The MEANS Procedure

Variable	Median	N Miss
newethnic	2.0	0
READMIT	2.0	0
BOIL	0.0	0
CIRR	0.0	0
CPD11	0.0	0
RENAL	0.0	0
CSBREAK9	0.0	0
HEART	0.0	0
CTD11	0.0	0
SMOKER	0.0	0
CVA	0.0	0
CYSTIC9	0.0	0
DULCER7	0.0	0
DEMENT9	0.0	0
DIABETES	0.0	0
HEMAP9	0.0	0
HIV	0.0	0
FLU	0.0	0
IVDU	0.0	0
TUMOR9	0.0	0
MI11	0.0	0
OBESITY	0.0	0
DRUG7	0.0	0
PEPTIC9	0.0	0
PVD	0.0	0
SSABC	0.0	0
SSAIW	0.0	0
SSBI	0.0	0
SSCEL	0.0	0
SSCHR	0.0	0
SSHER	0.0	0
SSINF	0.0	0
SSMAS	0.0	0
SSMYO	0.0	0
SSNF	0.0	0
SSPUS	0.0	0
CLINDA	1.0	0
OX	1.0	0
SXT	1.0	0
VANCO	1.0	0

Cluster	Variable	1 - RSquare Ratio
Cluster 1	CACASE	0.0482
	HACO_onset	0.2135
	anyhealth2	0.0717
	comm_onset	0.1321
	newhospyrd8	0.3442
Cluster 2	HO	0.0673
	PRIVRES9	0.7897
	TRANSF9	0.0873
	hosp_onset	0.0673
	onset_type	0.8293
Cluster 3	BLACK	0.0870
	WHITE	0.0676
	newBLACK	0.0315
	newWHITE	0.0308
Cluster 4	BLOOD	0.0191
	BSI	0.0178
Cluster 5	LTCF9	0.0615
	LTCYR	0.0593
Cluster 6	COLLECT	0.3302
	HOSPITAL	0.2941
Cluster 7	CAUSAL11	0.1270
	OUTCOME	0.1316
Cluster 8	IAB	0.0881
	PERITNL	0.0908
Cluster 9	LTACH10	0.1411
	LTACYR11	0.1311
Cluster 10	age_category	0.0338
	newage	0.0339
Cluster 11	CDIAL9	0.1623
	RENAL	0.1412
Cluster 12	DRUG7	0.3181
	SMOKER	0.3154
Cluster 13	OTHSITE	0.3609
	SSI	0.3903
Cluster 14	PREVSA12mo	0.2620
	priorinvasive	0.2476
Cluster 15	MI11	0.0000

Cluster	Variable	1 - RSquare Ratio
Cluster 16	UTIT	0.0000
Cluster 17	HEMAP9	0.0000
Cluster 18	ICU16	0.0000
Cluster 19	newethnic	0.0000
Cluster 20	PLEURAL	0.0000
Cluster 21	CSF	0.0000
Cluster 22	SEX	0.0000
Cluster 23	BOIL	0.0000
Cluster 24	mrsafinal	0.0000
Cluster 25	CSBREAK9	0.0000
Cluster 26	SST	0.0000
Cluster 27	CTD11	0.0000
Cluster 28	ethnicity_real	0.0000
Cluster 29	OTHERrace	0.0000
Cluster 30	HOMELESS9	0.0000
Cluster 31	INCERC9	0.0000
Cluster 32	OBESITY	0.0000
Cluster 33	HIV	0.0000
Cluster 34	OX	0.0000
Cluster 35	CNS	0.0000
Cluster 36	TUMOR9	0.0000
Cluster 37	BONE	0.0000
Cluster 38	UNKNOWN	0.0000
Cluster 39	UND	0.0000
Cluster 40	CIRR	0.0000
Cluster 41	BODYSITE	0.0000
Cluster 42	CPD11	0.0000
Cluster 43	SXT	0.0000
Cluster 44	SURGYR9	0.0000
Cluster 45	IVDU	0.0000
Cluster 46	UNIQUEID	0.0000
Cluster 47	PVD	0.0000
Cluster 48	CVI	0.0000
Cluster 49	DISLTACH10	0.0000
Cluster 50	PNE	0.0000
Cluster 51	PEPTIC9	0.0000
Cluster 52	HEMMALIG	0.0000

Cluster	Variable	1 - RSquare Ratio
Cluster 53	CVA	0.0000
Cluster 54	CVC9	0.0000
Cluster 55	NO_UI	0.0000
Cluster 56	DIABETES	0.0000
Cluster 57	CLINDA	0.0000
Cluster 58	DEMENT9	0.0000
Cluster 59	HEART	0.0000
Cluster 60	DISLTC8	0.0000
Cluster 61	VANCO	0.0000
Cluster 62	BJI	0.2288
	JOINT	0.2915
Cluster 63	UNKRACE	0.0000
Cluster 64	OTHPOS	0.0000
Cluster 65	DULCER7	0.0000

Variation Explained by Clusters

| Obs	Number of Clusters	Total Variation Explained by Clusters	Proportion of Variation Explained by Clusters	Minimum Proportion Explained by a Cluster	Maximum Second Eigenvalue in a Cluster	Minimum R-squared for a Variable	Maximum 1-R**2 Ratio for a Variable
1	1	7.826282	0.0889	0.0889	4.496439	0.0000	_
2	2	11.400563	0.1296	0.1113	4.007613	0.0001	0.9999
3	3	14.690978	0.1669	0.1262	3.727421	0.0001	1.0021
4	4	18.043838	0.2050	0.1284	2.816162	0.0001	1.1147
5	5	20.676847	0.2350	0.1842	1.979717	0.0001	1.1147
6	6	22.346766	0.2539	0.1842	1.976421	0.0001	1.1147
7	7	24.193717	0.2749	0.1788	1.787694	0.0004	1.1147
8	8	25.950531	0.2949	0.1933	1.762313	0.0004	1.1147
9	9	27.681910	0.3146	0.1933	1.697497	0.0004	1.1147
10	10	29.237321	0.3322	0.2428	1.579660	0.0004	1.1147
11	11	30.770753	0.3497	0.2428	1.472444	0.0004	1.1408
12	12	32.202047	0.3659	0.2428	1.422979	0.0004	1.1408
13	13	33.431390	0.3799	0.2769	1.364355	0.0004	1.1408
14	14	34.776784	0.3952	0.2769	1.252496	0.0004	1.1448
15	15	36.019352	0.4093	0.2616	1.241521	0.0004	1.1448
16	16	37.197131	0.4227	0.2616	1.220855	0.0004	1.1448
17	17	38.376514	0.4361	0.2616	1.169528	0.0004	1.1448
18	18	39.380766	0.4475	0.2616	1.168415	0.0004	1.1448
19	19	40.499701	0.4602	0.2616	1.167794	0.0004	1.1448
20	20	41.485512	0.4714	0.2616	1.137419	0.0004	1.1448
21	21	42.452779	0.4824	0.2616	1.126272	0.0004	1.4329
22	22	43.528777	0.4946	0.2616	1.064513	0.0004	1.4329
23	23	44.538767	0.5061	0.2616	1.048177	0.0004	1.4329
24	24	45.519647	0.5173	0.2616	1.044080	0.0004	1.4306
25	25	46.518618	0.5286	0.2616	1.037603	0.0004	1.4306
26	26	47.543814	0.5403	0.2616	1.035046	0.0004	1.4306
27	27	48.545231	0.5517	0.2616	1.034887	0.0004	1.4306
28	28	49.520609	0.5627	0.2616	1.027328	0.0004	1.4306
29	29	50.542513	0.5743	0.2616	1.026831	0.0004	1.4306
30	30	51.533298	0.5856	0.3121	1.002478	0.0004	1.4306
31	31	52.533107	0.5970	0.3121	1.001441	0.0151	1.4306
32	32	53.498300	0.6079	0.3725	0.999344	0.0151	1.4306
33	33	54.491474	0.6192	0.3725	0.996676	0.0151	1.4306
34	34	55.478543	0.6304	0.3725	0.994650	0.0151	1.4306
35	35	56.469414	0.6417	0.3725	0.993695	0.0681	1.1468

