Appendix A

The probability of switching in the different progression and biomarker groups at the three visits following disease progression for the base case scenario are presented in Table A1. Higher group numbers represent higher values for that group (that is, "time to progression group" 0 are the control group patients that had time-to-progression times in the lowest 33.3% of the control group). Note however that these groups only refer to patients who became "at-risk" of switching – that is, those control group patients that survived for longer than 21 days. Hence the lowest 33% represent the lowest third of the at-risk group, not the control group as a whole. As an example, a patient whose biomarker value at the point of disease progression was in the middle 33% of those values in control group patients who became at-risk of switching, and whose time to progression was in the upper 33% of these patients, had a probability of switching of 0.75 at the first visit after disease progression, 0.68 at the second visit (if they did not switch at the first visit), and 0.55 and the third visit after progression (if they did not switch at the first or second visits).

Table A1: Probability of treatment switching by prognostic groups and visit

Probability of switching at visit 1		Biomarker group at progression		
		0	1	2
Time to progression group	0	0.30	0.20	0.10
	1	0.51	0.41	0.30
	2	0.88	0.75	0.60
Probability of switching at visit 2		Biomarker group at progression		
		0	1	2
Time to progression group	0	0.23	0.15	0.07
	1	0.42	0.32	0.23
	2	0.84	0.68	0.51
Probability of switching at visit 3		Biomarker group at progression		
		0	1	2
Time to progression group	0	0.14	0.09	0.04
	1	0.29	0.22	0.15
	2	0.75	0.55	0.38

When switching is incorporated into Scenario 1 using the switching probabilities presented in Table A1, 156 (62%) control group patients switch treatments. The resulting Kaplan-Meier curve is as shown in Figure A1. As desired, the control group Kaplan-Meier curve moves towards the experimental group Kaplan-Meier curve (as can be seen by comparing Figure 1 (reproduced here) with Figure A1). The ITT HR in this instance increases to 0.743 from 0.719, demonstrating the effect of the switching.

Figure A1: Overall Survival Kaplan-Meier from simulated dataset Scenario 1: With switching

