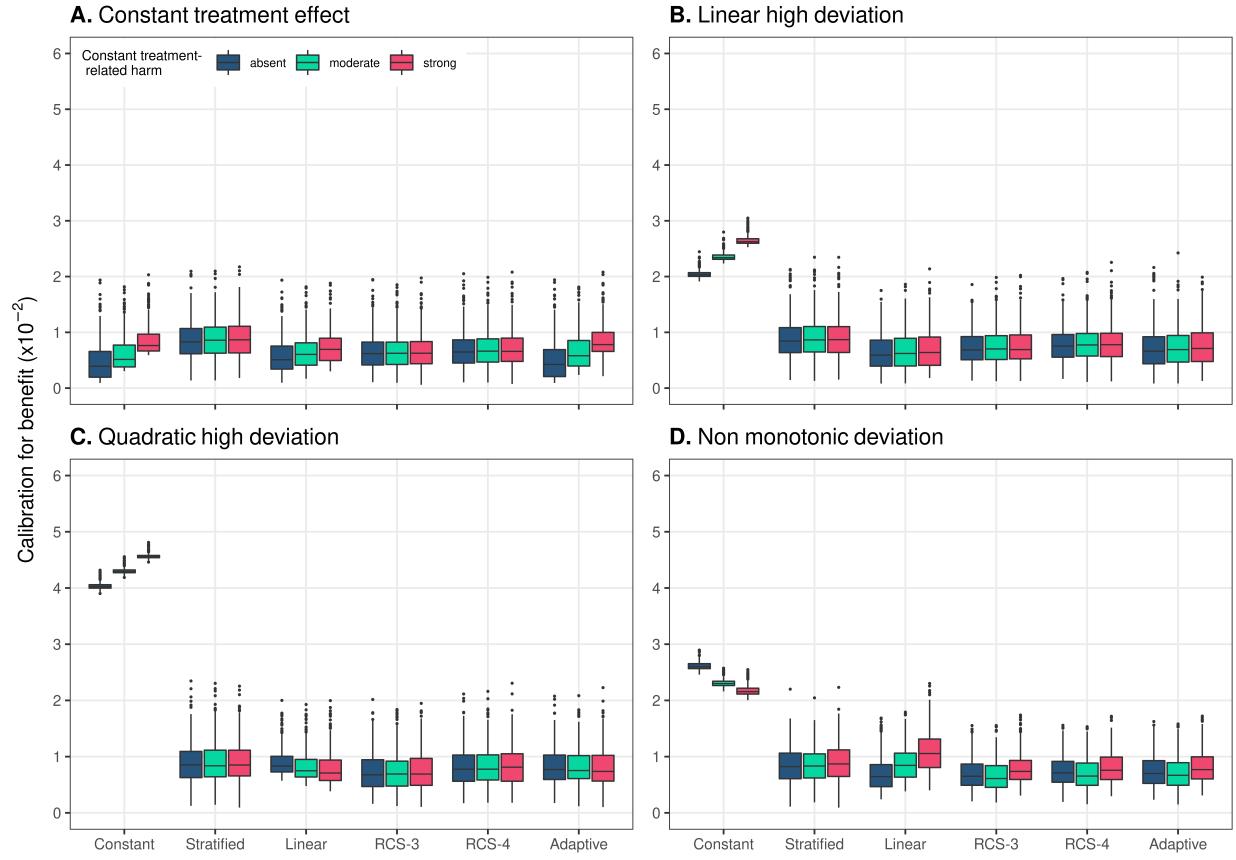


Figure S8: Calibration for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.75 and sample size of 17,000

