

Appendix to ‘A Simple Aggregation Rule for Penalized Regression
Coefficients after Multiple Imputation’

Contents

Simulation results	2
Exchangeable correlation setting with correlation = .5	2
AR(1) correlation setting with correlation = .1	6
Gaussian outcome	11
Application	17
Results across imputations	17
Approximate MALCoM linear probability models	19

Simulation results

Exchangeable correlation setting with correlation = .5

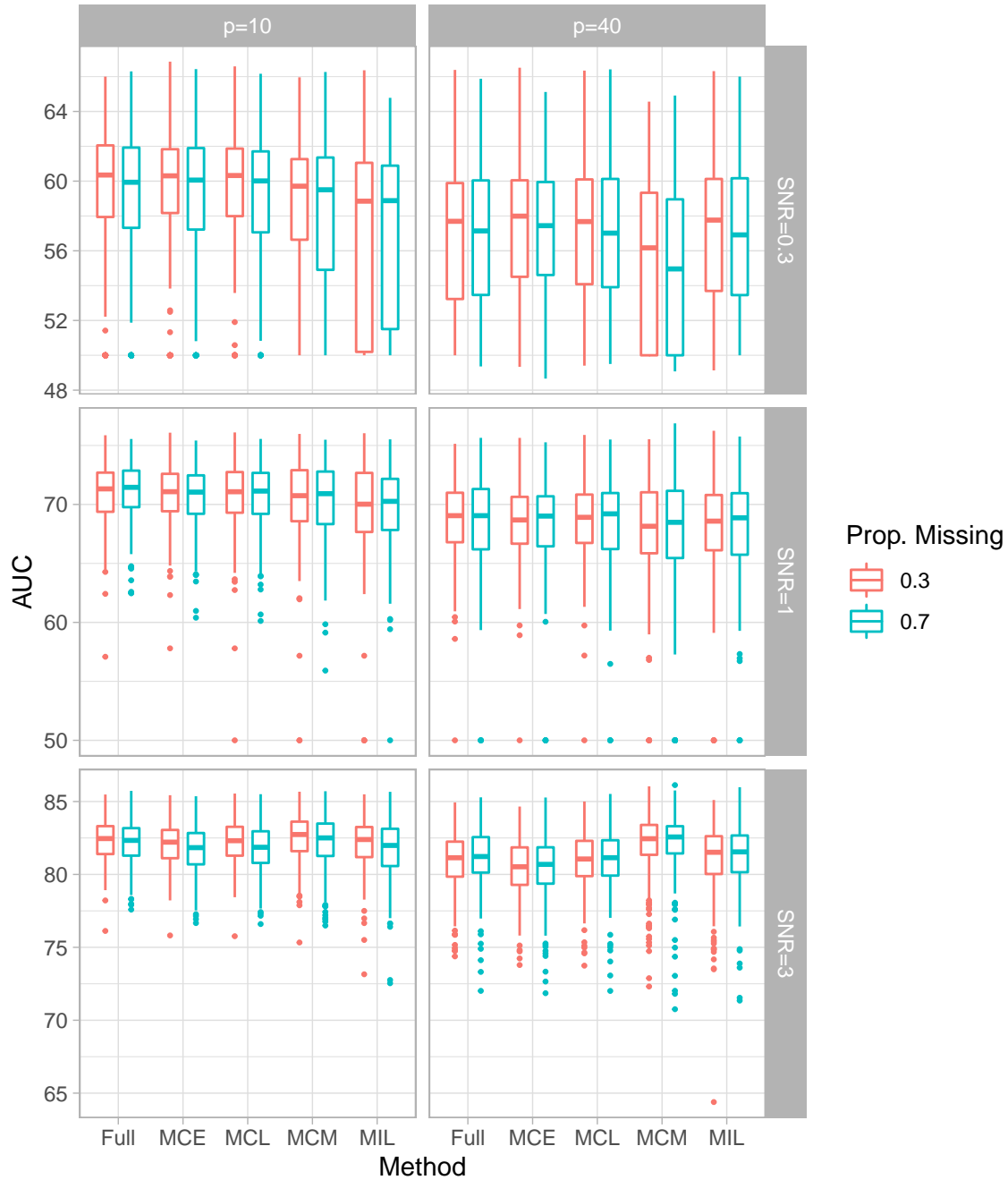


Figure A1: Raw simulation results for AUC using various methods and simulation settings with moderately high exchangeable feature correlation. Full: lasso model applied to full data set; MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

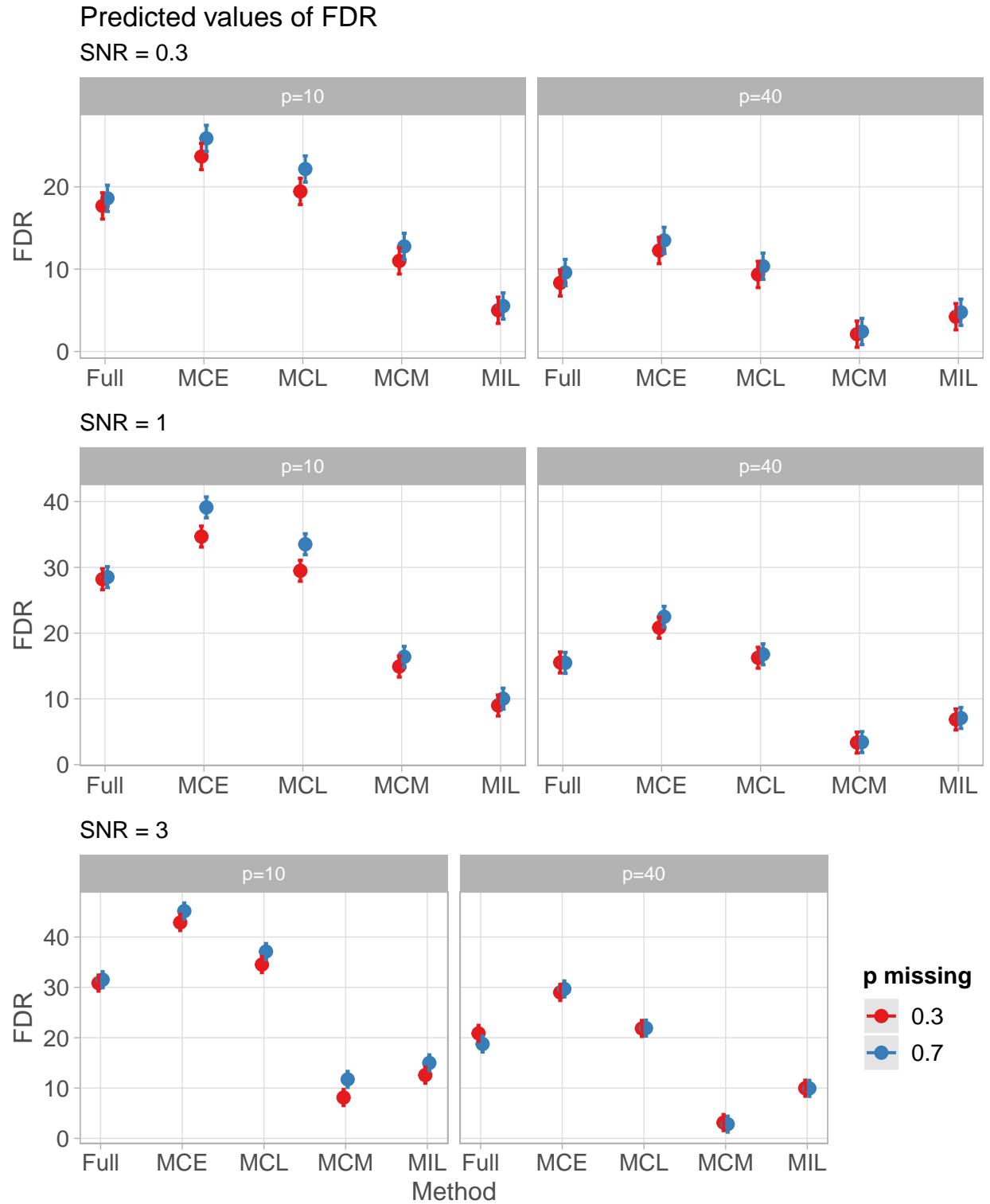


Figure A2: False discovery rate (FDR) for various methods under varied simulation settings with moderately high exchangeable feature correlation. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