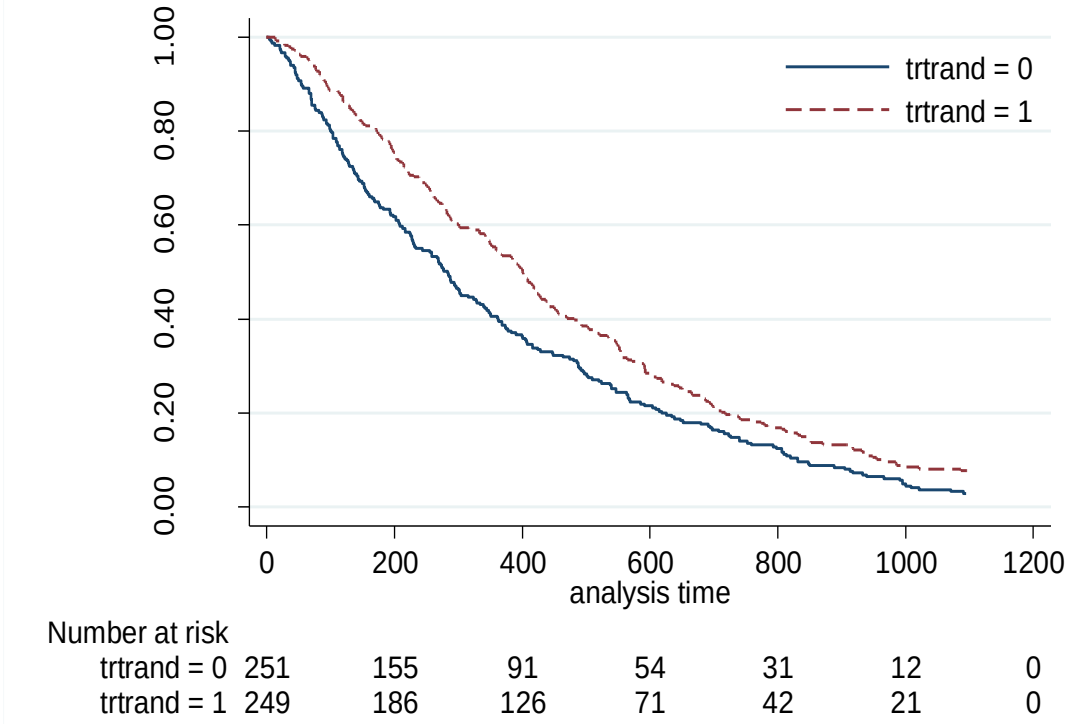
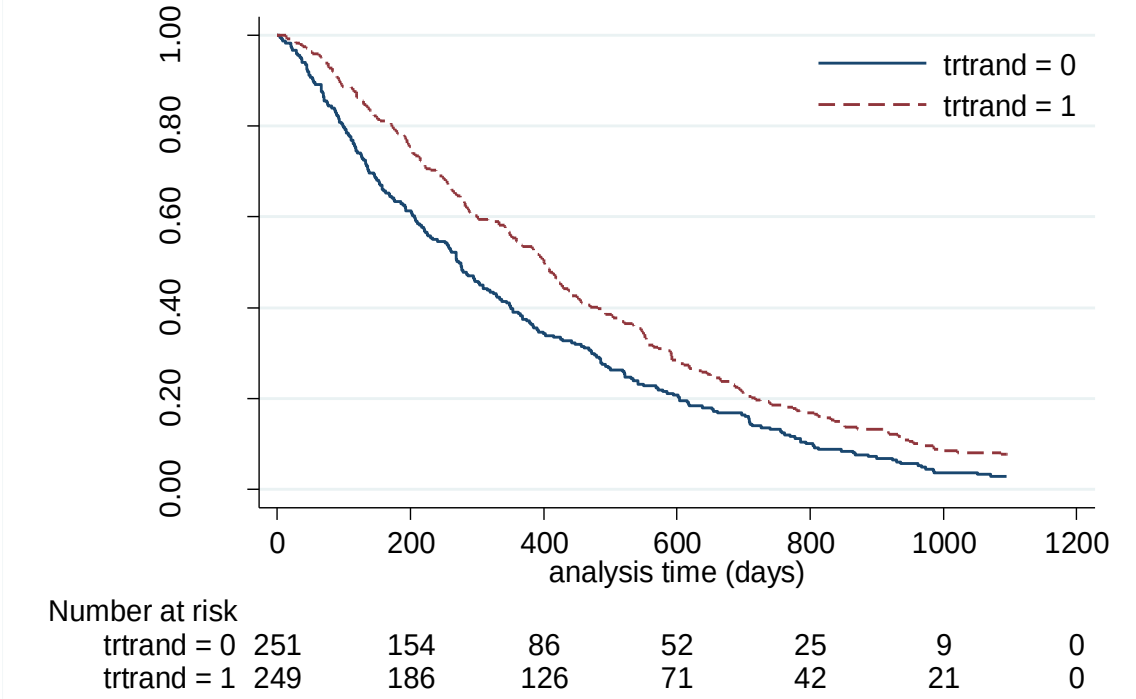


Figure 1: Overall Survival Kaplan-Meier from one simulated dataset Scenario 1: No switching



Appendix B

Values for each variable in Scenario 1 are quoted, as are alternative values for different scenarios.

