

Supplemental Information for:

Development of an ensemble machine learning prognostic model to predict 60-day risk of major adverse cardiac events in adults with chest pain

Version: March 8, 2021

A Appendix

A.1 Predictor summary

This table summarizes the predictors used in the model following missing value imputation and histogram condensing of unique values for high-cardinality variables. This table includes 71 total predictors but totals 74 predictors when the race predictor is decomposed from a categorical predictor into 4 indicator predictors.

Table A1. Summary of predictors

	Name	Unique values	Mode	Mean	Median	Minimum	Maximum
Biomarkers							
1	ECG	3	0	0	0	0	2
2	Peak troponin	6	0	0	0	0	0
3	Troponin 3HV	29	0	0	0	0	0
4	Troponin categories	2	1	1	1	1	2
Scores							
5	EDACS	41	8	11	12	-8	34
6	HEART	9	4	4	4	0	8
7	High EDACS	2	0	0	0	0	1
8	High HEART	2	1	1	1	0	1
Labs							
9	GFR	96	70	66	70	2	70
10	HbA1c	83	6	6	6	4	15
11	HDL	148	48	51	50	4	213
12	LDL	137	111	104	103	8	546
13	Total Cholesterol	104	186	181	181	33	1075
14	Triglycerides	86	5	5	5	3	9
Vitals							
15	BMI	100	30	30	28	12	82
16	Lowest SBP	174	120	121	120	42	244
17	O2 saturation	45	100	98	98	16	100
18	Obesity	2	0	0	0	0	1
19	Peak pulse	153	77	84	82	24	269
20	Pulse	168	74	80	78	12	256
21	Respiration	67	18	18	18	2	100
22	SBP	155	145	145	143	56	269
Demographics							
23	Age	87	53	59	59	18	104
24	Male	2	0	0	0	0	1

Table A1. Summary of predictors

	Name	Unique values	Mode	Mean	Median	Minimum	Maximum
25	Race	5	White	-	-	-	-
Clinical notes							
26	Diaphoresis	3	0	0	0	0	9
27	Exertion	3	0	0	0	0	9
28	Pain on inspiration	3	0	0	0	0	9
29	Pain on palpation	3	0	0	0	0	9
30	Radiating pain	6	0	1	0	0	9
31	Sharp pain	3	0	0	0	0	9
History							
32	Anxiety	2	0	0	0	0	1
33	Aortic athero.	2	0	0	0	0	1
34	APD	2	0	0	0	0	1
35	CAD	2	0	0	0	0	1
36	Coronary rev.	2	0	0	0	0	1
37	Diabetes	2	0	0	0	0	1
38	Family history of CAD	2	0	0	0	0	1
39	Hypercholesteremia	2	1	1	1	0	1
40	Hypertension	2	1	1	1	0	1
41	Myocardial infarction	2	0	0	0	0	1
42	PAD	2	0	0	0	0	1
43	Pre-diabetes	2	0	0	0	0	1
44	Prior catheter	2	0	0	0	0	1
45	Prior CT	2	0	0	0	0	1
46	Prior echo	2	0	0	0	0	1
47	Prior MPI	2	0	0	0	0	1
48	Prior treadmill	2	0	0	0	0	1
49	Prior UTL	2	0	0	0	0	1
50	Smoking	2	0	0	0	0	1
51	Stroke	2	0	0	0	0	1
Missingness indicators							
52	Missing BMI	2	0	0	0	0	1
53	Missing Diaphoresis	2	0	0	0	0	1
54	Missing ECG	2	0	0	0	0	1
55	Missing Exertion	2	1	1	1	0	1
56	Missing GFR	2	0	0	0	0	1
57	Missing HbA1c	2	1	1	1	0	1
58	Missing HDL	2	0	0	0	0	1
59	Missing LDL	2	0	0	0	0	1
60	Missing O2 Saturation	2	0	0	0	0	1
61	Missing Obesity	2	0	0	0	0	1
62	Missing Pain on Inspiration	2	1	1	1	0	1
63	Missing Pain on Palpation	2	1	1	1	0	1
64	Missing Pulse	2	0	0	0	0	1
65	Missing Radiating Pain	2	0	0	0	0	1
66	Missing Respiration	2	0	0	0	0	1
67	Missing SBP	2	0	0	0	0	1
68	Missing Sharp Pain	2	1	1	1	0	1
69	Missing Total Cholesterol	2	0	0	0	0	1

