The harm of SMOTE and other resampling techniques in clinical prediction models: a simulation study

Supplementary material

Ruben van den Goorbergh

TABLE S1 Performance of models on scenarios with $N = 5000$, $R = 3$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.73	0.02	0.95
		0.1	0.90	0.02	1.00	0.73	0.00	0.99
		0.3	0.74	0.31	0.92	0.74	-0.00	1.00
	RUS	0.01	0.66	0.65	0.66	0.72	-4.71	0.83
		0.1	0.66	0.67	0.66	0.73	-2.20	0.98
		0.3	0.67	0.68	0.67	0.74	-0.84	1.00
	ROS	0.01	0.67	0.65	0.67	0.72	-4.64	0.91
		0.1	0.66	0.67	0.66	0.73	-2.20	0.99
		0.3	0.67	0.68	0.67	0.74	-0.84	1.00
	SMOTE	0.01	0.68	0.64	0.68	0.72	-4.68	0.81
		0.1	0.68	0.64	0.69	0.73	-2.15	0.95
		0.3	0.69	0.61	0.73	0.74	-0.70	0.99
RID	Unadjusted	0.01	0.99	0.00	1.00	0.73	0.02	1.06
		0.1	0.90	0.02	1.00	0.73	0.00	1.01
		0.3	0.74	0.31	0.92	0.74	-0.00	1.01
	RUS	0.01	0.66	0.65	0.66	0.72	-4.60	1.12
		0.1	0.66	0.67	0.66	0.73	-2.19	1.02
		0.3	0.67	0.68	0.67	0.74	-0.84	1.01
	ROS	0.01	0.67	0.65	0.67	0.72	-4.64	0.91
		0.1	0.66	0.67	0.66	0.73	-2.20	1.00
		0.3	0.67	0.68	0.67	0.74	-0.84	1.01
	SMOTE	0.01	0.68	0.64	0.68	0.72	-4.68	0.81
		0.1	0.68	0.64	0.69	0.73	-2.15	0.96
		0.3	0.69	0.61	0.73	0.74	-0.69	1.00
RF	Unadjusted	0.01	0.99	0.00	1.00	0.60	-0.00	0.25
		0.1	0.90	0.04	0.99	0.67	-0.02	0.45
		0.3	0.71	0.34	0.87	0.69	-0.01	0.60
	RUS	0.01	0.61	0.61	0.61	0.66	-4.96	0.47
		0.1	0.63	0.63	0.63	0.68	-2.38	0.57
		0.3	0.64	0.64	0.64	0.69	-0.91	0.62
	ROS	0.01	0.99	0.00	1.00	0.60	-0.40	0.21
		0.1	0.88	0.10	0.97	0.66	-0.46	0.38
		0.3	0.69	0.44	0.80	0.69	-0.30	0.55
	SMOTE	0.01	0.91	0.14	0.91	0.62	-3.78	0.18
		0.1	0.78	0.34	0.83	0.66	-1.63	0.38
		0.3	0.68	0.50	0.76	0.69	-0.50	0.55

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S2 Performance of re-calibrated models on scenarios with $N = 5000$, $R = 3$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.69	0.62	0.70	0.73	0.02	0.95
		0.1	0.66	0.67	0.66	0.73	0.00	0.99
		0.3	0.67	0.68	0.67	0.74	-0.00	1.00
	RUS	0.01	0.99	0.00	1.00	0.72	0.02	0.83
		0.1	0.90	0.02	1.00	0.73	0.00	0.98
		0.3	0.74	0.31	0.92	0.74	-0.00	1.00
	ROS	0.01	0.99	0.00	1.00	0.72	0.02	0.91
		0.1	0.90	0.02	1.00	0.73	-0.00	0.99
		0.3	0.74	0.31	0.92	0.74	-0.00	1.00
	SMOTE	0.01	0.99	0.00	1.00	0.72	0.02	0.81
		0.1	0.90	0.02	1.00	0.73	-0.00	0.95
		0.3	0.74	0.32	0.92	0.74	-0.00	0.99
RID	Unadjusted	0.01	0.68	0.63	0.68	0.73	0.02	1.06
		0.1	0.66	0.67	0.66	0.73	0.00	1.01
		0.3	0.67	0.68	0.67	0.74	-0.00	1.01
	RUS	0.01	0.99	0.00	1.00	0.72	0.02	1.12
		0.1	0.90	0.01	1.00	0.73	0.00	1.02
		0.3	0.74	0.31	0.92	0.74	-0.00	1.01
	ROS	0.01	0.99	0.00	1.00	0.72	0.02	0.91
		0.1	0.90	0.02	1.00	0.73	-0.00	1.00
		0.3	0.74	0.31	0.92	0.74	-0.00	1.01
	SMOTE	0.01	0.99	0.00	1.00	0.72	0.02	0.81
		0.1	0.90	0.02	1.00	0.73	-0.00	0.96
		0.3	0.74	0.31	0.92	0.74	-0.00	1.00
RF	Unadjusted	0.01	0.81	0.32	0.82	0.60	-0.00	0.25
		0.1	0.67	0.57	0.68	0.67	-0.02	0.45
		0.3	0.64	0.63	0.65	0.69	-0.01	0.60
	RUS	0.01	0.99	0.00	1.00	0.66	0.05	0.47
		0.1	0.90	0.02	0.99	0.68	0.19	0.57
		0.3	0.72	0.25	0.92	0.69	0.25	0.62
	ROS	0.01	0.99	0.00	1.00	0.60	-4.37e+15	0.02
		0.1	0.90	0.00	1.00	0.66	2.26	0.25
		0.3	0.71	0.23	0.92	0.69	0.37	0.55
	SMOTE	0.01	0.99	0.01	1.00	0.62	-0.66	0.07
		0.1	0.88	0.07	0.98	0.66	-0.18	0.38
		0.3	0.70	0.35	0.85	0.69	-0.07	0.55

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S3 Performance of models on scenarios with $N = 5000$, $R = 6$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.03	0.93
		0.1	0.90	0.03	1.00	0.75	0.01	0.99
		0.3	0.73	0.32	0.91	0.74	0.00	0.98
	RUS	0.01	0.67	0.68	0.67	0.74	-4.69	0.75
		0.1	0.68	0.68	0.68	0.74	-2.15	0.97
		0.3	0.67	0.67	0.67	0.74	-0.83	0.98
	ROS	0.01	0.69	0.67	0.69	0.75	-4.54	0.88
		0.1	0.68	0.68	0.68	0.75	-2.14	0.99
		0.3	0.67	0.67	0.67	0.74	-0.83	0.98
	SMOTE	0.01	0.70	0.66	0.70	0.75	-4.62	0.76
		0.1	0.70	0.66	0.70	0.75	-2.10	0.91
		0.3	0.69	0.61	0.72	0.74	-0.69	0.94
RID	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.03	1.07
		0.1	0.90	0.03	1.00	0.75	0.01	1.02
		0.3	0.73	0.31	0.92	0.74	0.00	0.99
	RUS	0.01	0.67	0.68	0.67	0.74	-4.49	1.16
		0.1	0.68	0.68	0.68	0.74	-2.13	1.02
		0.3	0.67	0.67	0.67	0.74	-0.83	0.99
	ROS	0.01	0.69	0.67	0.69	0.75	-4.54	0.88
		0.1	0.68	0.68	0.68	0.75	-2.14	0.99
		0.3	0.67	0.67	0.67	0.74	-0.83	0.99
	SMOTE	0.01	0.70	0.66	0.70	0.75	-4.61	0.76
		0.1	0.70	0.66	0.70	0.75	-2.10	0.92
		0.3	0.69	0.61	0.72	0.74	-0.69	0.95
RF	Unadjusted	0.01	0.99	0.00	1.00	0.64	-0.08	0.36
		0.1	0.89	0.03	1.00	0.70	-0.07	0.73
		0.3	0.72	0.30	0.90	0.71	-0.04	0.85
	RUS	0.01	0.64	0.64	0.64	0.69	-4.63	0.74
		0.1	0.65	0.65	0.65	0.71	-2.18	0.84
		0.3	0.65	0.65	0.65	0.71	-0.85	0.87
	ROS	0.01	0.99	0.00	1.00	0.64	-0.38	0.31
		0.1	0.89	0.07	0.98	0.70	-0.42	0.65
		0.3	0.71	0.41	0.84	0.70	-0.28	0.82
	SMOTE	0.01	0.97	0.06	0.98	0.66	-2.52	0.30
		0.1	0.82	0.29	0.88	0.69	-1.31	0.59
		0.3	0.69	0.48	0.79	0.70	-0.45	0.79

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S4 Performance of re-calibrated models on scenarios with $N = 5000$, $R = 6$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.68	0.69	0.68	0.75	0.03	0.93
		0.1	0.67	0.70	0.66	0.75	0.01	0.99
		0.3	0.67	0.68	0.67	0.74	0.00	0.98
	RUS	0.01	0.99	0.00	1.00	0.74	0.03	0.75
		0.1	0.90	0.03	1.00	0.74	0.01	0.97
		0.3	0.73	0.32	0.91	0.74	0.00	0.98
	ROS	0.01	0.99	0.00	1.00	0.75	0.03	0.88
		0.1	0.90	0.03	1.00	0.75	0.01	0.99
		0.3	0.73	0.32	0.91	0.74	0.00	0.98
	SMOTE	0.01	0.99	0.00	1.00	0.75	0.03	0.76
		0.1	0.90	0.05	1.00	0.75	0.01	0.91
		0.3	0.73	0.33	0.91	0.74	0.00	0.94
RID	Unadjusted	0.01	0.65	0.72	0.65	0.75	0.03	1.07
		0.1	0.66	0.70	0.66	0.75	0.01	1.02
		0.3	0.67	0.68	0.67	0.74	0.00	0.99
	RUS	0.01	0.99	0.00	1.00	0.74	0.02	1.16
		0.1	0.90	0.03	1.00	0.74	0.01	1.02
		0.3	0.73	0.31	0.92	0.74	0.00	0.99
	ROS	0.01	0.99	0.00	1.00	0.75	0.03	0.88
		0.1	0.90	0.03	1.00	0.75	0.01	0.99
		0.3	0.73	0.31	0.92	0.74	0.00	0.99
	SMOTE	0.01	0.99	0.00	1.00	0.75	0.03	0.76
		0.1	0.90	0.05	1.00	0.75	0.01	0.92
		0.3	0.73	0.32	0.91	0.74	0.00	0.95
RF	Unadjusted	0.01	0.74	0.46	0.74	0.64	-0.08	0.36
		0.1	0.63	0.66	0.63	0.70	-0.07	0.73
		0.3	0.64	0.67	0.63	0.71	-0.04	0.85
	RUS	0.01	0.99	0.00	1.00	0.69	0.06	0.74
		0.1	0.89	0.01	1.00	0.71	0.22	0.84
		0.3	0.72	0.18	0.95	0.71	0.26	0.87
	ROS	0.01	0.99	0.00	1.00	0.64	-4.35e+15	0.03
		0.1	0.89	0.00	1.00	0.70	2.22	0.65
		0.3	0.71	0.15	0.96	0.70	0.38	0.82
	SMOTE	0.01	0.99	0.00	1.00	0.66	0.18	0.13
		0.1	0.89	0.02	0.99	0.69	0.11	0.59
		0.3	0.71	0.30	0.89	0.70	-0.02	0.79