Table B1: Simulated scenarios – Parameter values and alternatives tested

Variable	Value (Scenario 1)	Alternative Values
Sample size	500	-
Number of prognosis groups (prog)	2	-
Probability of good prognosis	0.5	-
Probability of poor prognosis	0.5	-
Maximum follow-up time	3 years (1095 days)	-
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	-
Survival time distribution	Weibull parameters when time-dependent effect included: Shape parameter 0.9 (mortality decreasing over time) Scale parameter 0.0005 Weibull parameters when time-dependent effect not included: Shape parameter 0.9 (mortality decreasing over time) Scale parameter 0.004	Alter scale parameter to 0.001 to represent a more severe disease (and hence less censoring) in scenarios with time-dependent effect Alter scale parameter to 0.007 to represent a more severe disease (and hence less censoring) in scenarios without time-dependent effect
Progression free survival	Overall survival time multiplied by a value from a beta distribution with shape parameters (5,5) – this implies the assumption that PFS is approximately half of OS. This is not an important assumption – PFS is only included because we model a situation where switching cannot occur before disease progression	-
Baseline treatment effect (note this is not the true treatment effect as this does not take into account the effect of the treatment that occurs through the time-dependent confounder, biomarker level, or the time-dependent part of the treatment effect, η)	Baseline log hazard ratio in scenarios that include an additional time-dependent effect = -0.7 Log hazard ratio in scenarios that do not include an additional time-dependent effect = -0.3	Alter log hazard ratio to -1.1 to represent a larger treatment effect in scenarios with time-dependent effect Alter log hazard ratio to -0.7 to represent a larger treatment effect in scenarios without time-dependent effect
Biomarker intercept	Calculated using a normal distribution with mean of 20 and standard deviation of 1	-
Biomarker value progression over time	As demonstrated by formula (8). $\beta_2 = -4$ to represent that the biomarker value increases more slowly in the experimental group, and $\beta_4 = 5$ so that bad prognosis patients start with higher levels of the biomarker	-
Impact of biomarker value on overall survival	As demonstrated by formulas (9). Increased biomarker value increases the risk of death. The strength of this relationship depends on the variable α, which equals 0.02 in Scenario 1	Remove impact of the biomarker value by setting α=0
Impact of biomarker value on treatment effect	As demonstrated by formulas (13). Because treatment reduces the progression of the biomarker value and increased biomarker values increase the	Remove impact of biomarker value by setting α=0 Model larger time-dependent effect

	risk of death, the treatment has an additional effect through the biomarker. The strength of this relationship depends on the variable α, which equals 0.02 in Scenario 1	by applying additional decrement multiplier to switching patients
Time-dependent portion of treatment effect, η	η =0.15 to generate a reduction in the treatment effect over time, in scenarios where a time-dependent treatment effect is assumed	Remove impact of η by setting η =0 in scenarios where the treatment effect is not time-dependent
Assumed frequency of visits	One every 3 weeks (21 days)	-
Probability of switching treatment over time	As shown in Table A1. This results in a switching proportion of approximately 63% in Scenario 1	Test a high switching scenario where all probabilities are increased to an extent where approximately 90% of patients that survive longer than 21 days switch
Prognosis of switching patients	As shown in Table A1. This makes switching more likely in good prognosis patients, via a mechanism that takes into account both time to progression and biomarker value at progression	Make switching more likely in poor prognosis patients. Test scenarios where switching is based on a simpler mechanism (only based on the biomarker value)
Treatment effect in switching patients	Equal to baseline treatment effect multiplied by ω. Set ω such that treatment effect received by switching patients is 85% of the average effect received by experimental group patients in base scenarios.	Alter ω such that treatment effect received by switching patients equals 75% - 78% of the average effect received by experimental group patients. In scenarios where the treatment effect is not time-dependent, set ω to 1 – such that the treatment effect received by switching patients is the baseline effect received by the experimental group.

Based on the variables and alternative values presented in Table B1, 72 scenarios were run.

Appendix C: Overview of simulation scenarios

Table C1 provides information on each of the scenarios simulated. The true restricted mean survival unconfounded by treatment switching is presented, along with the average treatment effect in terms of a hazard ratio (calculated using a Cox model) and an acceleration factor (calculated using a Weibull model). Where there is a time-dependent treatment effect this reflects only an approximation of the true treatment effect as proportional hazards/constant acceleration factor assumptions do not hold. In terms of a hazard ratio, the average treatment effect varied between 0.50 and 0.75.

Table C1 shows that the switching proportion varied between 52% and 94% of all control group patients. Scenarios 13-24, 37-48 and 61-72 were designed to result in higher levels of switching, although these levels are probabilistic and are reliant on other characteristics. This led the level of switching to vary between otherwise equivalent scenarios with switching proportions highest in Scenarios 25-48, followed by Scenarios 49-72 and 1-24. Table C1 also presents the switching proportion as a percentage of the control group patients that became "at-risk" of switching. Control group patients could only switch if they were alive at their first "visit" at 21 days. The proportion of patients that died before this point and never became at-risk depended upon disease severity. The proportion of switching patients as a percentage of patients that became at-risk ranged from 61% to 96%.

Table C1 shows that Scenarios 1-24 incorporated a complex switching probability mechanism in which better prognosis patients were generally more likely to switch. Scenarios 25-48 and 49-72 incorporated a simpler switching probability mechanism based only upon biomarker level at the time of disease progression. In Scenarios 25-48 patients with a relatively poor prognosis were more likely to switch. The opposite was true in Scenarios 49-72.