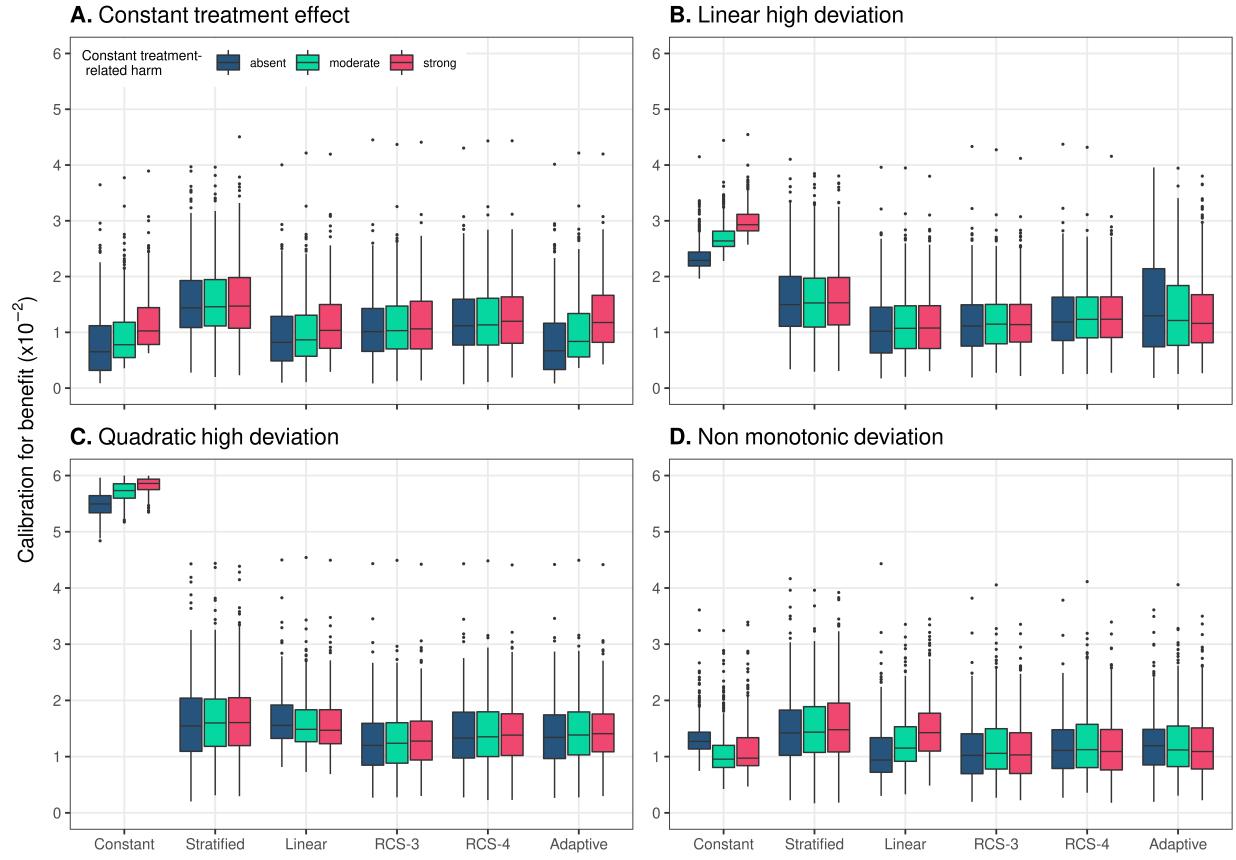


Figure S9: Calibration for benefit of the considered methods across 500 replications calculated in a simulated sample of size 500,000. True prediction AUC of 0.85 and sample size of 4,250

6 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/.