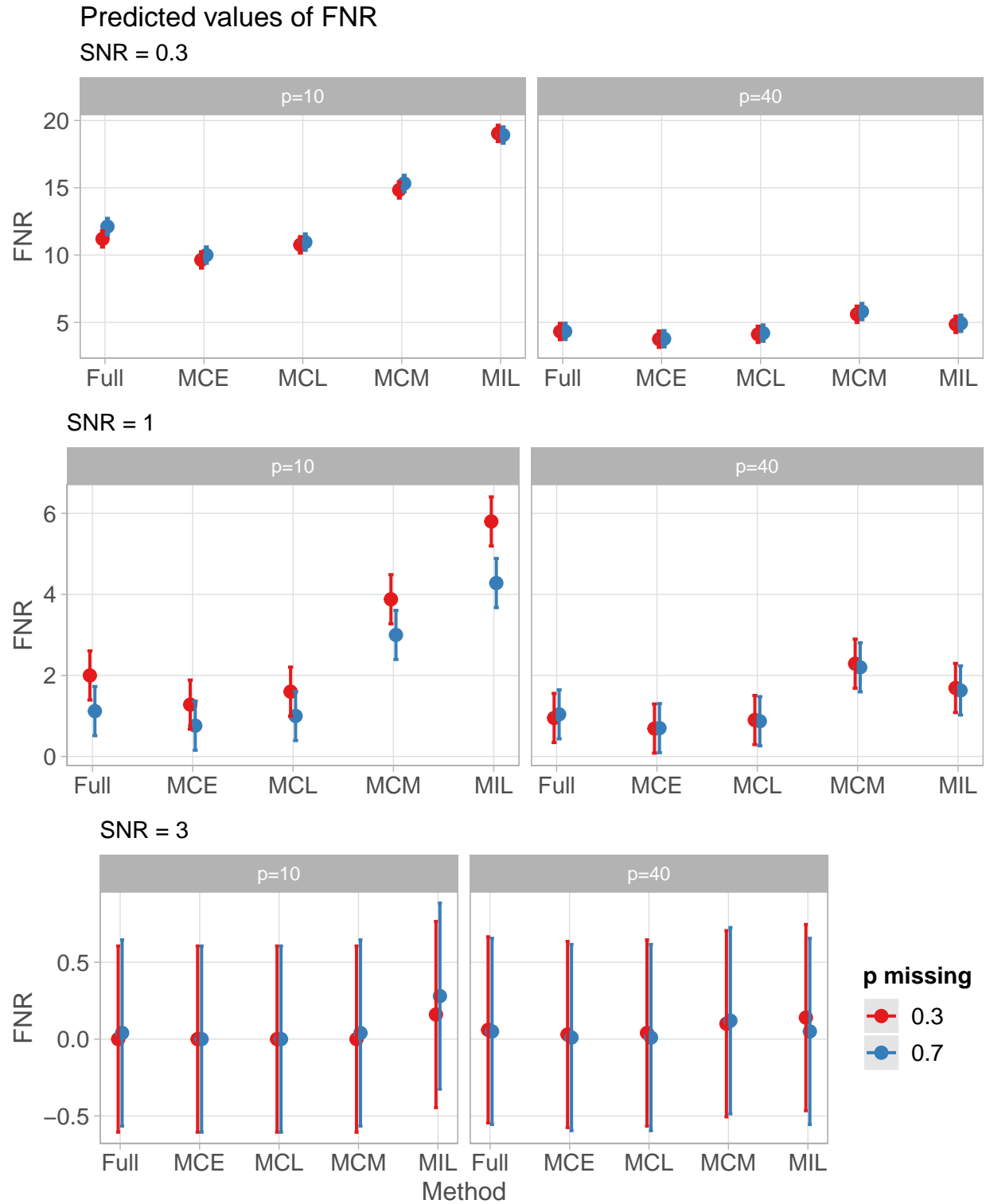


Figure A3: False negative rate (FNR) for various methods under varied simulation settings with moderately high exchangeable feature correlation. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

Table A1: False discovery rate (FDR) results for simulations. Intercept refers to the expected FDR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FDR (percentage points) if the corresponding method is used.

p.missing	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p					
0.3	10	0.3	17.68 (15.7, 19.7)	1.76 (0, 3.5)	-12.68 (-14.4, -11)
0.3	10	1.0	28.2 (26.2, 30.2)	1.28 (-0.4, 3)	-19.2 (-20.9, -17.5)
0.3	10	3.0	30.84 (29.1, 32.6)	3.68 (1.9, 5.4)	-18.28 (-20, -16.5)
High-missing, low-p					
0.3	40	0.3	8.33 (7.2, 9.4)	1.02 (-0.1, 2.1)	-4.11 (-5.2, -3)
0.3	40	1.0	15.54 (14.4, 16.7)	0.72 (-0.4, 1.8)	-8.67 (-9.8, -7.6)
0.3	40	3.0	20.86 (19.7, 22)	0.98 (-0.2, 2.1)	-10.89 (-12.1, -9.7)
Low-missing, high-p					
0.7	10	0.3	18.6 (16.5, 20.7)	3.56 (1.7, 5.4)	-13.08 (-15, -11.2)
0.7	10	1.0	28.52 (26.6, 30.5)	5 (3.2, 6.8)	-18.48 (-20.3, -16.7)
0.7	10	3.0	31.52 (29.7, 33.4)	5.6 (3.8, 7.4)	-16.52 (-18.3, -14.7)
High-missing, high-p					
0.7	40	0.3	9.59 (8.5, 10.7)	0.76 (-0.3, 1.9)	-4.83 (-5.9, -3.7)
0.7	40	1.0	15.48 (14.3, 16.7)	1.3 (0.1, 2.5)	-8.38 (-9.6, -7.2)
0.7	40	3.0	18.74 (17.6, 19.9)	3.17 (2, 4.3)	-8.8 (-10, -7.6)

Table A2: False negative rate (FNR) results for simulations. Intercept refers to the expected FNR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FNR (percentage points) if the corresponding method is used.