Table C1 shows that in Scenarios 1, 2, 5, 6, 13, 14, 17 and 18 the treatment effect received by experimental group patients was dependent upon the biomarker level and time and in these scenarios switching patients received a reduced treatment effect. In Scenarios 9, 10, 11, 12, 21, 22, 23 and 24 the treatment effect received by experimental group patients was time-dependent and related to the biomarker level, and an additional decrement (compared to scenarios 1, 2, 5, 6, 13, 14, 17 and 18) was applied to the effect received by switching patients (in Table C1 these are labelled as "Yes+" in the "Time-dependent treatment effect" column). In scenarios 3, 4, 7, 8, 15, 16, 19 and 20 the treatment effect was not time-dependent or related to the biomarker – in these scenarios the "common treatment effect" assumption held. This pattern across scenarios was repeated in scenarios 25-48 and 49-72 (i.e. scenarios 25 and 49 are equivalent to scenario 1, except with altered switching mechanisms, and so on). Further details on the effect size decrement applied to switching patients in each scenario are included ("Treatment effect in switching patients (AF)"), as well as on the effect size as a proportion of that received by the experimental group. This varied between 75% and 100%.

Table C1: Overview of simulated scenarios

Scenario	Truth		Average treatment effects		Mean switching % of total	Mean switching % of at risk	Mean censoring proportion (%)	Disease severity	Prognosis of switching patients	Time-dependent treatment effect	Treatment effect in switching patients (AF)	% of exp group treatment effect
	Restricted mean (Control group)	Restricted mean (Exp group)	HR	AF								
1	372.06	462.27	0.75	1.28	63.37%	65.33%	7.19%	Low	Complex - good	Yes	1.09	85%
2	372.06	579.28	0.52	1.75	61.54%	63.44%	13.42%	Low	Complex - good	Yes	1.48	85%
3	344.47	568.12	0.51	2.15	56.26%	61.07%	19.97%	Low	Complex - good	No	2.15	100%
4	344.47	437.88	0.75	1.39	58.46%	63.45%	11.65%	Low	Complex - good	No	1.39	100%
5	216.96	285.64	0.73	1.32	60.25%	64.04%	0.84%	High	Complex - good	Yes	1.12	85%
6	216.96	381.51	0.50	1.80	58.17%	61.82%	2.74%	High	Complex - good	Yes	1.53	85%
7	201.45	387.21	0.51	2.17	52.48%	60.56%	7.02%	High	Complex - good	No	2.17	100%
8	201.45	271.95	0.75	1.40	54.09%	62.41%	2.80%	High	Complex - good	No	1.40	100%
9	372.06	462.27	0.75	1.28	63.74%	65.71%	6.86%	Low	Complex - good	Yes +	1.00	78%
10	372.06	579.21	0.52	1.75	61.45%	63.36%	12.86%	Low	Complex - good	Yes +	1.31	75%
11	216.96	285.64	0.73	1.32	60.52%	64.32%	0.76%	High	Complex - good	Yes +	1.00	76%
12	216.96	381.51	0.50	1.80	58.02%	61.66%	2.58%	High	Complex - good	Yes +	1.36	75%
13	372.06	462.27	0.75	1.28	88.37%	91.10%	7.25%	Low	Complex - good	Yes	1.09	85%
14	372.06	579.21	0.52	1.75	87.96%	90.68%	13.80%	Low	Complex - good	Yes	1.48	85%
15	344.47	568.12	0.51	2.15	80.99%	87.90%	20.60%	Low	Complex - good	No	2.15	100%
16	344.47	437.88	0.75	1.39	81.29%	88.23%	11.80%	Low	Complex - good	No	1.39	100%
17	216.96	285.64	0.73	1.32	83.30%	88.53%	0.83%	High	Complex - good	Yes	1.12	85%
18	216.96	381.51	0.50	1.80	82.66%	87.85%	2.81%	High	Complex - good	Yes	1.53	85%
19	201.45	387.21	0.51	2.17	74.83%	86.34%	7.24%	High	Complex - good	No	2.17	100%
20	201.45	387.21	0.75	1.40	75.20%	86.77%	2.84%	High	Complex - good	No	1.40	100%
21	372.06	462.27	0.75	1.28	88.39%	91.12%	6.87%	Low	Complex - good	Yes +	1.00	78%
22	372.06	579.21	0.52	1.75	88.02%	90.75%	13.09%	Low	Complex - good	Yes +	1.31	75%
23	216.96	285.64	0.73	1.32	83.30%	88.53%	0.74%	High	Complex - good	Yes +	1.00	76%
24	216.96	381.51	0.50	1.80	82.77%	87.97%	2.65%	High	Complex - good	Yes +	1.36	75%
25	372.06	462.27	0.75	1.28	69.65%	71.80%	7.02%	Low	Simple - poor	Yes	1.09	85%