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S5 Performance of models on scenarios with $N = 5000$, $R = 12$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.01	0.86
		0.1	0.89	0.05	1.00	0.75	-0.00	0.97
		0.3	0.74	0.35	0.91	0.75	0.00	0.99
	RUS	0.01	0.66	0.67	0.66	0.73	-4.93	0.57
		0.1	0.68	0.69	0.68	0.75	-2.14	0.93
		0.3	0.67	0.67	0.67	0.74	-0.83	0.98
	ROS	0.01	0.71	0.65	0.71	0.75	-4.54	0.77
		0.1	0.69	0.69	0.69	0.75	-2.12	0.96
		0.3	0.68	0.68	0.68	0.75	-0.82	0.98
	SMOTE	0.01	0.72	0.64	0.72	0.75	-4.65	0.65
		0.1	0.70	0.66	0.71	0.75	-2.09	0.86
		0.3	0.70	0.63	0.72	0.75	-0.70	0.93
RID	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.01	1.07
		0.1	0.89	0.04	1.00	0.75	-0.00	1.01
		0.3	0.74	0.34	0.91	0.75	0.00	1.01
	RUS	0.01	0.66	0.67	0.66	0.74	-4.45	1.15
		0.1	0.68	0.69	0.68	0.75	-2.11	1.02
		0.3	0.68	0.68	0.68	0.75	-0.82	1.01
	ROS	0.01	0.71	0.65	0.71	0.75	-4.53	0.78
		0.1	0.69	0.69	0.69	0.75	-2.12	0.97
		0.3	0.68	0.68	0.68	0.75	-0.82	1.00
	SMOTE	0.01	0.72	0.64	0.72	0.75	-4.64	0.66
		0.1	0.70	0.66	0.71	0.75	-2.09	0.87
		0.3	0.70	0.63	0.72	0.75	-0.69	0.94
RF	Unadjusted	0.01	0.99	0.00	1.00	0.65	-0.16	0.41
		0.1	0.89	0.01	1.00	0.70	-0.12	0.90
		0.3	0.72	0.23	0.94	0.72	-0.04	1.21
	RUS	0.01	0.63	0.64	0.63	0.69	-4.48	1.02
		0.1	0.66	0.66	0.66	0.72	-2.09	1.20
		0.3	0.66	0.66	0.66	0.72	-0.80	1.27
	ROS	0.01	0.99	0.00	1.00	0.65	-0.41	0.38
		0.1	0.89	0.02	1.00	0.71	-0.42	0.97
		0.3	0.72	0.34	0.89	0.72	-0.27	1.22
	SMOTE	0.01	0.98	0.03	0.99	0.66	-2.16	0.40
		0.1	0.84	0.24	0.92	0.70	-1.25	0.76
		0.3	0.71	0.44	0.83	0.71	-0.44	1.10

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S6 Performance of re-calibrated models on scenarios with $N = 5000$, $R = 12$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.68	0.69	0.68	0.75	0.01	0.86
		0.1	0.66	0.71	0.66	0.75	-0.00	0.97
		0.3	0.68	0.69	0.67	0.75	0.00	0.99
	RUS	0.01	0.99	0.00	1.00	0.73	0.01	0.57
		0.1	0.89	0.05	1.00	0.75	-0.00	0.93
		0.3	0.74	0.35	0.91	0.75	0.00	0.98
	ROS	0.01	0.99	0.00	1.00	0.75	0.01	0.77
		0.1	0.89	0.05	1.00	0.75	-0.00	0.96
		0.3	0.74	0.35	0.91	0.75	0.00	0.98
	SMOTE	0.01	0.99	0.00	1.00	0.75	0.01	0.65
		0.1	0.89	0.07	0.99	0.75	-0.00	0.86
		0.3	0.74	0.36	0.90	0.75	0.00	0.93
RID	Unadjusted	0.01	0.64	0.73	0.63	0.75	0.01	1.07
		0.1	0.66	0.72	0.65	0.75	-0.00	1.01
		0.3	0.68	0.69	0.67	0.75	0.00	1.01
	RUS	0.01	0.99	0.00	1.00	0.74	0.01	1.15
		0.1	0.89	0.03	1.00	0.75	-0.00	1.02
		0.3	0.74	0.34	0.91	0.75	0.00	1.01
	ROS	0.01	0.99	0.00	1.00	0.75	0.01	0.78
		0.1	0.89	0.04	1.00	0.75	-0.00	0.97
		0.3	0.74	0.34	0.91	0.75	0.00	1.00
	SMOTE	0.01	0.99	0.00	1.00	0.75	0.01	0.66
		0.1	0.89	0.07	0.99	0.75	-0.00	0.87
		0.3	0.74	0.36	0.90	0.75	0.00	0.94
RF	Unadjusted	0.01	0.68	0.54	0.68	0.65	-0.16	0.41
		0.1	0.58	0.72	0.56	0.70	-0.12	0.90
		0.3	0.63	0.71	0.60	0.72	-0.04	1.21
	RUS	0.01	0.99	0.00	1.00	0.69	0.05	1.02
		0.1	0.89	0.00	1.00	0.72	0.21	1.20
		0.3	0.71	0.10	0.98	0.72	0.27	1.27
	ROS	0.01	0.99	0.00	1.00	0.65	-4.35e+15	0.04
		0.1	0.89	0.00	1.00	0.71	2.15	0.97
		0.3	0.71	0.08	0.98	0.72	0.38	1.22
	SMOTE	0.01	0.99	0.00	1.00	0.66	0.33	0.36
		0.1	0.89	0.01	1.00	0.70	0.11	0.76
		0.3	0.72	0.23	0.93	0.71	-0.01	1.10

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S7 Performance of models on scenarios with $N = 5000$, $R = 24$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.71	-0.03	0.65
		0.1	0.90	0.03	1.00	0.74	0.00	0.93
		0.3	0.73	0.33	0.91	0.74	-0.01	0.97
	RUS	0.01	0.62	0.62	0.62	0.67	-6.10	0.26
		0.1	0.67	0.67	0.67	0.73	-2.20	0.86
		0.3	0.67	0.68	0.67	0.74	-0.85	0.96
	ROS	0.01	0.72	0.57	0.72	0.70	-4.72	0.53
		0.1	0.68	0.67	0.68	0.74	-2.15	0.91
		0.3	0.67	0.68	0.67	0.74	-0.84	0.97
	SMOTE	0.01	0.73	0.55	0.74	0.70	-4.85	0.44
		0.1	0.70	0.64	0.70	0.74	-2.13	0.79
		0.3	0.69	0.62	0.72	0.74	-0.71	0.90
RID	Unadjusted	0.01	0.99	0.00	1.00	0.71	-0.02	1.04
		0.1	0.90	0.02	1.00	0.74	0.00	1.00
		0.3	0.74	0.31	0.92	0.74	-0.01	1.02
	RUS	0.01	0.63	0.63	0.63	0.68	-4.56	1.11
		0.1	0.67	0.67	0.67	0.73	-2.14	1.03
		0.3	0.67	0.68	0.67	0.74	-0.84	1.02
	ROS	0.01	0.72	0.57	0.72	0.70	-4.71	0.53
		0.1	0.68	0.67	0.68	0.74	-2.15	0.93
		0.3	0.67	0.68	0.67	0.74	-0.84	0.99
	SMOTE	0.01	0.73	0.55	0.74	0.70	-4.83	0.45
		0.1	0.70	0.64	0.70	0.74	-2.12	0.81
		0.3	0.69	0.62	0.72	0.74	-0.70	0.93
RF	Unadjusted	0.01	0.99	0.00	1.00	0.60	-0.24	0.29
		0.1	0.90	0.00	1.00	0.66	-0.13	0.92
		0.3	0.71	0.09	0.98	0.70	-0.06	1.62
	RUS	0.01	0.61	0.60	0.61	0.65	-4.55	1.12
		0.1	0.65	0.64	0.65	0.70	-2.10	1.61
		0.3	0.65	0.65	0.65	0.71	-0.81	1.80
	ROS	0.01	0.99	0.00	1.00	0.62	-0.45	0.35
		0.1	0.90	0.00	1.00	0.69	-0.41	1.27
		0.3	0.72	0.19	0.95	0.71	-0.28	1.75
	SMOTE	0.01	0.99	0.01	1.00	0.63	-2.03	0.43
		0.1	0.87	0.13	0.95	0.68	-1.21	0.87
		0.3	0.71	0.32	0.88	0.70	-0.44	1.44

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S8 Performance of re-calibrated models on scenarios with $N = 5000$, $R = 24$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.71	0.60	0.71	0.71	-0.03	0.65
		0.1	0.67	0.68	0.66	0.74	0.00	0.93
		0.3	0.67	0.68	0.67	0.74	-0.01	0.97
	RUS	0.01	0.99	0.01	1.00	0.67	-0.05	0.27
		0.1	0.90	0.04	1.00	0.73	-0.00	0.86
		0.3	0.73	0.33	0.91	0.74	-0.01	0.96
	ROS	0.01	0.99	0.00	1.00	0.70	-0.03	0.53
		0.1	0.90	0.03	1.00	0.74	0.00	0.91
		0.3	0.73	0.33	0.91	0.74	-0.01	0.97
	SMOTE	0.01	0.99	0.00	1.00	0.70	-0.04	0.44
		0.1	0.90	0.06	0.99	0.74	-0.00	0.79
		0.3	0.73	0.34	0.90	0.74	-0.01	0.90
RID	Unadjusted	0.01	0.63	0.68	0.63	0.71	-0.02	1.04
		0.1	0.66	0.69	0.65	0.74	0.00	1.00
		0.3	0.67	0.68	0.66	0.74	-0.01	1.02
	RUS	0.01	0.99	0.00	1.00	0.68	-0.02	1.11
		0.1	0.90	0.02	1.00	0.73	0.00	1.03
		0.3	0.73	0.31	0.92	0.74	-0.01	1.02
	ROS	0.01	0.99	0.00	1.00	0.70	-0.03	0.53
		0.1	0.90	0.03	1.00	0.74	0.00	0.93
		0.3	0.73	0.32	0.91	0.74	-0.01	0.99
	SMOTE	0.01	0.99	0.00	1.00	0.70	-0.04	0.45
		0.1	0.90	0.05	0.99	0.74	-0.00	0.81
		0.3	0.73	0.34	0.91	0.74	-0.01	0.93
RF	Unadjusted	0.01	0.64	0.50	0.65	0.60	-0.24	0.29
		0.1	0.52	0.72	0.50	0.66	-0.13	0.92
		0.3	0.61	0.72	0.56	0.70	-0.06	1.62
	RUS	0.01	0.99	0.00	1.00	0.65	0.01	1.12
		0.1	0.90	0.00	1.00	0.70	0.20	1.61
		0.3	0.70	0.02	1.00	0.71	0.26	1.80
	ROS	0.01	0.99	0.00	1.00	0.62	-0.30	0.04
		0.1	0.90	0.00	1.00	0.69	2.22	1.27
		0.3	0.70	0.01	1.00	0.71	0.38	1.75
	SMOTE	0.01	0.99	0.00	1.00	0.63	0.33	0.42
		0.1	0.89	0.00	1.00	0.68	0.10	0.87
		0.3	0.71	0.11	0.97	0.70	-0.03	1.44

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S9 Performance of models on scenarios with $N = 2500$, $R = 3$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.72	-0.04	0.89
		0.1	0.90	0.02	1.00	0.73	0.01	1.00
		0.3	0.73	0.31	0.92	0.74	0.01	0.99
	RUS	0.01	0.65	0.64	0.65	0.71	-4.84	0.71
		0.1	0.66	0.67	0.66	0.73	-2.20	0.98
		0.3	0.67	0.67	0.67	0.74	-0.84	0.99
	ROS	0.01	0.67	0.63	0.67	0.72	-4.71	0.83
		0.1	0.66	0.67	0.66	0.73	-2.19	0.99
		0.3	0.67	0.67	0.67	0.74	-0.84	0.99
	SMOTE	0.01	0.68	0.62	0.68	0.72	-4.78	0.71
		0.1	0.68	0.64	0.69	0.73	-2.15	0.95
		0.3	0.69	0.61	0.73	0.74	-0.69	0.98
RID	Unadjusted	0.01	0.99	0.00	1.00	0.72	-0.04	1.09
		0.1	0.90	0.01	1.00	0.73	0.01	1.03
		0.3	0.73	0.31	0.92	0.74	0.01	1.01
	RUS	0.01	0.65	0.64	0.65	0.71	-4.67	1.23
		0.1	0.66	0.67	0.66	0.73	-2.18	1.05
		0.3	0.67	0.67	0.67	0.74	-0.84	1.01
	ROS	0.01	0.67	0.63	0.67	0.72	-4.70	0.84
		0.1	0.66	0.67	0.66	0.73	-2.19	1.00
		0.3	0.67	0.67	0.67	0.74	-0.84	1.00
	SMOTE	0.01	0.68	0.62	0.68	0.72	-4.77	0.72
		0.1	0.68	0.64	0.69	0.73	-2.14	0.95
		0.3	0.69	0.61	0.73	0.74	-0.69	0.99
RF	Unadjusted	0.01	0.99	0.00	1.00	0.59	-0.06	0.23
		0.1	0.90	0.04	0.99	0.66	-0.01	0.43
		0.3	0.71	0.34	0.87	0.69	-0.01	0.59
	RUS	0.01	0.61	0.60	0.61	0.64	-5.01	0.45
		0.1	0.63	0.63	0.63	0.67	-2.39	0.55
		0.3	0.64	0.64	0.64	0.69	-0.91	0.60
	ROS	0.01	0.99	0.00	1.00	0.59	-0.46	0.20
		0.1	0.88	0.10	0.96	0.66	-0.46	0.36
		0.3	0.69	0.43	0.80	0.69	-0.30	0.54
	SMOTE	0.01	0.90	0.14	0.91	0.61	-3.90	0.16
		0.1	0.78	0.34	0.83	0.65	-1.63	0.36
		0.3	0.68	0.50	0.75	0.68	-0.49	0.54