p.missing	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p					
0.3	10	0.3	11.2 (9.9, 12.5)	-0.44 (-1.4, 0.6)	7.84 (6.8, 8.8)
0.3	10	1.0	2 (1.3, 2.7)	-0.4 (-1.1, 0.3)	3.8 (3.1, 4.5)
0.3	10	3.0	0 (-0.1, 0.1)	0 (-0.1, 0.1)	0.16 (0.1, 0.3)
High-missing, low-p					
0.3	40	0.3	4.32 (4, 4.6)	-0.21 (-0.4, 0)	0.53 (0.3, 0.8)
0.3	40	1.0	0.95 (0.7, 1.2)	-0.05 (-0.2, 0.1)	0.74 (0.5, 0.9)
0.3	40	3.0	0.06 (0, 0.1)	-0.02 (-0.1, 0)	0.08 (0, 0.1)
Low-missing, high-p					
0.7	10	0.3	12.12 (10.8, 13.4)	-1.16 (-2.2, -0.1)	6.8 (5.8, 7.8)
0.7	10	1.0	1.12 (0.5, 1.7)	-0.12 (-0.7, 0.5)	3.16 (2.6, 3.8)
0.7	10	3.0	0.04 (-0.1, 0.2)	-0.04 (-0.2, 0.1)	0.24 (0.1, 0.4)
High-missing, high-p					
0.7	40	0.3	4.33 (4, 4.6)	-0.13 (-0.4, 0.1)	0.6 (0.4, 0.8)
0.7	40	1.0	1.04 (0.8, 1.3)	-0.17 (-0.4, 0)	0.59 (0.4, 0.8)
0.7	40	3.0	0.05 (0, 0.1)	-0.04 (-0.1, 0)	0 (-0.1, 0.1)

AR(1) correlation setting with correlation = .1

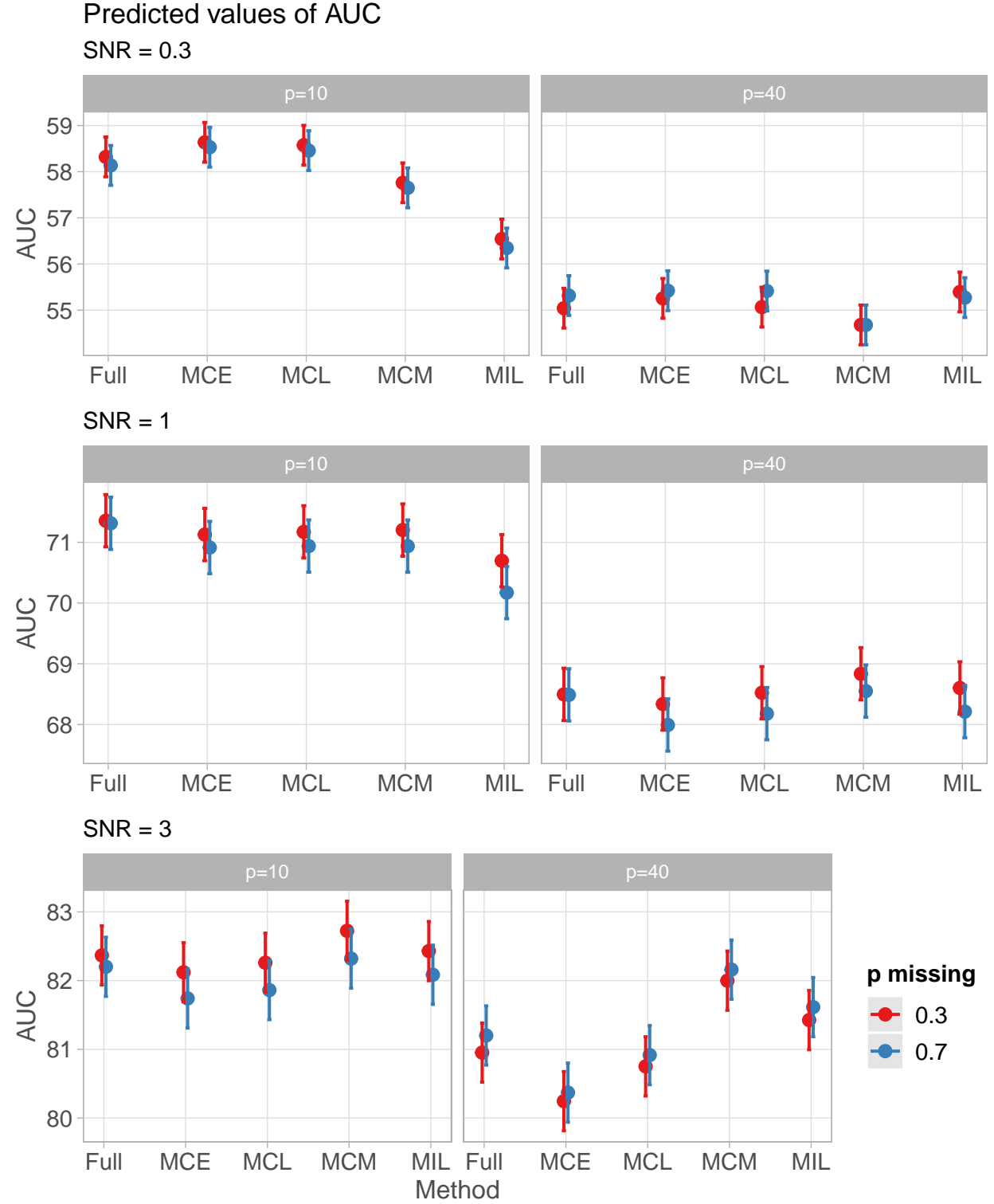


Figure A4: Predictive performance (area under ROC curve; AUC) for various methods under varied simulation settings with low AR(1) feature correlation. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