26	372.06	579.21	0.52	1.75	71.62%	73.84%	13.10%	Low	Simple - poor	Yes	1.48	85%
27	344.47	568.12	0.51	2.15	65.86%	71.48%	19.08%	Low	Simple - poor	No	2.15	100%
28	344.47	437.88	0.75	1.39	64.03%	69.50%	11.15%	Low	Simple - poor	No	1.39	100%
29	216.96	285.64	0.73	1.32	64.89%	68.96%	0.81%	High	Simple - poor	Yes	1.12	85%
30	216.96	381.51	0.50	1.80	66.71%	70.90%	2.60%	High	Simple - poor	Yes	1.53	85%
31	201.45	387.21	0.51	2.17	59.78%	68.98%	6.47%	High	Simple - poor	No	2.17	100%
32	201.45	387.21	0.75	1.40	57.88%	66.79%	2.61%	High	Simple - poor	No	1.40	100%
33	372.06	462.27	0.75	1.28	69.16%	71.30%	6.89%	Low	Simple - poor	Yes +	1.00	78%
34	372.06	579.21	0.52	1.75	71.64%	73.86%	12.52%	Low	Simple - poor	Yes +	1.31	75%
35	216.96	285.64	0.73	1.32	65.02%	69.10%	0.75%	High	Simple - poor	Yes +	1.00	76%
36	216.96	381.51	0.50	1.80	66.95%	71.15%	2.52%	High	Simple - poor	Yes +	1.36	75%
37	372.06	462.27	0.75	1.28	93.31%	96.20%	7.23%	Low	Simple - poor	Yes	1.09	85%
38	372.06	579.21	0.52	1.75	93.41%	96.30%	13.78%	Low	Simple - poor	Yes	1.48	85%
39	344.47	568.12	0.51	2.15	87.06%	94.50%	20.65%	Low	Simple - poor	No	2.15	100%
40	344.47	437.88	0.75	1.39	86.83%	94.25%	11.72%	Low	Simple - poor	No	1.39	100%
41	216.96	285.64	0.73	1.32	89.12%	94.71%	0.84%	High	Simple - poor	Yes	1.12	85%
42	216.96	381.51	0.50	1.80	89.54%	95.16%	2.80%	High	Simple - poor	Yes	1.53	85%
43	201.45	387.21	0.51	2.17	81.17%	93.65%	7.17%	High	Simple - poor	No	2.17	100%
44	201.45	387.21	0.75	1.40	80.77%	93.20%	2.81%	High	Simple - poor	No	1.40	100%
45	372.06	462.27	0.75	1.28	93.31%	96.20%	6.86%	Low	Simple - poor	Yes +	1.00	78%
46	372.06	579.21	0.52	1.75	93.51%	96.40%	12.96%	Low	Simple - poor	Yes +	1.31	75%
47	216.96	285.64	0.73	1.32	88.87%	94.44%	0.75%	High	Simple - poor	Yes +	1.00	76%
48	216.96	381.51	0.50	1.80	89.49%	95.11%	2.62%	High	Simple - poor	Yes +	1.36	75%
49	372.06	462.27	0.75	1.28	66.26%	68.31%	7.15%	Low	Simple -good	Yes	1.09	85%
50	372.06	579.21	0.52	1.75	64.57%	66.57%	13.45%	Low	Simple -good	Yes	1.48	85%
51	344.47	568.12	0.51	2.15	59.10%	64.15%	19.99%	Low	Simple -good	No	2.15	100%
52	344.47	437.88	0.75	1.39	60.78%	65.97%	11.46%	Low	Simple -good	No	1.39	100%
53	216.96	285.64	0.73	1.32	61.53%	65.39%	0.82%	High	Simple -good	Yes	1.12	85%
54	216.96	381.51	0.50	1.80	60.01%	63.78%	2.76%	High	Simple -good	Yes	1.53	85%