Variation Explained by Clusters

Obs	Number of Clusters	Total Variation Explained by Clusters	Proportion of Variation Explained by Clusters	Minimum Proportion Explained by a Cluster	Maximum Second Eigenvalue in a Cluster	Minimum R-squared for a Variable	Maximum 1-R**2 Ratio for a Variable
36	36	57.434517	0.6527	0.4320	0.981761	0.0681	1.1468
37	37	58.346028	0.6630	0.4320	0.974079	0.0681	1.1468
38	38	59.320108	0.6741	0.4320	0.972552	0.0681	1.1468
39	39	60.292659	0.6851	0.4320	0.967246	0.0681	1.1468
40	40	61.259826	0.6961	0.4320	0.960857	0.0844	1.1468
41	41	62.219367	0.7070	0.4320	0.960815	0.0844	1.1468
42	42	63.180116	0.7180	0.4320	0.956331	0.1807	1.1468
43	43	64.136447	0.7288	0.4320	0.955898	0.1807	1.1468
44	44	65.092345	0.7397	0.4320	0.942659	0.1807	1.1468
45	45	66.035004	0.7504	0.4320	0.902093	0.1807	1.1468
46	46	66.914858	0.7604	0.4501	0.894837	0.1807	1.1468
47	47	67.801834	0.7705	0.4996	0.891708	0.1807	1.1468
48	48	68.690133	0.7806	0.4996	0.882798	0.1807	1.1947
49	49	69.572930	0.7906	0.4996	0.874489	0.1807	1.1947
50	50	70.447419	0.8005	0.4996	0.871773	0.1807	1.1947
51	51	71.319192	0.8104	0.4996	0.869685	0.1807	1.1947
52	52	72.188877	0.8203	0.4996	0.856161	0.1807	1.1947
53	53	73.034558	0.8299	0.5174	0.856145	0.1807	1.1947
54	54	73.888628	0.8396	0.5174	0.838114	0.2559	0.8293
55	55	74.700295	0.8489	0.5174	0.824642	0.3118	0.8293
56	56	75.524937	0.8582	0.5174	0.824058	0.3118	0.8293
57	57	76.348995	0.8676	0.5174	0.804742	0.3118	0.8293
58	58	77.142831	0.8766	0.5174	0.762855	0.3748	0.8293
59	59	77.905686	0.8853	0.5174	0.759350	0.3748	0.8293
60	60	78.663872	0.8939	0.5174	0.759210	0.4804	0.8293
61	61	79.366905	0.9019	0.5174	0.755920	0.4804	0.8293
62	62	80.089686	0.9101	0.5174	0.743625	0.4804	0.8293
63	63	80.831735	0.9185	0.6471	0.705752	0.4804	0.8293
64	64	81.537487	0.9266	0.6482	0.703646	0.4804	0.8293
65	65	82.241133	0.9346	0.6553	0.689481	0.4804	0.8293

Rank of Spearman Correlations and Hoeffding Correlations

Obs	variable	Spearman rank of variables	Hoeffding rank of variables	Spearman Correlation	Spearman p-value	Hoeffding Correlation	Hoeffding p-value
1	BSI	1.0	1.0	0.51767	<.0001	0.02334	<.0001
2	newethnic	2.0	3.0	-0.45172	<.0001	0.01739	<.0001
3	CACASE	3.0	2.0	-0.44123	<.0001	0.02212	<.0001
4	ICU16	4.0	5.0	0.43647	<.0001	0.01156	<.0001
5	UNKRACE	5.0	7.0	-0.38760	<.0001	0.00978	<.0001
6	UND	6.0	6.0	-0.36655	<.0001	0.01055	<.0001
7	DISLTC8	7.0	11.0	0.31927	<.0001	0.00385	<.0001
8	hosp_onset	8.0	15.0	0.30836	<.0001	0.00286	<.0001
9	DIABETES	9.0	9.0	0.29768	<.0001	0.00715	<.0001
10	RENAL	10.0	12.0	0.27465	<.0001	0.00333	<.0001
11	CVC9	11.0	18.0	0.27221	<.0001	0.00214	0.0004
12	PNE	12.0	21.0	0.26829	<.0001	0.00152	0.0027
13	OTHPOS	13.0	13.0	-0.26725	<.0001	0.00324	<.0001
14	newage	14.0	4.0	0.26693	<.0001	0.01364	<.0001
15	SST	15.0	8.0	-0.26510	<.0001	0.00745	<.0001
16	HEART	16.0	22.0	0.21918	<.0001	0.00102	0.0119
17	mrsafinal	17.0	10.0	0.21820	<.0001	0.00430	<.0001
18	BJI	18.0	23.0	0.20729	<.0001	0.00102	0.0121
19	LTCYR	19.0	53.0	0.20719	<.0001	0.00047	0.0661
20	SURGYR9	20.0	19.0	0.20643	<.0001	0.00172	0.0015
21	SMOKER	21.0	20.0	0.19570	<.0001	0.00167	0.0018
22	CVA	22.0	59.0	0.19036	<.0001	0.00014	0.2103
23	VANCO	23.0	60.0	-0.18269	<.0001	0.00008	0.2612
24	DULCER7	24.0	56.0	0.17418	<.0001	-0.00023	0.8415
25	newWHITE	25.0	16.0	-0.16658	<.0001	0.00247	0.0002
26	OX	26.0	17.0	0.15355	<.0001	0.00238	0.0002
27	OBESITY	27.0	61.0	0.14506	<.0001	0.00004	0.3091
28	CVI	28.0	48.0	0.14460	<.0001	-0.00063	1.0000
29	DEMENT9	29.0	49.0	0.13686	<.0001	-0.00063	1.0000
30	MI11	30.0	55.0	0.13560	<.0001	-0.00039	1.0000
31	CIRR	31.0	51.0	0.13077	<.0001	-0.00058	1.0000
32	PVD	32.0	52.0	0.12901	<.0001	-0.00048	1.0000
33	CAUSAL11	33.0	46.0	0.11685	<.0001	-0.00071	1.0000
34	DISLTACH10	34.0	44.0	0.11550	<.0001	-0.00073	1.0000
35	HEMAP9	35.5	42.5	0.10565	<.0001	-0.00074	1.0000
36	IVDU	35.5	42.5	0.10565	<.0001	-0.00074	1.0000
37	HOMELESS9	37.0	45.0	0.10227	<.0001	-0.00073	1.0000