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S10 Performance of re-calibrated models on scenarios with $N = 2500$, $R = 3$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.70	0.60	0.70	0.72	-0.04	0.89
		0.1	0.67	0.67	0.67	0.73	0.01	1.00
		0.3	0.67	0.68	0.67	0.74	0.01	0.99
	RUS	0.01	0.99	0.00	1.00	0.71	-0.04	0.71
		0.1	0.90	0.02	1.00	0.73	0.01	0.98
		0.3	0.73	0.31	0.92	0.74	0.00	0.99
	ROS	0.01	0.99	0.00	1.00	0.72	-0.04	0.83
		0.1	0.90	0.02	1.00	0.73	0.01	0.99
		0.3	0.73	0.31	0.92	0.74	0.01	0.99
	SMOTE	0.01	0.99	0.00	1.00	0.72	-0.04	0.71
		0.1	0.90	0.02	1.00	0.73	0.01	0.95
		0.3	0.73	0.32	0.91	0.74	0.01	0.98
RID	Unadjusted	0.01	0.68	0.62	0.68	0.72	-0.04	1.09
		0.1	0.66	0.67	0.66	0.73	0.01	1.03
		0.3	0.67	0.68	0.67	0.74	0.01	1.01
	RUS	0.01	0.99	0.00	1.00	0.71	-0.04	1.23
		0.1	0.90	0.01	1.00	0.73	0.01	1.05
		0.3	0.73	0.31	0.92	0.74	0.01	1.01
	ROS	0.01	0.99	0.00	1.00	0.72	-0.04	0.84
		0.1	0.90	0.02	1.00	0.73	0.01	1.00
		0.3	0.73	0.31	0.92	0.74	0.01	1.00
	SMOTE	0.01	0.99	0.00	1.00	0.72	-0.04	0.72
		0.1	0.90	0.02	1.00	0.73	0.01	0.95
		0.3	0.73	0.32	0.92	0.74	0.01	0.99
RF	Unadjusted	0.01	0.82	0.30	0.83	0.59	-0.06	0.23
		0.1	0.67	0.56	0.68	0.66	-0.01	0.43
		0.3	0.64	0.63	0.65	0.69	-0.01	0.59
	RUS	0.01	0.99	0.00	1.00	0.64	-0.01	0.45
		0.1	0.90	0.02	0.99	0.67	0.19	0.55
		0.3	0.71	0.25	0.91	0.69	0.25	0.60
	ROS	0.01	0.99	0.00	1.00	0.59	-4.37e+15	0.02
		0.1	0.90	0.00	1.00	0.66	2.27	0.22
		0.3	0.71	0.23	0.92	0.69	0.37	0.53
	SMOTE	0.01	0.99	0.01	1.00	0.61	-0.88	0.05
		0.1	0.88	0.07	0.97	0.65	-0.17	0.35
		0.3	0.70	0.35	0.85	0.68	-0.06	0.54

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S11 Performance of models on scenarios with $N = 2500$, $R = 6$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.75	-0.01	0.85
		0.1	0.90	0.03	1.00	0.74	-0.00	0.96
		0.3	0.74	0.32	0.92	0.74	-0.00	0.99
	RUS	0.01	0.65	0.66	0.65	0.72	-5.01	0.54
		0.1	0.68	0.68	0.67	0.74	-2.17	0.93
		0.3	0.67	0.67	0.67	0.74	-0.84	0.99
	ROS	0.01	0.70	0.65	0.70	0.74	-4.61	0.75
		0.1	0.68	0.68	0.68	0.74	-2.15	0.95
		0.3	0.67	0.67	0.67	0.74	-0.84	0.99
	SMOTE	0.01	0.71	0.63	0.71	0.74	-4.71	0.63
		0.1	0.70	0.65	0.70	0.74	-2.12	0.87
		0.3	0.69	0.62	0.73	0.74	-0.70	0.95
RID	Unadjusted	0.01	0.99	0.00	1.00	0.75	-0.01	1.11
		0.1	0.90	0.03	1.00	0.74	-0.00	1.01
		0.3	0.74	0.32	0.92	0.74	-0.00	1.02
	RUS	0.01	0.65	0.67	0.65	0.73	-4.54	1.22
		0.1	0.68	0.68	0.67	0.74	-2.14	1.03
		0.3	0.67	0.67	0.67	0.74	-0.84	1.02
	ROS	0.01	0.70	0.65	0.70	0.74	-4.61	0.76
		0.1	0.68	0.68	0.68	0.74	-2.15	0.97
		0.3	0.67	0.67	0.67	0.74	-0.84	1.00
	SMOTE	0.01	0.71	0.63	0.71	0.74	-4.70	0.64
		0.1	0.70	0.65	0.70	0.74	-2.11	0.88
		0.3	0.69	0.62	0.73	0.74	-0.70	0.97
RF	Unadjusted	0.01	0.99	0.00	1.00	0.63	-0.12	0.33
		0.1	0.90	0.03	1.00	0.69	-0.09	0.67
		0.3	0.72	0.30	0.90	0.70	-0.04	0.85
	RUS	0.01	0.62	0.63	0.62	0.67	-4.69	0.68
		0.1	0.65	0.64	0.65	0.70	-2.20	0.80
		0.3	0.65	0.65	0.65	0.70	-0.86	0.86
	ROS	0.01	0.99	0.00	1.00	0.63	-0.43	0.27
		0.1	0.89	0.07	0.98	0.69	-0.44	0.60
		0.3	0.71	0.40	0.84	0.70	-0.29	0.81
	SMOTE	0.01	0.96	0.06	0.97	0.64	-2.55	0.25
		0.1	0.82	0.29	0.88	0.68	-1.32	0.55
		0.3	0.69	0.48	0.79	0.70	-0.45	0.78

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S12 Performance of re-calibrated models on scenarios with $N = 2500$, $R = 6$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.69	0.66	0.69	0.75	-0.01	0.85
		0.1	0.67	0.70	0.66	0.74	-0.00	0.96
		0.3	0.67	0.68	0.67	0.74	-0.00	0.99
	RUS	0.01	0.99	0.00	1.00	0.72	-0.01	0.54
		0.1	0.90	0.03	1.00	0.74	-0.00	0.93
		0.3	0.74	0.32	0.91	0.74	-0.00	0.99
	ROS	0.01	0.99	0.00	1.00	0.74	-0.01	0.75
		0.1	0.90	0.03	1.00	0.74	-0.00	0.95
		0.3	0.74	0.32	0.92	0.74	-0.00	0.99
	SMOTE	0.01	0.99	0.00	1.00	0.74	-0.02	0.63
		0.1	0.90	0.05	0.99	0.74	-0.00	0.87
		0.3	0.74	0.33	0.91	0.74	-0.00	0.95
RID	Unadjusted	0.01	0.65	0.71	0.65	0.75	-0.01	1.11
		0.1	0.66	0.70	0.66	0.74	-0.00	1.01
		0.3	0.67	0.68	0.67	0.74	-0.00	1.02
	RUS	0.01	0.99	0.00	1.00	0.73	-0.00	1.22
		0.1	0.90	0.02	1.00	0.74	-0.00	1.03
		0.3	0.74	0.31	0.92	0.74	-0.00	1.02
	ROS	0.01	0.99	0.00	1.00	0.74	-0.01	0.76
		0.1	0.90	0.03	1.00	0.74	-0.00	0.97
		0.3	0.74	0.32	0.92	0.74	-0.00	1.00
	SMOTE	0.01	0.99	0.00	1.00	0.74	-0.02	0.64
		0.1	0.90	0.04	1.00	0.74	-0.00	0.88
		0.3	0.74	0.33	0.91	0.74	-0.00	0.97
RF	Unadjusted	0.01	0.75	0.44	0.75	0.63	-0.12	0.33
		0.1	0.63	0.65	0.63	0.69	-0.09	0.67
		0.3	0.64	0.67	0.63	0.70	-0.04	0.85
	RUS	0.01	0.99	0.00	1.00	0.67	0.02	0.68
		0.1	0.90	0.01	1.00	0.70	0.20	0.80
		0.3	0.72	0.18	0.95	0.70	0.25	0.86
	ROS	0.01	0.99	0.00	1.00	0.63	-4.36e+15	0.03
		0.1	0.90	0.00	1.00	0.69	2.20	0.59
		0.3	0.72	0.15	0.96	0.70	0.37	0.81
	SMOTE	0.01	0.99	0.00	1.00	0.64	-0.00	0.07
		0.1	0.89	0.02	0.99	0.68	0.09	0.55
		0.3	0.72	0.30	0.89	0.70	-0.02	0.78

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S13 Performance of models on scenarios with $N = 2500$, $R = 12$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.02	0.74
		0.1	0.89	0.05	1.00	0.75	0.00	0.93
		0.3	0.74	0.35	0.91	0.75	-0.00	0.97
	RUS	0.01	0.63	0.65	0.63	0.70	-5.88	0.32
		0.1	0.68	0.68	0.68	0.75	-2.16	0.87
		0.3	0.68	0.68	0.67	0.74	-0.83	0.95
	ROS	0.01	0.72	0.62	0.73	0.74	-4.62	0.60
		0.1	0.69	0.68	0.69	0.75	-2.12	0.92
		0.3	0.68	0.68	0.68	0.75	-0.83	0.96
	SMOTE	0.01	0.74	0.60	0.74	0.74	-4.76	0.50
		0.1	0.70	0.66	0.71	0.75	-2.09	0.81
		0.3	0.69	0.63	0.72	0.75	-0.70	0.90
RID	Unadjusted	0.01	0.99	0.00	1.00	0.75	0.02	1.11
		0.1	0.89	0.03	1.00	0.75	0.00	1.01
		0.3	0.74	0.34	0.91	0.75	-0.00	1.01
	RUS	0.01	0.64	0.67	0.64	0.72	-4.45	1.23
		0.1	0.68	0.68	0.68	0.75	-2.10	1.04
		0.3	0.68	0.68	0.67	0.74	-0.82	1.01
	ROS	0.01	0.72	0.62	0.73	0.74	-4.60	0.61
		0.1	0.69	0.68	0.69	0.75	-2.11	0.94
		0.3	0.68	0.68	0.68	0.75	-0.82	0.98
	SMOTE	0.01	0.74	0.60	0.74	0.74	-4.73	0.51
		0.1	0.70	0.65	0.71	0.75	-2.09	0.83
		0.3	0.70	0.63	0.72	0.75	-0.70	0.93
RF	Unadjusted	0.01	0.99	0.00	1.00	0.63	-0.15	0.34
		0.1	0.89	0.01	1.00	0.69	-0.11	0.82
		0.3	0.72	0.22	0.94	0.71	-0.05	1.17
	RUS	0.01	0.62	0.63	0.62	0.67	-4.49	0.94
		0.1	0.65	0.65	0.65	0.71	-2.09	1.14
		0.3	0.66	0.66	0.66	0.71	-0.81	1.22
	ROS	0.01	0.99	0.00	1.00	0.64	-0.41	0.33
		0.1	0.89	0.03	0.99	0.70	-0.42	0.89
		0.3	0.72	0.33	0.89	0.71	-0.27	1.17
	SMOTE	0.01	0.98	0.03	0.99	0.65	-2.06	0.35
		0.1	0.84	0.22	0.92	0.69	-1.20	0.72
		0.3	0.70	0.42	0.83	0.71	-0.43	1.06