55	201.45	387.21	0.51	2.17	53.90%	62.19%	6.96%	High	Simple -good	No	2.17	100%
56	201.45	387.21	0.75	1.40	55.26%	63.76%	2.77%	High	Simple -good	No	1.40	100%
57	372.06	462.27	0.75	1.28	66.15%	68.20%	6.91%	Low	Simple -good	Yes +	1.00	78%
58	372.06	579.21	0.52	1.75	64.45%	66.45%	12.86%	Low	Simple -good	Yes +	1.31	75%
59	216.96	285.64	0.73	1.32	61.95%	65.84%	0.78%	High	Simple -good	Yes +	1.00	76%
60	216.96	381.51	0.50	1.80	60.00%	63.76%	2.58%	High	Simple -good	Yes +	1.36	75%
61	372.06	462.27	0.75	1.28	91.44%	94.27%	7.26%	Low	Simple -good	Yes	1.09	85%
62	372.06	579.21	0.52	1.75	90.87%	93.68%	13.88%	Low	Simple -good	Yes	1.48	85%
63	344.47	568.12	0.51	2.15	84.11%	91.30%	20.81%	Low	Simple -good	No	2.15	100%
64	344.47	437.88	0.75	1.39	84.13%	91.31%	11.77%	Low	Simple -good	No	1.39	100%
65	216.96	285.64	0.73	1.32	86.39%	91.81%	0.85%	High	Simple -good	Yes	1.12	85%
66	216.96	381.51	0.50	1.80	86.25%	91.67%	2.79%	High	Simple -good	Yes	1.53	85%
67	201.45	387.21	0.51	2.17	77.63%	89.57%	7.30%	High	Simple -good	No	2.17	100%
68	201.45	387.21	0.75	1.40	77.78%	89.75%	2.81%	High	Simple -good	No	1.40	100%
69	372.06	462.27	0.75	1.28	91.42%	94.25%	6.87%	Low	Simple -good	Yes +	1.00	78%
70	372.06	579.21	0.52	1.75	91.40%	94.23%	12.98%	Low	Simple -good	Yes +	1.31	75%
71	216.96	285.64	0.73	1.32	86.59%	92.03%	0.75%	High	Simple -good	Yes +	1.00	76%
72	216.96	381.51	0.50	1.80	86.27%	91.69%	2.59%	High	Simple -good	Yes +	1.36	75%

Appendix D: Percentage bias figures

Figures showing bias across scenarios are presented throughout this Appendix – care should be taken when comparing these because the y-axes use different scales.

Figure D1: Percentage bias (%) across scenarios – ITT

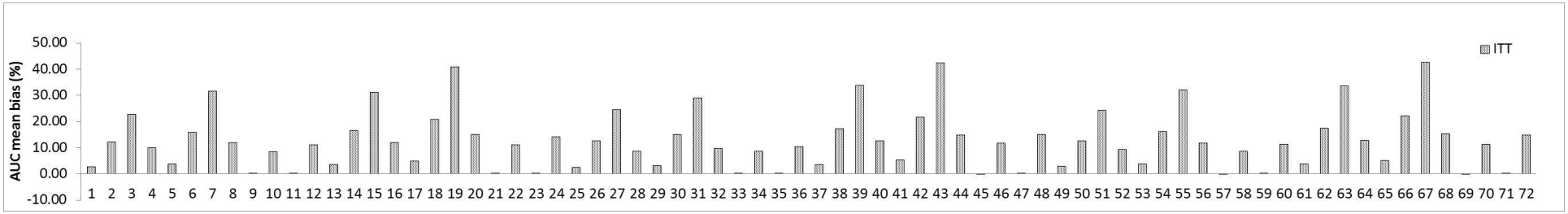


Figure D2: Percentage bias (%) across scenarios – Exclusion and censoring approaches