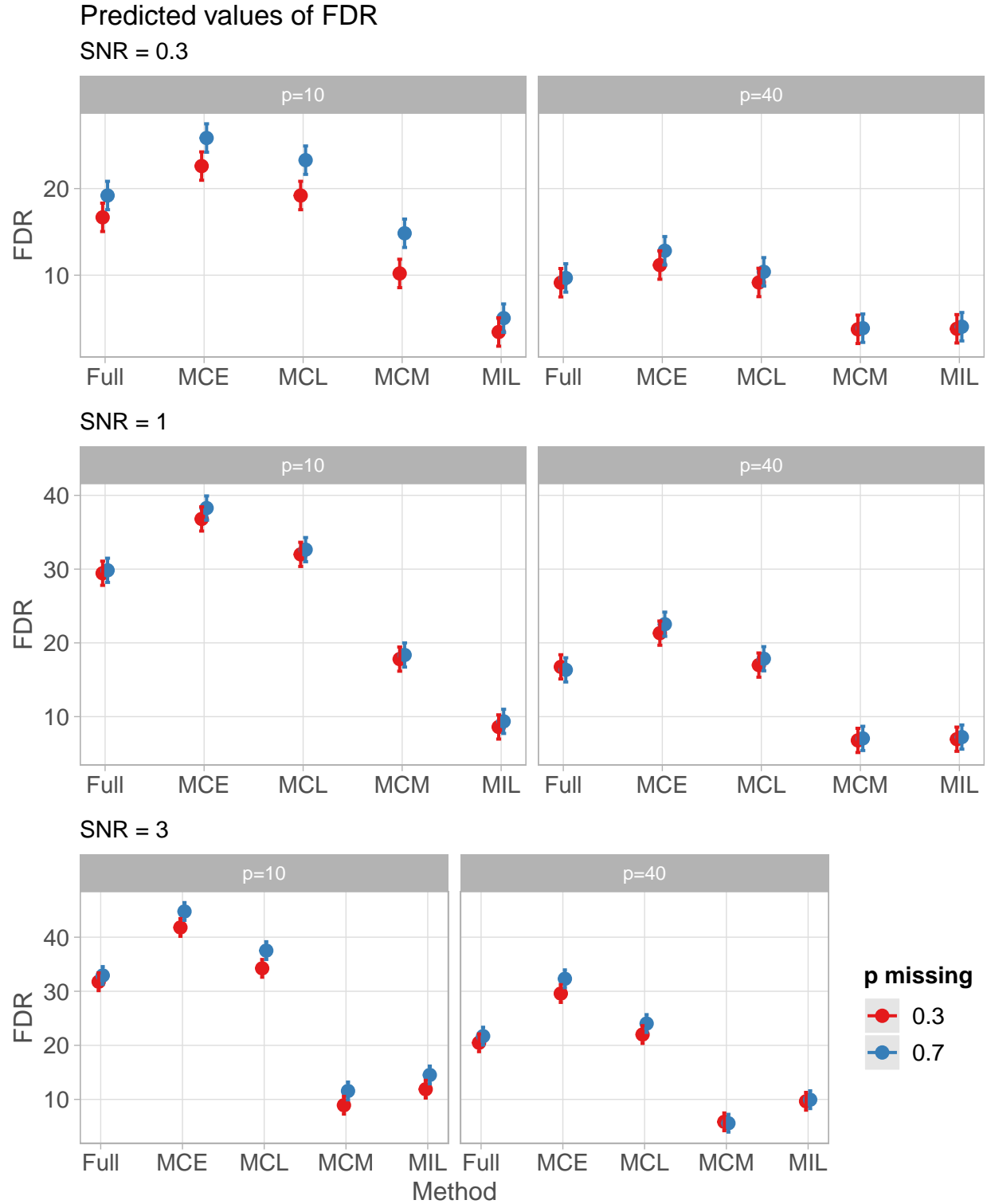


Figure A5: False discovery rate (FDR) for various methods under varied simulation settings with low AR(1) feature correlation. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

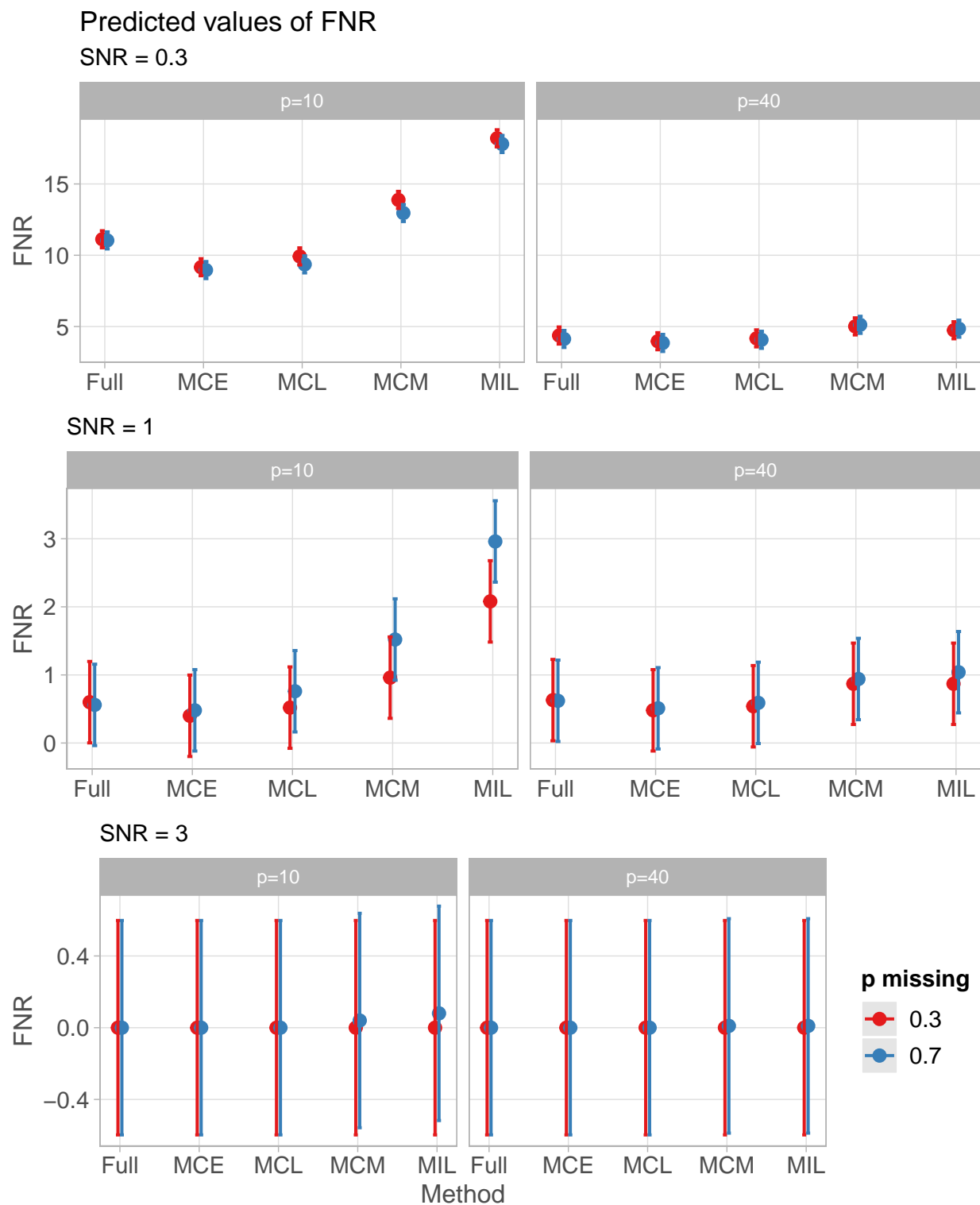


Figure A6: False negative rate (FNR) for various methods under varied simulation settings with low AR(1) feature correlation. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

Table A3: Simulation results. Intercept refers to the expected AUC for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in AUC (percentage points) if the corresponding method is used.

p.missing	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p					
0.3	10	0.3	58.32 (57.8, 58.9)	0.25 (-0.1, 0.6)	-1.78 (-2.1, -1.4)
0.3	10	1.0	71.36 (71, 71.7)	-0.18 (-0.3, 0)	-0.66 (-0.8, -0.5)
0.3	10	3.0	82.37 (82.2, 82.6)	-0.11 (-0.2, 0)	0.06 (0, 0.1)
High-missing, low-p					
0.3	40	0.3	55.04 (54.5, 55.6)	0.02 (-0.3, 0.3)	0.35 (0, 0.7)
0.3	40	1.0	68.5 (68, 69)	0.03 (-0.2, 0.3)	0.11 (-0.1, 0.3)
0.3	40	3.0	80.95 (80.7, 81.2)	-0.2 (-0.3, -0.1)	0.47 (0.4, 0.6)
Low-missing, high-p					
0.7	10	0.3	58.14 (57.6, 58.7)	0.32 (-0.1, 0.7)	-1.79 (-2.2, -1.4)
0.7	10	1.0	71.31 (71, 71.7)	-0.38 (-0.6, -0.1)	-1.14 (-1.4, -0.9)
0.7	10	3.0	82.2 (82, 82.4)	-0.34 (-0.4, -0.3)	-0.12 (-0.2, 0)
High-missing, high-p					
0.7	40	0.3	55.32 (54.8, 55.8)	0.1 (-0.2, 0.4)	-0.04 (-0.3, 0.2)
0.7	40	1.0	68.49 (67.9, 69)	-0.31 (-0.6, 0)	-0.27 (-0.5, 0)
0.7	40	3.0	81.2 (80.9, 81.5)	-0.29 (-0.4, -0.2)	0.41 (0.3, 0.5)

Table A4: False discovery rate (FDR) results for simulations. Intercept refers to the expected FDR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FDR (percentage points) if the corresponding method is used.