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S14 Performance of re-calibrated models on scenarios with $N = 2500$, $R = 12$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.70	0.67	0.70	0.75	0.02	0.74
		0.1	0.66	0.71	0.66	0.75	0.00	0.93
		0.3	0.67	0.69	0.67	0.75	-0.00	0.97
	RUS	0.01	0.99	0.01	1.00	0.70	0.01	0.32
		0.1	0.89	0.05	0.99	0.75	0.00	0.87
		0.3	0.74	0.35	0.91	0.74	-0.00	0.95
	ROS	0.01	0.99	0.00	1.00	0.74	0.02	0.60
		0.1	0.89	0.05	0.99	0.75	0.00	0.92
		0.3	0.74	0.35	0.91	0.75	-0.00	0.96
	SMOTE	0.01	0.99	0.01	1.00	0.74	0.01	0.50
		0.1	0.89	0.07	0.99	0.75	0.00	0.81
		0.3	0.74	0.37	0.90	0.75	-0.00	0.90
RID	Unadjusted	0.01	0.62	0.75	0.62	0.75	0.02	1.11
		0.1	0.66	0.72	0.65	0.75	0.00	1.01
		0.3	0.67	0.69	0.67	0.75	-0.00	1.01
	RUS	0.01	0.99	0.00	1.00	0.72	0.02	1.23
		0.1	0.89	0.03	1.00	0.75	0.00	1.04
		0.3	0.74	0.33	0.91	0.74	-0.00	1.01
	ROS	0.01	0.99	0.00	1.00	0.74	0.01	0.61
		0.1	0.89	0.04	1.00	0.75	0.00	0.94
		0.3	0.74	0.34	0.91	0.75	-0.00	0.98
	SMOTE	0.01	0.99	0.01	1.00	0.74	0.01	0.51
		0.1	0.89	0.07	0.99	0.75	0.00	0.83
		0.3	0.74	0.36	0.90	0.75	-0.00	0.93
RF	Unadjusted	0.01	0.69	0.50	0.69	0.63	-0.15	0.34
		0.1	0.57	0.71	0.56	0.69	-0.11	0.82
		0.3	0.63	0.71	0.59	0.71	-0.05	1.17
	RUS	0.01	0.99	0.00	1.00	0.67	0.05	0.94
		0.1	0.89	0.00	1.00	0.71	0.21	1.14
		0.71	0.09	0.98	0.71	0.26	1.22	
	ROS	0.01	0.99	0.00	1.00	0.64	-4.34e+15	0.03
		0.1	0.89	0.00	1.00	0.70	2.16	0.89
		0.3	0.71	0.07	0.98	0.71	0.37	1.17
	SMOTE	0.01	0.99	0.00	1.00	0.65	0.27	0.27
		0.1	0.89	0.01	1.00	0.69	0.14	0.72
		0.3	0.71	0.22	0.93	0.71	-0.01	1.06

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S15 Performance of models on scenarios with $N = 2500$, $R = 24$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.99	0.00	1.00	0.69	-0.05	0.48
		0.1	0.90	0.04	1.00	0.73	-0.02	0.85
		0.3	0.73	0.33	0.91	0.74	0.01	0.94
	RUS	0.01	0.58	0.59	0.58	0.62	-11.42	0.08
		0.1	0.66	0.66	0.66	0.72	-2.26	0.73
		0.3	0.67	0.67	0.66	0.73	-0.84	0.91
	ROS	0.01	0.75	0.48	0.75	0.68	-5.04	0.31
		0.1	0.68	0.65	0.68	0.73	-2.17	0.82
		0.3	0.67	0.67	0.67	0.73	-0.83	0.93
	SMOTE	0.01	0.77	0.46	0.77	0.68	-5.15	0.28
		0.1	0.70	0.62	0.70	0.73	-2.15	0.71
		0.3	0.69	0.61	0.72	0.73	-0.69	0.87
RID	Unadjusted	0.01	0.99	0.00	1.00	0.69	-0.02	1.07
		0.1	0.90	0.02	1.00	0.73	-0.02	0.99
		0.3	0.73	0.30	0.92	0.74	0.01	1.03
	RUS	0.01	0.61	0.62	0.61	0.66	-4.56	1.70
		0.1	0.66	0.66	0.66	0.72	-2.15	1.03
		0.3	0.67	0.67	0.66	0.73	-0.83	1.04
	ROS	0.01	0.75	0.48	0.75	0.68	-4.98	0.32
		0.1	0.68	0.65	0.68	0.73	-2.16	0.85
		0.3	0.67	0.67	0.67	0.73	-0.82	0.98
	SMOTE	0.01	0.77	0.46	0.77	0.68	-5.08	0.29
		0.1	0.70	0.62	0.70	0.73	-2.14	0.74
		0.3	0.69	0.61	0.72	0.73	-0.69	0.92
RF	Unadjusted	0.01	0.99	0.00	1.00	0.59	-0.25	0.26
		0.1	0.90	0.00	1.00	0.65	-0.14	0.81
		0.3	0.71	0.08	0.98	0.69	-0.04	1.54
	RUS	0.01	0.59	0.59	0.59	0.63	-4.57	0.99
		0.1	0.63	0.63	0.63	0.68	-2.12	1.45
		0.3	0.64	0.64	0.64	0.70	-0.81	1.70
	ROS	0.01	0.99	0.00	1.00	0.61	-0.48	0.28
		0.1	0.90	0.00	1.00	0.68	-0.43	1.11
		0.3	0.71	0.19	0.94	0.70	-0.27	1.64
	SMOTE	0.01	0.99	0.01	1.00	0.61	-1.87	0.33
		0.1	0.87	0.11	0.96	0.66	-1.18	0.82
		0.3	0.70	0.30	0.88	0.69	-0.42	1.39

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

TABLE S16 Performance of re-calibrated models on scenarios with $N = 2500$, $R = 24$

Model	Adjustment	Prevalence	Accuracy	Sensitivity	Specificity	C-statistic	Calibration intercept	Calibration slope
LR	Unadjusted	0.01	0.74	0.52	0.74	0.69	-0.05	0.48
		0.1	0.66	0.67	0.66	0.73	-0.02	0.85
		0.3	0.67	0.68	0.67	0.74	0.01	0.94
	RUS	0.01	0.98	0.02	0.99	0.62	-0.33	0.09
		0.1	0.90	0.05	0.99	0.72	-0.02	0.73
		0.3	0.73	0.33	0.91	0.73	0.01	0.91
	ROS	0.01	0.99	0.00	1.00	0.68	-0.07	0.31
		0.1	0.90	0.04	0.99	0.73	-0.02	0.82
		0.3	0.73	0.33	0.91	0.73	0.01	0.93
	SMOTE	0.01	0.99	0.01	1.00	0.68	-0.08	0.28
		0.1	0.89	0.07	0.99	0.73	-0.02	0.71
		0.3	0.73	0.35	0.90	0.73	0.01	0.87
RID	Unadjusted	0.01	0.61	0.67	0.61	0.69	-0.02	1.07
		0.1	0.65	0.69	0.64	0.73	-0.02	0.99
		0.3	0.67	0.68	0.66	0.74	0.01	1.03
	RUS	0.01	0.99	0.00	1.00	0.66	-0.16	1.70
		0.1	0.90	0.01	1.00	0.72	-0.02	1.03
		0.3	0.73	0.29	0.92	0.73	0.01	1.04
	ROS	0.01	0.99	0.00	1.00	0.68	-0.06	0.32
		0.1	0.90	0.04	1.00	0.73	-0.02	0.85
		0.3	0.73	0.31	0.91	0.73	0.01	0.98
	SMOTE	0.01	0.99	0.01	1.00	0.68	-0.07	0.29
		0.1	0.90	0.06	0.99	0.73	-0.02	0.74
		0.3	0.73	0.33	0.91	0.73	0.01	0.92
RF	Unadjusted	0.01	0.65	0.48	0.66	0.59	-0.25	0.26
		0.1	0.53	0.70	0.51	0.65	-0.14	0.81
		0.3	0.60	0.72	0.55	0.69	-0.04	1.54
	RUS	0.01	0.99	0.00	1.00	0.63	-0.13	0.99
		0.1	0.90	0.00	1.00	0.68	0.18	1.45
		0.3	0.70	0.02	1.00	0.70	0.27	1.70
	ROS	0.01	0.99	0.00	1.00	0.61	-0.47	0.03
		0.1	0.90	0.00	1.00	0.68	2.20	1.11
		0.3	0.70	0.01	1.00	0.70	0.39	1.64
	SMOTE	0.01	0.99	0.00	1.00	0.61	0.28	0.30
		0.1	0.90	0.00	1.00	0.66	0.12	0.82
		0.3	0.71	0.10	0.97	0.69	-0.01	1.39

Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

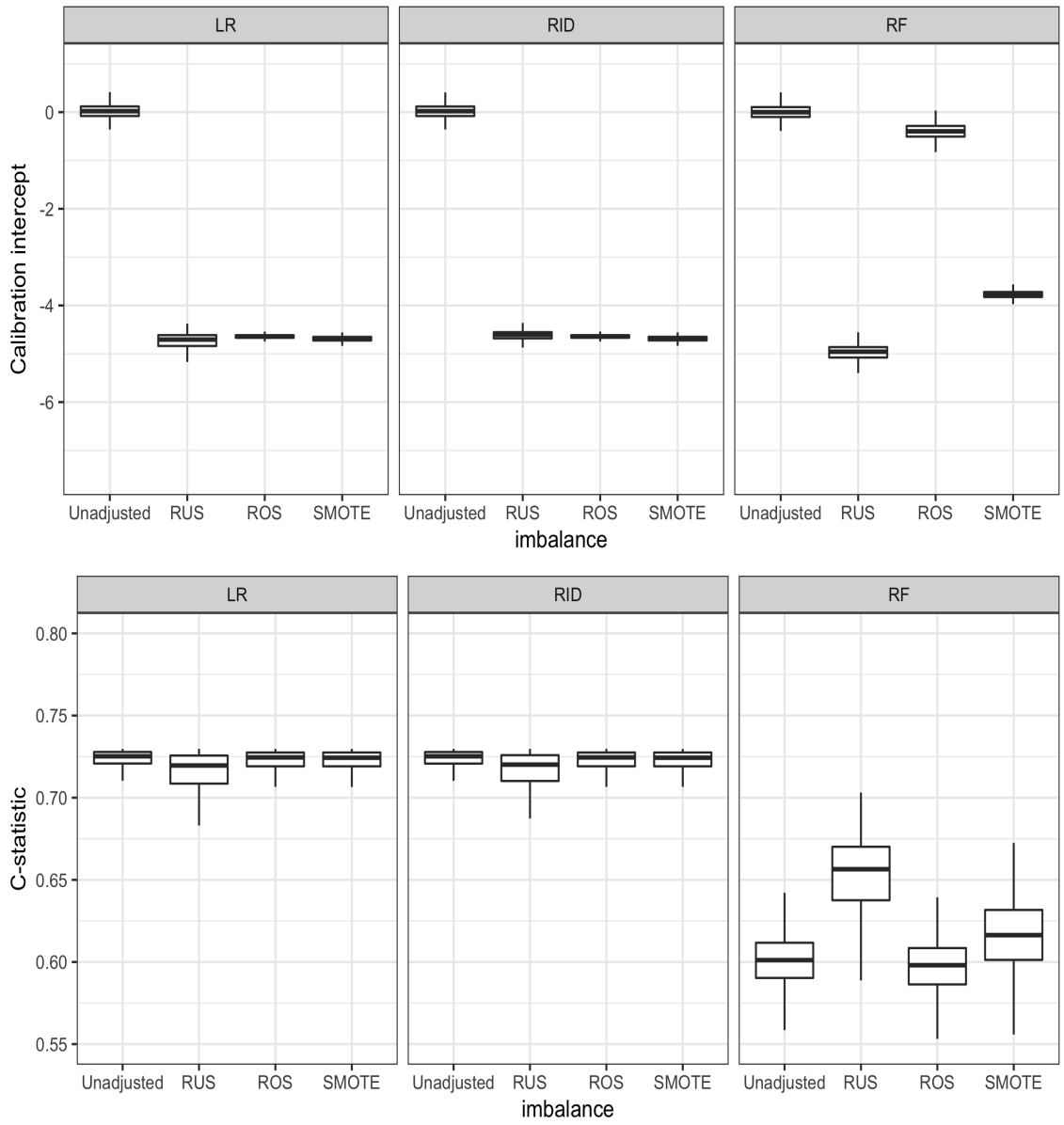

FIGURE S1 Boxplots of the C-statistics and calibration intercepts over 2000 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 3$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