Obs	variable	Spearman rank of variables	Hoeffding rank of variables	Spearman Correlation	Spearman p-value	Hoeffding Correlation	Hoeffding p-value
38	CPD11	38.0	58.0	0.10082	<.0001	-0.00020	0.7751
39	IAB	39.0	37.0	0.09772	<.0001	-0.00079	1.0000
40	TUMOR9	40.0	40.0	0.09407	0.0001	-0.00075	1.0000
41	priorinvasive	41.0	47.0	0.08157	0.0009	-0.00068	1.0000
42	CNS	42.0	33.0	0.08130	0.0010	-0.00082	1.0000
43	CTD11	43.0	36.0	0.07336	0.0029	-0.00079	1.0000
44	CSF	44.0	32.0	0.07274	0.0031	-0.00084	1.0000
45	OTHERrace	45.0	39.0	-0.06772	0.0060	-0.00076	1.0000
46	SXT	46.0	54.0	-0.06731	0.0063	-0.00045	1.0000
47	SEX	47.0	57.0	-0.06011	0.0147	-0.00022	0.8324
48	PEPTIC9	48.0	31.0	0.05551	0.0243	-0.00084	1.0000
49	PLEURAL	49.0	26.0	0.05415	0.0280	-0.00085	1.0000
50	OTHSITE	50.0	30.0	0.05065	0.0399	-0.00084	1.0000
51	HIV	51.0	38.0	0.04147	0.0925	-0.00076	1.0000
52	HEMMALIG	52.0	14.0	0.03963	0.3630	-0.00324	1.0000
53	BONE	53.0	29.0	0.03573	0.1472	-0.00085	1.0000
54	BODYSITE	54.0	24.0	0.02419	0.3266	-0.00086	1.0000
55	CSBREAK9	55.0	41.0	0.02130	0.3878	-0.00074	1.0000
56	LTACYR11	56.0	27.0	-0.01589	0.5192	-0.00085	1.0000
57	ethnicity_real	57.0	34.0	0.01275	0.6051	-0.00081	1.0000
58	CLINDA	58.0	50.0	-0.00623	0.8005	-0.00063	1.0000
59	BOIL	59.0	28.0	-0.00198	0.9360	-0.00085	1.0000
60	UTIT	60.0	35.0	0.00168	0.9456	-0.00080	1.0000
61	INCERC9	61.0	25.0	-0.00066	0.9788	-0.00086	1.0000

Scatter Plot of the Ranks of Spearman vs. Hoeffding
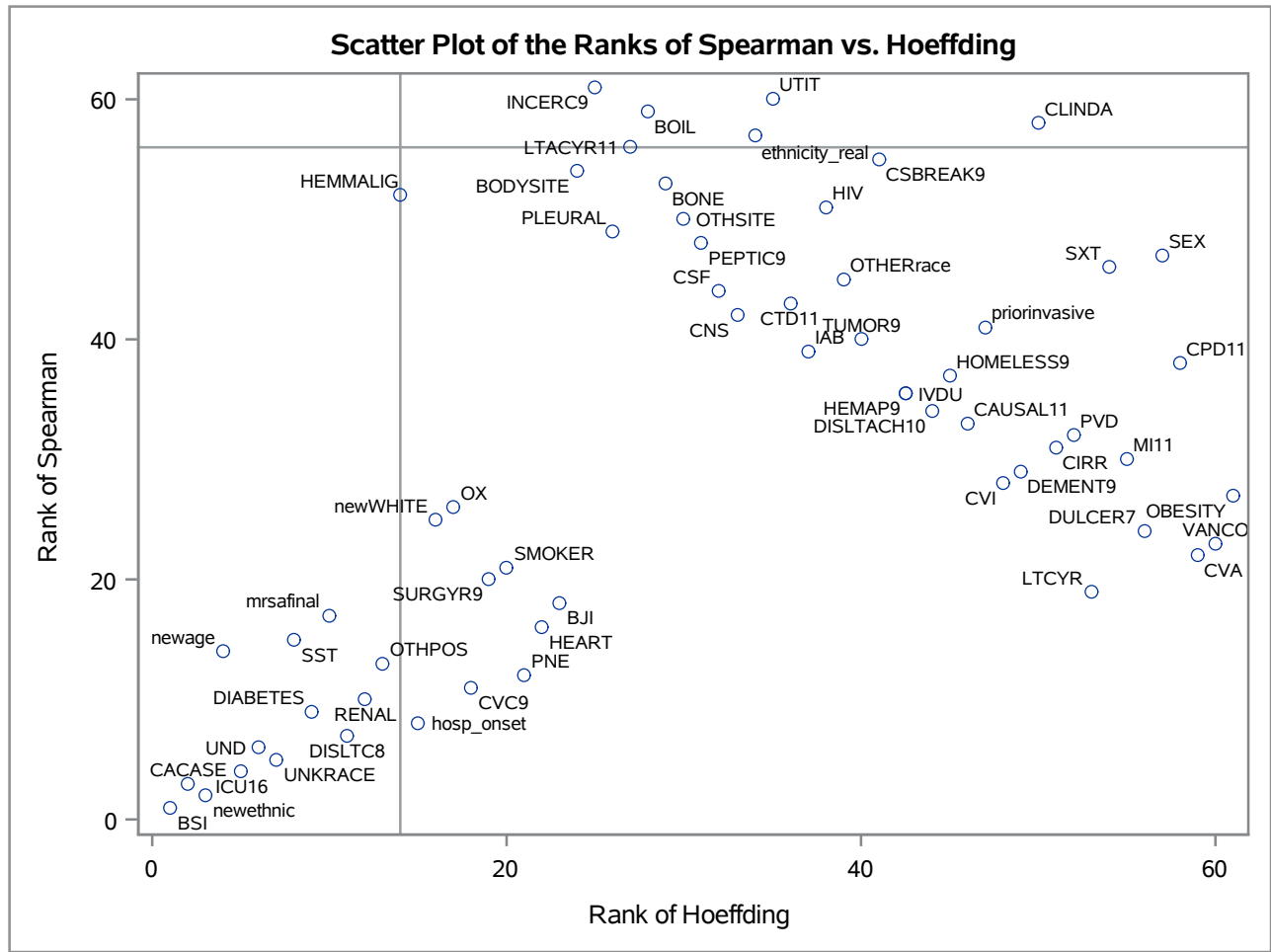
Row	variable
1	BSI
2	newethnic
3	CACASE
4	ICU16
5	UNKRACE
6	UND
7	DISLTC8
8	DIABETES
9	RENAL