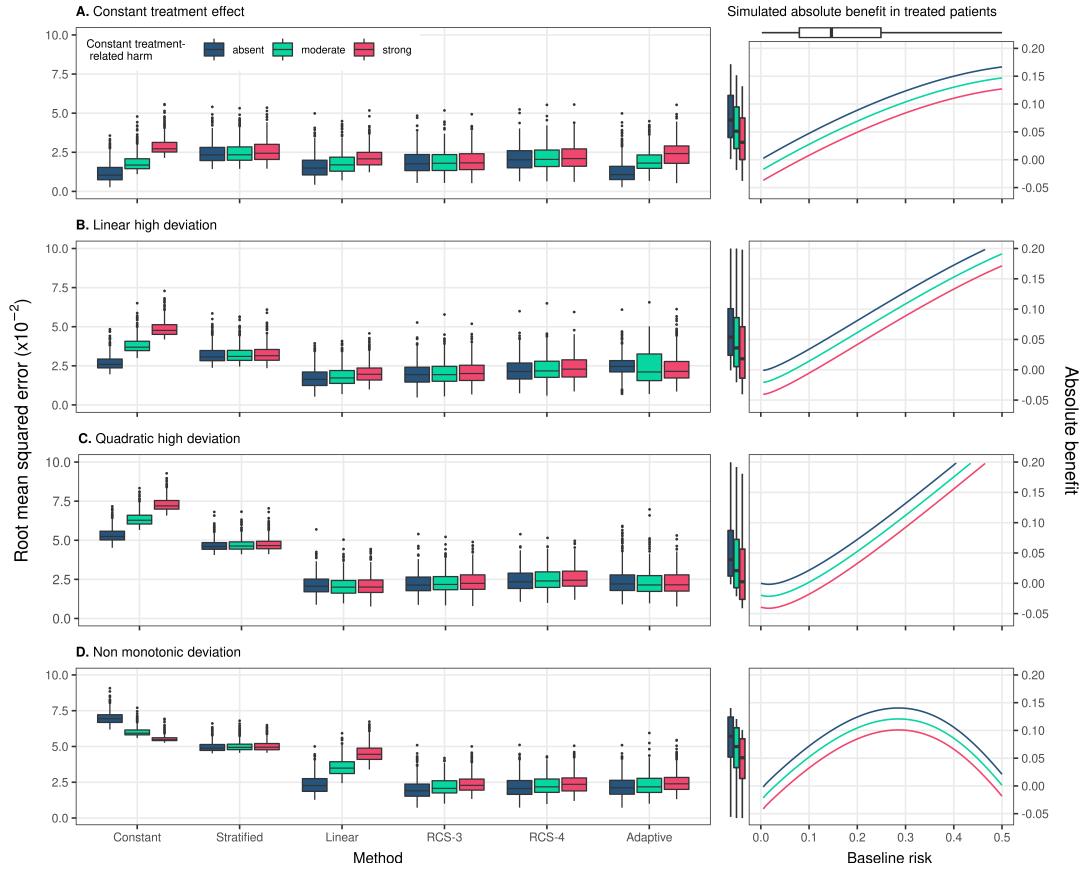


Figure S10: RMSE of the considered methods across 500 replications calculated in a simulated super-population of size 500,000. The scenario with true constant relative treatment effect (panel A) had a true prediction AUC of 0.75 and sample size of 4,250. The RMSE is also presented for strong linear (panel 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.