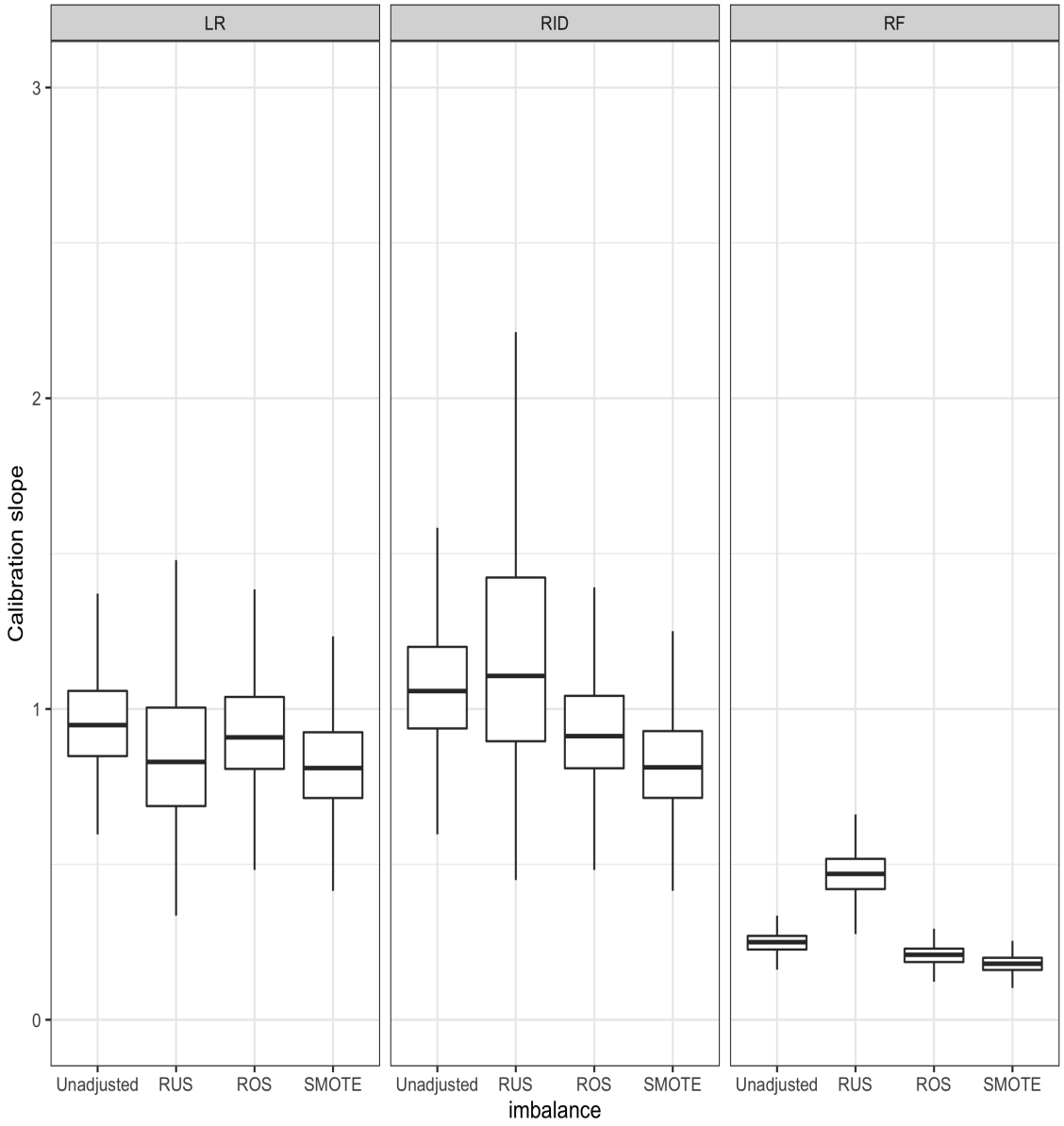

FIGURE S2 Boxplots of the Calibration slopes over 2000 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 3$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

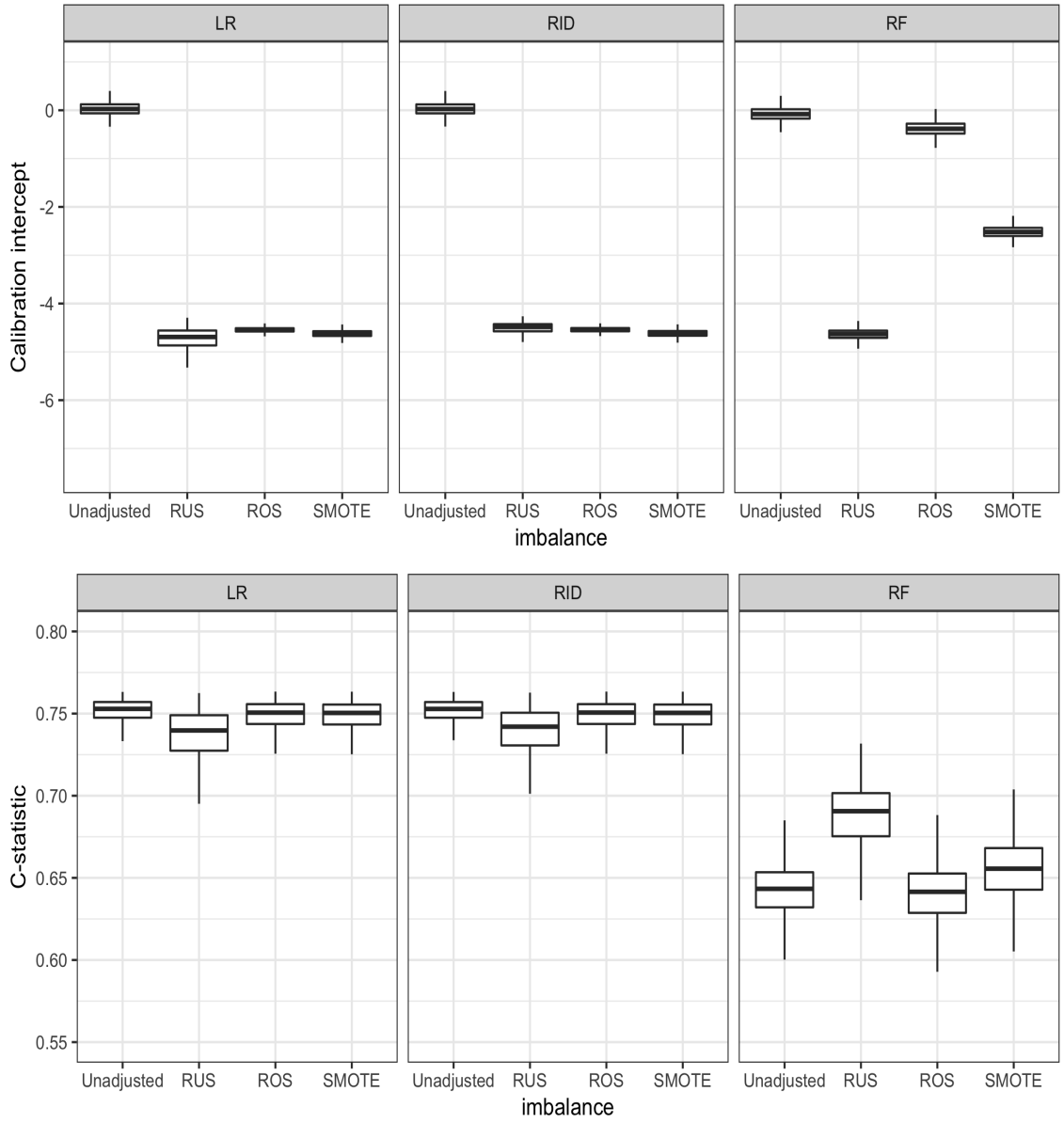

FIGURE S3 Boxplots of the C-statistics and calibration intercepts over 2000 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 6$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

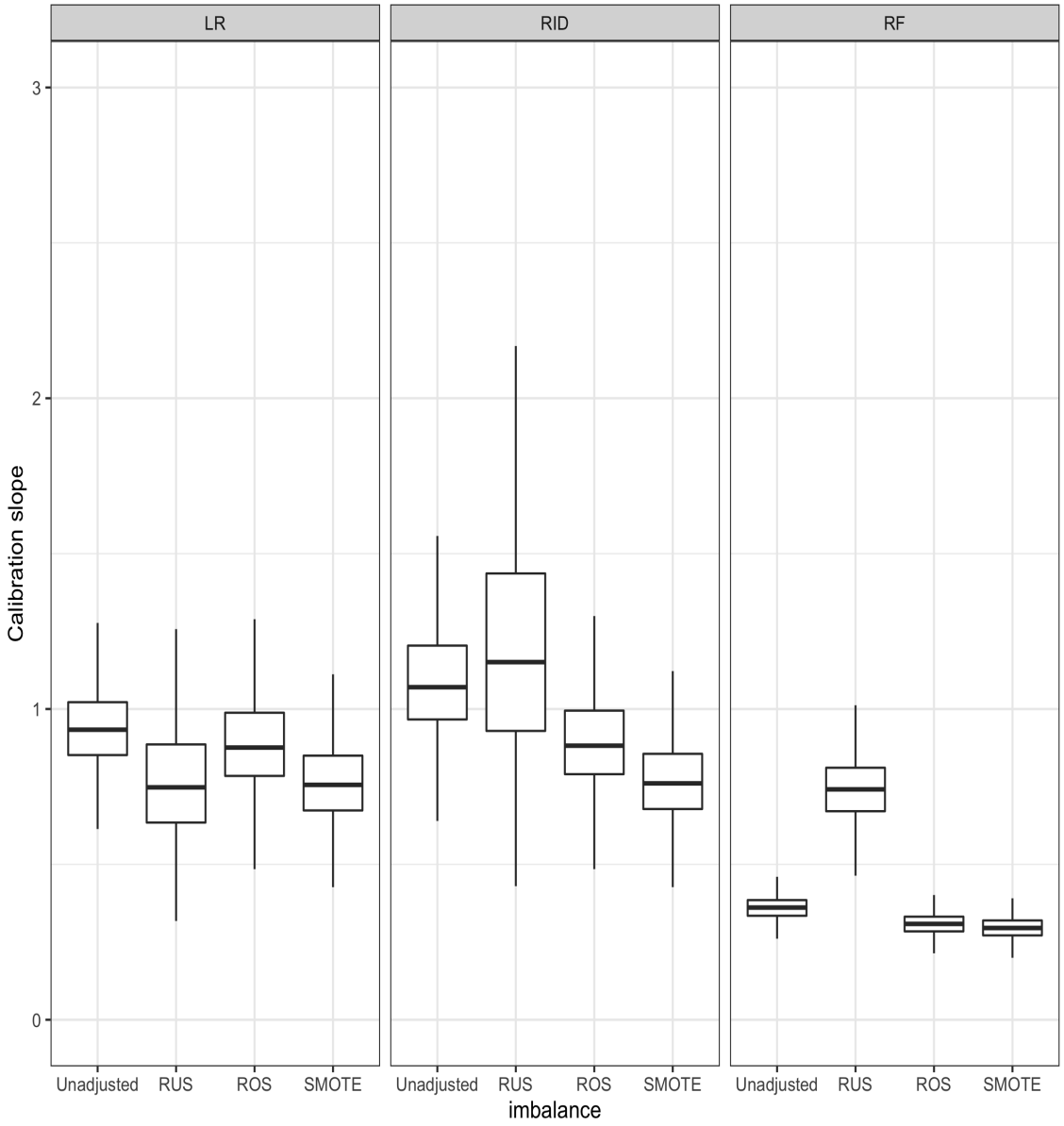

FIGURE S4 Boxplots of the Calibration slopes over 2000 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 6$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