p.missing	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p					
0.3	10	0.3	16.68 (14.8, 18.6)	2.52 (0.7, 4.3)	-13.24 (-15.1, -11.4)
0.3	10	1.0	29.44 (27.4, 31.5)	2.56 (0.8, 4.3)	-20.84 (-22.6, -19.1)
0.3	10	3.0	31.76 (29.9, 33.6)	2.48 (0.7, 4.3)	-19.88 (-21.7, -18.1)
High-missing, low-p					
0.3	40	0.3	9.13 (8, 10.3)	0.03 (-1, 1)	-5.32 (-6.3, -4.3)
0.3	40	1.0	16.75 (15.6, 17.9)	0.24 (-0.8, 1.3)	-9.81 (-10.9, -8.7)
0.3	40	3.0	20.47 (19.4, 21.6)	1.53 (0.5, 2.5)	-10.82 (-11.8, -9.8)
Low-missing, high-p					
0.7	10	0.3	19.2 (17, 21.4)	4.08 (2.1, 6)	-14.16 (-16.1, -12.2)
0.7	10	1.0	29.84 (27.8, 31.9)	2.8 (1, 4.6)	-20.48 (-22.3, -18.7)
0.7	10	3.0	32.92 (31.1, 34.7)	4.6 (2.8, 6.4)	-18.4 (-20.2, -16.6)
High-missing, high-p					
0.7	40	0.3	9.68 (8.4, 11)	0.71 (-0.4, 1.8)	-5.63 (-6.8, -4.5)
0.7	40	1.0	16.34 (15.1, 17.6)	1.51 (0.4, 2.6)	-9.11 (-10.2, -8)
0.7	40	3.0	21.7 (20.5, 22.9)	2.33 (1.2, 3.5)	-11.76 (-12.9, -10.6)

Table A5: False negative rate (FNR) results for simulations. Intercept refers to the expected FNR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FNR (percentage points) if the corresponding method is used.

p.missing	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p					
0.3	10	0.3	11.12 (9.8, 12.5)	-1.2 (-2.2, -0.2)	7.08 (6.1, 8.1)
0.3	10	1.0	0.6 (0.2, 1)	-0.08 (-0.5, 0.3)	1.48 (1.1, 1.9)
0.3	10	3.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
High-missing, low-p					
0.3	40	0.3	4.37 (4, 4.7)	-0.2 (-0.4, 0)	0.37 (0.1, 0.6)
0.3	40	1.0	0.63 (0.4, 0.8)	-0.09 (-0.2, 0)	0.24 (0.1, 0.4)
0.3	40	3.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
Low-missing, high-p					
0.7	10	0.3	11.04 (9.7, 12.4)	-1.68 (-2.7, -0.6)	6.76 (5.7, 7.8)
0.7	10	1.0	0.56 (0.1, 1)	0.2 (-0.3, 0.7)	2.4 (1.9, 2.9)
0.7	10	3.0	0 (-0.1, 0.1)	0 (-0.1, 0.1)	0.08 (0, 0.2)
High-missing, high-p					
0.7	40	0.3	4.13 (3.8, 4.5)	-0.06 (-0.3, 0.2)	0.72 (0.5, 1)
0.7	40	1.0	0.62 (0.4, 0.8)	-0.03 (-0.2, 0.1)	0.42 (0.3, 0.6)
0.7	40	3.0	0 (0, 0)	0 (0, 0)	0.01 (0, 0)

This concludes simulation results.

Gaussian outcome

Set up

An additional simulation was done to ensure that these results were also applicable in Gaussian outcome settings for which the MI-lasso was originally developed. The simulation parameters are the same as in the exchangeable correlation setting *except*: the “high-p” value was set to $p = 20$ instead of 40, and the SNR was set to either 0.1, 1, or 2. The Gaussian noise (σ) was set to 1.

Results

Generally, the Gaussian setting showed similar results to that in the binary outcome case. However, a few distinctions are worth noting. First, MALCoM methods saw a slightly larger loss in predictive accuracy when a higher percentage of observations were missing, though this difference was quite small. Second, there were very few false negatives in any of these methods, especially for higher SNR settings, and this led to model instability for the FNR evaluative models.

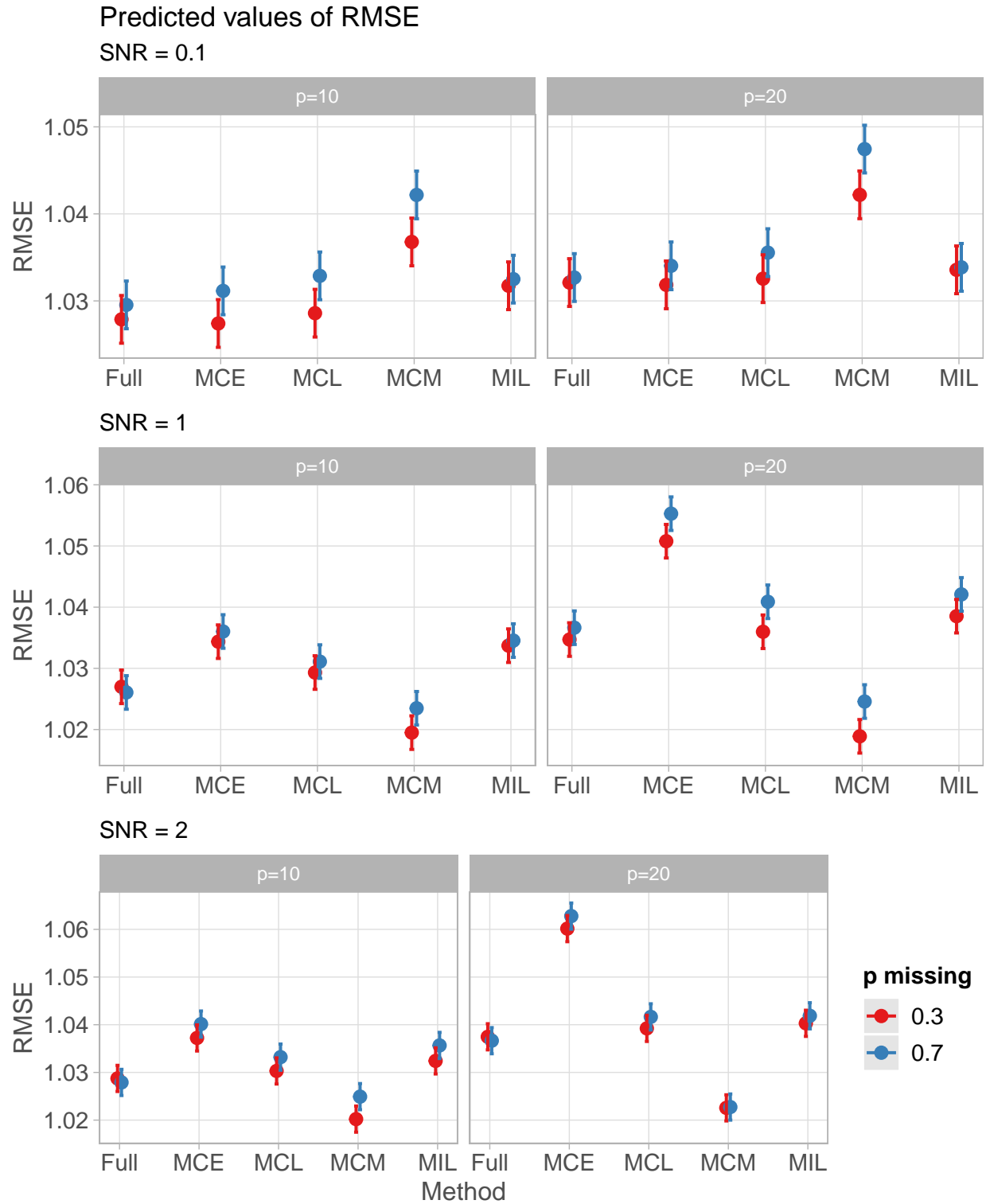


Figure A7: Predictive performance (root-mean-squared error; RMSE) for various methods under varied simulation settings and Gaussian outcome. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