The HPGENSELECT Procedure

Performance Information	
Execution Mode	Single-Machine
Number of Threads	2

Data Access Information			
Data	**Engine**	**Role**	**Path**
WORK.SAMPLE	V9	Input	On Client
WORK.PRED	V9	Output	On Client

The HPGENSELECT Procedure

CACASE=0

Model Information	
Data Source	WORK.SAMPLE
Response Variable	HOSPITAL
Weight Variable	WEIGHT
Class Parameterization	GLM
Distribution	Binary
Link Function	Logit
Optimization Technique	Nesterov

Selection Information	
Selection Method	LASSO
Stop Criterion	None
Choose Criterion	Validation ASE
Effect Hierarchy Enforced	None

Number of Observations			
Description	Total	Training	Validation
Number of Observations Read	794	557	237
Number of Observations Used	794	557	237

Response Profile				
Ordered Value	HOSPITAL	Total Frequency	Training	Validation
1	0	208	537	238
2	1	586	917	374

You are modeling the probability that HOSPITAL='1'.

Class Level Information		
Class	Levels	Values
UND	2	1 0
DISLTC8	2	1 0
DIABETES	2	1 0
RENAL	2	1 0
SST	2	1 0
mrsafinal	2	1 0
BSI	2	1 0

Model Selection: LASSO logistic regression

The HPGENSELECT Procedure

CACASE=0

Step	Description	Effects In Model	Lambda	AIC	AICC	BIC	ASE	Validation AIC	Validation AICC	Validation BIC	Validation ASE
0	Initial Model	1	1	1917.197	1917.204	1921.520	0.233	820.935	820.952	824.403	0.238
1	newage entered	2	0.8	1916.144	1916.165	1924.789	0.232	822.841	822.892	829.777	0.238
2		2	0.64	1913.508	1913.529	1922.153	0.232	822.770	822.821	829.706	0.238
3		2	0.512	1911.107	1911.128	1919.752	0.232	822.674	822.726	829.610	0.238
4		2	0.4096	1908.842	1908.864	1917.487	0.231	822.536	822.588	829.472	0.238
5		2	0.3277	1906.665	1906.686	1915.310	0.231	822.355	822.406	829.291	0.238
6		2	0.2621	1904.554	1904.575	1913.199	0.231	822.136	822.187	829.072	0.238
7		2	0.2097	1902.504	1902.526	1911.149	0.230	821.890	821.942	828.826	0.238
8		2	0.1678	1900.517	1900.539	1909.162	0.230	821.628	821.679	828.564	0.237
9		2	0.1342	1898.598	1898.619	1907.243	0.230	821.358	821.410	828.294	0.237
10		2	0.1074	1896.751	1896.773	1905.396	0.229	821.090	821.141	828.026	0.237
11		2	0.0859	1894.982	1895.004	1903.627	0.229	820.829	820.880	827.765	0.237
12		2	0.0687	1893.294	1893.315	1901.939	0.229	820.581	820.632	827.517	0.237
13		2	0.055	1891.688	1891.710	1900.333	0.228	820.350	820.401	827.286	0.237
14		2	0.044	1890.158	1890.180	1898.803	0.228	820.136	820.187	827.072	0.237
15		2	0.0352	1888.712	1888.733	1897.357	0.228	819.943	819.994	826.879	0.237
16	BSI entered	3	0.0281	1888.972	1889.045	1906.262	0.227	822.875	823.047	836.747	0.236
17	UND entered	8	0.0225	1893.776	1894.550	1954.291	0.225	837.822	839.714	886.375	0.234
	DISLTC8 entered	8	0.0225	1893.776	1894.550	1954.291	0.225	837.822	839.714	886.375	0.234
	DIABETES entered	8	0.0225	1893.776	1894.550	1954.291	0.225	837.822	839.714	886.375	0.234
	RENAL entered	8	0.0225	1893.776	1894.550	1954.291	0.225	837.822	839.714	886.375	0.234
	mrsafinal entered	8	0.0225	1893.776	1894.550	1954.291	0.225	837.822	839.714	886.375	0.234
18		8	0.018	1867.880	1868.655	1928.396	0.221	828.627	830.518	877.179	0.231
19		8	0.0144	1837.579	1838.354	1898.095	0.216	817.718	819.610	866.271	0.227
20		8	0.0115	1806.241	1807.016	1866.757	0.211	806.375	808.267	854.928	0.223*

* Optimal Value of Criterion

Maximum Regularization Parameter	4.912624
Chosen Regularization Parameter	0.056639

The HPGENSELECT Procedure

Selected Model

CACASE=0

Selected Effects:	Intercept UND DISLTC8 DIABETES RENAL mrsafinal newage BSI

Dimensions	
Number of Effects	8
Number of Parameters	14
Columns in X	14

Fit Statistics		
	Training	Validation
-2 Log Likelihood	1778.24	778.38
AIC (smaller is better)	1806.24	806.38
AICC (smaller is better)	1807.02	808.27
BIC (smaller is better)	1866.76	854.93
Pearson Chi-Square	1353.02	596.07
Pearson Chi-Square/DF	2.4918	2.6972
Average Square Error	0.2111	0.2227

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	0.133321
UND 1	1	-0.053741
UND 0	1	0.041276
DISLTC8 1	1	0.032164
DISLTC8 0	1	-0.042794
DIABETES 1	1	0.051804
DIABETES 0	1	-0.068195
RENAL 1	1	0.060869
RENAL 0	1	-0.079084
mrsafinal 1	1	0.057472
mrsafinal 0	1	-0.074550
newage	1	0.012053
BSI 1	1	0.095228
BSI 0	1	-0.120818

The HPGENSELECT Procedure

CACASE=1

Model Information	
Data Source	WORK.SAMPLE
Response Variable	HOSPITAL
Weight Variable	WEIGHT
Class Parameterization	GLM
Distribution	Binary
Link Function	Logit
Optimization Technique	Nesterov

Selection Information	
Selection Method	LASSO
Stop Criterion	None
Choose Criterion	Validation ASE
Effect Hierarchy Enforced	None

Number of Observations			
Description	Total	Training	Validation
Number of Observations Read	853	598	255
Number of Observations Used	853	598	255

Response Profile				
Ordered Value	HOSPITAL	Total Frequency	Training	Validation
1	0	600	1632	693
2	1	253	424	186

You are modeling the probability that HOSPITAL='1'.