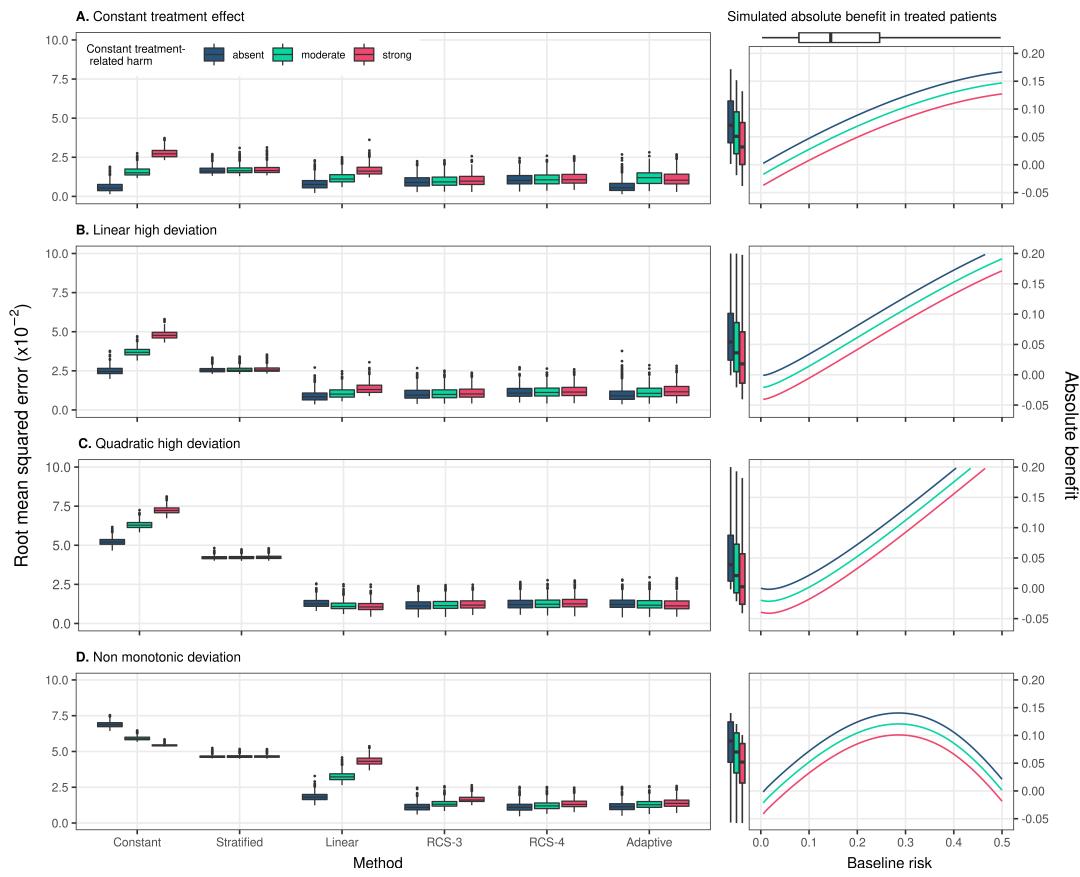


Figure S11: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000. Sample size is 17,000 rather than 4,250 in Figure S10.