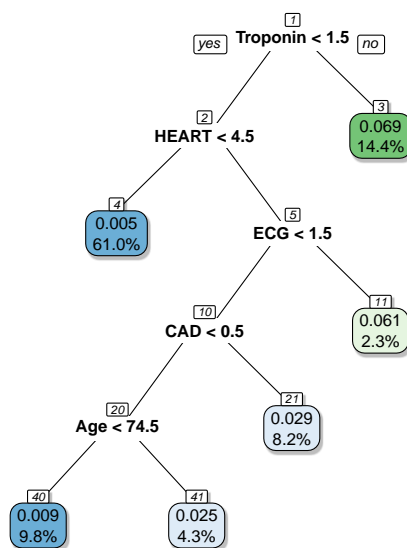
Table A1. Summary of predictors

	Name	Unique values	Mode	Mean	Median	Minimum	Maximum
70	Missing Triglycerides	2	0	0	0	0	1
71	Missing Troponin 3HV	2	0	0	0	0	1

A.2 Decision tree benchmark

A single decision tree is a helpful benchmark approach to risk prediction. It is easy to explain and to visualize, and represents a decision-making process that can be conducted by medical staff. It is limited by its simplicity though: decision trees rarely have competitive predictive performance and the selected variables & cutpoints are known to be sensitive to small changes to the dataset. Figure A1 shows an example tree as estimated on our data.

Figure A1. Single decision tree, where each leaf node shows 1) predicted risk, and 2) the percentage of patients that fall into that node. Here Troponin refers to peak Troponin.

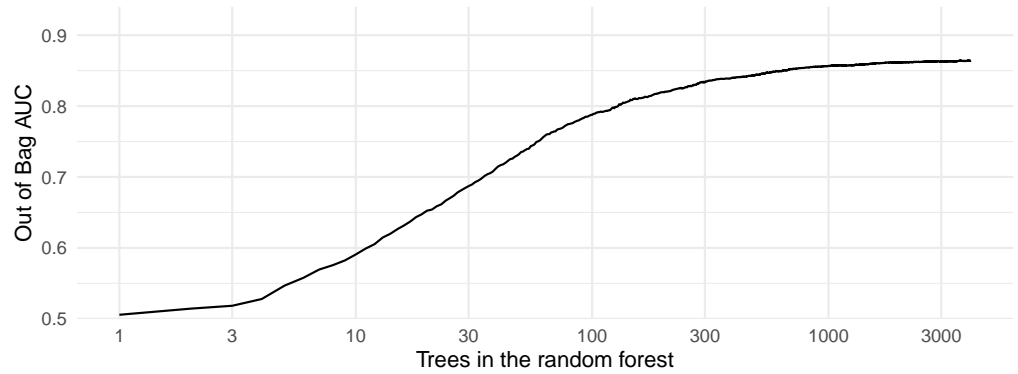


Of particular clinical interest is the blue leaf node (#4) showing a predicted risk of 0.5% and encompassing 61% of the sample. That node represents the lowest-risk group as identified by the decision tree. On first glance the identification of the low-risk group as comprising 61% of patients would seem to improve upon the 50% low-risk result from our complex nested ensemble. However, the 0.5% estimated risk for that leaf node is an average across the large proportion of patients that met the two simple decisions: non-elevated troponin and HEART score ≤ 4 . We know that every such patient doesn't have exactly 0.5% risk - there is a distribution around that average, with some patients have lower risk and others having higher risk. Thus we can see that this low-risk node must suffer from some miscalibration: an unknown proportion of those patients have higher risk than our target 0.5% threshold, and should not be discharged early despite the recommendation of the decision tree.

A.3 Random forest convergence plot

The number of trees is a key hyperparameter for random forest. Breiman (2001) proved that adding additional trees does not lead to overfitting due to the independence of the tree estimators, unlike the sequential fitting of boosting approaches. Yet it remains beneficial to examine the convergence of the random forest's performance to ensure that an adequate number of trees is being used Probst et al. (2017). To assess performance convergence we analyzed the out-of-bag ROC-AUC for up to 3,000 trees, shown in Figure A2. From visual inspection we can see that only 100 trees would be inadequate, whereas 1,000 - 3,000 trees achieves an ROC-AUC above 0.85.

Figure A2. Random forest convergence plot shows ROC-AUC on out-of-bag data as forest size is increased exponentially.