Class Level Information		
Class	Levels	Values
UND	2	1 0
DISLTC8	2	1 0
DIABETES	2	1 0
RENAL	2	1 0
SST	2	1 0
mrsafinal	2	1 0
BSI	2	1 0

The HPGENSELECT Procedure

CACASE=1

									Validation	Validation	Validation	Validation
Step	Description	Effects In Model	Lambda	AIC	AICC	BIC	ASE	Validation AIC	Validation AICC	Validation BIC	Validation ASE
0	Initial Model	1	1	2094.650	2094.656	2099.043	0.164	909.413	909.429	912.954	0.167
1	newage entered	2	0.8	2091.491	2091.511	2100.278	0.163	909.182	909.230	916.264	0.166
2		2	0.64	2087.349	2087.369	2096.136	0.163	907.457	907.504	914.539	0.166
3		2	0.512	2083.808	2083.828	2092.595	0.163	906.042	906.090	913.125	0.166
4		2	0.4096	2080.652	2080.672	2089.439	0.162	904.832	904.879	911.914	0.166
5		2	0.3277	2077.770	2077.790	2086.558	0.162	903.767	903.815	910.850	0.165
6		2	0.2621	2075.108	2075.128	2083.895	0.162	902.816	902.864	909.898	0.165
7		2	0.2097	2072.637	2072.658	2081.425	0.162	901.960	902.008	909.043	0.165
8		2	0.1678	2070.345	2070.365	2079.132	0.162	901.189	901.237	908.272	0.165
9		2	0.1342	2068.223	2068.243	2077.010	0.161	900.495	900.542	907.577	0.165
10		2	0.1074	2066.265	2066.285	2075.052	0.161	899.871	899.919	906.954	0.165
11		2	0.0859	2062.857	2062.877	2071.644	0.161	898.834	898.881	905.916	0.165
12		2	0.0687	2060.033	2060.053	2068.820	0.161	898.021	898.068	905.103	0.164
13		2	0.055	2057.697	2057.717	2066.484	0.161	897.390	897.438	904.472	0.164
14		2	0.044	2055.767	2055.787	2064.554	0.160	896.907	896.955	903.989	0.164
15		2	0.0352	2054.174	2054.194	2062.961	0.160	896.543	896.590	903.625	0.164
16	UND entered	4	0.0281	2052.191	2052.333	2078.552	0.160	901.142	901.481	922.390	0.164
	BSI entered	4	0.0281	2052.191	2052.333	2078.552	0.160	901.142	901.481	922.390	0.164
17	DIABETES entered	5	0.0225	2043.048	2043.292	2078.196	0.158	900.297	900.883	928.628	0.163
18	mrsafinal entered	6	0.018	2026.302	2026.676	2070.238	0.157	896.593	897.495	932.006	0.161
19	DISLTC8 entered	7	0.0144	2004.149	2004.683	2056.872	0.154	891.008	892.298	933.504	0.159
20		7	0.0115	1974.349	1974.882	2027.072	0.152	880.283	881.572	922.778	0.157*

* Optimal Value of Criterion

Maximum Regularization Parameter	4.852855
Chosen Regularization Parameter	0.05595

The HPGENSELECT Procedure

Selected Model

CACASE=1

Selected Effects:	Intercept UND DISLTC8 DIABETES mrsafinal newage BSI

Dimensions	
Number of Effects	7
Number of Parameters	12
Columns in X	12

Fit Statistics		
	Training	Validation
-2 Log Likelihood	1950.35	856.28
AIC (smaller is better)	1974.35	880.28
AICC (smaller is better)	1974.88	881.57
BIC (smaller is better)	2027.07	922.78
Pearson Chi-Square	1845.09	830.26
Pearson Chi-Square/DF	3.1486	3.4739
Average Square Error	0.1519	0.1573

Parameter Estimates		
Parameter	**DF**	**Estimate**
Intercept	1	-1.818762
UND 1	1	-0.145528
UND 0	1	0.125829
DISLTC8 1	1	0.006658
DISLTC8 0	1	-0.008017
DIABETES 1	1	0.051218
DIABETES 0	1	-0.060262
mrsafinal 1	1	0.024260
mrsafinal 0	1	-0.028861
newage	1	0.014599
BSI 1	1	0.087698
BSI 0	1	-0.101538

The HPGENSELECT Procedure

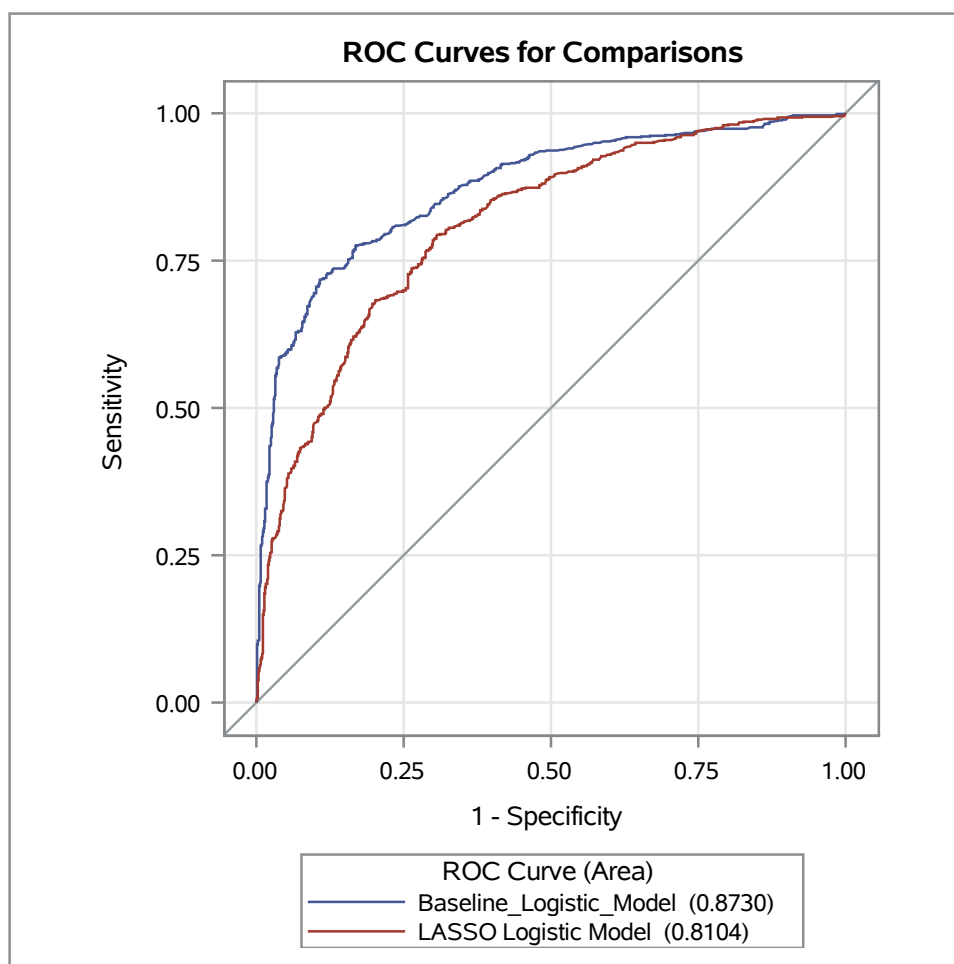| Procedure Task Timing		
Task	Seconds	Percent
Reading and Levelizing Data	0.00	0.51%
Candidate model fit	0.00	0.57%
Performing Model Selection	0.37	98.83%
Producing Output Data Set	0.00	0.09%