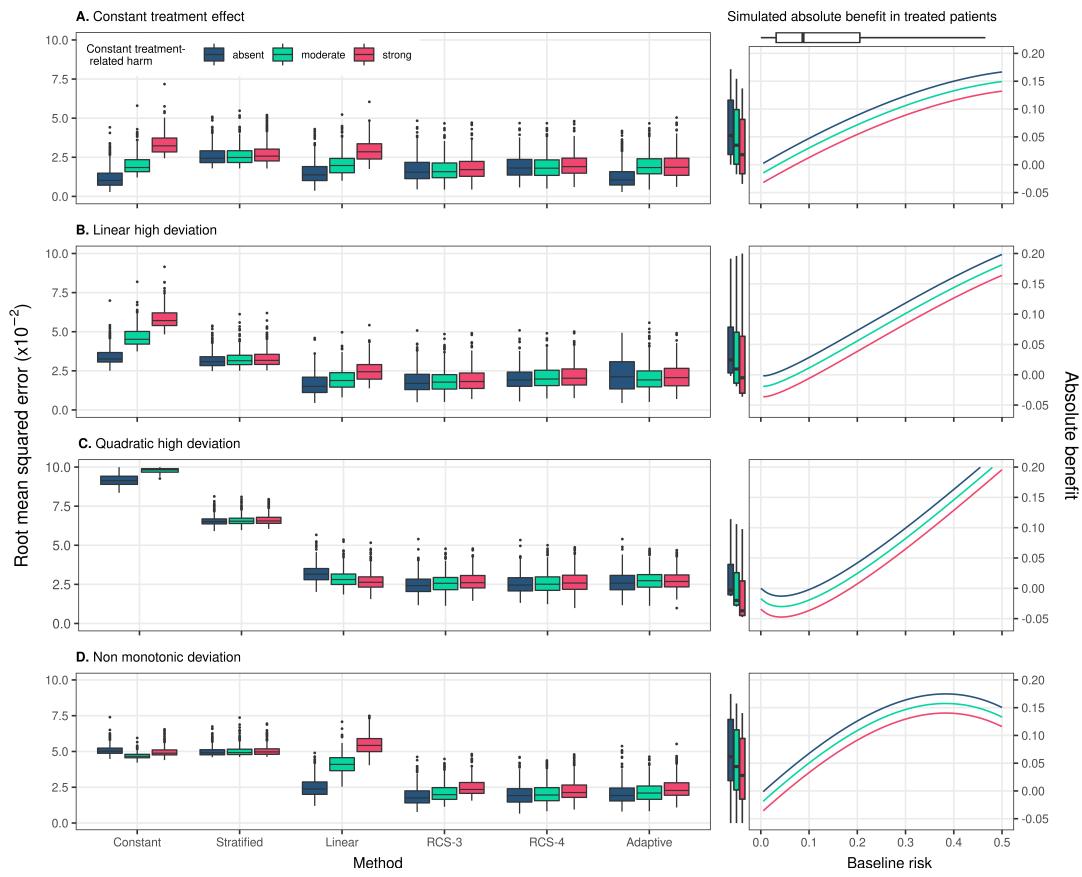


Figure S12: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000. AUC is 0.85 rather than in Figure S10.

7 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 treatment-covariate interaction scenarios.