A.4 Precision-Recall AUC table

Table A2. Cross-validated PR-AUC discrimination performance

Learner	PR-AUC	Std. Err.	CI Lower	CI Upper
Marginal mean	0.0188	0.0000	0.0187	0.0188
Stratification (EDACS)	0.0456	0.0018	0.0422	0.0491
Stratification (ECG)	0.0468	0.0019	0.0430	0.0506
Decision tree (trop.)	0.0509	0.0015	0.0479	0.0539
Decision tree (trop., ECG)	0.0509	0.0015	0.0479	0.0539
Decision tree pruned	0.0509	0.0015	0.0479	0.0539
Decision tree (trop., ECG, scores)	0.0583	0.0012	0.0560	0.0607
Decision tree	0.0587	0.0012	0.0563	0.0611
Logistic reg. (trop.)	0.0635	0.0030	0.0576	0.0695
Logistic reg. (trop., ECG)	0.0635	0.0030	0.0576	0.0695
Stratification (trop.)	0.0639	0.0030	0.0581	0.0698
Stratification (HEART)	0.0643	0.0030	0.0585	0.0700
Stratification (trop., ECG, EDACS)	0.0871	0.0030	0.0812	0.0930
Stratification (trop., ECG)	0.0887	0.0048	0.0792	0.0982
Decision tree tuned	0.0944	0.0063	0.0820	0.1068
Logistic reg. (trop., ECG, scores)	0.0960	0.0034	0.0894	0.1026
Stratification (trop., ECG, HEART)	0.1011	0.0069	0.0876	0.1146
Linear reg.	0.1119	0.0066	0.0990	0.1248
Lasso	0.1195	0.0080	0.1038	0.1351
Logistic reg.	0.1199	0.0087	0.1029	0.1369
XGBoost	0.1239	0.0086	0.1071	0.1407
Splines	0.1250	0.0091	0.1072	0.1428
BART	0.1321	0.0090	0.1145	0.1496
XGBoost tuned	0.1387	0.0115	0.1161	0.1613
Random forest	0.1438	0.0103	0.1235	0.1641
Random forest tuned	0.1457	0.0105	0.1252	0.1662
SuperLearner	0.1479	0.0117	0.1250	0.1709

A.5 Receiver Operating Characteristic AUC table**A.6 Index of prediction accuracy table**

Table A3. Cross-validated ROC-AUC discrimination performance

Learner	ROC-AUC	Std. Err.	CI Lower	CI Upper
Marginal mean	0.5000	0.0108	0.4789	0.5211
Stratification (ECG)	0.6864	0.0084	0.6699	0.7029
Decision tree (trop.)	0.6983	0.0105	0.6777	0.7190
Decision tree (trop., ECG)	0.6983	0.0105	0.6777	0.7190
Decision tree pruned	0.6983	0.0105	0.6777	0.7190
Logistic reg. (trop.)	0.7001	0.0084	0.6837	0.7166
Logistic reg. (trop., ECG)	0.7001	0.0084	0.6837	0.7166
Stratification (trop.)	0.7082	0.0084	0.6918	0.7246
Stratification (EDACS)	0.7325	0.0053	0.7221	0.7428
Stratification (HEART)	0.7647	0.0054	0.7541	0.7753
Stratification (trop., ECG)	0.7707	0.0067	0.7577	0.7838
Stratification (trop., ECG, EDACS)	0.7715	0.0058	0.7601	0.7829
Decision tree (trop., ECG, scores)	0.7820	0.0069	0.7685	0.7956
Decision tree	0.7835	0.0069	0.7699	0.7971
Decision tree tuned	0.8052	0.0053	0.7948	0.8156
Logistic reg. (trop., ECG, scores)	0.8058	0.0045	0.7970	0.8146
Stratification (trop., ECG, HEART)	0.8117	0.0048	0.8023	0.8212
Linear reg.	0.8379	0.0044	0.8292	0.8467
Logistic reg.	0.8422	0.0040	0.8344	0.8500
Lasso	0.8427	0.0039	0.8350	0.8504
XGBoost	0.8430	0.0039	0.8353	0.8507
Splines	0.8486	0.0039	0.8411	0.8562
Random forest	0.8571	0.0038	0.8498	0.8645
XGBoost tuned	0.8594	0.0036	0.8523	0.8665
BART	0.8609	0.0036	0.8539	0.8679
Random forest tuned	0.8618	0.0036	0.8547	0.8688
SuperLearner	0.8665	0.0035	0.8596	0.8734

Table A4. Cross-validated index of prediction accuracy for each learner and the ensemble