Predictors of Community-onset Infection: Training vs. Validation ASE from LASSO model
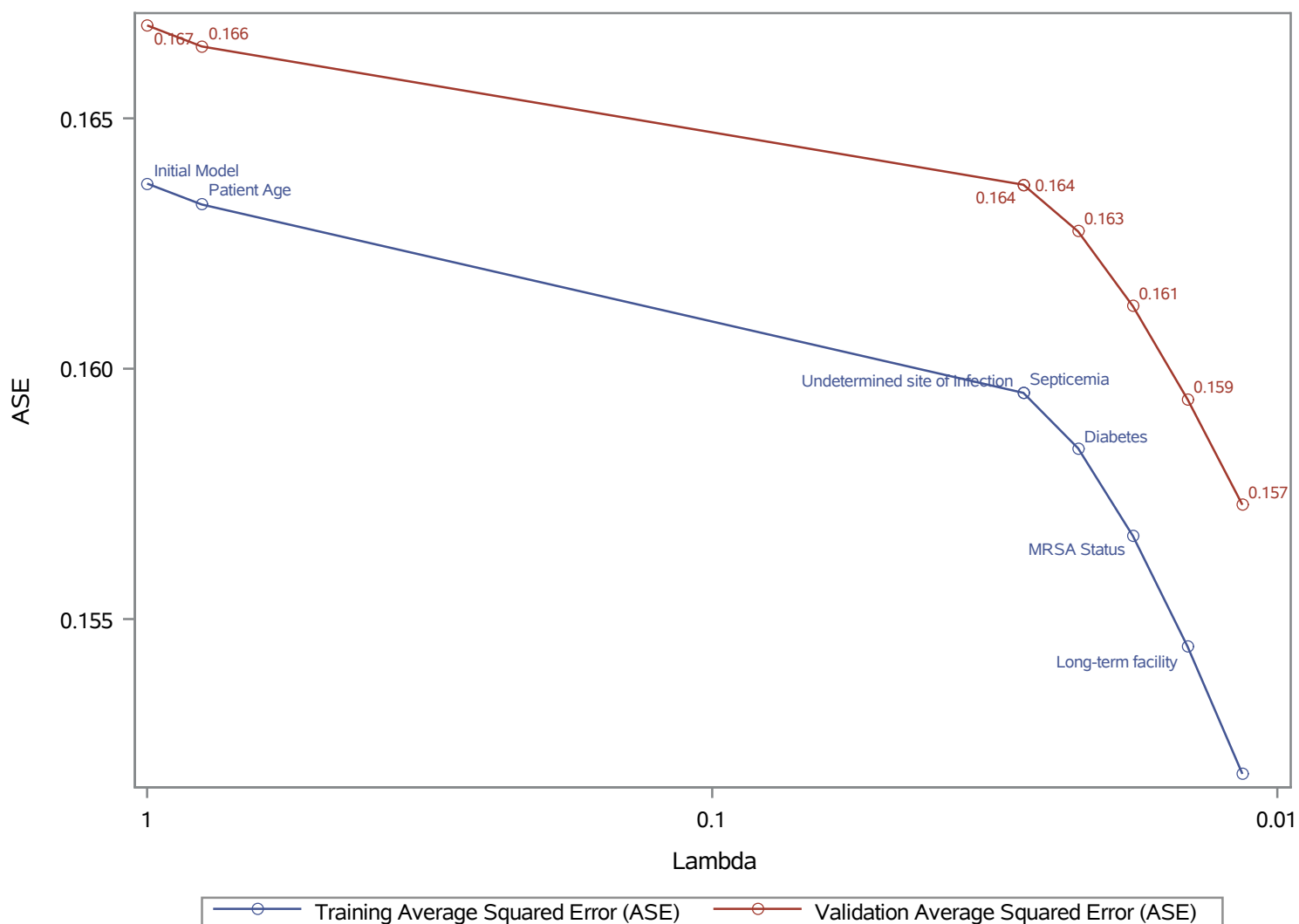
Predictors of Prior Healthcare Infection: Training vs. Validation ASE from LASSO model