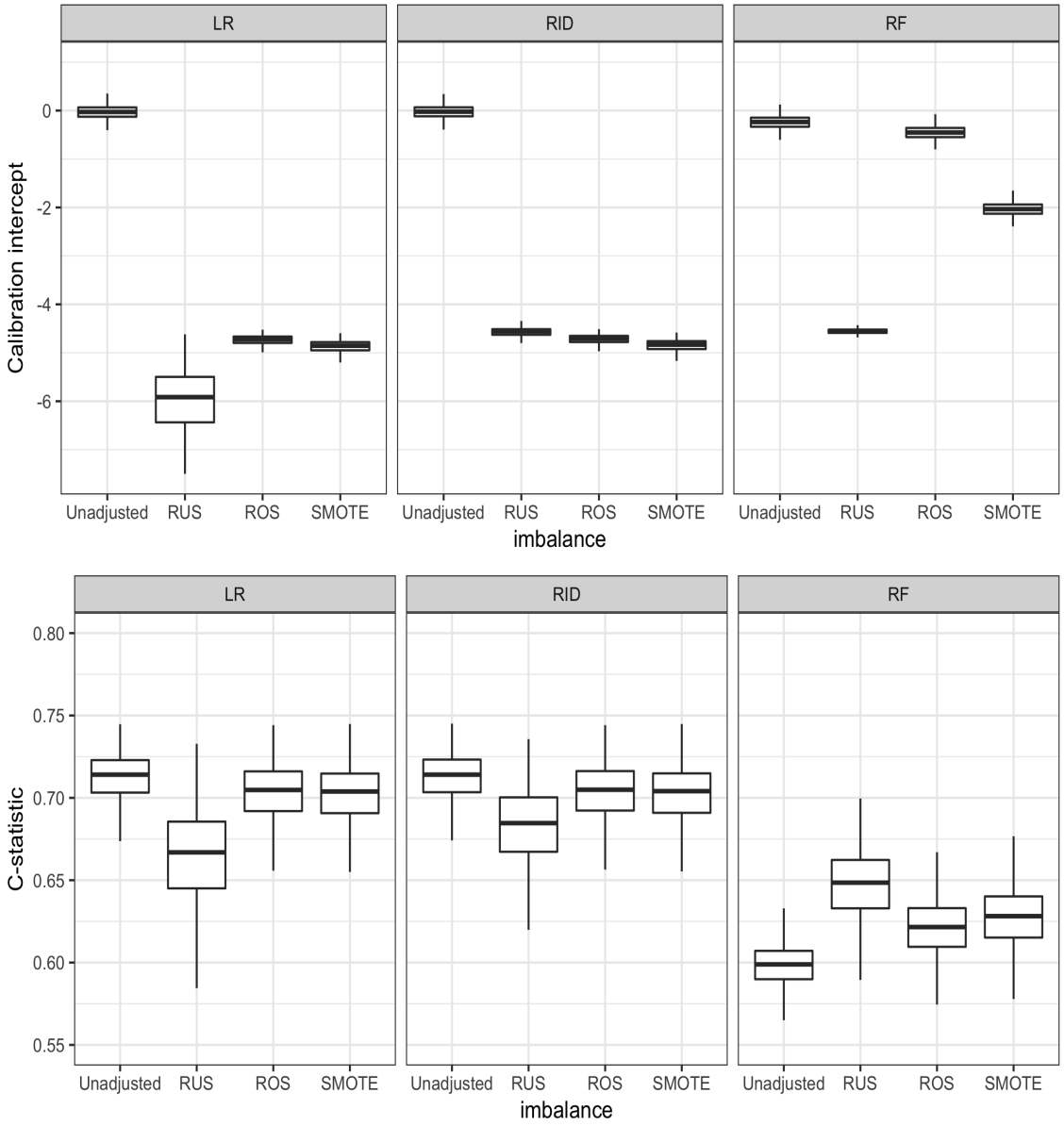

FIGURE S5 Boxplots of the C-statistics and calibration intercepts over 1991 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 24$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

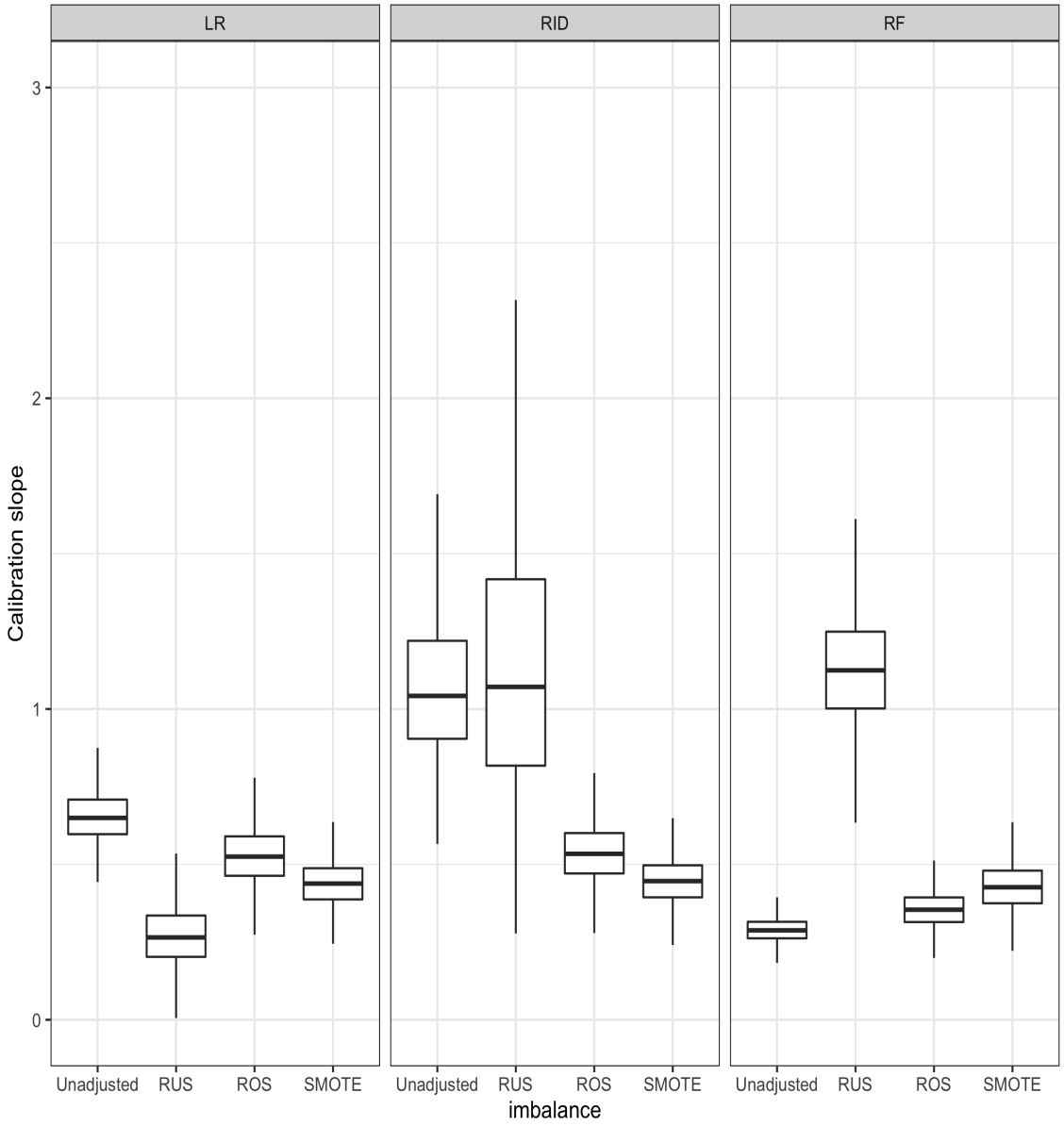

FIGURE S6 Boxplots of the Calibration slopes over 1991 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 5000$, number of predictors $R = 24$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

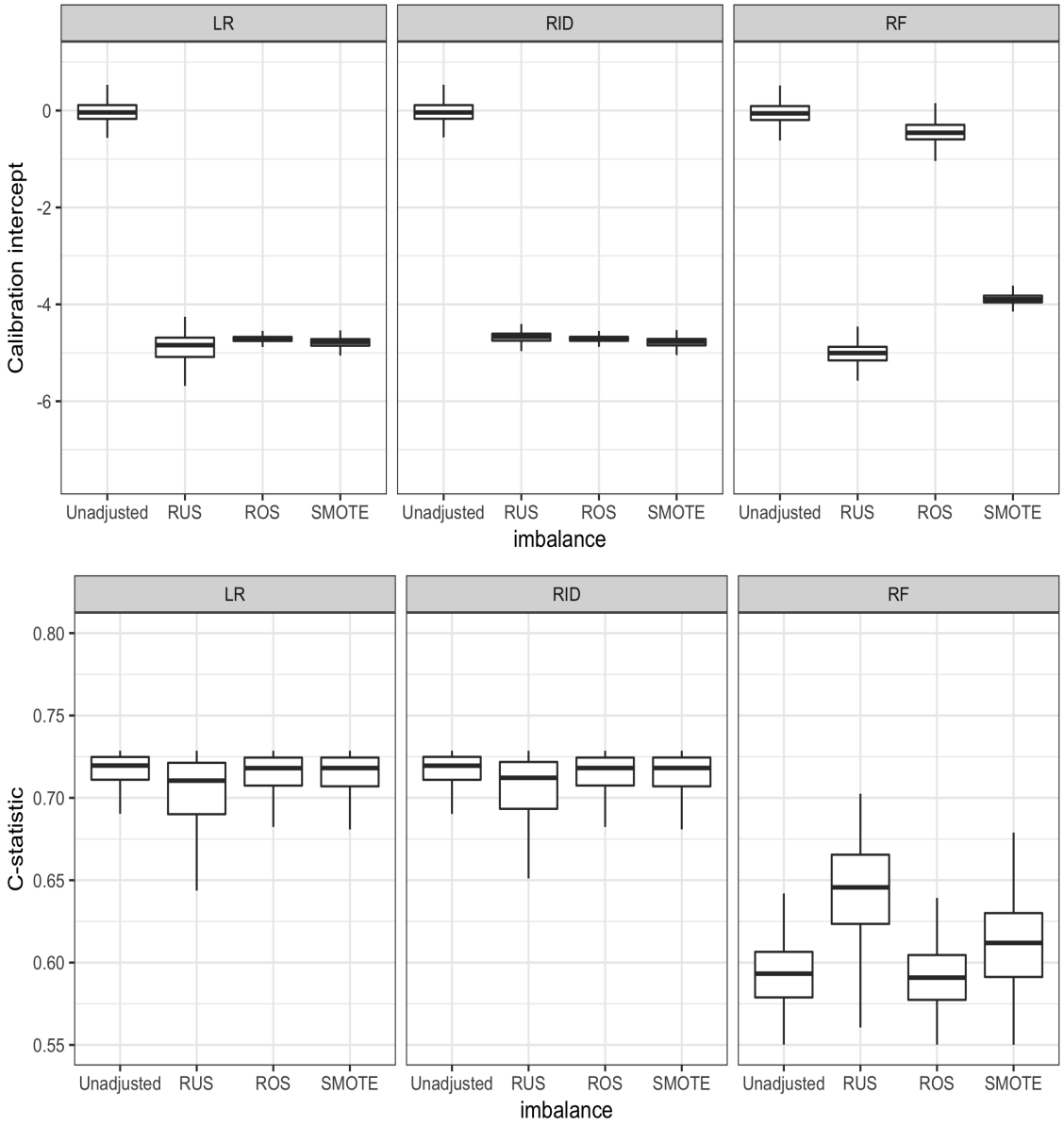

FIGURE S7 Boxplots of the C-statistics and calibration intercepts over 2000 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 3$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

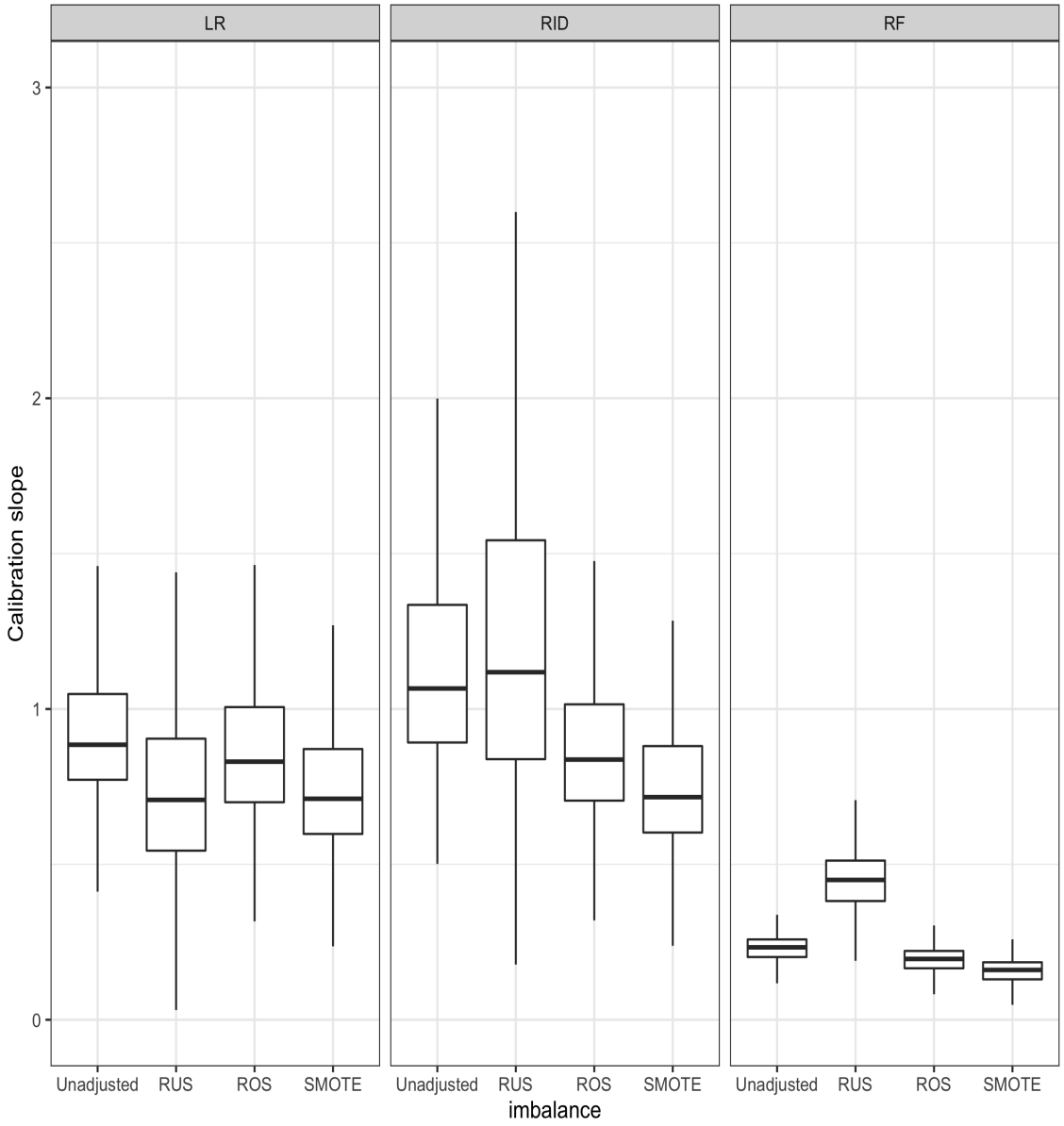

FIGURE S8 Boxplots of the Calibration slopes over 2000 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 3$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