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

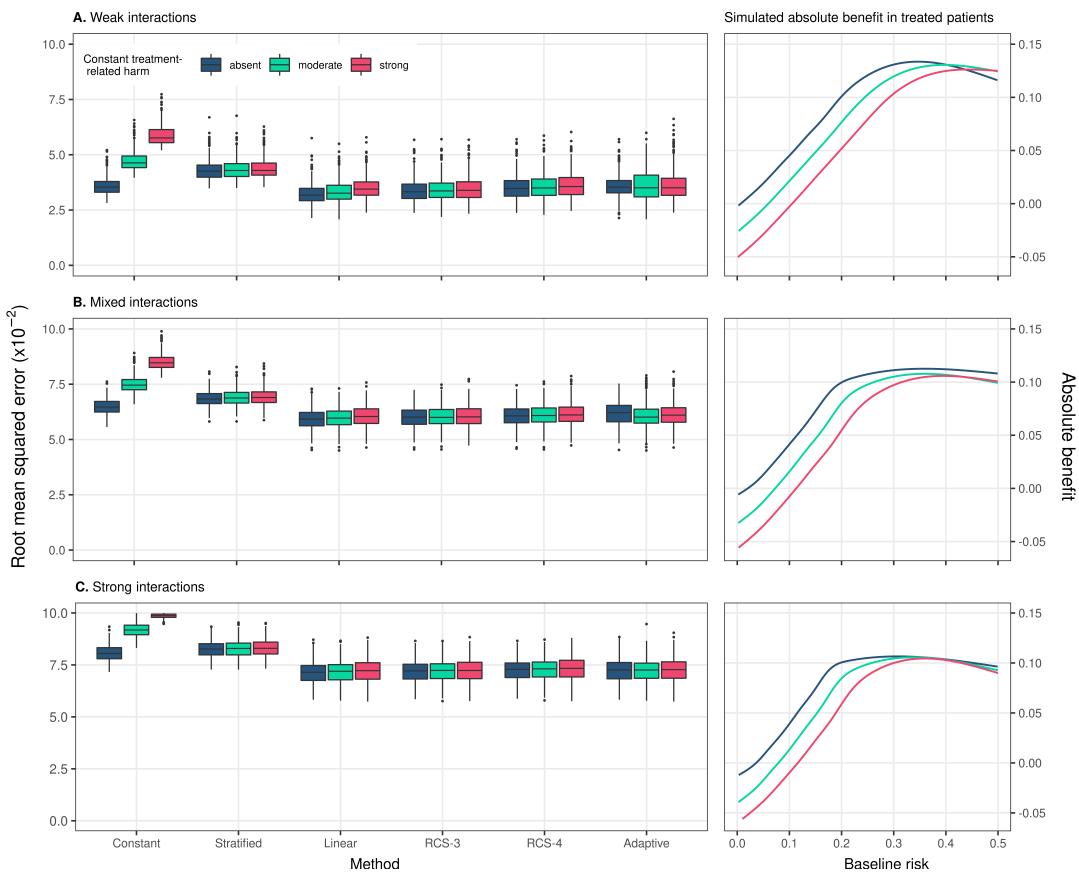


Figure S13: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions all favoring treatment were considered.

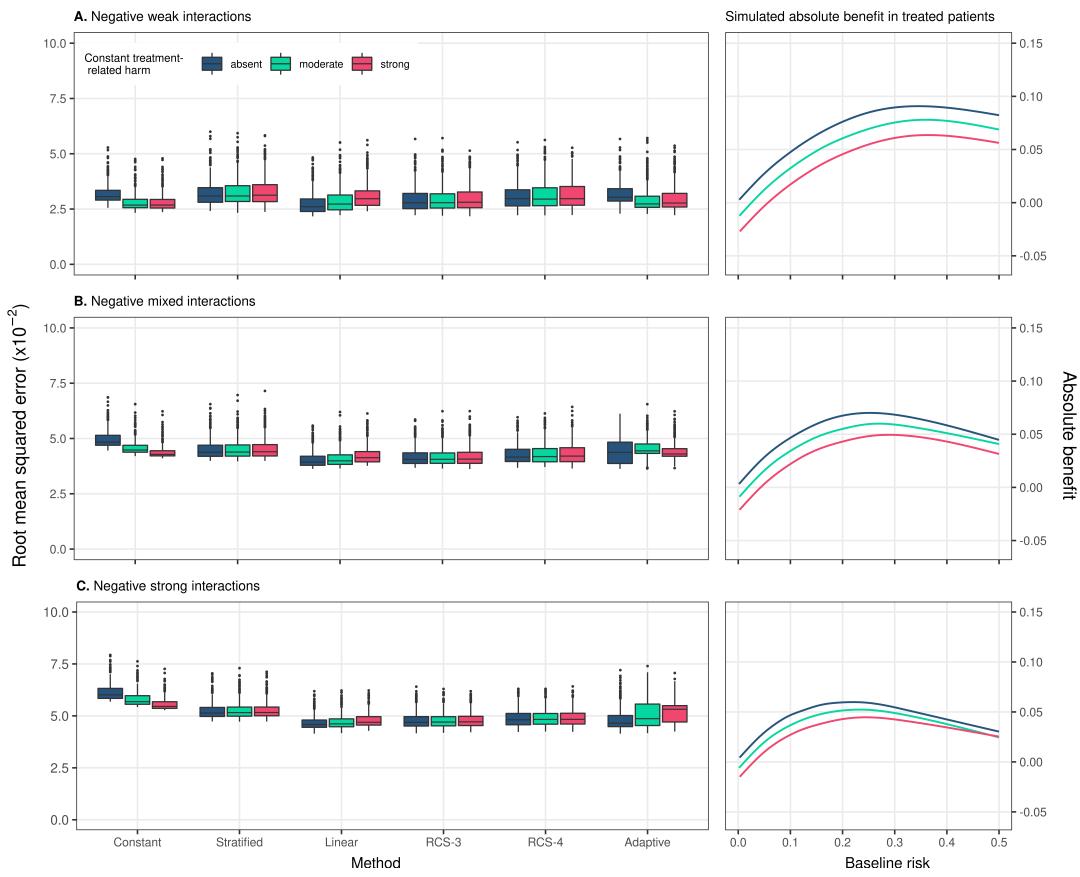


Figure S14: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions all favoring the control were considered.