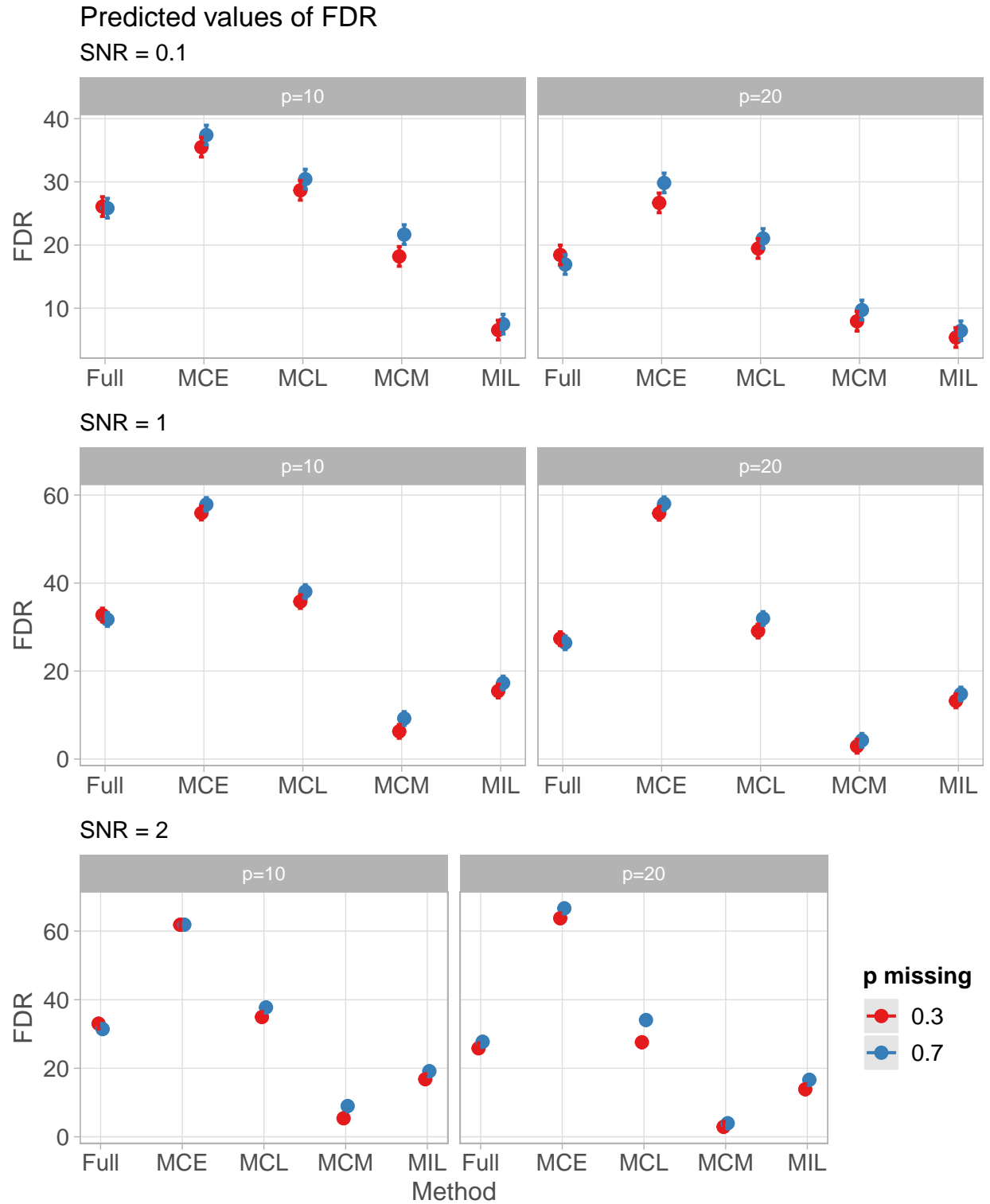


Figure A8: False discovery rate (FDR) for various methods under varied simulation settings and a Gaussian outcome. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

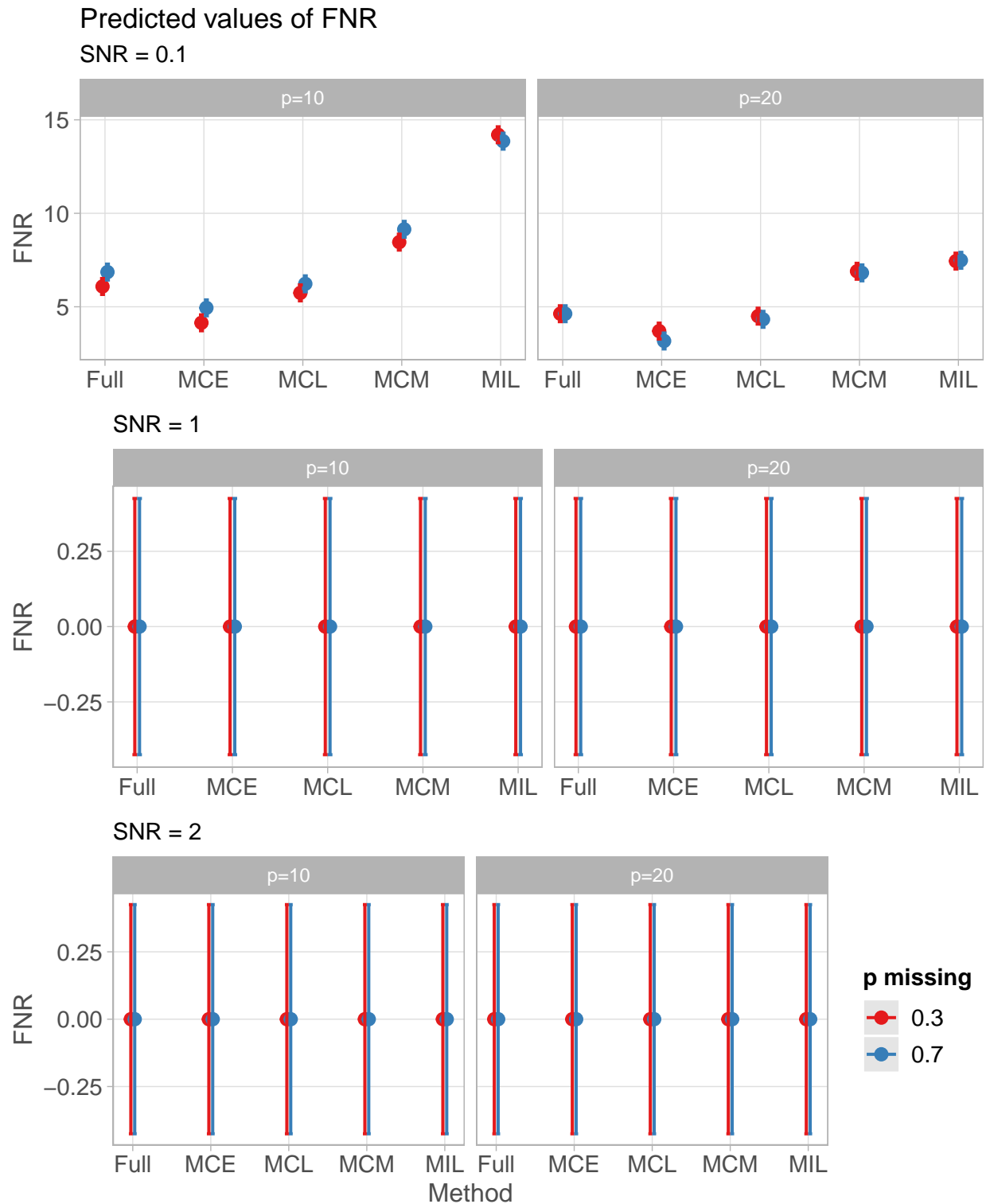


Figure A9: False negative rate (FNR) for various methods under varied simulation settings and a Gaussian outcome. MIL: Multiple imputation lasso; MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty.