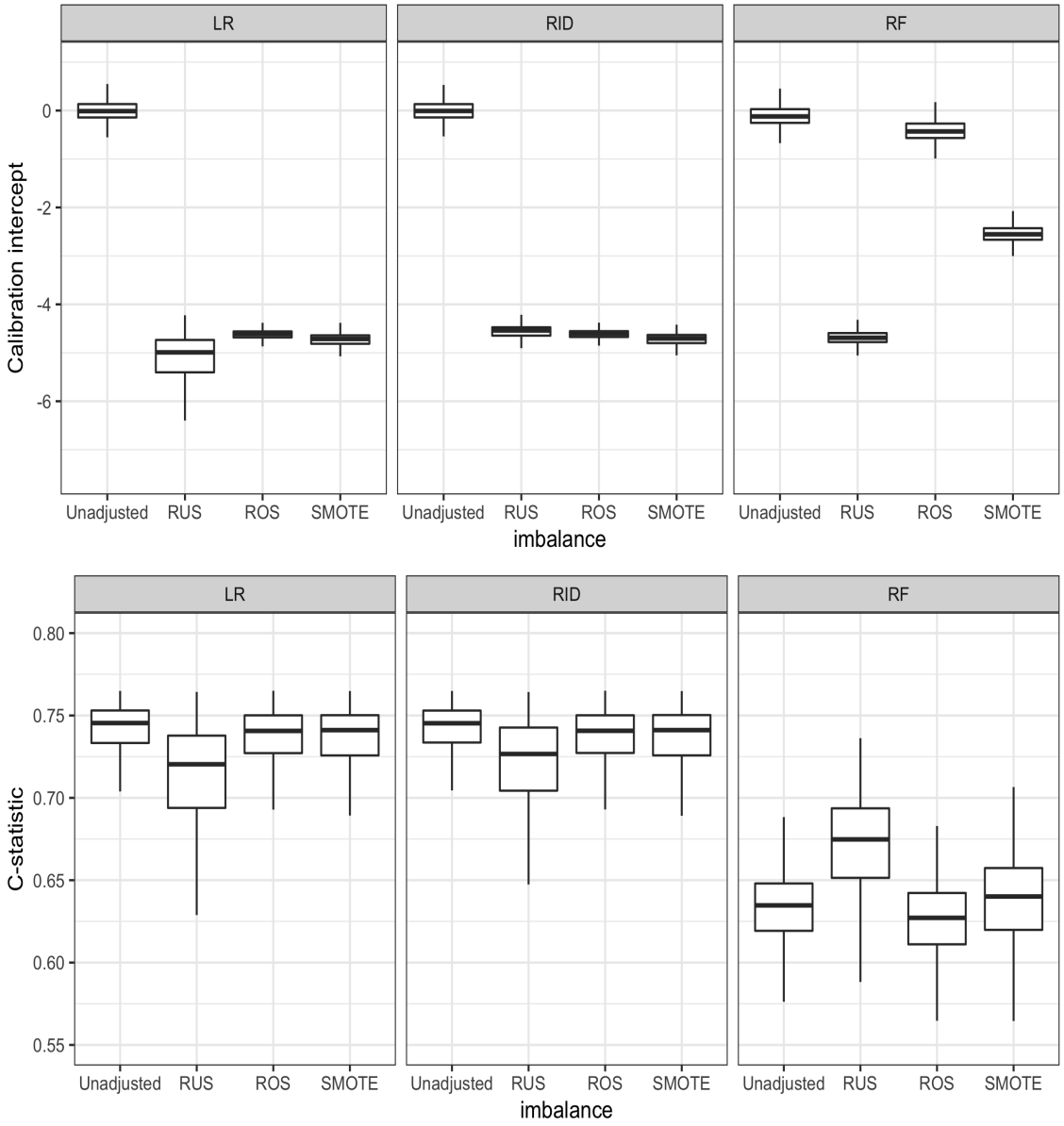

FIGURE S9 Boxplots of the C-statistics and calibration intercepts over 1997 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 6$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

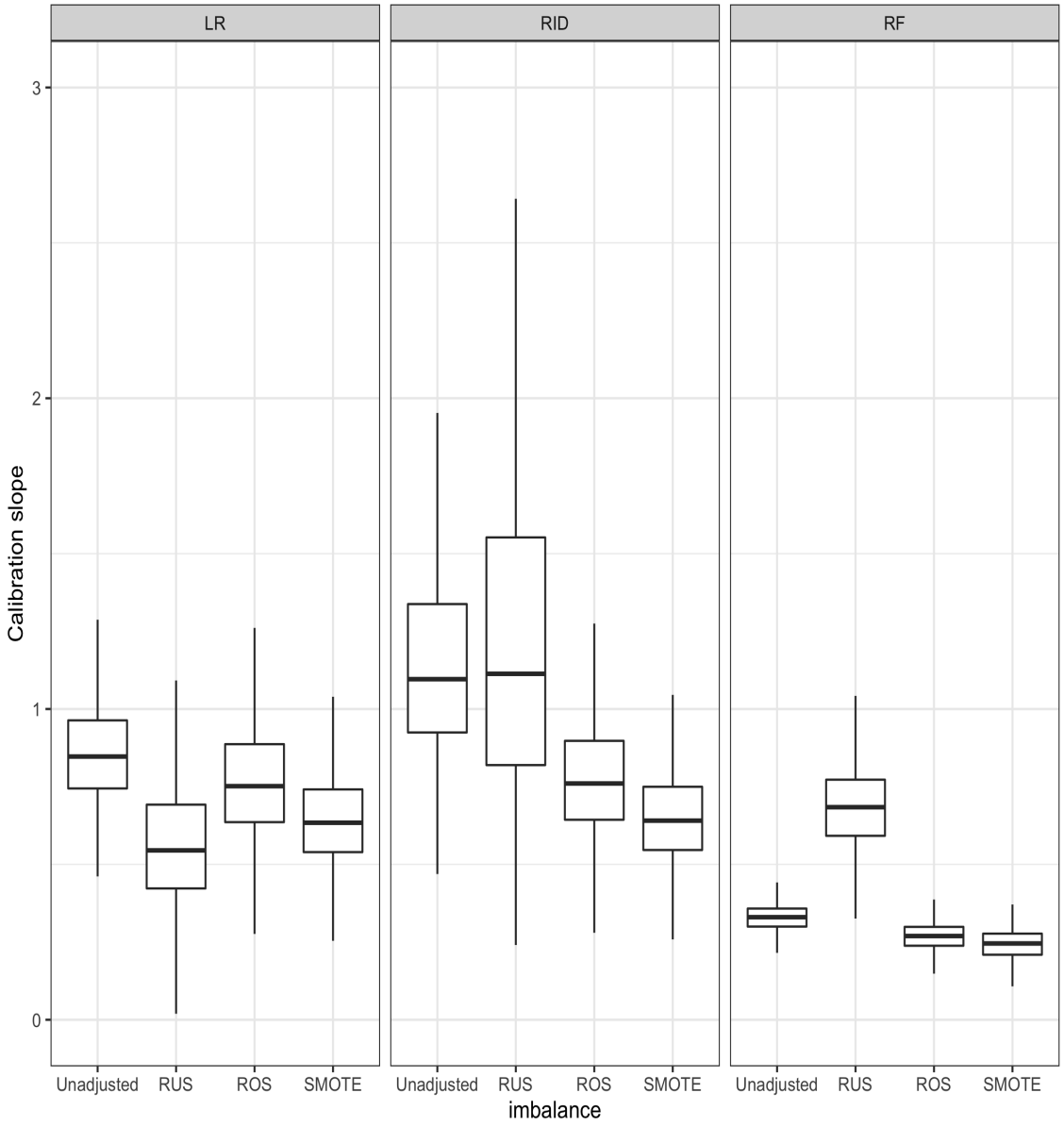

FIGURE S10 Boxplots of the Calibration slopes over 1997 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 6$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

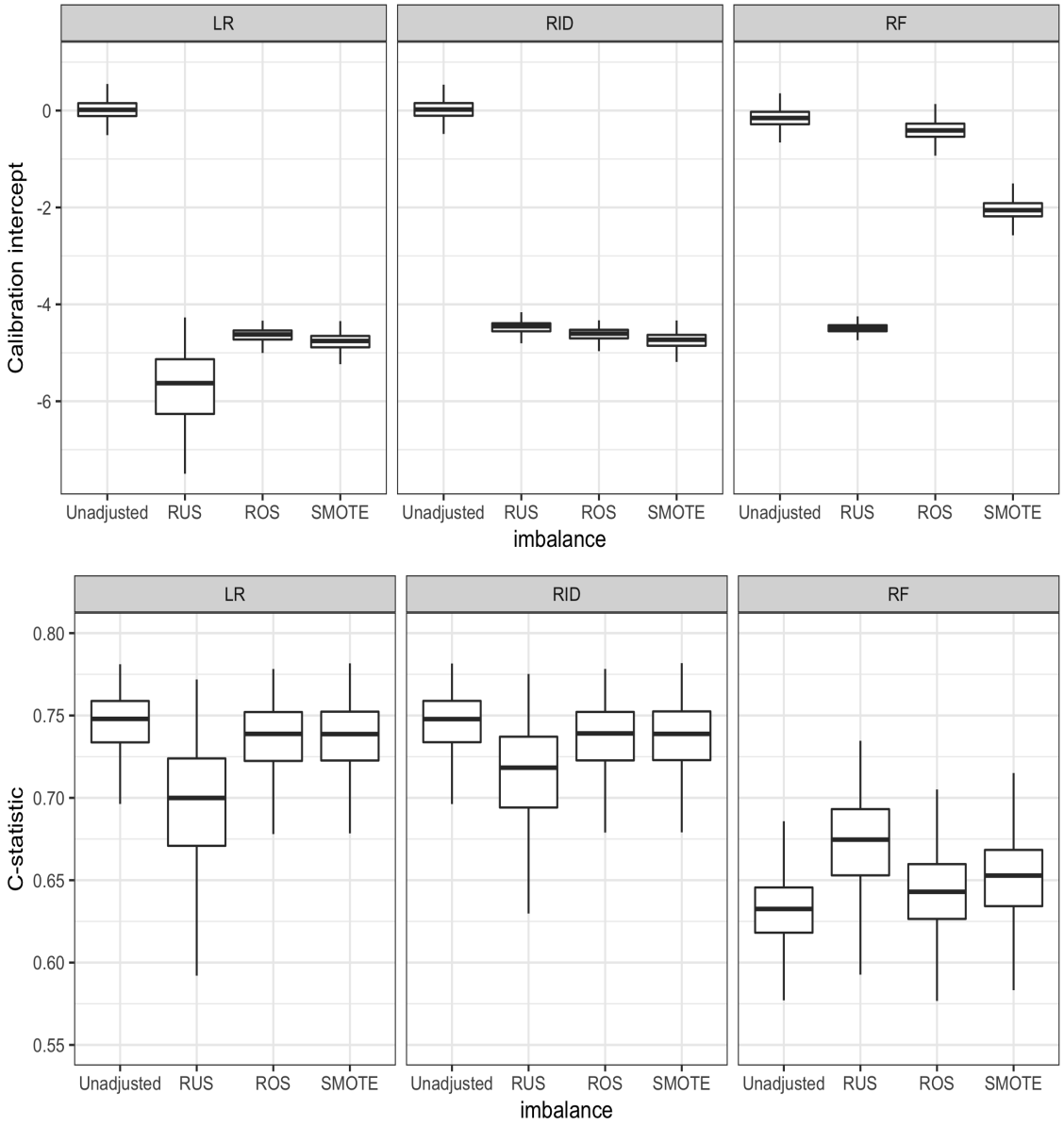

FIGURE S11 Boxplots of the C-statistics and calibration intercepts over 1948 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 12$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