The LOGISTIC Procedure

ROC Curves for Comparisons

ROC Curve (Area)
- Baseline_Logistic_Model (0.8730)
- LASSO Logistic Model (0.8104)
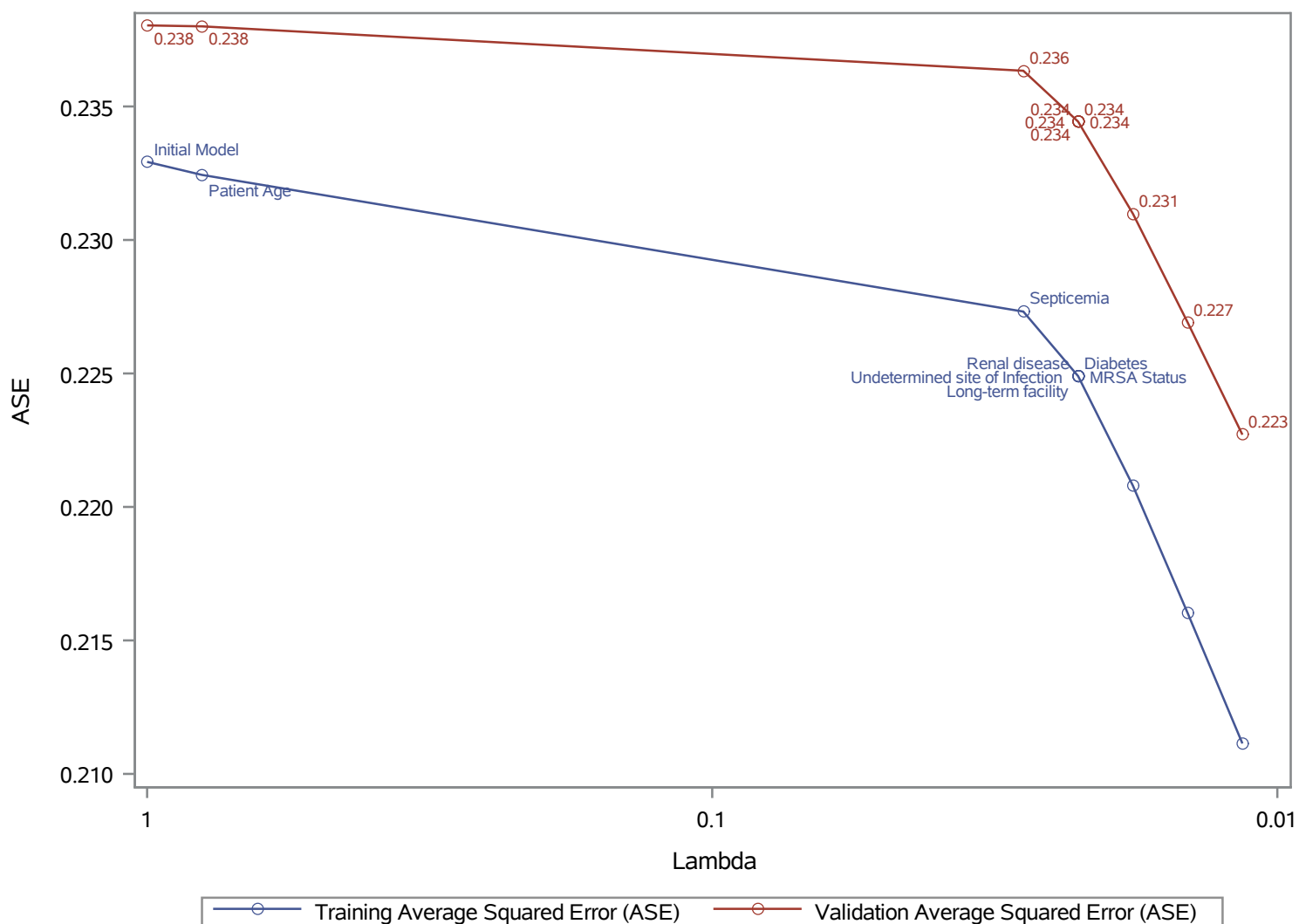
ROC Association Statistics							
		Mann-Whitney					
ROC Model	Area	Standard Error	95% Wald Confidence Limits		Somers' D	Gamma	Tau-a
Baseline_Logistic_Model	0.8730	0.00867	0.8560	0.8900	0.7459	0.7462	0.3730
LASSO Logistic Model	0.8104	0.0105	0.7899	0.8310	0.6209	0.6211	0.3105

The LOGISTIC Procedure

Model Information	
Data Set	WORK.TRAINING
Response Variable	HOSPITAL
Number of Response Levels	2
Weight Variable	WEIGHT
Model	binary logit
Optimization Technique	Newton-Raphson

Number of Observations Read	1155
Number of Observations Used	1155
Sum of Weights Read	3510
Sum of Weights Used	3510

Response Profile			
Ordered Value	HOSPITAL	Total Frequency	Total Weight
1	0	567	2169.0000
2	1	588	1341.0000

Probability modeled is HOSPITAL='1'.

Class Level Information			
Class	Value	Design Variables	
UND	1	1	0
	0	0	1
DISLTC8	1	1	0
	0	0	1
DIABETES	1	1	0
	0	0	1
RENAL	1	1	0
	0	0	1
SST	1	1	0
	0	0	1
mrsafinal	1	1	0
	0	0	1
BSI	1	1	0
	0	0	1

The LOGISTIC Procedure

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	4670.717	3330.289
SC	4675.769	3375.756
-2 Log L	4668.717	3312.289

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1356.4278	8	<.0001
Score	1147.9990	8	<.0001
Wald	685.2433	8	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
UND	1	93.9653	<.0001
DISLTC8	1	66.4630	<.0001
DIABETES	1	57.3384	<.0001
RENAL	1	59.3029	<.0001
SST	1	64.4613	<.0001
mrsafinal	1	81.6576	<.0001
newage	1	14.8212	0.0001
BSI	1	110.8282	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1101	0.1331	69.5132	<.0001
UND	1	1	-1.0933	0.1128	93.9653	<.0001
UND	0	0	0	.	.	.
DISLTC8	1	1	2.0516	0.2517	66.4630	<.0001
DISLTC8	0	0	0	.	.	.
DIABETES	1	1	0.8131	0.1074	57.3384	<.0001
DIABETES	0	0	0	.	.	.

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
RENAL	1	1	1.2007	0.1559	59.3029	<.0001
RENAL	0	0	0	.	.	.
SST	1	1	-0.7378	0.0919	64.4613	<.0001
SST	0	0	0	.	.	.
mrsafinal	1	1	0.8302	0.0919	81.6576	<.0001
mrsafinal	0	0	0	.	.	.
newage		1	0.00818	0.00212	14.8212	0.0001
BSI	1	1	2.7088	0.2573	110.8282	<.0001
BSI	0	0	0	.	.	.