Table A6: Simulation results. Intercept refers to the expected RMSE for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in RMSE if the corresponding method is used. Estimates are multiplied by 100.

p.missing	rho	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p						
0.3	0.45	10	0.1	102.79 (102.6, 103)	0.07 (-0.1, 0.2)	0.39 (0.3, 0.5)
0.3	0.45	10	1.0	102.7 (102.4, 103)	0.23 (0.1, 0.4)	0.67 (0.5, 0.8)
0.3	0.45	10	2.0	102.87 (102.6, 103.1)	0.16 (0, 0.3)	0.37 (0.2, 0.5)
High-missing, low-p						
0.3	0.45	20	0.1	103.21 (103, 103.5)	0.04 (-0.1, 0.2)	0.15 (0, 0.3)
0.3	0.45	20	1.0	103.47 (103.2, 103.8)	0.13 (0, 0.3)	0.38 (0.2, 0.5)
0.3	0.45	20	2.0	103.75 (103.5, 104)	0.18 (0, 0.3)	0.28 (0.1, 0.5)
Low-missing, high-p						
0.7	0.45	10	0.1	102.95 (102.7, 103.2)	0.33 (0.2, 0.5)	0.3 (0.2, 0.4)
0.7	0.45	10	1.0	102.61 (102.3, 102.9)	0.5 (0.4, 0.7)	0.85 (0.7, 1)
0.7	0.45	10	2.0	102.79 (102.5, 103.1)	0.53 (0.4, 0.7)	0.78 (0.6, 0.9)
High-missing, high-p						
0.7	0.45	20	0.1	103.27 (103, 103.5)	0.29 (0.1, 0.5)	0.12 (-0.1, 0.3)
0.7	0.45	20	1.0	103.66 (103.4, 104)	0.42 (0.2, 0.6)	0.54 (0.4, 0.7)
0.7	0.45	20	2.0	103.67 (103.4, 104)	0.5 (0.3, 0.7)	0.52 (0.3, 0.7)

Table A7: False discovery rate (FDR) results for simulations. Intercept refers to the expected FDR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FDR (percentage points) if the corresponding method is used.

p.missing	rho	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p						
0.3	0.45	10	0.1	26.09 (24.2, 28)	2.57 (0.8, 4.4)	-19.54 (-21.3, -17.8)
0.3	0.45	10	1.0	32.74 (31.2, 34.3)	3.06 (1.4, 4.7)	-17.29 (-18.9, -15.6)
0.3	0.45	10	2.0	33 (31.5, 34.5)	1.94 (0.3, 3.6)	-16.2 (-17.9, -14.5)
High-missing, low-p						
0.3	0.45	20	0.1	18.43 (16.9, 20)	1.03 (-0.4, 2.4)	-13.06 (-14.5, -11.6)
0.3	0.45	20	1.0	27.37 (26.1, 28.7)	1.71 (0.3, 3.1)	-14.14 (-15.6, -12.7)
0.3	0.45	20	2.0	25.83 (24.5, 27.1)	1.76 (0.3, 3.2)	-11.96 (-13.4, -10.5)
Low-missing, high-p						
0.7	0.45	10	0.1	25.83 (23.8, 27.8)	4.6 (2.8, 6.4)	-18.34 (-20.1, -16.5)
0.7	0.45	10	1.0	31.71 (30.1, 33.3)	6.34 (4.7, 8)	-14.43 (-16.1, -12.7)
0.7	0.45	10	2.0	31.43 (29.9, 33)	6.31 (4.6, 8)	-12.29 (-14, -10.6)
High-missing, high-p						
0.7	0.45	20	0.1	16.93 (15.3, 18.5)	4.11 (2.7, 5.5)	-10.5 (-11.9, -9.1)
0.7	0.45	20	1.0	26.41 (25.1, 27.7)	5.53 (4.1, 7)	-11.63 (-13.1, -10.2)
0.7	0.45	20	2.0	27.73 (26.4, 29.1)	6.34 (4.8, 7.8)	-11.1 (-12.6, -9.6)

Table A8: False negative rate (FNR) results for simulations. Intercept refers to the expected FNR for a complete-data lasso model tuned via 10-fold cross-validation. Beta coefficients represent the expected change in FNR (percentage points) if the corresponding method is used.

p.missing	rho	p	SNR	Intercept	beta.MCL	beta.MIL
Low-missing, low-p						
0.3	0.45	10	0.1	6.09 (5.2, 7)	-0.34 (-1.2, 0.5)	8.11 (7.3, 8.9)
0.3	0.45	10	1.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
0.3	0.45	10	2.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
High-missing, low-p						
0.3	0.45	20	0.1	4.63 (4.1, 5.1)	-0.13 (-0.5, 0.3)	2.81 (2.4, 3.2)
0.3	0.45	20	1.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
0.3	0.45	20	2.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
Low-missing, high-p						
0.7	0.45	10	0.1	6.86 (5.9, 7.8)	-0.63 (-1.4, 0.2)	7 (6.2, 7.8)
0.7	0.45	10	1.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
0.7	0.45	10	2.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
High-missing, high-p						
0.7	0.45	20	0.1	4.63 (4.1, 5.1)	-0.3 (-0.7, 0.1)	2.86 (2.4, 3.3)
0.7	0.45	20	1.0	0 (0, 0)	0 (0, 0)	0 (0, 0)
0.7	0.45	20	2.0	0 (0, 0)	0 (0, 0)	0 (0, 0)

This concludes simulation results.

Application

Results across imputations

The following tables present the coefficients from all imputed analysis for our application.

Table A9: Imputation-specific results for MALCoM-lasso (MCL). The final MCL model was selected in 20 percent of single imputations. SBA: Symptoms duration before admission; NW: Nonwhite.