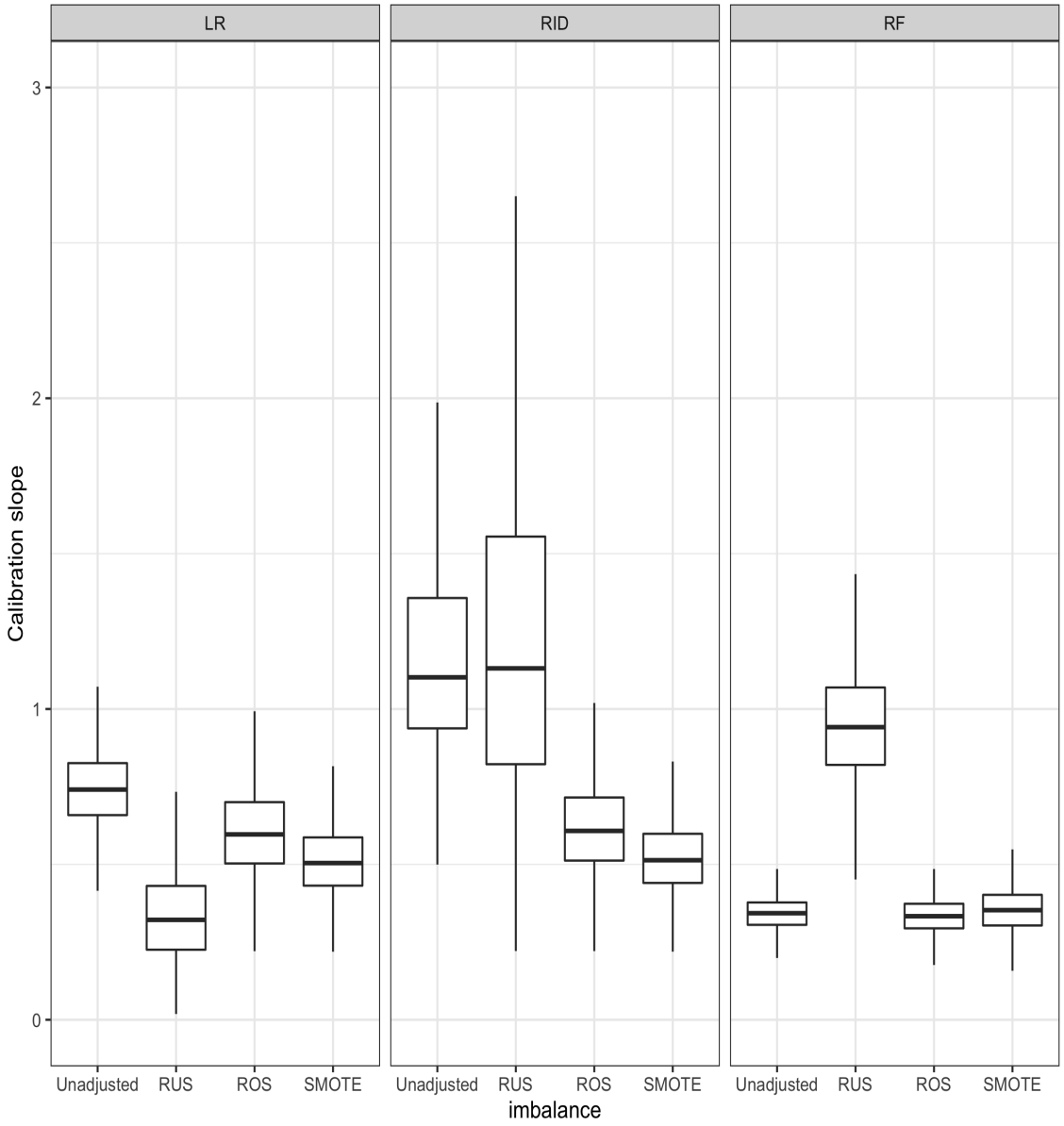

FIGURE S12 Boxplots of the Calibration slopes over 1948 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 12$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

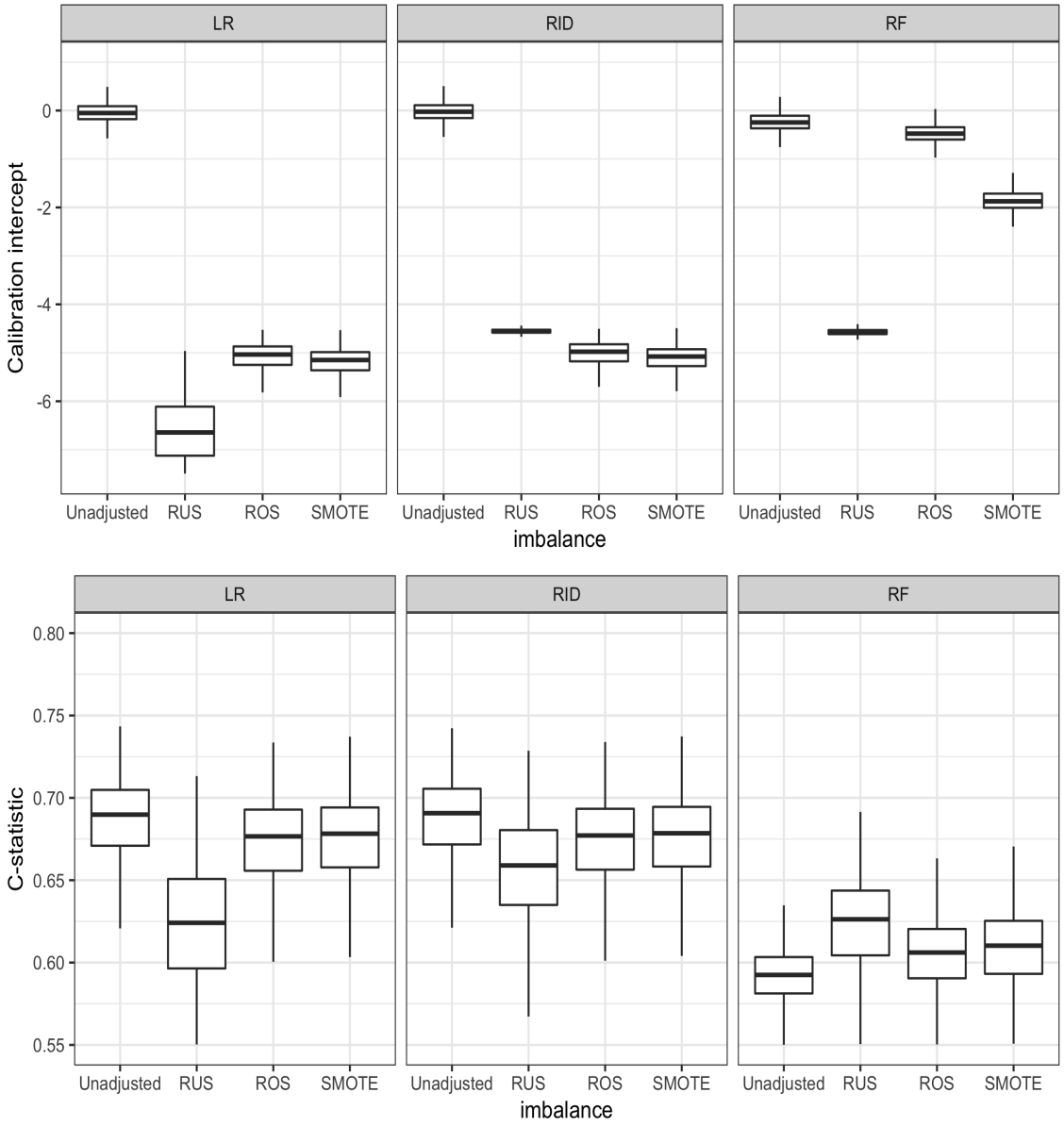

FIGURE S13 Boxplots of the C-statistics and calibration intercepts over 762 simulation iterations. The length of the whiskers is at most 1.5 times the interquartile range. Calibration intercepts are winsorized at -7.5 and 1 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 24$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique

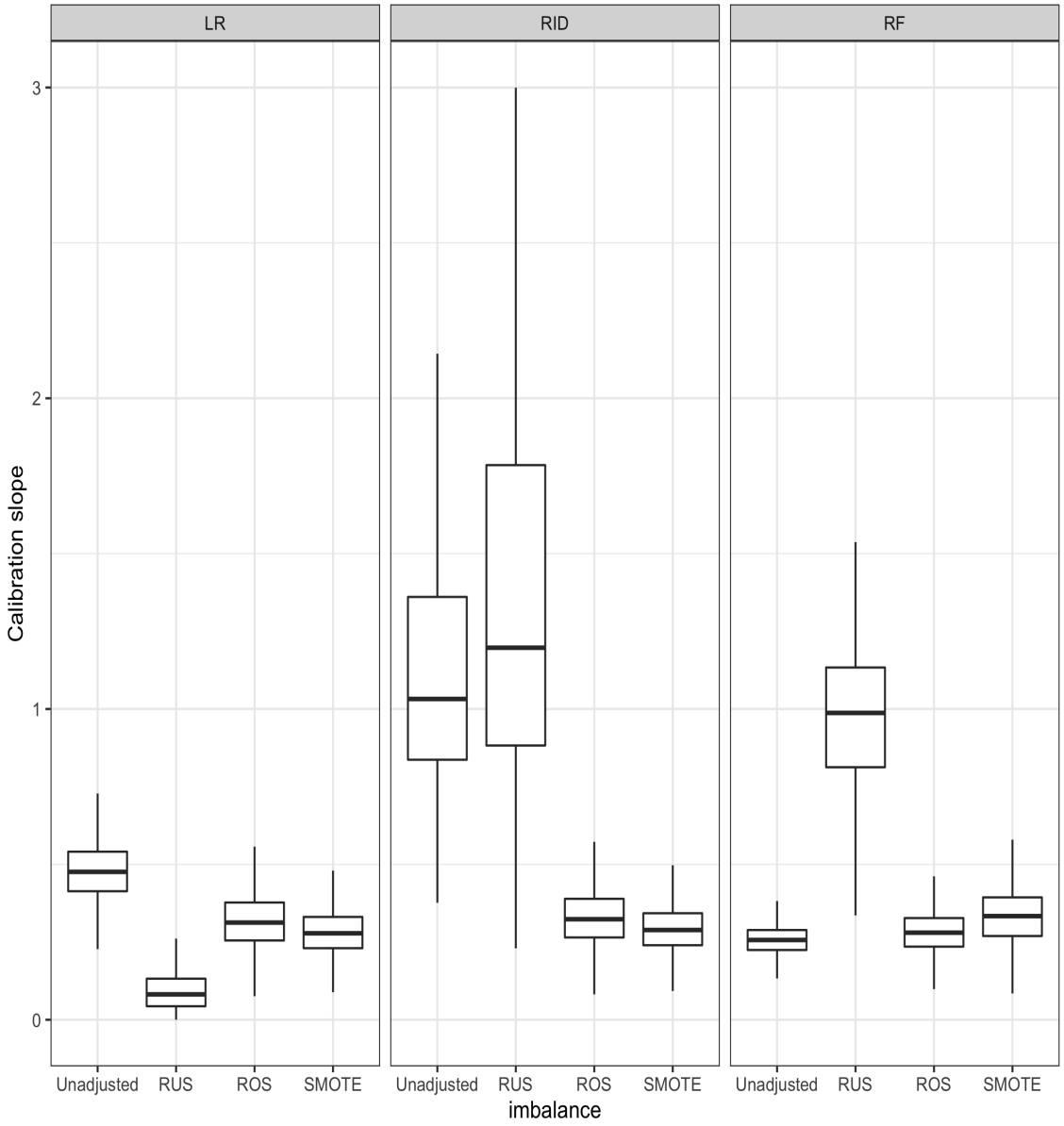

FIGURE S14 Boxplots of the Calibration slopes over 762 simulation iterations. The length of the whiskers is 1.5 at most times the interquartile range. Calibration slopes are winsorized at 0 and 3 for visualization purposes. Sample size $N = 2500$, number of predictors $R = 24$, prevalence = 0.01. Abbreviations: LR = Maximum likelihood logistic regression, RID = Ridge logistic regression, RF = Random forest, RUS = Random undersampling, ROS = Random oversampling, SMOTE = Synthetic Minority Oversampling Technique