Odds Ratio Estimates		
Effect	Point Estimate	95% Wald Confidence Limits
UND 1 vs 0	0.335	0.269 0.418
DISLTC8 1 vs 0	7.780	4.751 12.741
DIABETES 1 vs 0	2.255	1.827 2.783
RENAL 1 vs 0	3.322	2.448 4.510
SST 1 vs 0	0.478	0.399 0.573
mrsafinal 1 vs 0	2.294	1.916 2.746
newage	1.008	1.004 1.012
BSI 1 vs 0	15.012	9.066 24.857

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	87.8	Somers' D	0.756
Percent Discordant	12.2	Gamma	0.756
Percent Tied	0.0	Tau-a	0.378
Pairs	333396	c	0.878

The LOGISTIC Procedure

The LOGISTIC Procedure

ROC Curve for WORK.VALIDATE
Area Under the Curve = 0.8529
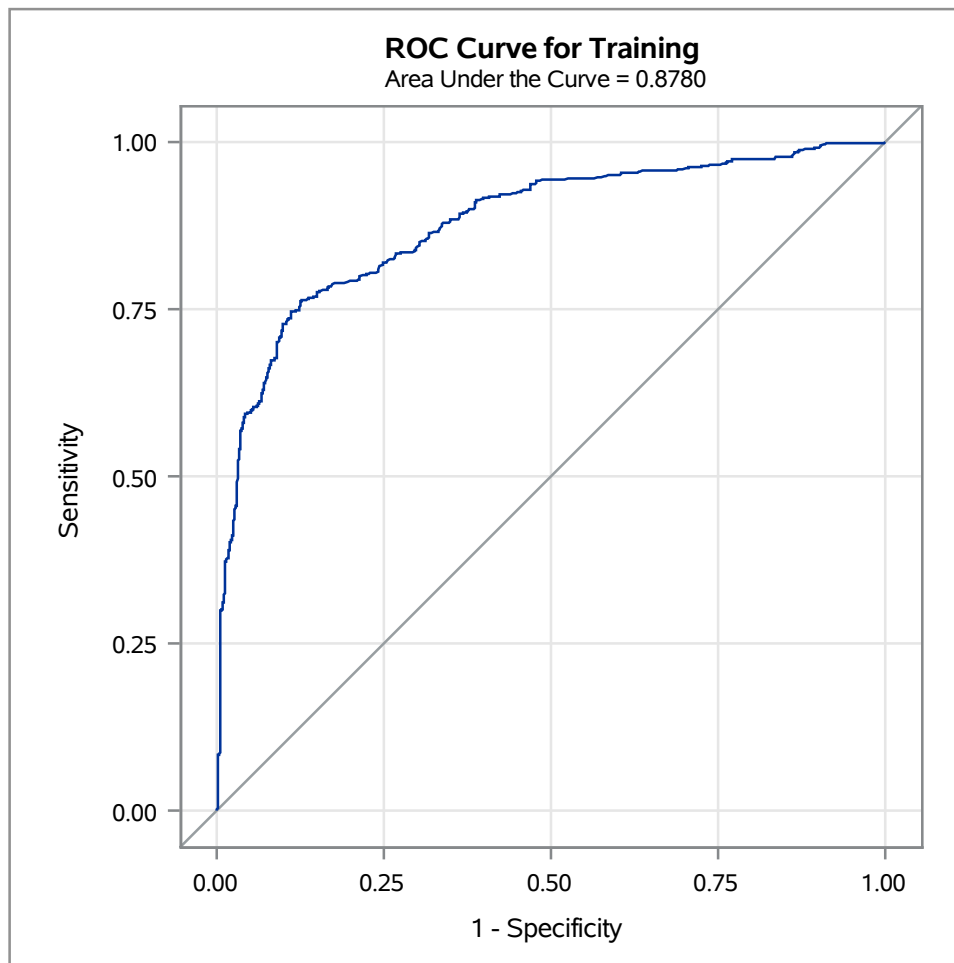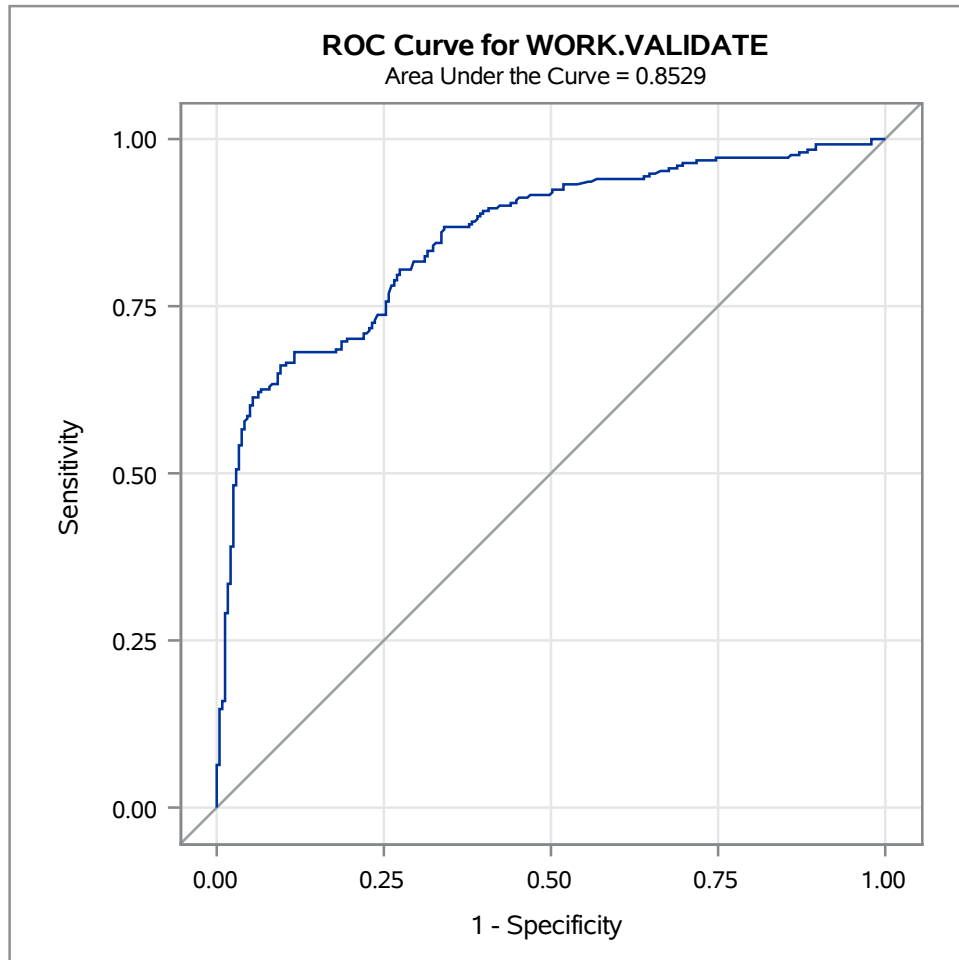
Weighted Logistic Model: Risk Factors of Hospitalization from S. aureus Infection
Odds Ratios with 95% Confidence Intervals

Risk Factor	aOR 95% CI
Septicemia	15.01 (9.07, 24.86)
Long-term facility	7.78 (4.75, 12.74)
Renal disease	3.32 (2.45, 4.51)
MRSA Status	2.29 (1.92, 2.75)
Diabetes	2.25 (1.83, 2.78)
Patient Age	1.01 (1.00, 1.01)
Skin-Soft Tissue	0.48 (0.40, 0.57)
Undetermined site of Infection	0.34 (0.27, 0.42)

Adjusted Odds Ratio