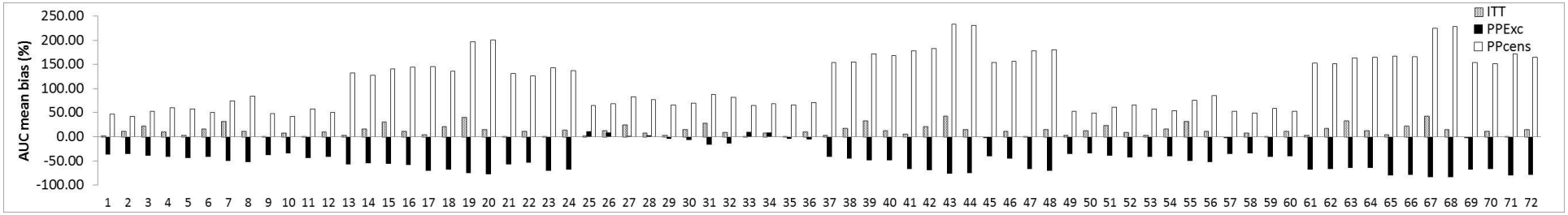


Figure D3: Percentage bias (%) across scenarios – IPCW

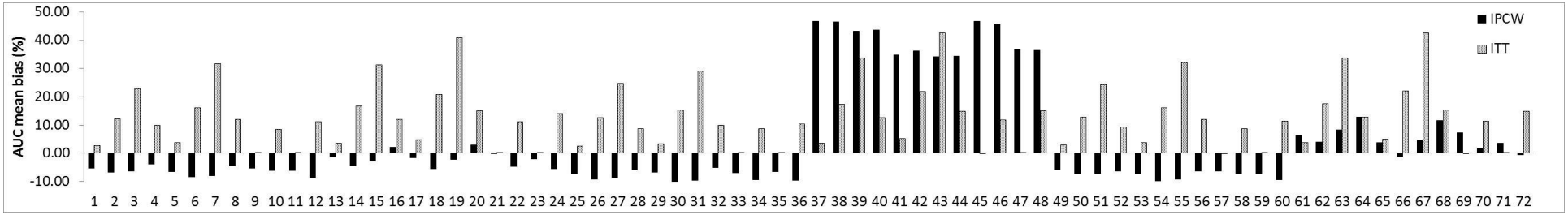


Figure D4: Percentage bias (%) across scenarios – RPSFTM and IPE Weibull approaches

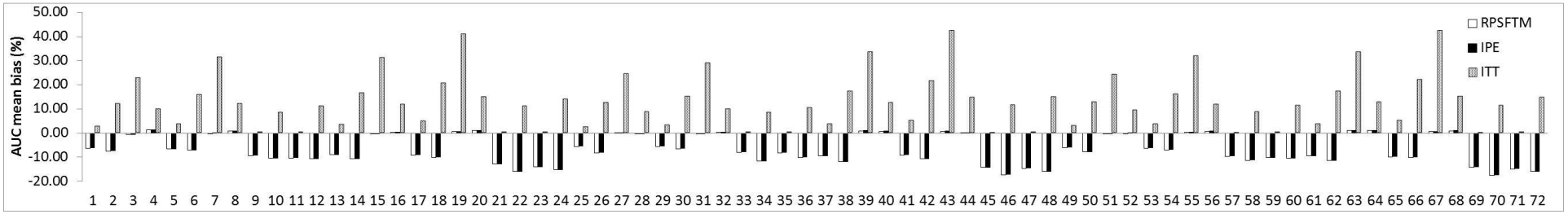


Figure D5: Percentage bias (%) across scenarios – SNM

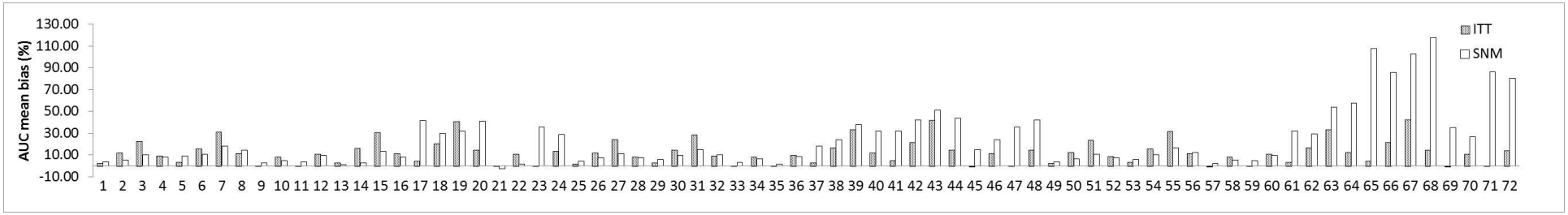


Figure D6: Percentage bias (%) across scenarios – Two-stage Weibull