	Intercept	Z(crp)	Z(ldh)	Z(LDD)	Z(LF)	SBA	Z(age)	Z(bmi)	NW	dm	male	Z(nlr)
I.1	-0.63	0.61	0.26	-	-	-	-	-	0.10	0.51	-	0.49
I.2	-0.61	0.68	0.28	-	0.03	-	-	-	0.01	0.66	-	0.46
I.3	-0.54	0.67	0.22	-	-	-	-	-	-	0.44	-	0.22
I.4	-0.85	0.69	0.30	-	0.08	-	-	-	0.22	0.61	0.13	0.31
I.5	-0.46	0.75	0.40	0.11	0.01	-	-	-	-	0.26	-	0.39
I.6	-0.48	0.61	0.40	-	-	-	-	-	-	0.23	-	0.16
I.7	-0.54	0.66	0.36	-	-	-	-	-	0.02	0.39	-	0.12
I.8	-0.52	0.60	0.35	-	-	-	-	-	-	0.41	-	0.35
I.9	-0.60	0.69	0.34	-	-	-	-	-	-	0.61	-	-
I.10	-0.55	0.77	0.23	-	-	-	-	-	0.00	0.47	0.01	0.28
I.11	-0.58	0.84	0.01	0.16	0.04	-	-	-	0.03	0.42	0.03	-
I.12	-0.81	0.77	0.37	-	-	0	-	0.00	0.20	0.58	0.15	0.22
I.13	-0.54	0.71	0.30	-	-	-	-	-	-	0.43	-	0.25
I.14	-0.85	0.69	0.25	0.18	0.09	-	-	0.08	0.23	0.59	0.14	0.22
I.15	-0.56	0.74	0.34	-	-	-	-	-	-	0.52	-	0.21
I.16	-0.54	0.57	0.29	0.05	-	-	-	-	0.01	0.42	-	0.16
I.17	-0.58	0.71	0.35	-	0.01	-	-	-	-	0.51	0.03	0.29
I.18	-0.57	0.68	0.35	-	0.14	-	-	-	-	0.51	-	0.26
I.19	-0.52	0.71	0.26	-	-	-	-	-	-	0.36	-	-
I.20	-0.64	0.71	0.23	-	-	-	-	-	0.05	0.53	0.06	0.19
MCL	-0.56	0.69	0.30	-	-	-	-	-	0.00	0.49	-	0.22

Table A10: Imputation-specific results for MALCoM-MCP (MCM). The final MCM model was selected in 35 percent of single imputations. SBA: Symptoms duration before admission; NW: Nonwhite.

	Intercept	Z(crp)	Z(ldh)	Z(LDD)	Z(LF)	SBA	Z(age)	Z(bmi)	NW	dm	male	Z(nlr)
Imp.1	-0.58	0.83	0.14	-	-	-	-	-	-	0.60	-	0.45
Imp.2	-0.74	0.91	0.30	-	-	-	-	-	-	1.10	0.01	0.68
Imp.3	-0.59	0.94	0.18	-	-	-	-	-	-	0.60	-	0.24
Imp.4	-0.58	0.89	0.15	-	-	-	-	-	-	0.54	-	0.19
Imp.5	-0.65	1.00	0.44	0.52	0.20	-0.06	-0.34	0.17	0.38	0.64	0.30	0.83
Imp.6	-0.41	0.72	0.30	-	-	-	-	-	-	-	-	-
Imp.7	-0.55	0.89	0.34	-	-	-	-	-	-	0.43	-	-
Imp.8	-0.56	0.81	0.39	-	-	-	-	-	-	0.54	-	0.42
Imp.9	-0.73	0.94	0.37	-	-	-	-	-	-	0.96	-	-
Imp.10	-0.66	1.01	0.24	-	-	-	-	-	-	0.75	0.07	0.35
Imp.11	-0.45	0.95	-	-	-	-	-	-	-	0.14	-	-
Imp.12	-0.48	0.89	0.15	-	-	-	-	-	-	0.25	-	-
Imp.13	-0.61	0.97	0.31	-	-	-	-	-	-	0.65	-	0.28
Imp.14	-0.51	0.75	0.18	-	-	-	-	-	-	0.35	-	-
Imp.15	-0.66	0.93	0.42	-	-	-	-	-	-	0.82	0.01	0.31
Imp.16	-0.59	0.83	0.23	-	-	-	-	-	-	0.57	-	0.12
Imp.17	-0.59	0.95	0.23	-	-	-	-	-	-	0.56	-	0.20
Imp.18	-0.69	0.83	0.47	-	0.16	-	-	-	-	0.81	0.05	0.35
Imp.19	-0.68	0.99	0.30	-	-	-	-	-	-	0.69	0.08	-
Imp.20	-0.51	0.93	-	-	-	-	-	-	-	0.32	-	-
MCM	-0.59	0.92	0.27	-	-	-	-	-	-	0.58	-	0.19

Table A11: Imputation-specific results for MALCoM-elastic net (MCE). The final MCE model was selected in 5 percent of single imputations. SBA: Symptoms duration before admission; NW: Nonwhite.

	Intercept	Z(crp)	Z(ldh)	Z(LDD)	Z(LF)	SBA	Z(age)	Z(bmi)	NW	dm	male	Z(nlr)
Imp.1	-0.65	0.57	0.26	-	-	-	-	-	0.13	0.48	-	0.47
Imp.2	-0.63	0.64	0.28	-	0.04	-	-	-	0.05	0.62	-	0.45
Imp.3	-0.62	0.64	0.25	-	-	-	-	-	0.08	0.47	0.02	0.25
Imp.4	-0.57	0.57	0.24	-	0.06	-	-	-	0.03	0.42	-	0.21
Imp.5	-0.46	0.70	0.39	0.12	0.03	0.00	-	-	-	0.27	-	0.38
Imp.6	-0.53	0.59	0.41	-	0.03	-	-	-	0.04	0.29	-	0.21
Imp.7	-0.75	0.68	0.42	-	-	-0.01	-	-	0.22	0.50	0.13	0.20
Imp.8	-0.52	0.57	0.36	-	-	-	-	-	0.00	0.42	-	0.35
Imp.9	-0.59	0.64	0.33	-	0.01	-	-	-	-	0.58	-	-
Imp.10	-0.55	0.70	0.23	-	-	-	-	-	0.02	0.43	0.01	0.27
Imp.11	-0.65	0.78	0.04	0.17	0.05	-	-	-	0.10	0.41	0.07	0.00
Imp.12	-0.71	0.69	0.34	-	-	-	-	-	0.14	0.49	0.09	0.19
Imp.13	-0.54	0.67	0.31	-	0.03	-	-	-	0.01	0.43	-	0.27
Imp.14	-0.76	0.63	0.25	0.15	0.09	-	-	0.05	0.18	0.52	0.08	0.19
Imp.15	-0.61	0.72	0.35	-	-	-	-	-	0.02	0.54	0.05	0.23
Imp.16	-0.73	0.59	0.33	0.12	-	-	-	-	0.21	0.50	0.02	0.24
Imp.17	-0.68	0.68	0.35	0.00	0.04	-	-	-	0.08	0.52	0.09	0.32
Imp.18	-0.53	0.60	0.31	-	0.11	-	-	-	-	0.41	-	0.21
Imp.19	-0.65	0.71	0.30	-	-	-	-	-	0.09	0.44	0.06	-
Imp.20	-0.86	0.72	0.28	-	-	-	-	-	0.22	0.60	0.18	0.26
MCE	-0.63	0.66	0.31	-	0.01	-	-	-	0.06	0.48	0.01	0.23

Approximate MALCoM linear probability models

Table A12: Estimated aggregated coefficients from each modeling framework using linear probability models (estimates from MCE, MCM, and MCL models are approximating the models presented in Table 4 of the main manuscript, and all represent changes in estimated probability). MCL: Median coefficient lasso; MCE: Median coefficient elastic net; MCM: Median coefficient minimax concave penalty; BMI: Body Mass Index; LDH: Lactic acid dehydrogenase; CRP: C-reactive protein; NLR: Neutrophil-lymphocyte ratio.

	MCE	MCM	MCL	MI-lasso*
Intercept	0.37	0.38	0.38	0.39
CRP (Z-score)	0.14	0.20	0.14	0.13
LDH (Z-score)	0.06	0.01	0.06	0.04
Log D-dimer	-	-	-	-
Log ferritin	0.00	-	-	-
Symptom duration	-	-	-	-
Age	-	-	-	-
BMI	-	-	-	-
Nonwhite	0.00	-	-	-
Diabetes	0.09	0.09	0.09	0.06
Male	-	-	-	-
NLR	0.04	-	0.04	0.02

*