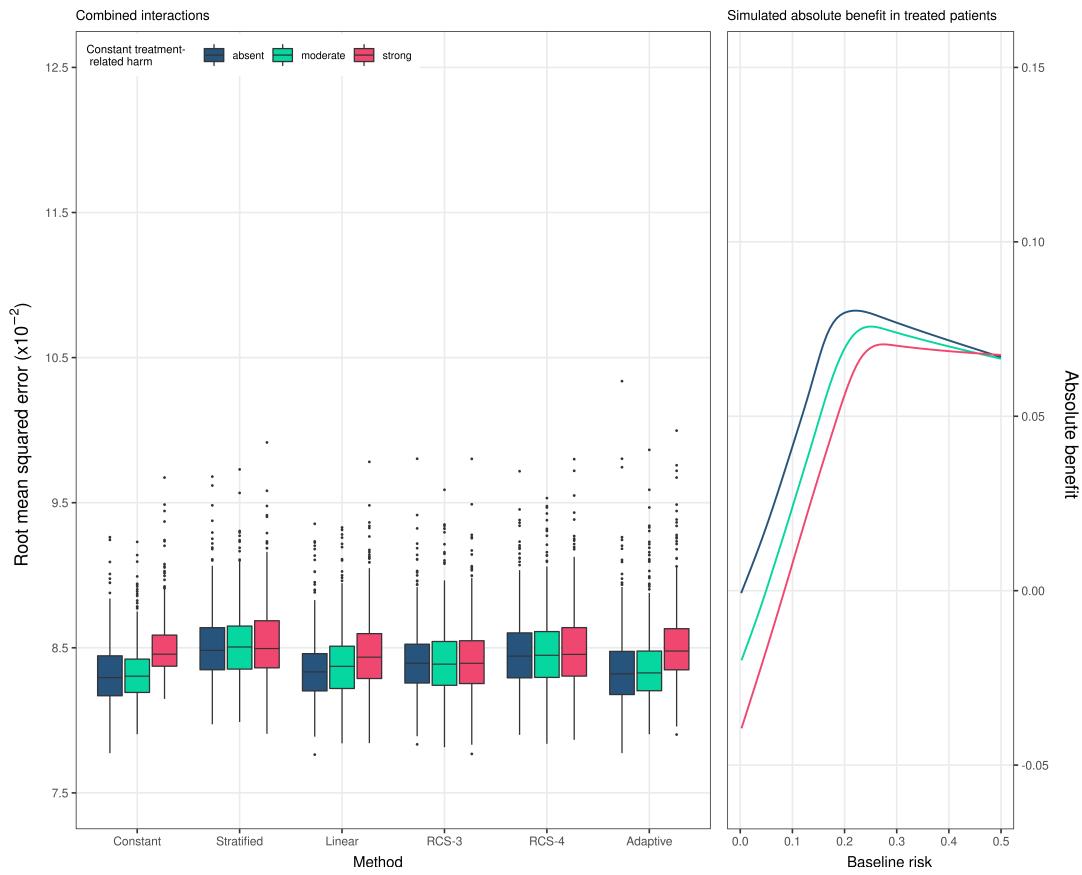


Figure S15: RMSE of the considered methods across 500 replications calculated in a simulated sample of size 500,000 where treatment-covariate interactions 2 favoring treatment and 2 favoring the control were considered.

8 Empirical illustration

For predicting baseline risk of 30-day mortality we fitted a logistic regression model with age, Killip class (*Killip*), systolic blood pressure (*sysbp*), pulse rate (*pulse*), prior myocardial infarction (*pmi*), location of myocardial infarction (*miloc*) and treatment as the covariates. Baseline predictions were made setting treatment to 0.

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

where

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

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

9 References

- [1] Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: Re-analysis of individual participant data from 32 large clinical trials. International Journal of Epidemiology 2016:dyw118. <https://doi.org/10.1093/ije/dyw118>.