Learner	IPA	Std. Err.	CI Lower	CI Upper
Marginal mean	0.0000	0.0000	0.0000	0.0000
Stratification (EDACS)	0.0138	0.0010	0.0118	0.0158
Stratification (ECG)	0.0153	0.0015	0.0125	0.0182
Logistic reg. (trop.)	0.0189	0.0020	0.0149	0.0229
Logistic reg. (trop., ECG)	0.0189	0.0020	0.0149	0.0229
Decision tree (trop.)	0.0234	0.0016	0.0203	0.0266
Decision tree (trop., ECG)	0.0234	0.0016	0.0203	0.0266
Decision tree pruned	0.0234	0.0016	0.0203	0.0266
Stratification (HEART)	0.0243	0.0017	0.0210	0.0277
Stratification (trop., ECG, EDACS)	0.0253	0.0038	0.0179	0.0327
Stratification (trop.)	0.0273	0.0022	0.0230	0.0317
Decision tree (trop., ECG, scores)	0.0278	0.0011	0.0256	0.0301
Decision tree	0.0283	0.0010	0.0263	0.0302
Stratification (trop., ECG)	0.0384	0.0029	0.0327	0.0440
Decision tree tuned	0.0393	0.0024	0.0345	0.0441
Logistic reg. (trop., ECG, scores)	0.0408	0.0018	0.0371	0.0444
XGBoost	0.0416	0.0092	0.0237	0.0596
Stratification (trop., ECG, HEART)	0.0433	0.0041	0.0352	0.0514
Linear reg.	0.0458	0.0018	0.0423	0.0492
Logistic reg.	0.0539	0.0047	0.0447	0.0631
Lasso	0.0544	0.0041	0.0463	0.0625
Splines	0.0565	0.0054	0.0459	0.0672
BART	0.0641	0.0047	0.0548	0.0733
XGBoost tuned	0.0660	0.0062	0.0539	0.0782
DiscreteSL	0.0661	0.0040	0.0582	0.0740
Random forest	0.0673	0.0035	0.0603	0.0742
Random forest tuned	0.0682	0.0029	0.0626	0.0738
SuperLearner	0.0725	0.0045	0.0637	0.0813

A.7 Brier score table

Table A5. Cross-validated Brier score for each learner and the ensemble

Learner	Brier score	Std. Err.	CI Lower	CI Upper
Marginal mean	0.01840	0.00001	0.01839	0.01842
Stratification (EDACS)	0.01815	0.00001	0.01812	0.01818
Stratification (ECG)	0.01812	0.00003	0.01807	0.01817
Logistic reg. (trop.)	0.01806	0.00004	0.01797	0.01814
Logistic reg. (trop., ECG)	0.01806	0.00004	0.01797	0.01814
Decision tree (trop.)	0.01797	0.00003	0.01791	0.01804
Decision tree (trop., ECG)	0.01797	0.00003	0.01791	0.01804
Decision tree pruned	0.01797	0.00003	0.01791	0.01804
Stratification (HEART)	0.01796	0.00003	0.01790	0.01801
Stratification (trop., ECG, EDACS)	0.01794	0.00006	0.01781	0.01807
Stratification (trop.)	0.01790	0.00004	0.01781	0.01799
Decision tree (trop., ECG, scores)	0.01789	0.00002	0.01785	0.01793
Decision tree	0.01788	0.00002	0.01785	0.01792
Stratification (trop., ECG)	0.01770	0.00006	0.01759	0.01781
Decision tree tuned	0.01768	0.00005	0.01759	0.01777
Logistic reg. (trop., ECG, scores)	0.01765	0.00003	0.01759	0.01772
XGBoost	0.01764	0.00017	0.01730	0.01798
Stratification (trop., ECG, HEART)	0.01761	0.00008	0.01745	0.01776
Linear reg.	0.01756	0.00003	0.01750	0.01763
Logistic reg.	0.01741	0.00009	0.01724	0.01758
Lasso	0.01740	0.00008	0.01725	0.01755
Splines	0.01736	0.00010	0.01717	0.01756
BART	0.01722	0.00009	0.01705	0.01740
XGBoost tuned	0.01719	0.00012	0.01696	0.01742
DiscreteSL	0.01719	0.00008	0.01703	0.01734
Random forest	0.01717	0.00007	0.01703	0.01730
Random forest tuned	0.01715	0.00006	0.01704	0.01726
SuperLearner	0.01707	0.00009	0.01690	0.01724

A.8 Grouped calibration table

Decile	Predicted risk	Observed risk	Predicted / Actual	Predicted - Actual
1	0.0005	0.0002	2.7813	0.0003
2	0.0011	0.0008	1.3707	0.0003
3	0.0018	0.0015	1.2685	0.0004
4	0.0031	0.0025	1.2329	0.0006
5	0.0050	0.0039	1.3012	0.0012
6	0.0080	0.0071	1.1301	0.0009
7	0.0129	0.0116	1.1115	0.0013
8	0.0211	0.0198	1.0660	0.0013
9	0.0376	0.0337	1.1159	0.0039
10	0.1041	0.1067	0.9759	-0.0026

References

Breiman, Leo (2001). “Random forests”. In: *Machine learning* 45.1, pp. 5–32.

Probst, Philipp and Anne-Laure Boulesteix (2017). “To tune or not to tune the number of trees in random forest”. In: *The Journal of Machine Learning Research* 18.1, pp. 6673–6690.